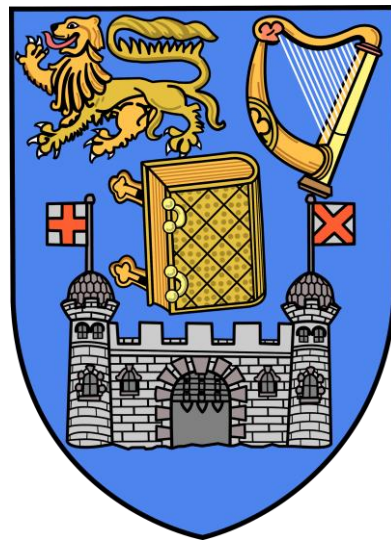


A Genomic Compendium of an Island

Documenting Continuity and Change across Irish Human Prehistory

Lara M. Cassidy



Smurfit Institute of Genetics
Trinity College Dublin

A thesis submitted for the degree of
Doctor of Philosophy

October 2017

Declaration and online access

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

Signed

Lara M. Cassidy

© 2017

Lara M. Cassidy

ALL RIGHTS RESERVED

For my father who gave me all his curiosity, my mother
for her unending support, and my sister, who never
ceased to make me laugh at myself.

‘Is maith an scéaláí an aimsir.’

Table of Contents

Acknowledgements	i
Summary	v
1. Introduction	1
Overview	1
A Brief Prehistory of Genetics	2
The Initial Genetic Scaffolding of Human Evolutionary History	3
What's in a Genome?	6
Detecting Human Population Structure in Genomic Data	7
Next Generation Sequencing and the Genomics Era	10
Ancient DNA: The Early Years	13
A Palaeogenomic Revolution	15
References	19
2. The Takings of Ireland: Punctuated population replacement followed by long term continuity on Europe's Atlantic edge	26
Overview	26
Introduction	27
Methods	38
Results	41
Conclusions	52
References	52
3. The First Arrivals: A genetic insight into Ireland's Mesolithic inhabitants	61
Overview	61
Introduction	62
Methods	70
Results	76
Conclusions	93
References	98
4. The Genomics of Megaliths: Origins and structure of Irish Neolithic societies	101
Overview	101
Introduction	102
Methods	113
Results	120
Conclusions	140
References	147
5. Bronze Age Beginnings: Signals of continuity across the Irish Metal Ages and the establishment of the Insular Atlantic Genome	150
Overview	150
Introduction	151
Methods	165
Results	172
Conclusions	197
References	201
6. Final Discussion	205
Appendix I: Archaeological Contexts and Sampling Information.	209
Appendix II: Molecular and Bioinformatic Methodology	273

Electronic Data Tables S1-S7 are available at

<https://docs.google.com/spreadsheets/d/1mk9pMMUbChzyW8CwVUYgokVL4iv83WBAKdIf3pWXJnw/edit?usp=sharing>

Acknowledgements

Firstly, I would like to thank my supervisor Dan Bradley for the opportunities, the encouragement, the patience, the knowledge and the pints. I grew up in your lab and I can't express my appreciation for all you've done for me over the past four years. Secondly, I'd like to thank Valeria Mattiangeli, without whom our lab would not run and this thesis would certainly not be done. You've guided me since my undergraduate, averted many a crisis, drilled bones with me from dawn to dusk, and always kept morale high. We don't deserve you. Rui Martiniano, you put up with me sitting opposite you for three years and that in itself needs acknowledgement. Thank you so much for your continual support, both emotional and bioinformatic, your time, which you gave so generously, and your contributions to this thesis, of which there are too many to name. You've been a true friend. I'd also like to thank Matthew Teasdale, who was there from the old goat beginnings, always interested, always ready with advice, and always always able to resuscitate the server. You are much missed in the lab.

Marta Verdugo and Victoria Mullin, you have been comrades and confidants on this very long road. I'm so glad we did this together and I'm so proud of us. I could not have chosen better travelling companions and I would have been lost without you. I wish you both every success wherever your paths take you next. Eppie Jones, from my first day you've been so kind and clear-headed. You would listen to my rambling worries about adapter trimming with such patience and provide amazingly well organised notes on any topic imaginable. Your advice in all matters has been truly invaluable. Russell McLaughlin, you have a real knack for making a person see the best in a situation and in themselves. Thank you for your unwavering encouragement.

Kevin Daly, thank you for the goat memes. They never disappointed and brightened many a dreary day. And thank you for listening to my many rants. I hope I can give you some of the support next summer that you've given me. Andrew Hare, you've be an excellent desk neighbour and such a calm presence. Thank you for putting up with my grouchy fourth year persona for so long. Mark Doherty, without whom this thesis would probably have been done a bit quicker, thank you for the chats, which helped keep me sane during those long nights in the lab. Ross Byrne, there was great comfort in having someone sitting opposite me who was also frantically writing about Irish genetics. Thank you for bouncing all those ideas around and thank you for the crash course in ChromoPainter. I couldn't have put together Chapter Five without you. Pierpaolo Maisano Delser, you are such an asset to our lab and it's been an absolute pleasure working with you this past year. Thank you for the always candid advice and support. Two honorary members of the Bradley Lab, Lucy Scott and Ciarán Campbell, deserve a special acknowledgement. You were both excellent students and your work on the Burren sites has contributed much to this thesis. I'd also like to give my wholehearted thanks to all other members of the Bradley and McLaughlin Labs, new and old. Working with you has felt like being part of a slightly unusual, but very caring, extended family and I'm so fond of you all.

This sentiment can be extended to the Smurfit Institute of Genetics as a whole. The warm and welcoming atmosphere was palpable from the moment I entered the building and is a testament to the people who work tirelessly to run it. I want to thank Brenda Campbell, David Sullivan, Sue Holahan, Paul McDermott and other members of the technical, administrative and research staff, without whom the whole place would obviously collapse. You are the pillars of the department and go above and beyond the call of duty every day without complaint to make sure all of our research runs smoothly.

There have been so many professors and lecturers who have inspired me over the years. My thesis committee, Aoife McLysaght and David McConnell, were mentors to me long before I began my PhD and I value so much the advice and encouragement you've always given me. Your passion for teaching and public engagement is infectious and has instilled in me and many others a deep love of biology. Ken Wolfe, my undergrad literature review supervisor, you were equally encouraging and gave me a huge amount of confidence in my writing skills. More than anything I want to acknowledge the late Mario Fares, whose death this year was a sudden and tremendous loss to us all. I'll always remember your course on protein evolution, where we sat in a circle and you encouraged us to think and discuss and debate. I came out with the firm belief that molecular chaperones were the most fascinating things to have ever existed and was actually excited to sit the exam, a possibly unique occurrence. You had an extraordinary gift and inspired so many of us with your warmth and passion. You will be sorely missed.

I want to say thank you to the other PhDs and postdocs of the department, many of whom were also fellow undergraduates. The solidarity and emotional support through the years has been invaluable and many good friendships have come from the shared trauma and the shared celebrations. The community spirit and goodwill has made the department feel like a second home and I want to wish you all the best of luck in whatever future endeavours you undertake.

I'd like to now turn my attention away from the geneticists and onto the archaeologists, without whom this thesis would be vanishingly thin. So many of you have given so much time and support to this project, not to mention samples, that it's hard to know where to begin. Ros Ó Maoldúin, you spent weeks digging through bone coffins with me, hunting high and low for petrous bones. The sampling strategy upon which this thesis is built owes much to your knowledge, as do the interpretations of results. Thank you for your constant encouragement and insight. Your enthusiasm is an absolute inspiration. Mary Cahill, you gave us the opportunity to carry out this project and for that I'm so grateful. None of this would have been achievable without your support, which you gave so generously. Thank you for your time and your trust, and for ferrying me back and forth across Dublin with boxes of bones. Maeve Sikora, you have also been beyond generous with your time and energy, no matter how little you had to spare. Your dedication to your work is truly inspiring and I appreciate so much your help and support. Thank you also to Eamonn Kelly, who preceded Mary and Maeve as Keeper of Irish Antiquities. I'd also like to thank the other NMI staff, Eamonn McLoughlin, Nessa O'Connor, Eimear Ashe and many more, who assisted with sampling and permission, and made the museum such a welcoming place during my

visits. Greer Ramsey, Mike Simms and the other NMNI staff, you were equally accommodating and went above and beyond to help with my sampling. Thank you as well to Marta Mirazon Lahr and Maggie Bellatti at the Duckworth Laboratory in Cambridge, who again went out of their way to help me collect samples.

Thomas Kador, you've been so enthusiastic and so supportive with this project and it's been a real pleasure collaborating with you. I'd like to thank both you and the other members of the 'Carrowkeel Team', Robert Hensey, Jonny Geber, Pdraig Meehan and Sam Moore, for showing me some of Neolithic Ireland and giving me a deeper appreciation and understanding of the archaeology. Carleton Jones, you and Ros also brought me on a megalithic walking-tour, this time around the Burren, an invaluable experience and a highly enjoyable one. You've been a wonderful collaborator and a real fount of knowledge. Thank you so much for your constant support over the years. Thank you to Eileen Murphy, whose enthusiasm helped get this project off the ground. Needless to say, without the hard work and dedication of you and our co-author Barrie Hartwell the foundations of this thesis would never have been formed. Ann Lynch, I appreciate so much you taking the time to help me sample Poulmabrone and your encouragement and interest in the project. Thank you as well to Edward Bourke, also at the Department of Culture, Heritage and the Gaeltacht, who helped with this sampling. Thank you also to Stephen Davis, Abigail Ash and James Eogan for your sampling contributions, support and enthusiasm. Elizabeth O'Brien also deserves special mention for providing some of the earliest samples of this project.

Jim Mallory, your contributions to this thesis, both direct and indirect, have been many. Thank you for providing the knowledge and insight we needed to write the first paper of this project. I read your book *Origins of the Irish* in my first months as a PhD, and it's been a cornerstone reference for me ever since. I'd also like to extend my appreciation to the wider archaeological community in Ireland, past and present. Over a century's worth of books, papers and excavations have formed the basis of this thesis and it is overwhelming to think of the number of individuals who have contributed in some way to the results presented here. I only hope it can do some small justice to their work and dedication. In particular, I would especially like to acknowledge the late Peter Woodman, who tragically passed away this year. Much of what is understood today about the Irish Mesolithic can be attributed to his passion and persistence and I am so grateful to have had the opportunity to learn from him. Chapter Three, which would not have been possible without his contributions, is dedicated to his memory.

Most importantly, I want to acknowledge my funding body, the Irish Research Council, and its staff, who have provided much more than monetary support to this project. Thank you so much for seeing the worth in my research proposal and providing me with the tools I needed to carry it through to the end. I also want to thank the Irish Centre for High-End Computing (ICHEC) for giving me access to their excellent resources and any technical support I required. Besides from my own project, I have also worked with many other research groups, inside and outside of Trinity, and want to thank them all for

the opportunities they have given me. This includes the Old Irish Goat Society; the Palaeogenetics Group in Mainz; Hannes Schroeder, Ashot Margaryan and other collaborators at the University of Copenhagen; the Campbell and McLaughlin Labs in Trinity; and the Carrowkeel Team. I also want to give special thanks to the Kitano Lab in the National Institute of Genetics in Mishima, who gave me my first taste of population genetics on a placement with the NIGINTERN program.

Finally, I'd like to thank my friends, my flatmates and my family members. I can't start naming you or I won't stop, but you know who you are. You've shared in my successes, mopped up a lot of tears, shown genuine interest as I tried to explain PCA plots (badly) and continued to feign interest after we moved onto genomic coverage. You've shown extraordinary empathy and have cheered me on every step of the way. You have forgiven me for my growing neglect the past year, the unanswered texts, the constant rescheduling of Skype calls and the unattended parties, despite all being madly busy yourselves. You have reminded me of my worth in a career path where imposter syndrome can creep into all aspects of life. You have kept me grounded and brought me back to the land of the living once in a while, turning the seemingly all-consuming frustrations and disappointments of ancient DNA research into laughable absurdities.

Writing these acknowledgements, it is impossible not to appreciate the huge collective effort that is found behind every human endeavour and seldom reflected in a linear list of authors. At its best scientific research should represent the pinnacle of societal rather than individual achievement, an ideal we all lose sight of from time to time as we compete to stay in a career we love. On that final collaborative note, I feel my last thanks should go to the 140 Irish humans who contributed their remains to this study, and with whom I've probably spent more days than any living person over the past four years. Studying the past makes you appreciate the present, specifically how brief a human life is. Time is the most precious of all commodities and for that reason I'd like to once again thank from the bottom of my heart everyone, both on and off this list, who has contributed some of theirs to me.

Summary

The thesis submitted here concerns the palaeogenomic analysis of 140 ancient individuals from all periods of Irish prehistory, with a view to providing a working demographic framework for the entirety of the island's human occupation. This was achieved through the use of Illumina next generation sequencing (NGS) technology, which when combined with skeletal sampling of the petrous temporal bone gives unprecedented access to the surviving endogenous DNA present in archaeological remains. The 93 successful samples were sequenced to an average of 1X coverage, and data was processed following standard NGS pipelines adapted for aDNA research. Diploid genotype calls were imputed for all samples and utilised alongside pseudo-haploid calls for population genetic analyses.

Chapter Two creates an initial demographic scaffold for Irish prehistory based on this dataset, established with respect to the larger palaeogenomic narrative that has emerged for the European continent. ADMIXTURE and principal component analysis identify three ancestrally distinct Irish populations, whose inhabitation of the island corresponds closely to the Mesolithic, Neolithic and Chalcolithic/Early Bronze Age eras, with large scale migration to the island implied during the transitional periods. Haplotype-based sharing methods and Y chromosome analysis demonstrate strong continuity between the Early Bronze Age and modern Irish populations, suggesting no substantial population replacement has occurred on the island since this point in time. Chapters Three, Four and Five respectively provide more detailed analysis of the Mesolithic, Neolithic and Chalcolithic to Iron Age periods.

Chapter Three uses D - and f -statistics to demonstrate high shared genetic drift between Irish hunter-gatherers and contemporaries from France and Luxembourg. Allelic affinities further suggest that these northwestern hunter-gatherer populations find their origins in more eastern glacial refugia, such as Italy, rather than Iberia. Runs of Homozygosity (ROH) analysis demonstrate the Irish population underwent a severe inbreeding bottleneck, indicating some level of demographic isolation occurred after initial colonisation of the island. Phenotypic and polygenic trait analyses were also carried out, revealing the individuals studied to be dark-skinned and blue-eyed, with relatively inflated estimates of genomic height.

Chapter Four utilises both allelic and haplotype-sharing methods to establish substantial contributions from both Mediterranean farming groups, whose origins lie in Anatolia, and northwestern hunter-gatherers to the Neolithic Irish population. Moreover, evidence for local Mesolithic survival and introgression in southwestern Ireland, long after the commencement of the Neolithic, is also implied in haplotype-analysis. Societal complexity during the Neolithic is suggested in patterns of Y chromosome and autosomal structure, while the identification of a highly inbred individual through ROH analysis, retrieved from an elite burial context, strongly suggests that the elaboration and expansion of megalithic monuments over the course of the Neolithic was accompanied in some regions by dynastic hierarchies.

Chapter Five addresses the nature of the Chalcolithic and Early Bronze Age transitions in Ireland. Haplotypic affinities and distributions of steppe-related introgression among samples suggest a potentially bimodal introduction of Beaker culture to the island from both Atlantic and northern European sources, with southwestern individuals showing inflated levels of Neolithic ancestry relative to individualised burials from the north and east. Signals of genetic continuity and change after this initial establishment of the Irish population are also explored, with haplotypic diversification evident between both the Bronze Age and Iron Age, and the Iron Age and present day. Across these intervals selection pressures related to nutrition appear to have acted, with variants involved in lactase persistence and skin depigmentation showing steady increases in frequency through time.

1. Introduction

Overview

This introduction provides a summary of the strands of genetic research that have been gradually woven together over the past century to make possible the thesis on ancient Irish genomics presented here. Progress can be bracketed into four main areas, with much overlap in between.

1. The crucial advances made in molecular biology that allowed the material of inheritance, DNA, to be extracted, isolated, characterised, manipulated, amplified and eventually sequenced at high efficiency, unlocking the wealth of genetic variation hidden within organisms.
2. The development of statistical procedures with which to visualise and describe the distribution of this variation among populations, and the construction of models which could explain how such patterns emerge.
3. The rapid improvements in robotics and information technology over the past several decades, which have provided the means to produce, store and edit huge quantities of genetic data, allowing these phylogenetic and population genetic analyses to be applied on a scale hitherto unimaginable.
4. The tailored application of the above methodologies to the study of human evolutionary and demographic history, which could inform and in turn be informed by developments in other fields related to our species' past, including archaeology, linguistics and anthropology.

As key progressions within these four different areas are deeply entwined with one another, they will not be discussed in separate sections, but instead presented together in a chronological fashion. The final sections will then consider the impact these advances have collectively had on the study of ancient DNA (aDNA), a niche field, which in recent years has been transformed into a core pillar of human evolutionary and population genetic research.

A Brief Prehistory of Genetics

Genetics is, in its essence, the science of inheritance, a concept deeply intertwined with the study of human history and identity. The field itself has collectively enthralled over a century's worth of researchers dedicated to demystifying the origins of human populations. Indeed, these efforts had begun long before the establishment of what we know today as the modern field of molecular genetics, which can perhaps be dated to the identification of DNA as the hereditary material (Avery et al. 1944) and the subsequent decoding of its structure (Watson & Crick 1953). It was many decades beforehand, at the start of the 20th century, that the rediscovery of Mendelian genetics had ignited heavy debate among those attempting to ground the fledgling field of evolutionary biology within a practical explanatory framework. This need to somehow reconcile the work of Mendel and Darwin was addressed through the development of mathematical models, built on statistical reasoning, which would go on to form the basis of modern population genetics.

Building on Mendel's principles, the giants of this emerging field, Fisher, Haldane and Wright, identified four key phenomena - mutation, drift, selection and migration - by which the genetic variation of a population could be shaped and maintained, providing the fodder needed for adaptation and evolution to occur (Hartl & Clark 1997). According to their models, reproductive isolation between populations would lead to genetic divergence and substructure, and admixture to homogenisation, detectable by comparison of observed allelic frequencies to those expected under Hardy-Weinberg equilibrium, described by a set of statistics known as Wright's fixation indices. Given that these processes through which populations diverge are implicitly dependent on both generation time and population size, the potential of these models as a vehicle to study both the deeper evolutionary history and more recent demography of species was clear.

However, progress was inhibited by the elusive nature of the molecule of inheritance itself. Researchers were restricted to investigating genetic variation indirectly, through its phenotypic effects. In humans, one of most famously studied traits was blood group (Landsteiner 1901; Bernstein 1924). Indeed, the demonstration that blood type frequencies varied greatly from region to region, with distinct geographical trends (Hirszfeld & Hirszfeld 1919), marked the beginning of the application of genetics to the study of human history. By the late 1940s other classical markers, such as enzyme polymorphisms and blood serum proteins, had been identified and were used to establish genetic relationships between populations based on differences in allele frequencies. Previous anthropological categories of discrete human races were dismantled, as it became clear that it was not only genetic isolation that had played a significant role in the shaping of modern human populations, but also admixture driven by migration and demic diffusion.

A key proponent of this view was Cavalli-Sforza, whose seminal work, built on decades of research on classical genetic markers in global populations, demonstrated human variation was a series of clines (Cavalli-Sforza et al. 1994), the proposed result of these successive mixing events. The work involved

pioneering usage of principal component analysis (PCA), a statistical method used to deconstruct highly dimensional data into linear components in order to explore overarching trends in variation over large numbers of markers. Explanations for the many gradients of human variation described by these PCs were sought in archaeological and linguistic phenomena. In the 1970s, he proposed that Europeans were in part descended from West Asian farming populations who diffused into the region during the Neolithic, mixing with Mesolithic groups, setting up a southeast to northwest gradient of variation (Ammerman & Cavalli-Sforza 1984; Sokal et al. 1991). This was in turn linked to the Anatolian Hypothesis of Indo-European language spread (Renfrew 1990), going against the grain of anti-migrationist archaeological thought at the time (Zvelebil & Zvelebil 1988). Other clines in European variation were attributed to separate demographic events, such as that described by the third principal component, which peaked in populations of the Pontic Steppe and was proposed to represent an alternative or additional spread of Indo-European language into Europe through the pastoralist Kurgan culture (Piazza et al. 1995). The potential power of statistics to elucidate population relationships when applied to large numbers of genetic markers was becoming clear.

However, several key developments in molecular genetics were required before such methods could be applied to the vast bank of variation present in the human genome. Even as the field progressed and direct detection of variation in the DNA itself became possible, the majority of early research focused on the non-recombining mitochondrial genome (mtDNA), and later the Y chromosome. The use of such singular markers in the study of human prehistory moved focus from population genetic to phylogenetic methods, a situation not fully rectified until after the publication of the human genome (Lander et al. 2001). That being said, human evolutionary biology greatly benefited from the construction and fine-tuning of such phylogenies, which succeeded in sketching a broad picture of the migrations undertaken by *Homo sapiens* since their emergence in Africa (Underhill & Kivisild 2007).

The Initial Genetic Scaffolding of Human Evolutionary History

With the publication of its structure in 1953, molecular biologists had soon turned their attention to detecting variation within DNA itself. This proved to be a more arduous task than earlier work on proteins, given the long and chemically monotonous nature of the molecule. However, it was soon seen to be relatively simple to segregate DNA molecules based on their length. The discovery of restriction enzymes (Danna & Nathans 1971) allowed researchers to make use of this fact in the detection of genetic variation through restriction fragment length polymorphism analysis (RFLP). The compact and easily purified mtDNA was the obvious target for these early studies, which soon demonstrated the organelle's fast mutation rate (Brown et al. 1979) and characteristic maternal inheritance (Hutchison et al. 1974; Giles et al. 1980), precluding recombination.

These traits allowed for the creation of phylogenies at much shallower time depths than was possible using differences in amino acid sequences, making the mtDNA ideal for the study of recent human evolution. Advances in tree building algorithms (Saitou & Nei 1987) and the application of the molecular

clock technique (Zuckerkandl & Pauling 1962) resulted in the construction of a global maternal phylogeny for humans, which demonstrated the ancestor of all human mtDNA lineages originated in Africa (Cann et al. 1987), a theory supported by Darwin over 100 years earlier (Darwin 1871). Moreover, the most recent maternal ancestor of all humans was estimated to have lived as late as 140,000 to 200,000 years ago, effectively disproving the ‘Candelabra’ hypothesis of independent parallel evolution of *Homo sapiens* from separate *Homo erectus* groups, an issue the fossil record had been unable to resolve (Xinzhi 1981).

The Out-of-Africa (OoA) model became the most popular theory of human origins, emphasising recent expansion of *Homo sapiens* from Africa with little or no admixture between the newcomers and older Eurasian species of *Homo*. Tracing and timing the subsequent migrations of humans across the continents became a key focus of mitochondrial studies (Torroni et al. 1993; Richards et al. 1996; Watson et al. 1996). The potential of other loci for RFLP analysis was also explored at this time using hybridisation probes (Southern 1975), including the Y chromosome, the largest non-recombining block in the genome (Casanova et al. 1985; Lucotte & Ngo 1985; Jakubiczka et al. 1989). However, variant discovery was inefficient, with those mutations that were discovered of limited use for phylogenetic purposes (Jobling & Tyler-Smith 1995). The mtDNA remained the dominant marker, though an upper-limit was gradually being reached in terms of the resolution available from RFLP comparisons.

Direct inference of the exact base pair sequence of a DNA molecule was the obvious next step for studies of genetic variation. Concerted efforts towards this goal cumulated in the invention of ‘Sanger Sequencing’ in 1977 (Sanger et al. 1977), which remained the most widely used method of DNA sequencing for almost 30 years. The technique, like RFLP, also made use of DNA fragment length patterns, which were created through the interruption of DNA replication *in vitro* with specific A, C, G and T termination nucleotides, allowing for the detection of specific base pairs at known points in the sequence. While initially carried out in four separate reactions, the development of fluorescent dye labeling allowed their combination, bringing increased speed, efficacy, and eventually automation to the process.

The sequencing process required large, purified quantities of the target DNA fragment and this was initially achieved through bacterial cloning, made possible through newly developed recombinant DNA technology, which took advantage of exponential bacterial propagation. However, this was a long and cumbersome process. Given the amount of time and money required to sequence relatively small lengths of DNA, the ability to isolate the exact DNA fragments of interest was crucial. Artificially produced DNA primers were developed to initiate replication, and consequently sequencing, at specific sites of interest, mimicking the *in vivo* process. However, this could only be successful if the targeted region had been taken up at the cloning stage, an event dependent on random chance. Work progressed slowly, with research first focusing on the small genomes of viruses (Fiers et al. 1978), before tackling the slightly larger genomes of eukaryotic organelles, including the human mitochondrion (Anderson et al. 1981).

However, despite the early inference of its complete sequence, large-scale surveys of mitochondrial sequence information remained entirely unfeasible without effective targeting techniques.

The milestone development of the polymerase chain reaction (PCR) in 1983 (Mullis et al. 1987) offered an excellent way to combine the two protracted preparation steps for Sanger sequencing, targeting and amplification, into a single rapid one, revolutionising the field of genetics in the process. Through the use of a pair of primers, rather than a single one, a runaway amplification of a specific genomic region could be set up, using thermocycling to initiate multiple rounds of replication. This resulted in such exponentially high concentrations of the targeted fragment relative to other genomic material that it could be considered in effect purified. While sequencing was still an expensive process, the workload had now massively decreased. The potential applications of PCR in the discovery and assaying of genetic variation were immense, both through direct sequencing, as well as indirect methods such as tandem repeat size separation and selective allelic amplification.

Y Chromosome studies gradually rose to prominence, offering a view of male lineage history, complementary to mtDNA research (Hammer 1995; Jobling & Tyler-Smith 1995; Jobling & Tyler-Smith 2003). This was achieved in part by the abundance of satellite DNA discovered on the chromosome, variation in which could be easily detected with new PCR methods. Mitochondrial sequencing for large cohorts of individuals was now also an achievable vista, with many studies focusing on the highly variable D-loop control region. Together, both markers provided a broad map of the routes and timings of early human migrations (Wells et al. 2001; Underhill & Kivisild 2007), as well as insight into the impact of more recent historical events such as the Viking migrations, the Arab and Trans-Atlantic slave trades, Jewish diaspora and Mongolian invasions (Richards et al. 2003; Zerjal et al. 2003; Salas et al. 2005; Behar et al. 2006; McEvoy et al. 2006).

However, while the non-recombining nature of both markers had provided highly accurate genealogical reconstruction of both male and female lineage history, this trait also rendered each in effect a single genetic locus. Thus they could only ever provide a small portion of the full genealogical compendium available from the human genome. Moreover, given the highly stochastic process of coalescence, estimated divergence timings for populations based on such singular phylogenies were viewed with caution (Novembre & Ramachandran 2011). For these reasons, much genetic research into population history had continued to rely on protein markers, which, though relatively few in number (several hundred at most), provided a more nuanced picture of human population structure. This was soon to change however, as rapidly advancing sequencing technology, spurred on by the invention of PCR, was put to work on one of the central challenges of modern biology: the elucidation of the entire human genome.

What's in a Genome?

Medical geneticists had long been preoccupied with the cataloguing of autosomal genome variation as a means to identify causative disease alleles which, through the aforementioned advances in molecular techniques, had culminated in the human genome project (HGP); biology's largest public collaborative effort to date. A draft sequence of the complete human genome was published in 2001 (Lander et al. 2001) and the project declared complete in 2003. It had taken 13 years and cost approximately 16 to 30 cents per base pair, roughly 3 billion positions in total. The project had built on many decades worth of genetic and physical mapping of the human genome, with the key goal of locating disease genes. Genetic maps, based on recombination frequencies, provided some order to the genome through the identification of linkage groups, which violated Mendel's law of independent segregation. These were based first on inheritance pedigrees of phenotypic traits, and later on direct markers, such as RFLPs and microsatellites (Botstein et al. 1980). Such linkage maps could be anchored onto physical scaffolds of the chromosomes, constructed using both cytogenetic and sequence mapping techniques. The latter involved the systematic arrangement of recombinant clone overlaps, achievable through the use of both restriction fragment fingerprints and uniquely mapped sequence-tagged sites (Olson et al. 1989). Over the course of the HGP these methods collectively produced mappable contigs of BAC clones, which were subsequently fragmented, shotgun sequenced using Sanger technology, and ordered using developing bioinformatic techniques, eventually culminating in the full sequence of the human genome.

Throughout the HGP, identification of human genetic variation remained a central focus, with single nucleotide polymorphisms (SNPs), the most common type of genetic variant, becoming key targets. Primers for many newly identified sequenced-tagged sites (Hudson et al. 1995) were made available for use in resequencing projects, paving the way for large scale SNP identification (Wang et al. 1998), while the mosaic, not to mention diploid, nature of the first human genome, also unearthed some 800,000 SNPs within its overlapping sequences (Kwok & Chen 2003). Whole genome shotgun sequencing of individuals, followed by comparison to the newly published reference sequence, soon became the most efficient method of variant discovery, and these projects were spurred on by parallel improvements in both sequencing technology and bioinformatics tools developed to handle increasingly large amounts of data. By the time of the human genome draft publication, more than 1.4 million SNPs had been identified (Sachidanandam et al. 2001).

However, a full understanding of human genome diversity and its role in disease, required not only efficient methods of SNP discovery, but also accurate and inexpensive techniques for genotyping vast numbers of known SNPs in large cohorts of individuals. These needs were met through the development of microarray technology (Gershon 2002), based on older DNA hybridisation techniques for detecting specific sequence motifs, which were now adapted through fluorescence microscopy and solid surface DNA capture for simultaneous genotype inference across thousands of variant sites (LaFramboise 2009). It was by methods such as these that common variation in human populations was catalogued, which could then go on to further inform the design of commercial arrays. These early efforts were guided by

the International HapMap Project, which succeeded in producing a comprehensive database of SNPs that captured the common patterns of haplotypic variation across human populations (The International HapMap Consortium 2005). In doing so, knowledge of linkage disequilibrium (LD) patterns and recombination hotspots across the genome greatly increased, revealing that a few hundred thousand well-chosen tag SNPs were all that was required for robust assays of variation at the genome-scale. Such a resource proved essential in the identification of genes involved in complex traits and disease through genome-wide association studies (GWAS), which the ever decreasing cost of SNP arrays made all the more feasible (McCarthy et al. 2008).

While the key impetus for these early genomic studies remained medical in nature, an obvious side application for the plethora of newly uncovered autosomal variation was the understanding of human population structure and history, which in itself could help inform studies of disease variation. However, the International HapMap Project, while aiming to sample the vast majority of common human genetic variation, had chosen to survey large numbers of individuals belonging to a small number of diverse groups living in easily accessible urban centers. It was somewhat the reverse that would prove effective for studies of human evolutionary history, namely the genotyping of small numbers of individuals belonging to a large variety of indigenous populations from across the globe. This mantle was taken up early on by the Human Genome Diversity Project (HGDP) (Cann et al. 2002), successfully culminating in the HGDP-CEPH cell line panel, sampled from over 1,000 individuals from 52 diverse populations.

Detecting Human Population Structure in Genomic Data

The first genome-wide population genetic analysis of the HGDP dataset was based on several hundred microsatellite loci (Rosenberg et al. 2002). This revealed the vast majority of human variation lay within, rather than between, populations, confirming much earlier work on classical markers (Boyd 1950; Lewontin 1972), a testament to the recent shared history and small effective population size of all modern humans. Nonetheless, genetic structure was still detectable based on cumulative allele frequency differences across many loci, and subpopulations could be identified correlating strongly with broad geographical regions and linguistic groupings. This level of fine-scale genetic differentiation was visualised through the use of a novel model-based clustering algorithm, implemented by the program STRUCTURE (Pritchard et al. 2000), which was designed to identify a discrete predefined number of populations based on allele frequencies, with which each individual's ancestry could then be described. Crucially, the model allowed for admixed individuals, who possess proportions of ancestry from multiple distinct sources.

The identification of such distinguishable geographical clusters, corresponding closely with continental groupings, found support in previous phylogenetic studies of autosomal markers (Bowcock et al. 1994), but somewhat contradicted observations that allele frequencies formed gradients of continuous variation across geographical space (Cavalli-Sforza et al. 1994; Serre & Pääbo 2004). These two apparently opposing perspectives were reconciled by way of geographical barriers to gene flow, such as the Sahara

Desert or Himalayas, which could cause sharp discontinuities in the typical gradients expected under a simple isolation-by-distance model (Rosenberg et al. 2005). Moreover, populations inhabiting these boundary regions tended to present with divergent ancestries from both sides of the divide. In this way, both clines and clusters were required to fully describe the global patterns of human genetic variation, though it was still somewhat beyond the scope of research to explain exactly when and how these patterns had formed. For this reason, a cautious approach was taken in making any inferences on human prehistory or ancestral population groups, based purely on observable modern structure alone.

Other computational methods for describing human population structure using large multi-loci datasets were also developed (Novembre & Ramachandran 2011). Aside from the above admixture modelling, multi-dimensional summary statistics provided the main mode of data visualisation for fledgling genome-wide studies. This strategy had first been implemented decades earlier, in PCAs of classical markers. However, while these foundational works had been based on population level allele frequencies and visualised through geographical maps, the new methods were now adapted to dense genotype data on the individual-level (Patterson et al. 2006; Price et al. 2006). PCA proved less computationally intensive than model-based techniques, which contributed to the method's growing popularity as the density of SNP arrays increased. To address these limitations, admixture modelling, though still exponentially slower than PCA, has been optimised in more recent years in the program ADMIXTURE (Alexander et al. 2009), and the complementary methods have become staples of population genomics research.

Plotting of principal components provided a less rigid way to visualise relationships between individuals compared to the predefined nature of admixture modelling, allowing for the identification of both discrete clusters, as well as more clinal forms of structure. This ability was demonstrated most strikingly in the clear mirroring of genes with geography in PCA plots of European variation (Novembre et al. 2008), in which definite, though overlapping, regional clusters could be distinguished, despite the overall homogeneity of European populations. This was seen as an elegant confirmation that, in the absence of geographical barriers to gene flow, genetic divergence strongly correlated with distance, with the main components of variation representing approximate perpendicular axes of geographical space. Moreover, it was emphasised that no long-distance migratory expansions or diffusions were required for such gradients to be set up over time, as they could be adequately modelled as the result of a constant homogeneous short-range migration process (Novembre & Stephens 2008), calling into question the archaeological and linguistic interpretations drawn from the earlier PCAs of classical markers. In truth, both interpretations could be seen as equally valid, given there was no adequate way to distinguish whether a cline was the result of a single homogenous population gradually diverging over time through isolation-by-distance, or the reverse, namely multiple divergent populations gradually homogenising over time through migration and admixture.

To address the issues of interpretation that plagued both PCA and ADMIXTURE analyses, formal statistical tests of admixture were developed, which could be used to test multiple possible demographic

histories and build up a picture of population relationships that fit the observed genetic data (Reich et al. 2009; Patterson et al. 2012). These relied on estimations of shared genetic drift between populations of interest, summarised using newly defined f -statistics and D -statistics (based on squared allele frequency differences). The statistics corresponded to the traditional notion of branch length on a phylogeny, the most widely utilised being the shared branch length of two populations with a third (three-population test), or a pair of populations with another pair (four-population or ABBA BABA test), which was used to demonstrate Neanderthal introgression into modern Eurasian populations (Green et al. 2010). Importantly, in traditional phylogenies only one path exists across the tree between any two populations, while if past admixture events have occurred multiple pathways will exist. Using this core concept, it is possible to test whether observed patterns of shared drift violated a predefined simple population phylogeny. If so, an admixture event could be assumed to have taken place. The degree of violation could in turn inform on the magnitude of such an event. Moreover, the four-population test offered some insight into the directionality of gene flow. Notably, as these tests require prespecified populations as input, some previous inference of population structure and clustering is preferable, usually gleaned from the above described PCA and model-based clustering analyses. Moreover, some prior knowledge of population relationships, if possible, can help to inform the construction and interpretation of test phylogenies, such as known outgroups. Taken altogether, these different approaches provide a powerful suite of tools with which to interrogate genomic datasets and have provided the core framework for the majority of palaeogenomic papers to date, given their robustness in dealing with the pseudo-haploid genotype calls generated from low-coverage data (Gamba et al. 2014; Lazaridis et al. 2014; Skoglund et al. 2014a; Allentoft et al. 2015; Haak et al. 2015).

Other methods used to explore population structure and admixture, based on haplotypic sharing, specifically require phased diploid calls (Li & Stephens 2003; Price et al. 2009; Lawson et al. 2012). These likelihood-based models are complex and computationally challenging, but, by harnessing the wealth of information hidden within patterns of linkage disequilibrium (LD) across the genome, can provide higher resolution of subtle population structure than that achievable using unlinked methods. Such approaches have been recently applied to large samples from relatively homogenous populations with great success (Leslie et al. 2015; Byrne et al. submitted). Given that the rate of LD decay, which is driven by recombination and mutation, is dependent on generation time, the size of shared haplotypic chunks can also be used to date demographic events, such as episodes of admixture (Hellenthal et al. 2014).

Diploid data also allows the identification of runs of homozygosity (ROH) within individual genomes, the numbers and length of which can vary widely between populations due to both recent and ancient inbreeding events. Levels of smaller ROH tend to increase with distance from Africa, while larger ROH indicate recent inbreeding (Kirin et al. 2010; Pemberton et al. 2012). However, dense whole genome data is required to fit these patterns of homozygosity to precise models of past ancestral population sizes (Li & Durbin 2011; MacLeod et al. 2013). Whole genome sequences provide not only exponentially more information than SNP array data, including the identification of rare variation, but also avoid

ascertainment bias, which can confound results depending on the populations in which SNP discovery took place (Albrechtsen et al. 2010). Model-based analyses, used to estimate demographic parameters such as divergence times, migration rates and population size, are particularly vulnerable to ascertainment as they require completely unbiased observations of the site-frequency spectrum (Novembre & Ramachandran 2011). Some successful demographic inference has been achieved based on SNP array data, through either the use of haplotypic methods (Lohmueller et al. 2009; Reich et al. 2009; Hellenthal et al. 2014), noted above, which are somewhat more robust to ascertainment bias, or by incorporating specific ascertainment parameters into the demographic model itself (Wollstein et al. 2010). However, for a full understanding of human demographic history it has become apparent that diverse whole genome sequences are a non-negotiable requirement.

Next Generation Sequencing and the Genomics Era

As population geneticists were reaching the upper limit of what could be inferred through genotype data alone, medical geneticists were also coming to the conclusion that rare variation played a central role in many common human diseases (Manolio et al. 2009; Cirulli & Goldstein 2010). Whole genome sequencing of large numbers of individuals would be necessary to capture such diversity, a hugely ambitious task that would require the development of high-throughput, inexpensive sequencing technology, as well as efficient bioinformatic tools and algorithms needed to put order on such large amounts of sequence data.

These needs were met with the invention of a number of new next-generation sequencing (NGS) technologies in the mid-2000s (reviewed in Metzker 2010; Goodwin et al. 2016; Mardis 2017), including Roche 454, SOLiD, Helicos, and Illumina, the latter currently the most popular for palaeogenomic studies. By miniaturisation and mass parallelisation of sequencing reactions these technologies allowed the harvesting of hundreds of millions of short DNA sequences over a relatively short period of time. While the exact chemistry of the sequencing reactions vary from platform to platform, they all tend to work on a sequencing-by-synthesis basis, with nucleotide incorporation digitally detected in real-time across a lawn of DNA templates (typically clonally amplified) anchored on a solid surface or ‘flow cell’. NGS library preparation also proved a much speedier affair relative to traditional Sanger shotgun sequencing approaches, which required the cultivation of DNA libraries in microbial cultures. NGS methods instead make use of artificial adapter molecules, which are ligated to the ends of fragmented DNA from the target source. PCR primers, complementary to these universal adapters, can then be used to amplify the entire sequencing library in a single efficient step. The input DNA for library creation can come from diverse sources, including unmodified genomic DNA, pooled PCR products, DNA released after chromatin immunoprecipitation (ChIP) or tissue-specific cDNA.

Data-handling and computational methods kept pace with these rapid advances in molecular and sequencing techniques (Pop & Salzberg 2008). The enormous volumes of short read data that could be outputted by NGS platforms within a day (gigabases in order) were stored in FASTQ format, previously

developed during the automation of Sanger sequencing in the HGP to combine both sequence information and the PHRED base quality score within the same file (Cock et al. 2010). Quality control and adapter trimming could then be performed using any one of a range of newly developed programs (Andrews 2010; Cox et al. 2010; Martin 2011; Bolger et al. 2014; Chen et al. 2014). The subsequent reconstruction of individual genomes from filtered short read data proved an intense computational bottleneck, which was addressed with the development of alignment algorithms based on index data structures, implemented in programs such as BWA and Bowtie, which could map reads to a known reference genome sequence on a scale of 7 Gbp per CPU per day (Langmead et al. 2009; Li & Durbin 2009; Li & Homer 2010). Alignment outputs required their own quality filtering, with the SAM/BAM format proving the most popular form of data storage, easily manipulated and edited using SAMtools software (Li et al. 2009).

The overall impact of NGS technologies was an unprecedentedly rapid drop in the price of sequencing a human genome, beginning in 2008 and far out-pacing hypothetical predictions based on Moore's Law, which trends in sequencing costs had previously followed (Wetterstrand 2016). As of 2016 an entire high quality human genome sequence costs approximately \$1,000, with further reductions in price expected. Such profound technological leaps spurred forward the 1000 Genomes Project, which ran between 2008 and 2015, a natural follow-on from the work initiated by the International HapMap Project, aiming to provide of comprehensive catalogue of human genetic variation. By the end of the project, the genomes of over 2,500 individuals from 26 populations had been published through a combination of both low and high coverage whole-genome sequencing, deep exome sequencing and dense microarray genotyping (1000 Genomes Project Consortium 2015).

The discovery and verification of variants from such large quantities of NGS read data required robust genotype calling algorithms, which were developed alongside the project, culminating in tools such as UnifiedGenotyper and HaplotypeCaller from the Genome Analysis Tool Kit (GATK) (McKenna et al. 2010), and mpileup from SAMtools (Li et al. 2009). Over 88 million variants were discovered, genotyped and phased in the dataset, the most abundant category being SNPs, which numbered over 84 million in total. The majority of these variants were rare, with only 8 million being present in over 5% of individuals, though within any given individual genome the vast majority of variants present were common ones. Rarer variants, which tend to be more recent in origin, were also typically restricted to individuals from the same population or continental group (McVean et al. 2012). Africans were seen to harbour the highest numbers of variant sites, as predicted by OaA. Analysis of variants that differed greatly in frequency between closely related populations, using the F_{ST} -based population branch statistic (Yi et al. 2010), could provide evidence for localised adaptation. Only a small number of loci involved in pigmentation, diet and immunity showed strong evidence of selection, emphasising the rarity of such selective sweeps in recent human history (Hernandez et al. 2011).

While groundbreaking in their own right, the immediate findings of the 1000 Genomes project were minor relative to the long term impact the resource would have on further studies of human disease and variation. Besides from enabling effective study and array designs, the dataset crucially provided a dense panel of phased haplotypes, with which improved genotype imputation could be carried out, replacing the previous HapMap reference dataset (~3.8 million variants). Such robust statistical imputation of missing genotypes, not included within the typical commercial SNP array, prompted the discovery of vast numbers of new functional variants involved in disease (Zheng-Bradley & Flicek 2017), still ongoing today. However, imputation of rare variation remains a challenge, encouraging regional whole genome sequencing projects, such as the UK10K (UK10K Consortium 2015), aimed at cataloguing recent rare variation within specific populations, as well as clinical exome sequencing of large patient cohorts (Brown & Meloche 2016).

However, these unparalleled new datasets, while optimal for studies of human disease, suffer from the same pitfall as the International HapMap Project before them. By focusing on large urban populations these studies can never capture the full picture of human genetic diversity, necessary for a complete understanding of our species' history. With similar motivations to the HGDP over a decade beforehand, the EGD (Pagani et al. 2016) and SGDP (Mallick et al. 2016) datasets sought to address this deficit, through the retrieval of high quality whole genome sequences from almost 700 individuals from over 270 geographically, culturally and linguistically diverse populations. The initial explorations of these datasets, as well as a third comprising Australasian populations (Malaspinas et al. 2016), have attempted to address some of the longstanding questions surrounding early human migrations, specifically the number of OoA dispersals and the timing and order of population splits upon entry into Eurasia. However these three studies were unable to reach a consensus. Pagani et al. propose multiple dispersals, with an earlier OoA contributing minorly to Australasians, while Malaspinas et al. and Mallick et al. put more emphasis on a single dispersal, though they differ on the branching pattern of Eurasian populations. Malaspinas et al. support previous arguments for the early separation of Australasians from all other Europeans (Rasmussen et al. 2011), while Mallick et al. suggest the earliest split occurred between west and east Eurasians, Australasians included.

Overall, these somewhat conflicting conclusions emphasise the difficulty of elucidating past demographic events from even the most high quality modern data. Signatures of previous migrations, expansions and admixtures can be obscured by subsequent events, with ever more complex models required to account for all such possibilities. Moreover, multiple demographic scenarios can give rise to the same patterns of modern variation, which likelihood models, no matter how well informed, may not be able to distinguish between. Finally, numerous diverse human and hominid populations may have existed in the past, which have left no discernible trace on modern genomes. Overall, it is clear that a complete understanding of the species' history will require not only a full geographical range of human whole genome diversity, but also a temporal one, achievable only through the retrieval of genetic material

from ancient remains. Fortunately, numerous researchers, aware of such theoretical upper limits to modern population genetics, had been working steadily towards this goal for many decades.

Ancient DNA: The Early Years

The field of ancient DNA had, until recently, been developing quietly alongside that of population genetics. There were, however, huge obstacles to surmount before this research could reach the mainstream. For one, the amount of endogenous DNA was seen to be miniscule in most archaeological and taxidermic remains compared to the levels of DNA from the microenvironment. Moreover, surviving ancient DNA (aDNA) is of a damaged, fragmented nature, with DNA molecules rarely more than a hundred or so base pairs in length (Hagelberg et al. 2015). For early researchers, attempting to isolate and sequence these molecules appeared a Sisyphean task. The first ancient DNA sequences reported were in 1984 from a museum specimen of the quagga, an extinct equid, retrieved from the sample through bacterial cloning (Higuchi et al. 1984). The vast majority of sequenced DNA was found to belong to environmental contaminants, with only two small fragments of apparent quagga mitochondrial DNA retrieved. Despite this low yield these results represented a paradigm shift: aDNA retrieval was indeed possible. The potential applications of aDNA research, in fields as diverse as archaeology, evolutionary biology, conservation, forensics, archaeogenetics, linguistics and anthropology, could not be ignored. Pioneers of the field, such as Svante Pääbo (Pääbo 1985a; Pääbo 1985b; Pääbo 1986), redoubled their efforts and were soon bolstered by the development of PCR in the late 1980s. Now it was possible to target sequences of interest, rather than shooting in the dark, hoping to hit an endogenous molecule in a haystack of microbial and environmental material. These new PCR-based studies focused almost exclusively on the mitochondrial DNA (mtDNA) for its high copy number; a single cell can possess as many as 100 to 10,000 mitochondria. This locus, for reasons discussed above, was also the most popular target for contemporary studies of modern genetic variation.

With the advent of PCR the field progressed rapidly, indeed some might say almost hysterically. Papers soon emerged reporting DNA sequences extracted from specimens tens of millions of years in age (Cano et al. 1992; DeSalle et al. 1992). The crowning glory at this time was the publication of mitochondrial sequences from dinosaur bones (Woodward et al. 1994). However, these antediluvian studies were soon shown to be irreproducible and unreliable (Austin et al. 1997), with the upper limit of usable DNA survival now estimated roughly as 1.5 million years (Allentoft et al. 2012). Projects on more recent organisms proved to be extremely successful and shed insight into population histories of both extinct and extant species (Thomas et al. 1989; Thomas et al. 1990; Cooper et al. 1992). The demonstration that aDNA could be retrieved, not only from scarce soft tissue remains, such as taxidermic, frozen and mummified specimens, but also from bone (Hagelberg et al. 1989; Horai et al. 1989), further expanded the field's horizons and had resounding implications both for human population genetics and forensics.

However, in spite of these advances, the post-PCR era brought with it a new and more pronounced concern in regards to contamination, a fear realised through the growing number of debunked aDNA

studies. Even a minute amount of contaminating human or environmental DNA, similar in sequence to that being targeted by PCR, could amplify to large quantities and confound results (Stoneking 1995; Cooper 1997). For the sequencing of ancient humans this was a major issue: how could one be sure that the amplified sequences belonged to the ancient human in question rather than one of its modern counterparts in the lab? For this reason, many researchers believed aDNA studies were simply not suited to human remains and turned their attention to more amenable types of fauna and flora (Goloubinoff et al. 1993; Hagelberg et al. 1994; Hänni et al. 1994; Höss et al. 1994; Höss et al. 1996; Yang et al. 1996; Dumolin-Lapègue et al. 1999; Greenwood et al. 1999; Leonard et al. 2000; Cooper et al. 2001; Paxinos et al. 2002), though several human studies were produced (Hagelberg & Clegg 1993; Stone & Stoneking 1993).

The successful sequencing of the first fragments of Neanderthal DNA in the late 1990s began to change these attitudes (Krings et al. 1997; Krings et al. 1999), bringing with it a renewed interest in the history of our own species, as well as a stringent set of criteria for aDNA extraction and sequencing, formulated to safeguard against contamination during the project. The new slogan became ‘Ancient DNA: do it right or not at all’ (Cooper & Poinar 2000). Ancient DNA laboratories came to resemble those used in forensics. All work was to be carried out in sterile, cleanroom environments, while wearing full body anti-contamination suits, and punctuated by copious amounts of cleaning. These new extreme standards brought heightened credibility to the field and were followed by a flurry of human studies, nearly all aimed at the mtDNA (Adcock et al. 2001; Endicott et al. 2003; Keyser-Tracqui et al. 2003; Vernesi et al. 2004; Sampietro et al. 2005). A reminder of the potential for aDNA to address longstanding questions within European archaeology was provided in a population level study of Neolithic Europeans (Haak et al. 2005), which rejected continuity between modern-day Europeans and these groups. Further studies also demonstrated discontinuity between previous hunter-gatherer populations and both Neolithic and Modern Europeans (Bramanti et al. 2009), suggesting migration had played a recurring role in the continent’s prehistory.

Another confounding factor that became apparent with the advent of PCR was the issue of post-mortem damage to ancient DNA molecules, which accumulates over time (Hansen et al. 2001). The ability to sequence a small mtDNA region to high coverage revealed a high degree of heterogeneity between overlapping molecules, unexplainable by sequencing error or contamination alone. The most prominent of these base modifications was an excess of C to T mutations, which could be reduced through uracil DNA glycosylase (UDG) treatment, identifying the causative process to be cytosine deamination to uracil (methylated thymine) (Hofreiter et al. 2001). These changes were later demonstrated to occur mainly at single-stranded overhanging ends of molecules (Briggs et al. 2007). Despite the problematic implications such phenomena had for the investigation of ancient variation, they also provided a definitive signal that differentiated modern contaminant DNA from true ancient molecules, with such patterns still used today in the verification of aDNA authenticity (Jónsson et al. 2013; Skoglund et al. 2014b; Orlando et al. 2015).

However, it could be argued that the biggest hurdle facing the field was not one of data quality, undermined by damage and contamination, but of data quantity. Due to the low survival rates of DNA in ancient specimens, the vast majority of research remained restricted to small regions of the mtDNA. Indeed, while the continued molecular and technical breakthroughs discussed in previous sections had given modern population genetic surveys access to large numbers of Y chromosome and autosomal loci, aDNA research could not benefit from such advances. Extraction techniques were gradually improved upon to maximise the amount of endogenous DNA retrieved from ancient remains, as well as minimising the co-extraction of PCR inhibitors, with the added caveat of avoiding overly aggressive treatments which can further damage the already degraded aDNA (Rohland & Hofreiter 2007b). Silica-binding procedures were shown to have increased PCR success compared to other methods (Rohland & Hofreiter 2007a), particularly those based on silica columns (Yang et al. 1998; MacHugh et al. 2000; Dabney et al. 2013; Gamba et al. 2016). For bone material, given the lack of intact cell membranes, few chemicals were seen to be required for aDNA extraction, with only EDTA and proteinase K producing a positive effect on DNA yields (Rohland & Hofreiter 2007b). Fine powder was preferable to coarser material and incubation times and temperatures could also be adjusted for optimum yields. A combination of well-preserved remains and well-chosen extraction methods made nuclear and Y chromosome DNA retrieval possible in a number of cases (Keyser-Tracqui et al. 2003; Römpler et al. 2006; Lacan et al. 2011). However, a near unimaginable paradigm shift was required before large-scale genomic surveys of ancient variation could become a reality.

A Palaeogenomic Revolution

In a review of NGS technologies from 2010, it was said “the potential of NGS is akin to the early days of PCR, with one's imagination being the primary limitation to its use” (Metzker 2010). Perhaps it is then no wonder that both technologies triggered veritable revolutions in aDNA research, a field which, for all its deficits, never suffered from a lack of imagination. However, while PCR methods had allowed aDNA studies to develop into a credible, though somewhat niche, scientific pursuit, NGS succeeded in reinventing the field entirely, allowing the retrieval of entire ancient genomes for the first time. This new era of palaeogenomics has brought aDNA analysis to the very forefront of evolutionary and population genetic research, providing unfathomably rapid resolutions to the previously unanswerable questions of human prehistory.

In addition to the more general advantages of NGS over traditional sequencing methods, such as the high levels of data obtained for low cost in a short amount of time, NGS has a number of qualities that make it extraordinarily well suited for aDNA sequencing in particular (Knapp & Hofreiter 2010). Indeed, one of the main criticisms levelled at NGS technology - the shorter read lengths produced in comparison to Sanger Sequencing - is a feature perfectly suited to the heavily fragmented nature of aDNA. Moreover, the use of universal adapter ligation in NGS allows for the amplification of fragmented aDNA molecules too short for traditional PCR methods, vastly increasing the amount of raw data extractable from ancient

specimens. Universal adapters also allow for the incorporation of sample-specific barcodes during library amplification. While initially developed to allow bioinformatic differentiation of distinct libraries pooled on a single sequencing lane, for ancient samples of minimal endogenous content such barcodes also provided a safeguard against any subsequent contamination events, allowing further work to take place outside sterile environments.

Most importantly, the use of universal adapters circumvented the need for PCR targeting of specific sequences, substantially decreasing contamination risks. Overall, this had the effect of transforming contamination into a factor to measure and control, rather than one fatal to an experiment. The new strategy was to sequence absolutely everything retrievable from a specimen and segregate endogenous sequences from environmental DNA at a later bioinformatic stage, during read alignment. Owing to the enormous amount of information produced by NGS technology, substantial numbers of endogenous DNA fragments may be sequenced from samples with very little surviving aDNA. Moreover, nuclear aDNA was soon seen to be substantially less prone to degradation than that retrieved from the mitochondria, possibly the result of increased protection by proteins. This allowed relatively longer intact strands to be recovered and further bolstered early palaeogenomic research (Rizzi et al. 2012).

The first NGS technology was released in 2005 (Margulies et al. 2005) and was soon implemented in the sequencing of thirteen million base pairs of the woolly mammoth genome (Poinar et al. 2006). In the next five years, this was followed by the draft nuclear genomes of the mammoth (Miller et al. 2008), the Neanderthal (Green et al. 2010) and a 4,500-year-old Palaeo-Eskimo (Rasmussen et al. 2010). Remarkably, this first ancient human genome sequence was achieved less than ten years after the first modern one, with Eurasian ancient genomes soon following (Keller et al. 2012). However, despite the successful sequencing of these first palaeogenomes, researchers were still facing the persistent issue of uneconomically low levels of endogenous content in ancient samples, with the majority of these early projects focusing on samples retrieved from permafrost or ice, conditions believed to encourage aDNA survival.

NGS technology actually allowed for the first time a full appraisal of the problem, by providing direct ratios of endogenous to environmental DNA in any given sequenced library. Such estimates were used to further hone extraction techniques. Multiple rounds of pre-digestions or extractions on the same bone powder were seen to increase the percentage of endogenous DNA, while decreasing overall DNA concentrations, likely through the removal of outer microbial contaminants and allowing full release of aDNA from the bone matrix (Der Sarkissian et al. 2014; Damgaard et al. 2015; Orlando et al. 2015). Most crucially, aDNA survival rates in different tissues could be tested, leading to the identification of the petrous temporal bone as an excellent preserver of aDNA molecules (Gamba et al. 2014), yielding exponentially higher endogenous contents relative to other skeletal elements, which tend to fall lower than 1%. Target-enrichment methods have also been used to great effect in aDNA studies (Orlando et al. 2015), either through the creation of selective bias towards damaged molecules during library

construction (Gansauge & Meyer 2014), or through hybridisation capture after library construction, either in solution or on microarrays. Targeted regions for capture have included mitochondrial genomes, microbial genomes, exomes and whole nuclear genomes (Briggs et al. 2009; Burbano et al. 2010; Bos et al. 2011; Carpenter et al. 2013).

SNP captures have also proved effective for studies of ancient human variation, particularly from older samples or those retrieved from climates unfavourable for DNA survival (Haak et al. 2015; Mathieson et al. 2015; Fu et al. 2016; Lazaridis et al. 2016; Skoglund et al. 2016). Studies such as these, alongside a smaller number of projects focused on low coverage whole genome shotgun sequencing of ancient populations, have provided genome-wide data from upwards of 1,000 ancient humans to date (reviewed in Slatkin & Racimo 2016; Marciniak & Perry 2017). These have provided unprecedented insight into human prehistory, although until recently the focus has remained fairly Eurocentric. However, it is beyond the scope of this introduction to detail such a broad spectrum of research into European prehistory. Instead, the findings of aDNA studies relevant to the current thesis will be discussed alongside their archaeological contexts in the introductory section to each chapter.

Finally, it must be noted that the currently popular SNP capture approaches can fall prey to the same pitfalls as modern SNP arrays, including ascertainment bias and loss of rare variation, issues that would be particularly pronounced for ancient individuals harbouring diversity no longer present in modern populations. For low coverage whole genome sequence data, these problems are less pronounced, though the assaying of novel variation in such datasets is still not feasible. This issue was highlighted in a recent analysis of a 57X genome from a Mesolithic Scandinavian, which discovered ~10,000 SNPs not known in modern populations, 17% of which were common among Mesolithic Scandinavians (Günther et al. 2017), suggesting a substantial fraction of variation present 9,000 years ago has disappeared today.

A handful of studies have produced ancient genomes of high enough coverage for robust diploid genotype calling (Gamba et al. 2014; Lazaridis et al. 2014; Broushaki et al. 2016; Cassidy et al. 2016; Hofmanová et al. 2016; Jones et al. 2017), highlighting the wide vista of analyses available to such datasets, including haplotypic sharing (Lazaridis et al. 2014; Broushaki et al. 2016; Cassidy et al. 2016), whole genome coalescent modelling (Lazaridis et al. 2014; Jones et al. 2015) and ROH analysis (Gamba et al. 2014; Jones et al. 2015; Broushaki et al. 2016; Cassidy et al. 2016; Hofmanová et al. 2016). Other studies have gone a step further, imputing diploid calls from low coverage ancient data (~1X) using the 1000 Genomes haplotype panel as a reference (Gamba et al. 2014; Martiniano et al. 2017). This not only allows ROH and haplotypic analyses to be performed on large ancient datasets using common SNPs, but also allows accurate assessment of variants involved in phenotypes of interest. While, the majority of ancient genomic data available today (low coverage SNP capture) has levels of missingness too high for accurate imputation, this will likely change in the near future, as emphasis returns to whole genome shotgun sequencing and perhaps targeted nuclear capture.

Clearly, the only way to fully utilise ancient specimens, of which there are a finite amount, is through whole genome sequencing to coverages high enough for robust variant discovery. There will always be obstacles to working with degraded aDNA, including low endogenous contents, damage and short fragment sizes, which can decrease alignment qualities, confound variant calls, or inhibit the characterisation structural variation. However, new molecular and bioinformatic techniques to overcome such issues are in constant development (Briggs et al. 2010; Jónsson et al. 2013; Kerpedjiev et al. 2014; Orlando et al. 2015; Link et al. 2017) and ancient genomes are fast becoming as accessible as modern ones. Indeed, the field of ancient DNA as a whole appears to operate on the basis of making the impossible possible. The publication of the complete genome of a 700,000 year old horse from permafrost (Orlando et al. 2013) and nuclear sequences from 430,000 year old Neanderthal cave remains provides a testament to this (Meyer et al. 2016), with future milestones constantly on the horizon.

References

- 1000 Genomes Project Consortium, 2015. A global reference for human genetic variation. *Nature*, 526(7571), pp.68–74.
- Adcock, G.J. et al., 2001. Mitochondrial DNA sequences in ancient Australians: Implications for modern human origins. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2), pp.537–542.
- Albrechtsen, A., Nielsen, F.C. & Nielsen, R., 2010. Ascertainment biases in SNP chips affect measures of population divergence. *Molecular Biology and Evolution*, 27(11), pp.2534–2547.
- Alexander, D.H., Novembre, J. & Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), pp.1655–1664.
- Allentoft, M.E. et al., 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society B: Biological Sciences*, 279(1748), pp.4724–4733.
- Allentoft, M.E. et al., 2015. Population genomics of Bronze Age Eurasia. *Nature*, 522(7555), pp.167–172.
- Ammerman, A.J. & Cavalli-Sforza, L.L., 1984. *The Neolithic transition and the population genetics of Europe*. New Jersey: Princeton University Press.
- Anderson, S. et al., 1981. Sequence and organisation of the human mitochondrial genome. *Nature*, 290(5806), pp.457–465.
- Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data. Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Austin, J.J. et al., 1997. Problems of reproducibility--does geologically ancient DNA survive in amber-preserved insects? *Proceedings of the Royal Society B: Biological Sciences*, 264(1381), pp.467–474.
- Avery, O.T., Macleod, C.M. & McCarty, M., 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types : Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *The Journal of experimental medicine*, 79(2), pp.137–158.
- Behar, D.M. et al., 2006. The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. *American Journal of Human Genetics*, 78(3), pp.487–497.
- Bernstein, F., 1924. Ergebnisse Einer Biostatistischen Zusam-Menfassenden Betrachtung über die Erblichen Blutstrukturen des Menschen. *Klinische Wochenschrift*, 3(33), pp.1495–1497.
- Bolger, A.M., Lohse, M. & Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), pp.2114–2120.
- Bos, K.I. et al., 2011. A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature*, 478(7370), pp.506–510.
- Botstein, D. et al., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32(3), pp.314–331.
- Bowcock, A.M. et al., 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, 368(6470), pp.455–457.
- Boyd, W.C., 1950. *Genetics and the Races of Man: An Introd. to Modern Physical Anthropology*. New York: Little, Brown and Company.
- Bramanti, B. et al., 2009. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science*, 326(5949), pp.137–140.
- Briggs, A.W. et al., 2007. Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, 104(37), pp.14616–14621.
- Briggs, A.W. et al., 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science*, 325(5938), pp.318–321.
- Briggs, A.W. et al., 2010. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Research*, 38(6), p.e87.
- Broushaki, F. et al., 2016. Early Neolithic genomes from the eastern Fertile Crescent. *Science*, 353(6298), pp.499–503.
- Brown, T.L. & Meloche, T.M., 2016. Exome sequencing a review of new strategies for rare genomic disease research. *Genomics*, 108(3-4), pp.109–114.
- Brown, W.M., George, M., Jr & Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 76(4), pp.1967–1971.
- Burbano, H.A. et al., 2010. Targeted investigation of the Neandertal genome by array-based sequence capture. *Science*, 328(5979), pp.723–725.
- Byrne, R.P. et al., 2017 Celtic population structure and genomic footprints of migration. Manuscript submitted for publication.
- Cann, H.M. et al., 2002. A human genome diversity cell line panel. *Science*, 296(5566), pp.261–262.
- Cann, R.L., Stoneking, M. & Wilson, A.C., 1987. Mitochondrial DNA and human evolution. *Nature*, 325(6099), pp.31–36.
- Cano, R.J., Poinar, H. & Go, P.J., 1992. Isolation and partial characterisation of DNA from the bee *Problebeia dominicana*

- (Apidae:Hymenoptera) in 25–40 million year old amber. *Medical Science Research*, 20, pp.249–251.
- Carpenter, M.L. et al., 2013. Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. *American Journal of Human Genetics*, 93(5), pp.852–864.
- Casanova, M. et al., 1985. A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science*, 230(4732), pp.1403–1406.
- Cassidy, L.M. et al., 2016. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proceedings of the National Academy of Sciences of the United States of America*, 113(2), pp.368–373.
- Cavalli-Sforza, L.L., Menozzi, P. & Piazza, A., 1994. *The History and Geography of Human Genes*. New Jersey: Princeton University Press.
- Chen, C. et al., 2014. Software for pre-processing Illumina next-generation sequencing short read sequences. *Source Code for Biology and Medicine*, 9, p.8.
- Cirulli, E.T. & Goldstein, D.B., 2010. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews Genetics*, 11(6), pp.415–425.
- Cock, P.J.A. et al., 2010. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, 38(6), pp.1767–1771.
- Cooper, A. et al., 2001. Complete mitochondrial genome sequences of two extinct moas clarify ratite evolution. *Nature*, 409(6821), pp.704–707.
- Cooper, A. et al., 1992. Independent origins of New Zealand moas and kiwis. *Proceedings of the National Academy of Sciences of the United States of America*, 89(18), pp.8741–8744.
- Cooper, A., 1997. Reply to Stoneking: ancient DNA--how do you really know when you have it? *American Journal of Human Genetics*, 60(4), pp.1001–1003.
- Cooper, A. & Poinar, H.N., 2000. Ancient DNA: do it right or not at all. *Science*, 289(5482), p.1139.
- Cox, M.P., Peterson, D.A. & Biggs, P.J., 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics*, 11, p.485.
- Dabney, J. et al., 2013. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences of the United States of America*, 110(39), pp.15758–15763.
- Damgaard, P.B. et al., 2015. Improving access to endogenous DNA in ancient bones and teeth. *Scientific Reports*, 5, p.11184.
- Danna, K. & Nathans, D., 1971. Specific cleavage of simian virus 40 DNA by restriction endonuclease of *Hemophilus influenzae*. *Proceedings of the National Academy of Sciences of the United States of America*, 68(12), pp.2913–2917.
- Darwin, C., 1871. *The descent of man, and selection in relation to sex*. London: J. Murray.
- Der Sarkissian, C. et al., 2014. Shotgun microbial profiling of fossil remains. *Molecular Ecology*, 23(7), pp.1780–1798.
- DeSalle, R. et al., 1992. DNA sequences from a fossil termite in Oligo-Miocene amber and their phylogenetic implications. *Science*, 257(5078), pp.1933–1936.
- Dumolin-Lapègue, S. et al., 1999. Amplification of oak DNA from ancient and modern wood. *Molecular Ecology*, 8(12), pp.2137–2140.
- Endicott, P. et al., 2003. The genetic origins of the Andaman Islanders. *American Journal of Human Genetics*, 72(1), pp.178–184.
- Fiers, W. et al., 1978. Complete nucleotide sequence of SV40 DNA. *Nature*, 273(5658), pp.113–120.
- Fu, Q. et al., 2016. The genetic history of Ice Age Europe. *Nature*, 534(7606), pp.200–205.
- Gamba, C. et al., 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications*, 5, p.5257.
- Gamba, C., Hanghøj, K. & Gaunitz, C., 2016. Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. *Molecular Ecology*, 16(2), pp.459–469.
- Gansauge, M.-T. & Meyer, M., 2014. Selective enrichment of damaged DNA molecules for ancient genome sequencing. *Genome Research*, 24(9), pp.1543–1549.
- Gershon, D., 2002. Microarray technology: an array of opportunities. *Nature*, 416(6883), pp.885–891.
- Giles, R.E. et al., 1980. Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 77(11), pp.6715–6719.
- Goloubinoff, P., Pääbo, S. & Wilson, A.C., 1993. Evolution of maize inferred from sequence diversity of an *Adh2* gene segment from archaeological specimens. *Proceedings of the National Academy of Sciences of the United States of America*, 90(5), pp.1997–2001.
- Goodwin, S., McPherson, J.D. & McCombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17(6), pp.333–351.
- Green, R.E. et al., 2010. A draft sequence of the Neandertal genome. *Science*, 328(5979), pp.710–722.

- Greenwood, A.D. et al., 1999. Nuclear DNA sequences from late Pleistocene megafauna. *Molecular Biology and Evolution*, 16(11), pp.1466–1473.
- Günther, T. et al., 2017. Genomics of Mesolithic Scandinavia reveal colonisation routes and high-latitude adaptation. *bioRxiv*, p.164400.
- Haak, W. et al., 2005. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science*, 310(5750), pp.1016–1018.
- Haak, W. et al., 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522(7555), pp.207–211.
- Hagelberg, E., Sykes, B. & Hedges, R., 1989. Ancient bone DNA amplified. *Nature*, 342(6249), p.485.
- Hagelberg, E. & Clegg, J.B., 1993. Genetic polymorphisms in prehistoric Pacific islanders determined by analysis of ancient bone DNA. *Proceedings of the Royal Society B: Biological Sciences*, 252(1334), pp.163–170.
- Hagelberg, E. et al., 1994. DNA from ancient mammoth bones. *Nature*, 370(6488), pp.333–334.
- Hagelberg, E., Hofreiter, M. & Keyser, C., 2015. Introduction. Ancient DNA: the first three decades. *Philosophical Transactions of the Royal Society B: Biological sciences*, 370(1660), p.20130371.
- Hammer, M.F., 1995. A recent common ancestry for human Y chromosomes. *Nature*, 378(6555), pp.376–378.
- Hänni, C. et al., 1994. Tracking the origins of the cave bear (*Ursus spelaeus*) by mitochondrial DNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, 91(25), pp.12336–12340.
- Hansen, A. et al., 2001. Statistical evidence for miscoding lesions in ancient DNA templates. *Molecular Biology and Evolution*, 18(2), pp.262–265.
- Hartl, D.L. & Clark, A.G., 2007. *Principles of population genetics (4th ed.)*. Massachusetts: Sinauer and Associates.
- Hellenthal, G. et al., 2014. A genetic atlas of human admixture history. *Science*, 343(6172), pp.747–751.
- Hernandez, R.D. et al., 2011. Classic selective sweeps were rare in recent human evolution. *Science*, 331(6019), pp.920–924.
- Higuchi, R. et al., 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature*, 312(5991), pp.282–284.
- Hirszfeld, L. & Hirszfeld, H., 1919. Serological differences between the blood of different races. *The Lancet*, 2, pp.675–679.
- Hofmanová, Z. et al., 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proceedings of the National Academy of Sciences of the United States of America*, 113(25), pp.6886–6891.
- Hofreiter, M. et al., 2001. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Research*, 29(23), pp.4793–4799.
- Horai, S. et al., 1989. DNA Amplification from Ancient Human Skeletal Remains and Their Sequence Analysis. *Proceedings of the Japan Academy B: Physical and Biological Sciences*, 65(10), pp.229–233.
- Höss, M., Pääbo, S. & Vereshchagin, N.K., 1994. Mammoth DNA sequences. *Nature*, 370(6488), p.333.
- Höss, M. et al., 1996. Molecular phylogeny of the extinct ground sloth *Mylodon darwini*. *Proceedings of the National Academy of Sciences of the United States of America*, 93(1), pp.181–185.
- Hudson, T.J. et al., 1995. An STS-based map of the human genome. *Science*, 270(5244), pp.1945–1954.
- Hutchison, C.A., 3rd et al., 1974. Maternal inheritance of mammalian mitochondrial DNA. *Nature*, 251(5475), pp.536–538.
- The International HapMap Consortium, 2005. A haplotype map of the human genome. *Nature*, 437(7063), p.1299.
- Jakubiczka, S. et al., 1989. A search for restriction fragment length polymorphism on the human Y chromosome. *Human Genetics*, 84(1), pp.86–88.
- Jobling, M.A. & Tyler-Smith, C., 1995. Fathers and sons: the Y chromosome and human evolution. *Trends in Genetics*, 11(11), pp.449–456.
- Jobling, M.A. & Tyler-Smith, C., 2003. The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics*, 4(8), pp.598–612.
- Jones, E.R. et al., 2015. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications*, 6, p.8912.
- Jones, E.R. et al., 2017. The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Current Biology*, 27(4), pp.576–582.
- Jónsson, H. et al., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, 29(13), pp.1682–1684.
- Keller, A. et al., 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nature Communications*, 3, p.698.
- Kerpedjiev, P. et al., 2014. Adaptable probabilistic mapping of short reads using position specific scoring matrices. *BMC Bioinformatics*, 15, p.100.

- Keyser-Tracqui, C., Crubézy, E. & Ludes, B., 2003. Nuclear and mitochondrial DNA analysis of a 2,000-year-old necropolis in the Egyin Gol Valley of Mongolia. *American Journal of Human Genetics*, 73(2), pp.247–260.
- Kirin, M. et al., 2010. Genomic runs of homozygosity record population history and consanguinity. *PLoS One*, 5(11), p.e13996.
- Knapp, M. & Hofreiter, M., 2010. Next Generation Sequencing of Ancient DNA: Requirements, Strategies and Perspectives. *Genes*, 1(2), pp.227–243.
- Krings, M. et al., 1997. Neandertal DNA sequences and the origin of modern humans. *Cell*, 90(1), pp.19–30.
- Krings, M. et al., 1999. DNA sequence of the mitochondrial hypervariable region II from the neandertal type specimen. *Proceedings of the National Academy of Sciences of the United States of America*, 96(10), pp.5581–5585.
- Kwok, P.-Y. & Chen, X., 2003. Detection of single nucleotide polymorphisms. *Current Issues in Molecular Biology*, 5(2), pp.43–60.
- Lacan, M. et al., 2011. Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *Proceedings of the National Academy of Sciences of the United States of America*, 108(45), pp.18255–18259.
- LaFramboise, T., 2009. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Research*, 37(13), pp.4181–4193.
- Lander, E.S. et al., 2001. Initial sequencing and analysis of the human genome. *Nature*, 409(6822), pp.860–921.
- Landsteiner, K., 1901. Agglutination phenomena in normal human blood. *Wiener klinische Wochenschrift*, 14, pp.1132–1134.
- Langmead, B. et al., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10(3), p.R25.
- Lawson, D.J. et al., 2012. Inference of population structure using dense haplotype data. *PLoS Genetics*, 8(1), p.e1002453.
- Lazaridis, I. et al., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 513(7518), pp.409–413.
- Lazaridis, I. et al., 2016. Genomic insights into the origin of farming in the ancient Near East. *Nature*, 536(7617), pp.419–424.
- Leonard, J.A., Wayne, R.K. & Cooper, A., 2000. Population genetics of ice age brown bears. *Proceedings of the National Academy of Sciences of the United States of America*, 97(4), pp.1651–1654.
- Leslie, S. et al., 2015. The fine-scale genetic structure of the British population. *Nature*, 519(7543), pp.309–314.
- Lewontin, R.C., 1972. The apportionment of human diversity. *Evolutionary Biology*, 6(381), p.e398.
- Li, H. et al., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), pp.2078–2079.
- Li, H. & Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), pp.1754–1760.
- Li, H. & Homer, N., 2010. A survey of sequence alignment algorithms for next-generation sequencing. *Briefings in Bioinformatics*, 11(5), pp.473–483.
- Li, H. & Durbin, R., 2011. Inference of human population history from individual whole-genome sequences. *Nature*, 475(7357), pp.493–496.
- Li, N. & Stephens, M., 2003. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, 165(4), pp.2213–2233.
- Link, V. et al., 2017. Atlas: analysis tools for low-depth and ancient samples. *bioRxiv*, p.105346.
- Lohmueller, K.E., Bustamante, C.D. & Clark, A.G., 2009. Methods for human demographic inference using haplotype patterns from genomewide single-nucleotide polymorphism data. *Genetics*, 182(1), pp.217–231.
- Lucotte, G. & Ngo, N.Y., 1985. p49f, A highly polymorphic probe, that detects Taq1 RFLPs on the human Y chromosome. *Nucleic acids research*, 13(22), p.8285.
- MacHugh, D.E. et al., 2000. The extraction and analysis of ancient DNA from bone and teeth: a survey of current methodologies. *Ancient biomolecules*, 3(2), pp.81–103.
- MacLeod, I.M. et al., 2013. Inferring demography from runs of homozygosity in whole-genome sequence, with correction for sequence errors. *Molecular biology and evolution*, 30(9), pp.2209–2223.
- Malaspinas, A.-S. et al., 2016. A genomic history of Aboriginal Australia. *Nature*, 538(7624), pp.207–214.
- Mallick, S. et al., 2016. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature*, 538(7624), pp.201–206.
- Manolio, T.A. et al., 2009. Finding the missing heritability of complex diseases. *Nature*, 461(7265), pp.747–753.
- Marciniak, S. & Perry, G.H., 2017. Harnessing ancient genomes to study the history of human adaptation. *Nature Reviews Genetics*, doi:10.1038/nrg.2017.65. Epub ahead of print.
- Mardis, E.R., 2017. DNA sequencing technologies: 2006–2016. *Nature Protocols*, 12(2), pp.213–218.
- Margulies, M. et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057), pp.376–

380.

- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), pp.10–12.
- Martiniano, R. et al., 2017. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genetics*, 13(7), p.e1006852.
- Mathieson, I. et al., 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528(7583), pp.499–503.
- McCarthy, M.I. et al., 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews Genetics*, 9(5), pp.356–369.
- McEvoy, B. et al., 2006. The scale and nature of Viking settlement in Ireland from Y-chromosome admixture analysis. *European Journal of Human Genetics*, 14(12), pp.1288–1294.
- McKenna, A. et al., 2010. The Genome Analysis Toolkit: a MapReduce framework for analysing next-generation DNA sequencing data. *Genome Research*, 20(9), pp.1297–1303.
- McVean, G.A. et al., 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature*, 491(7422), pp.56–65.
- Metzker, M.L., 2010. Sequencing technologies - the next generation. *Nature Reviews Genetics*, 11(1), pp.31–46.
- Meyer, M. et al., 2016. Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature*, 531(7595), pp.504–507.
- Miller, W. et al., 2008. Sequencing the nuclear genome of the extinct woolly mammoth. *Nature*, 456(7220), pp.387–390.
- Mullis, K.B. et al., 1987. Process for amplifying, detecting, and/or-cloning nucleic acid sequences. *US Patent*, US4683195 A.
- Novembre, J. et al., 2008. Genes mirror geography within Europe. *Nature*, 456(7218), pp.98–101.
- Novembre, J. & Stephens, M., 2008. Interpreting principal component analyses of spatial population genetic variation. *Nature Genetics*, 40(5), pp.646–649.
- Novembre, J. & Ramachandran, S., 2011. Perspectives on human population structure at the cusp of the sequencing era. *Annual Review of Genomics and Human genetics*, 12, pp.245–274.
- Olson, M. et al., 1989. A common language for physical mapping of the human genome. *Science*, 245(4925), pp.1434–1435.
- Orlando, L. et al., 2013. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, 499(7456), pp.74–78.
- Orlando, L., Gilbert, M.T.P. & Willerslev, E., 2015. Reconstructing ancient genomes and epigenomes. *Nature Reviews Genetics*, 16(7), pp.395–408.
- Pääbo, S., 1985a. Molecular cloning of Ancient Egyptian mummy DNA. *Nature*, 314(6012), pp.644–645.
- Pääbo, S., 1985b. Preservation of DNA in ancient Egyptian mummies. *Journal of Archaeological Science*, 12(6), pp.411–417.
- Pääbo, S., 1986. Molecular genetic investigations of ancient human remains. *Cold Spring Harbor Symposia on Quantitative Biology*, 51 Pt 1, pp.441–446.
- Pagani, L. et al., 2016. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature*, 538(7624), pp.238–242.
- Patterson, N., Price, A.L. & Reich, D., 2006. Population structure and eigenanalysis. *PLoS Genetics*, 2(12), p.e190.
- Patterson, N. et al., 2012. Ancient admixture in human history. *Genetics*, 192(3), pp.1065–1093.
- Paxinos, E.E. et al., 2002. mtDNA from fossils reveals a radiation of Hawaiian geese recently derived from the Canada goose (*Branta canadensis*). *Proceedings of the National Academy of Sciences of the United States of America*, 99(3), pp.1399–1404.
- Pemberton, T.J. et al., 2012. Genomic patterns of homozygosity in worldwide human populations. *American Journal of Human Genetics*, 91(2), pp.275–292.
- Piazza, A. et al., 1995. Genetics and the origin of European languages. *Proceedings of the National Academy of Sciences of the United States of America*, 92(13), pp.5836–5840.
- Poinar, H.N. et al., 2006. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science*, 311(5759), pp.392–394.
- Pop, M. & Salzberg, S.L., 2008. Bioinformatics challenges of new sequencing technology. *Trends in Genetics*, 24(3), pp.142–149.
- Price, A.L. et al., 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8), pp.904–909.
- Price, A.L. et al., 2009. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genetics*, 5(6), p.e1000519.
- Pritchard, J.K., Stephens, M. & Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155(2), pp.945–959.
- Rasmussen, M. et al., 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature*, 463(7282), pp.757–762.

- Rasmussen, M. et al., 2011. An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science*, 334(6052), pp.94–98.
- Reich, D. et al., 2009. Reconstructing Indian population history. *Nature*, 461(7263), pp.489–494.
- Renfrew, C., 1990. *Archaeology and Language: The Puzzle of Indo-European Origins*. Cambridge: CUP Archive.
- Richards, M. et al., 1996. Paleolithic and neolithic lineages in the European mitochondrial gene pool. *American Journal of Human Genetics*, 59(1), pp.185–203.
- Richards, M. et al., 2003. Extensive female-mediated gene flow from sub-Saharan Africa into near eastern Arab populations. *American Journal of Human Genetics*, 72(4), pp.1058–1064.
- Rizzi, E. et al., 2012. Ancient DNA studies: new perspectives on old samples. *Genetics, Selection, Evolution*, 44, p.21.
- Rohland, N. & Hofreiter, M., 2007a. Ancient DNA extraction from bones and teeth. *Nature Protocols*, 2(7), pp.1756–1762.
- Rohland, N. & Hofreiter, M., 2007b. Comparison and optimization of ancient DNA extraction. *BioTechniques*, 42(3), pp.343–352.
- Römpler, H. et al., 2006. Nuclear gene indicates coat-color polymorphism in mammoths. *Science*, 313(5783), p.62.
- Rosenberg, N.A. et al., 2002. Genetic structure of human populations. *Science*, 298(5602), pp.2381–2385.
- Rosenberg, N.A. et al., 2005. Clines, clusters, and the effect of study design on the inference of human population structure. *PLoS Genetics*, 1(6), p.e70.
- Sachidanandam, R. et al., 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 409(6822), pp.928–933.
- Saitou, N. & Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), pp.406–425.
- Salas, A. et al., 2005. Charting the ancestry of African Americans. *American Journal of Human Genetics*, 77(4), pp.676–680.
- Sampietro, M.L. et al., 2005. The genetics of the pre-Roman Iberian Peninsula: a mtDNA study of ancient Iberians. *Annals of Human Genetics*, 69(Pt 5), pp.535–548.
- Sanger, F., Nicklen, S. & Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp.5463–5467.
- Serre, D. & Pääbo, S., 2004. Evidence for gradients of human genetic diversity within and among continents. *Genome Research*, 14(9), pp.1679–1685.
- Skoglund, P. et al., 2014a. Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science*, 344(6185), pp.747–750.
- Skoglund, P. et al., 2014b. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, 111(6), pp.2229–2234.
- Skoglund, P. et al., 2016. Genomic insights into the peopling of the Southwest Pacific. *Nature*, 538(7626), pp.510–513.
- Slatkin, M. & Racimo, F., 2016. Ancient DNA and human history. *Proceedings of the National Academy of Sciences of the United States of America*, 113(23), pp.6380–6387.
- Sokal, R.R., Oden, N.L. & Wilson, C., 1991. Genetic evidence for the spread of agriculture in Europe by demic diffusion. *Nature*, 351(6322), pp.143–145.
- Southern, E.M., 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, 98(3), pp.503–517.
- Stone, A.C. & Stoneking, M., 1993. Ancient DNA from a pre-Columbian Amerindian population. *American Journal of Physical Anthropology*, 92(4), pp.463–471.
- Stoneking, M., 1995. Ancient DNA: how do you know when you have it and what can you do with it? *American Journal of Human Genetics*, 57(6), pp.1259–1262.
- Thomas, R.H. et al., 1989. DNA phylogeny of the extinct marsupial wolf. *Nature*, 340(6233), pp.465–467.
- Thomas, W.K. et al., 1990. Spatial and temporal continuity of kangaroo rat populations shown by sequencing mitochondrial DNA from museum specimens. *Journal of Molecular Evolution*, 31(2), pp.101–112.
- Torroni, A. et al., 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. *American Journal of Human Genetics*, 53(3), pp.563–590.
- UK10K Consortium, 2015. The UK10K project identifies rare variants in health and disease. *Nature*, 526(7571), pp.82–90.
- Underhill, P.A. & Kivisild, T., 2007. Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annual Review of Genetics*, 41, pp.539–564.
- Vernesi, C. et al., 2004. The Etruscans: a population-genetic study. *American Journal of Human Genetics*, 74(4), pp.694–704.
- Wang, D.G. et al., 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science*, 280(5366), pp.1077–1082.

- Watson, E. et al., 1996. mtDNA sequence diversity in Africa. *American Journal of Human Genetics*, 59(2), pp.437–444.
- Watson, J.D. & Crick, F.H., 1953. The structure of DNA. *Cold Spring Harbor Symposia on Quantitative Biology*, 18, pp.123–131.
- Wells, R.S. et al., 2001. The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18), pp.10244–10249.
- Wetterstrand, K.A., 2016. DNA sequencing costs: data from the NHGRI Genome Sequencing Program (GSP). *National Human Genome Research Institute*. Available at: <http://www.genome.gov/sequencingcosts>.
- Wollstein, A. et al., 2010. Demographic history of Oceania inferred from genome-wide data. *Current Biology*, 20(22), pp.1983–1992.
- Woodward, S.R., Weyand, N.J. & Bunnell, M., 1994. DNA sequence from Cretaceous period bone fragments. *Science*, 266(5188), pp.1229–1232.
- Xinzhi, W., 1981. The well preserved cranium of an early Homo sapiens from Dali, Shanxi. *Scientia Sinica*, 2, pp.200–206.
- Yang, D.Y. et al., 1998. Technical note: improved DNA extraction from ancient bones using silica-based spin columns. *American Journal of Physical Anthropology*, 105(4), pp.539–543.
- Yang, H., Golenberg, E.M. & Shoshani, J., 1996. Phylogenetic resolution within the Elephantidae using fossil DNA sequence from the American mastodon (*Mammot americanum*) as an outgroup. *Proceedings of the National Academy of Sciences of the United States of America*, 93(3), pp.1190–1194.
- Yi, X. et al., 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. *Science*, 329(5987), pp.75–78.
- Zerjal, T. et al., 2003. The genetic legacy of the Mongols. *American Journal of Human Genetics*, 72(3), pp.717–721.
- Zheng-Bradley, X. & Flicek, P., 2017. Applications of the 1000 Genomes Project resources. *Briefings in Functional Genomics*, 16(3), pp.163–170.
- Zuckerkandl, E. & Pauling, L., 1962. Molecular Evolution. In M. Kasha & B. Pullman (eds.), *Horizons in Biochemistry*. New York: Academic Press, pp. 189–225.
- Zvelebil, M. & Zvelebil, K.V., 1988. Agricultural transition and Indo-European dispersals. *Antiquity*, 62(236), pp.574–583.

2. The Takings of Ireland

Punctuated population replacement followed by long term continuity on Europe's Atlantic edge

Overview

This chapter is based on the 2016 paper (Cassidy et al. 2016), published in the initial phase of this project, which created the first holistic demographic framework for Irish prehistory. In addition to the four original published samples, this chapter has been updated to include the entire dataset of 93 individuals sequenced for the current thesis. The introduction takes the reader through the main schools of thought regarding the origins and histories of Irish populations, which have influenced both academic work and mainstream consciousness well into the 21st century. These have drawn on classical, early Christian and medieval texts; philology and linguistics; ethnography; antiquarian and later archaeological studies; and finally the genetics of modern populations. The results presented here add the major contribution of ancient DNA (aDNA) research to this list.

Through principal component and ADMIXTURE analysis, three discrete populations are identified that have inhabited the island of Ireland through its major prehistoric periods and the origins of each are established with respect to the newly emerging European palaeogenomic narrative. The timings of demographic upheavals can be closely correlated with two major events in the island's human prehistory, the arrival of agriculture (~3,750 BC) followed by the onset of metallurgy (~2,500 BC). Both events were catalysed by mass migration into the island, the earlier originating in early farming populations of West Asia, and the later in Bronze Age pastoralist groups of the Pontic steppe. Importantly, strong signals of continuity are observed from the Chalcolithic onwards, evidenced in both Y chromosome R1b haplotype frequencies, and inflated autosomal haplotypic sharing with modern Irish, Scottish and Welsh populations. This provides new avenues of interpretation for ongoing debate surrounding the origins and spread of the both the Celtic language family and its speakers.

Introduction

An Island of Invasions: Historical Attempts at Solving Irish Origins

The human population of Ireland was first pulled out of the darkness of prehistory and into the written record by Christian monks in the 5th century, who succeeded in creating the earliest vernacular literature of Western Europe (Fig. 1.1A and C). From this point onwards, the comings and goings of peoples on the island, including Vikings, Normans, English and Scottish settlers; demographic collapses, such as those triggered by the Black Death and Great Famine; as well as linguistic and cultural turnovers associated with these events, could all be gleaned by direct account, in addition to archaeological means. However, Ireland had been permanently inhabited for almost 8,000 years by the time the first monk put his pen to parchment (Bayliss & Woodman 2009). Indeed, faced with a largely unknown past, one of the main occupations of these early scribes was to provide a suitable origin story for the people of the island, known then as the Gaels, that wove together the prevailing biblical narratives with native pagan mythologies. This culminated in the famous *Lebor Gabála Éirenn* or “The Book of the Takings of Ireland”, in which Ireland saw successive waves of colonisers arrive on her shores from regions as far removed as Iberia, Scythia and Egypt. Despite the eventual 20th century academic dismissal of these early accounts as a pseudo-history (Carey & Dumville 1994; Mallory 2016), the theme of invasions has remained a recurrent one in the study of Irish prehistory, the most unshakable being perhaps that of “Celtic Peoples” of the Iron Age (Mallory 2013).

Before Ireland’s inhabitants had begun recording their own affairs, classical explorers had taken note of the far westerly isle of *Hibernia*. Ptolemy’s *Geography* (150 AD) went as far as to list tribal and place-names, recognised today as belonging to the Celtic language family. Both main branches of the family, Goidelic (Irish, Scots Gaelic, Manx) and Brittonic (Welsh, Cumbric, Cornish, Breton), have been recognised in Ptolemy’s list, as well as several tribal names with analogues in Britain and Gaul. This opens the possibility of British and continental movement into the island, particularly the south (De Bernardo Stempel 2007). However, these classical writers had in fact reserved the terms *Keltoi* (Greek) or *Celtae* (Latin) for continental groups, never actually extending the definition to include the peoples of Britain and Ireland. The so-called ‘Celts’ appeared in classical Greek literature around the 6th century BC, with earliest accounts placing their homelands on the Atlantic seaboard in the extreme west of the continent (Cunliffe 1997). Later in Roman literature, the word Gaul began to be used synonymously with Celt, and the term was extended to include the tribes of much of western and central Europe, and was used to contrast them with the Scythians of the east. Celtic tribes were also documented in Iberia, living to the west of pre-Indo-European speakers who inhabited the eastern coast. Celtic expansions southwards and eastwards, into Italy, Greece and Anatolia were all well-documented by these early historians.

It was through the link with the Gauls that the term ‘Celtic’ found its way into the modern vernacular and began its long term association with northwest insular Europe. Post-renaissance studies of the European past focused primarily on direct documentary and philological evidence from ancient historical

texts. The latter of these was based on the biblical concept that all languages must descend from a single ancestor (Rowley-Conwy 2007), and thus their relatedness and distributions in modern times were believed to provide clues to the history of their speakers, echoing later population genetic theory. Similarities noted between ancient Gallic languages and those spoken in Britain and Ireland led to proposition of the ‘Celtic Language family’ in the early 18th century, subsequently placed into the larger Indo-European classification system (Prichard 1831). Its key proponent, Edward Lhuyd, imagined these tongues had arrived on the islands through continental Gaulish invasions, despite the lack of written historical evidence (Lhuyd 1707). The Celtic Revival movement seized on this origin story, marrying it together with early Christian and Medieval literature and art from Ireland and Britain; while romantic antiquarians soon began to ascribe the scattered assemblage of prehistoric monuments and megaliths found across Britain, Ireland and France to the mythical Celtic Druids of classical texts (Waddell 2005; Cunliffe 1997). This narrative of Celticism took hold, subsumed into the nationalist politics of the 19th and 20th centuries, and heavily influenced generations of linguistic, archaeological and ethnographic research into Ireland’s population.

The Curious Case of the Celts: Evolving Archaeological Perspectives

Archaeology as a field had been maturing from hobbyist antiquarianism into a scientific discipline across the 19th century. Important cognitive leaps in the understanding of geological stratigraphy and the concept of deep time had led to the ordering of human prehistory into a series of ages based on artefacts - the Stone Age, later divided into Palaeo- Meso- and Neolithic periods, the Bronze Age and the Iron Age. This categorization has persisted into modern usage, providing a consistent chronological framework for most regions of western Eurasia. However, many early archaeologists balked at the notion that such an expansive human ‘prehistory’ had occurred prior to the advent of written records, preferring established biblical calendars, such as that of the Irish Bishop, James Ussher, who placed the date of creation at 4,004 BC.

It took Ireland almost 50 years to accept this three-age system, after its establishment in Scandinavia in the 1830s, the last archaeological tradition of the then British Isles to do so (Rowley-Conwy 2007). This was in part due to the island’s comparatively older written records, which prompted historians, rather than antiquarians, to provide the main chronologies for island, with golden ages of Celtic heroism and Early Christian scholarship taking precedence over stone, bronze and iron (Waddell 2005). Despite the growing awareness of deep time, many Irish archaeologists refused to accept that material assemblages could be used to accurately date sites (O’Curry 1878). Others looked to anthropometric measurements, such as skull shape, to classify the ages of human remains uncovered (Wilde 1850), and these were then linked to the waves of invaders described by the early manuscripts, some of whom in turn were postulated to represent earlier pre-Celtic “Finn” and “Basque” populations, borrowing from linguistic studies (Rowley-Conwy 2007). All this fed into the racist theories emerging from ethnographers of the time, and the ideological desire to define a ‘Celtic Race’ in Ireland, which pervaded well past the eventual

establishment of the deeper three-age chronology at the end of the century (Wood-Martin 1895; Wood-Martin 1886), perhaps peaking in the 1930s with the Harvard Mission (Carew 2012).

Meanwhile, the search for the origins of the continental Celts had been ongoing in pre-Roman Iron Age archaeological sites of northern and western Europe. They were soon identified at La Tène in Switzerland, whose material culture was also seen in an Etruscan site, and thus could be crucially linked to documented Celtic migrations southward into Northern Italy. Further systematic archaeological divisioning of the Iron Age defined a preceding Hallstatt period, from which La Tène culture was seen as a direct derivative. Both cultures had widespread distributions across western Europe and soon became synonymous with the classical concept of Celts (Cunliffe 1997). In particular, the spread of La Tène material could be convincingly related to historical migrations of Celts into central and eastern Europe, providing fodder for popular invasionist models of archaeological cultural spread. In the west of the continent, the confirmed historic presence of Celtic languages across Ireland, Britain and Iberia put an impetus on archaeologists to find evidence of similar undocumented Celtic expansions into these regions, deriving from the core zones of Hallstatt and La Tène cultures in Central Europe.

The required evidence was soon provided, as many antiquities from parts of both Britain and Ireland were seen to have broadly Hallstatt or La Tène affinities. This was especially true for the La Tène style of art, spread across much of Britain and the northern two-thirds of Ireland; and from which inspiration could also be seen in later early Christian and medieval works of the region. A series of continental invasions, again echoing the pervasive *Lebor Gabála Éirenn*, was then invoked to explain the varying distributions of the Hallstatt and La Tène cultures across the islands, as well as the different branches of the Celtic language family present. Moreover, the construction of hillforts, which is now known to have begun in the preceding Bronze Age period, were seen to represent the militaristic nature of such migrations (Hawkes 1931). It seemed the Gallic invasion hypothesis, put forward by the linguist Edward Lhuyd almost 200 years previously, had been successfully confirmed by the fledgling field of scientific archaeology, and has remained a dominant theory on Irish origins to this day.

However, in more recent years, this simplistic attempt at packaging the Celts as an invasive ethnic group, defined by shared language and material culture, has been gradually dismantled (Cunliffe & Koch 2012; Mallory 2013). This was brought on in part by reactionary questioning of the popular invasionist models of cultural change during the 1960s, but also through the required reassessment of established chronologies with the advent of radiocarbon dating, as well as continued archaeological discovery. Indeed, the now apparent continuity across the Bronze Age to Iron Age transition in Ireland (Mallory 2013) has led many experts to conclude that there is little evidence to suggest large-scale migration accompanied the appearance of La Tène and Hallstatt artefacts on the island (O'Donnabháin 2000; Raftery 1994), necessitating a radical rethink of the origins of both proto-Celtic and its historic speakers, including the population of Ireland.

Pots, People and Proto-Celtic: Questioning Migration's Role in Cultural and Linguistic Change

A fresh approach to Celtic origins has been provided by Cunliffe and Koch in a series of books and essays (Cunliffe 2001; Cunliffe & Koch 2012), who proposed an Atlantic origin for the language family, likely as a lingua franca during the Later Bronze Age. The possibility of an earlier Chalcolithic spread along the seaboard, through the Beaker Culture, is also considered. Others have envisioned a still earlier Neolithic dispersal (Renfrew 1990). Though still heavily debated, these Atlantic theories find archaeological support in the extensive cultural networks that have existed between the extreme western populations of Europe since at least the Neolithic period, and which reached their height during the Late Bronze Age. Linguistic evidence is provided in Tartessian inscriptions from the 8th century BC in southwest Iberia, argued by Koch to be the earliest attested form of Celtic; and in the known western distribution of the language family during the classical period.

At this juncture it is important to note that the search for the origins of a language family and those of its speakers are not necessarily one and the same. Indeed, while linguistic change does tend to involve some level of population movement, the spread of proto-Celtic may not have necessitated substantial population turnover, and alternative vectors, such as trade or imperialism, can also be hypothesised. Indeed, in avoidance of the simplistic and circular cross-disciplinary reasoning that led to earlier invasionist models, contemporary archaeological and linguistic thought is cautious in any endeavours to link together language, cultures and peoples. It is stressed that language can spread without a severe break in culture, while many historical cultural and societal shifts may not have induced language change. Crucially, neither field requires that population migration be the main catalyst for either type of change, though most concede a mass migration event, if demonstrable, would induce both.

Given that neither the Atlantic Bronze Age, nor Hallstatt and La Tene horizons have produced archaeological evidence of such a demographic upheaval in Ireland, both arguments for proto-Celtic entry rest more heavily on the intensity of trade, cultural and elite networks, and militaristic sites during the periods. The potential of ancient genomics in providing a clear demographic framework is thus key for the progression of such debates, as it can pinpoint clear horizons where mass migration, and thus language change *must* have occurred. These reference points can then be used for further speculation on language spread based on archaeological trends and known linguistic distributions.

Theories of large-scale migration as the main driver of cultural and linguistic spread, only now testable through aDNA, have been out of vogue since the backlash of the 1960s, suggested to have been brought on in part by the politics of the time (Lozny 2011). New generations of thinkers have invoked acculturation, the gradual adoption of exogenous culture through contact and trade, and diffusionism, gradual demic expansion with no particular migratory impetus, to explain observable discontinuities in the archaeological record. A “Pots versus People” dichotomy hardened between these views and more traditional invasionist models, with the battleground falling across the boundary of almost every observed cultural transition in Europe, right up to historical Anglo-Saxon settlement of Britain (Hills

1993; Esmonde-Cleary 2002). In Ireland, as across Europe, several profound cultural shifts became the subject of particularly intense debate. These correlated well with the major transition points of the aforementioned three-age system proposed in the previous century.

The latest of these, the Iron Age transition discussed above, as catalysed in Western Europe by Hallstatt and La Tene culture, was clearly a point of friction and heavily influenced by modern ethnic identities. However, two more dramatic cultural upheavals had occurred in the more distant past, that from hunting and gathering to agriculture (Mesolithic to Neolithic), and that from stone tools to metallurgy (Neolithic to Bronze Age). The demographic nature of these transitions would become just as contentious, as researchers attempted to address the deeper origins of European populations. Theories ranging from complete continuity between Palaeolithic and present-day peoples, to complete replacement of indigenous groups by farming and metal-working populations were suggested, each selectively drawing support from observed distributions of Indo-European languages and patterns of modern genetic variation.

Major Horizons of European Archaeology

The single most important development in these debates was the advent and progression of radiocarbon dating from the 1950s onwards (Libby 1954), allowing for the first time a full appreciation of the true scale of human prehistory. Crucially, the technology removed archaeologists' dependence on time-consuming interpretation of relative chronologies, freeing up collective headspace in which to consider the mechanisms of cultural, demographic and linguistic change. The Pleistocene to Holocene boundary, signalling the end of the last glacial period, was soon tied down to approximately 10,000 BC (Taylor & Bar-Yosef 2014) and numerous Palaeolithic (Old Stone Age) assemblages from Europe were securely dated to before this point, confirming the ancient association humans have had with the continent. The Holocene period can be used to mark the beginning of the Mesolithic (Middle Stone Age) in Europe, whose post-glacial, pre-pottery populations were, as in the preceding Palaeolithic, completely dependent on a hunting and gathering economy.

This mode of subsistence persisted until the later Neolithic (New Stone Age), which had been defined early in its study, not only by the presence of pottery, but also the appearance of domesticated animals, cereals, farming tools and new forms of settlement structure. Overall it was interpreted as a complete revolution of lifestyle, with agriculture now taking precedence over hunting and gathering, leading to rapid population growth. Prior to radiocarbon technology, the main theories of Neolithic spread into Europe involved rapid invasions traced through craniometric and ceramic data (Coon 1972). By the 1960s, a comprehensive model of agricultural expansion from the Levant to Europe had been proposed, detailing a much older and longer timescale for the Neolithic period than had previously been suggested (Clark 1965). This encouraged more gradual demic diffusion models of agricultural spread (Renfrew 1990), supported by Cavalli Sforza's famous principal components plots, detailing gentle gradients of European genetic variation (See Introductory Chapter; Ammerman & Cavalli-Sforza 1984). In the past

several decades Neolithic chronology has been increasingly expanded on, from its entry into the Balkans from Anatolia circa 6,500 BC, to its arrival in Britain and Ireland 2500 years later. However, the demographic nature of the transition has been continuously debated, and made all the more contentious by the apparent diversification of the Neolithic package in different regions, arguably impacted by variable interactions with indigenous groups. Recently, ancient DNA research has brought migrationist theories firmly back into the limelight (Gamba et al. 2014; Lazaridis et al. 2014; Skoglund et al. 2014; Haak et al. 2015), with large scale movements into many regions apparent, but studies have up to this point neglected the northwest of the continent. Moreover, a number of regions have demonstrated continuity across the transition (Jones et al. 2017), calling into question the universality of such models.

Importantly, a linguistic dimension to the spread of agriculture has also been proposed, summarised in the ‘Anatolian Hypothesis’, which describes the entry of Indo-European language to Europe from West Asia at this horizon (Renfrew 1990). In this scenarios the isolate Basque language can be viewed as a relic language from the Mesolithic period. However, this conflicts with an older model of Indo-European introduction to Europe, the Kurgan Hypothesis, which was developed in the 1950s (Gimbutas 1956). Here, the proposed urheimat of the language family was the Pontic Steppe, among Chalcolithic and Early Bronze Age groups unified by their usage of Kurgan burial mounds. Through this lens, pre-Indo-European languages of the south, including Basque, were now seen as surviving Neolithic languages. The theory was expanded on in the 1970s (Anthony 2010) pinpointing the pastoralist Yamna culture (4,000-2,300 BC) as the key horizon, which subsequently spread into central Europe through the vector of Corded Ware culture (2,900-2,350 BC). Anthony speculated that Bell Beaker sites, discussed below, may have then aided in the dispersal of Yamna dialects west into southern Germany, where Proto-Celtic could have developed, somewhat at odds with the proposal of an Atlantic origin for the subfamily. While debated, migration was generally favoured over diffusion as the model of this cultural spread, which finds strong confirmation in recent aDNA studies (Allentoft et al. 2015; Haak et al. 2015). These studies emphasise mass migration into north and central Europe from the steppe alongside near complete turnover of male line lineages. However, once again the northwest has remained unsampled, while certain regions such as Iberia and Hungary have demonstrated continuity across the metal-working horizon (Gamba et al. 2014; Martiniano et al. 2017). Despite, or perhaps in part due to this progress in demographic understanding for both the agricultural and metallurgical transitions in Europe, the linguistic debate remains as lively as ever, though opinion is gradually shifting towards the Kurgan hypothesis, at least when considering the major northern and western European branches of Indo-European.

It is not surprising that the Copper and Bronze Ages were pinpointed early on as horizons for mass linguistic and demographic change, given the extensive cultural upheavals that occurred across Europe with the introduction of these technologies. In the west, antiquarians had long been aware of the pervasive Bell Beaker Culture (2,900-1,800 BC), defined by its distinctive vessels, which tended to be associated with single burials under barrows, echoing the contemporaneous Corded Ware culture of the

east. Copper and gold artefacts, along with archery equipment (Fig. 1.1C), were also characteristic of these Beaker assemblages, which hinted at intensified social stratification (Cunliffe 2008). However, on the Atlantic coast, Beaker-ware was also found associated within older collective megalithic tombs of the Neolithic, while in Ireland its appearance is accompanied by new types of megalithic construction, wedge tombs. While its origin(s) remain contentious, and its mode of spread even more so, one consensus that can be reached on Bell Beaker culture is the unprecedented level of connectivity it introduced across much of western and central Europe during the third millennium BC. Beaker culture gradually gave way to a diversity of regional Bronze Ages. This includes the sequential Únětice, Tumulus and Urnfield cultures of central Europe (2,300-700 BC), which are succeeded by the potentially Celtic Hallstatt Iron Age. Further west, the Atlantic Bronze Age (1,300–700 BC), another suggested incubator for proto-Celtic, emerged through heightened connectivity along its seaways, with influence from the further inland Urnfield and Hallstatt cultures also apparent.

The Irish Perspective

From the above sections it is clear that a complex array of candidate progenitors for Europe's (and by proxy Ireland's) peoples and languages has been assembled, including, among others, indigenous hunter-gatherers, Anatolian farmers, steppe pastoralists, western Beaker-folk and Hallstatt Celts, all with their own advocates and opponents. The Irish archaeological tradition, which had blossomed across the 20th century, did not escape such debates. The three-age system in Ireland had been fully fleshed out by the 1930s, which saw the initiation of systematic archiving schemes and extensive excavation projects across the island (Mahr 1937). The oppressive narratives of Celticism were shaken, and the field was moved towards non-speculative, scientific reasoning, based on the meticulous collection of data (Cooney 1995), which could then be used to ask the correct questions, rather than to provide the correct answers. Through these efforts large datasets were built and a new chronology of Irish prehistory (Fig. 1.1A), was gradually fitted.

The earliest periods of Irish prehistory, the Palaeolithic and Mesolithic, have remained throughout their study the most elusive. Indeed, the existence of an Irish Palaeolithic remains contentious, though recent evidence suggests there may indeed have been some human presence during the last glacial period (Dowd & Carden 2016). The possibility of an Irish Mesolithic was treated with either skepticism or indifference until a series of ground-breaking excavations at Mount Sandel in the 1970s, which provided radiocarbon determinations dating back almost 10,000 years, well before the onset of Neolithic technologies (Woodman 2015). There is now little doubt that a prolonged Mesolithic occupation of the island occurred (comprising over a third of Irish prehistory), raising the question of what genetic input these people contributed to both the succeeding Neolithic era, as well as to modern Irish populations. Moreover, a break in material culture between the Early and Late stages of the Mesolithic calls into question the extent of demographic continuity across the period.

The search for Ireland's Neolithic, based on that observed on the continent, was also not entirely successful at first, with scant evidence of characteristic rectangular dwellings (Smyth 2014). Moreover, the island's most enduring Neolithic legacy, its megalithic monuments (Fig. 1.1C), were incorrectly interpreted as Bronze Age structures until the advent of radiocarbon dating. However, gradually a robust picture of the Irish Neolithic has come to light. Beginning approximately 4,000 BC and stretching across 1,500 years, it was not the brief period of colonisation followed by an indigenous comeback that some archaeologists had proposed (Piggott 1954) but an entirely more enduring affair. Ireland's Neolithic can be now appreciated as belonging to a wider Atlantic sphere of influence, defined by megalithic construction, likely brought about through some level of migration. However, models of acculturation have not been fully rejected, despite the abruptness with which the entire package of Neolithic technologies arrives in Ireland and the apparent lack of contribution of Irish Mesolithic peoples to the new material culture (Mallory 2013; Woodman 2015). An important consideration for colonisation models has been the nature of the source population, who themselves could be result the of Mesolithic acculturation, as opposed to the descendants of a stream of colonists from further afield. The density of Mesolithic populations along the Atlantic littoral is a key factor in this (Cunliffe 2008). Finally, it is important to emphasise that the Irish Neolithic cannot be considered a wholly homogenous entity, but had witnessed over its duration various transformations and diversifications. This leaves open the possibility of multiple colonisation or acculturation events, conceivably from different source populations (Sheridan 2010).

The Bell Beaker phenomenon of western Europe proved to be another elusive horizon in Ireland's archaeological record, and, for a long time, was not believed to have extended to the island at all (Mallory 2013). However, by the 1970s new discoveries had overturned this view entirely, with the Beaker culture now believed to represent the last great 'invasion' of Ireland's prehistory, arriving on the island in two separate waves, one into the southwest and the other, via Britain, into the east, displacing Neolithic communities (Herity & Eogan 1977; Harbison 1975). Indeed, for some scholars this was the only break in the archaeological record dramatic enough of to warrant a complete change of language (Macalister 1949). Again, the *modus operandi* for the spread of Beaker culture to Ireland followed the popular trends of the time. Earlier models were rooted in migration, either by whole communities, or smaller groups of social elites or craftsmen, which were later countered with acculturation hypotheses (Waddell 2010). What was agreed was that the earliest introduction of metallurgy to the island (~2,500 BC) was associated with the Beaker horizon, perhaps most strikingly at the southern copper mines of Ross island. From 2,200 BC onwards, beaker vessels give way to new series of ceramics (Irish Bowls, Vases and Urns) and bronze working emerges. However, contrarily, it is at this point that burials in Ireland begin to resemble the more classic individual Beaker graves of northwest Europe, though some use of older megaliths persisted (Mallory 2013).

The Early Bronze Age is seen to develop into the Middle and Late Bronze Ages, during which time Ireland is increasingly subsumed into the wider Atlantic network. From a linguistic perspective, the key

question across these periods is whether any evidence for migration into the island exists that could provide an avenue for Celtic language entry. However, as the Bronze Age progresses, persistent and intense interactions between Ireland and Britain, and to a lesser extent the continent, make clear windows for such a specific event hard to define. The parallel developments of structural, cultural and ritual features across both islands, such as hillforts, the abandonment of ceramics, hoard deposition and burial rites, may warrant the treatment of the two islands as a fluid demographic unit (Mallory 2013). The unprecedented level of interdependency between both islands and the continent forms part of the appeal of the 'Atlantic lingua franca' model of proto-Celtic formation (Cunliffe & Koch 2012).

By the Iron Age the Atlantic Bronze Age system had begun to fragment, with Iberia entering into a Phoenician sphere of influence (circa 900 BC), and Ireland becoming relatively isolated after 700 BC (Henderson 2007). The limited archaeological finds from the period, particularly in the south, suggested economic stagnation and demographic collapse on the island, though recent dating schemes have begun to challenge the idea of such an extensive Dark Age (Corlett & Potterton 2012). Unlike the majority of continental Europe, bronze-working remained the dominant metal industry in Ireland and other Atlantic regions, feeding into the more generalised view of staticity and continuity in Europe's western peripheries during the period (Henderson 2007). However, broader continental phenomena, such as Hallstatt and La Tène artistic styles did not leave these regions untouched, although, as discussed above, the relative contribution of migrants and local populations to their development remains debated (Mallory 2013)

The Contribution of Genetics

Despite continued archaeological developments, no full consensus could be reached on the demographic nature of any transition, no matter how disruptive. New isotopic technologies offered some hope through the identification of geographical outliers, who could be considered as recent migrants into a region. However, as isotopic signals of long distance migration do not survive past a generation, these lack the resolution to confirm long term, large-scale demographic change. A more durable marker of demographic discontinuity is the human genome, which has been used to address questions of human prehistory since the early 20th century. Ireland has been included in such investigations, though up until now virtually all research has been restricted to the modern population, which opens up its own interpretive pitfalls.

The island formed a key reference point in Cavalli-Sforza's famous principal component analyses of European allozyme data, for which the first major component of variation was revealed to fall along a southeast to northwest geographical cline, which many saw as evidence for demic diffusion into the continent during the Neolithic revolution (Ammerman & Cavalli-Sforza 1984). Moreover, in his *History and Geography of Human Genes* an F_{ST} tree, of European populations, based on 22 genes, placed Ireland in a cluster with Scotland, which together formed an outgroup to the majority of other populations tested (notable exceptions being Basque, Icelandic, Sardinian, Finnic and Balkan groups). This tree emphasised

the segregation of genes with language, and served to distinguish Celtic and Germanic speakers of Britain and Ireland (Cavalli-Sforza et al. 1994).

Later work on uniparental markers, placed emphasis on the Mesolithic input into modern Irish populations against this backdrop of Neolithic diffusion (Hill et al. 2000; McEvoy et al. 2004). Studies of mtDNA diversity focused on perceived patterns of expansion of hunter-gatherers from post-glacial southwestern refugia and later of farming populations from Anatolia, although the overall lack of maternal structure seen in Europe called into question the extent of inference which could be drawn (Richards et al. 1996; Torroni et al. 1998; Simoni et al. 2000; Torroni et al. 2001). Y chromosome haplotypes revealed more substantial frequency differences among populations. In particular “haplogroup 1” (Jobling & Tyler-Smith 1995), now known to closely correlate with R1b-M269, was seen to fall along an approximate southeast to northwest gradient in Europe, with highest frequencies nearing 100% seen in the west of Ireland (Hill et al. 2000). Moreover, within Ireland, the haplogroup was seen to associate closely with surnames of Gaelic origin.

A coalescence date of 4,200 BP was estimated for haplogroup 1 within Ireland, falling almost directly on the Neolithic to Bronze Age transition. However, with the observed gradient mirroring that of autosomal data, and agriculture remaining the main interpretative focus for such trends, the most parsimonious explanation appeared to be migration of Neolithic farmers. Specifically, lower penetrance of these groups and higher survival of Mesolithic male lineages, such as haplogroup 1, in the extreme west of the continent was suggested. Notably, high frequencies were also seen in the non-Indo-European Basque population, an outlier on PCAs of Y chromosomal diversity, though falling closer to Insular Celtic populations than to other Iberians (Rosser et al. 2000). While later studies rejected such an ancient origin for the haplogroup based on narrower coalescence times, explanations were still centered on Neolithic migrations, with R1b-M269 lineages now suggested to have surfed to high frequencies on the peripheral waves of agricultural expansion (Balaesque et al. 2010; Myres et al. 2011).

Regardless of exact interpretation, the theme of ancient and shared ancestry between the Atlantic populations of Europe became a recurring one. For instance, the Y chromosome ‘Atlantic Modal Haplotype’ (Y Chromosome Consortium 2002), a subclade of R1b-M269 that segregates with the downstream mutation R1b-L151, was identified and so named due to its peak frequencies in Portugal, Northern Spain and western regions of Ireland and Britain. Fitted with similar trends in mtDNA frequencies (McEvoy et al. 2004), the observations were taken to support a primarily Atlantic, rather than central European origin for the speakers of Celtic languages, though any attempt to ground these associations in time remained purely speculative. Large autosomal SNP datasets did nothing to elucidate matters, revealing only that genetic variation in Europe mirrored the continent’s geography (Novembre et al. 2008) with Ireland falling, as expected, near Britain, to northern and western extremes of variation. Further surveys highlighted the population’s relatively low levels of genomic diversity, which were ascribed to the island’s geographic isolation, rural history and episode of famine (O’Dushlaine et al.

2010). Meanwhile, studies of Y chromosome distributions moved away from deeper prehistory to address more recent phenomena, examining the impact of Viking migrations (McEvoy et al. 2006), patriarchal Gaelic Dynasties (Moore et al. 2006; McEvoy et al. 2008), and their correlations with surname. It was not until the advent of ancient genomic technology that accurate time depths of European Y chromosome diversity were revealed, and the source was, for the most part, entirely unexpected. High frequencies of R1b chromosomes had in fact swept through Europe from the east, much more recently than the Neolithic, during mass steppe migrations of the Bronze Age (Allentoft et al. 2015; Haak et al. 2015). While some modern genetic evidence, such as the third principal component of European variation identified by Cavalli-Sforza (Piazza et al. 1995), had hinted at some dispersal into Europe from above the Caspian Sea, the scale of the inferred migration was entirely unprecedented, agreeing strongly with the Kurgan Hypothesis of Indo-European spread. Interestingly this component also saw a second lesser peak in the west of Ireland.

The potential of ancient genomes in addressing the long-standing controversies of prehistory has been effectively demonstrated. However, inconsistencies in interpretation, still exist, such as the near fixation of R1b in Basque populations, despite their pre-Indo-European language. Indeed, aDNA research has the potential to create new questions as fast as it answers old ones, as any good science should. For this reason, focused regional studies are critical for the progression of the field. Ireland, and the northwest of the continent as a whole, had been neglected from surveys of ancient genetic variation. This chapter seeks to rectify this, aiming for the first time to provide a holistic demographic framework spanning the island's prehistory, which can be used to inform both archaeological and linguistic interpretations and debate. To this end, the main trends in ancient Irish genetic variation are presented and discussed here, using dataset of 93 individuals sampled from across all major periods of Irish prehistory (Fig. 1.1).

Methods

Sample Processing

A full description of molecular methodology and bioinformatic data processing is provided in Appendix II. Briefly, we extracted DNA using a modified silica column based method (Yang et al. 1998; MacHugh et al. 2000). DNA fragments were incorporated into next-generation sequencing libraries (Meyer & Kircher 2010; Gamba et al. 2014), amplified with distinct indexing oligos, purified (MinElute PCR Purification kit, Qiagen), pooled, and sequenced with Illumina technology. Sample reads were subsequently trimmed of adapter sequences by using cutadapt (Version. 1.2.1) (Martin 2011), filtered for a length of 34 bp and aligned to the human reference genome (hg19/GRCh37), with the mitochondrial genome replaced by the revised Cambridge reference sequence using BWA (Version 0.7.5) (Li & Durbin 2009). Reads were filtered for mapping quality, and clonal amplification products were removed with SAMtools (Li et al. 2009). Indel realignment was performed with GATK (McKenna et al. 2010) and the qualities of the final two base pairs at read ends were soft-clipped. Alignment or realignment and filtering of published BAM or FASTQ file data was carried out in a similar manner for the datasets released in Keller et al. 2012, Gamba et al. 2014, Lazaridis et al. 2014, Olalde et al. 2014, Skoglund et al. 2014, Allentoft et al. 2015, Günther et al. 2015, Haak et al. 2015, Jones et al. 2015, Mathieson et al. 2015, Olalde et al. 2015, Broushaki et al. 2016, Fu et al. 2016, Hofmanova et al. 2016, Kılınç et al. 2016, Lazaridis et al. 2016, Martiniano et al. 2016, Schiffels et al. 2016, González-Forbes et al. 2017, Jones et al. 2017 and Martiniano et al. 2017.

Data authenticity was established through molecular control analysis, examination of molecular damage patterns, and levels of mtDNA contamination. Molecular sex determination, mtDNA, and Y chromosome analysis were carried out following established methods.

Variant Calling and Dataset Preparation

For ADMIXTURE and PCA analyses, pseudo-haploid genotypes at 594,896 known SNP positions from the Human Origins Panel (Lazaridis et al. 2014) were called in all ancient samples using the Pileup tool from GATK (McKenna et al. 2010) (See Appendix II). For PCA analysis ancient sample genotypes were merged with those of 604 modern west Eurasian individuals from the Human Origins dataset, representing 71 populations. Ancient samples with less than 20,000 genotypes represented were discarded from analysis.

For ADMIXTURE analysis ancient genotypes were merged with the entire modern Human Origins dataset (1941 individuals from 198 populations). In addition, 64 modern individuals from the SGDP, unrepresented in Human Origins, and 400 from the EGDP datasets (2 removed due to duplication in Human Origins dataset) were included to provide the largest sample of global diversity available. Finally,

the lack of modern Irish individuals from the above datasets warranted the inclusion of 15 newly-genotyped Irish individuals to be released in Byrne et al. (submitted), bringing the total number of modern individuals to 2420.

For FineSTRUCTURE analysis a second modern SNP dataset was used based on samples from (Li et al. 2008; Behar et al. 2010; Henn et al. 2011; Busby et al. 2012) genotyped using the Illumina 660W array in (Hellenthal et al. 2014). This consists of a total of 474,491 analysed autosomal SNPs in 1530 individuals from 95 different populations. Crucially this included Ireland, Wales, Scotland and England. SNP positions were mapped to build 37 (GRCh37) of the human genome and genotypes with a minor allele frequency below 0.5% or a genotyping rate below 99% across all samples were excluded. Diploid genotypes at these sites were then called using the UnifiedGenotyper tool from GATK (McKenna et al. 2010) in six high coverage ancient genomes. These included two Irish individuals presented here, two Hungarians from the Neolithic and Bronze periods (Gamba et al. 2014), and two individuals from Neolithic Germany and Mesolithic Luxembourg (Lazaridis et al. 2014). These were filtered for a genotype quality of 30 or above and a depth of coverage of 10X or above. In order to maximize the number of SNPs for analysis, each high coverage sample was merged separately with a set of modern West Eurasian populations from the modern dataset, keeping only SNPs confidently called across all individuals within each merge (238,762-356,240 SNPs) for analysis.

All dataset manipulation and merging was carried out in PLINK v1.90 (Chang et al. 2015).

ADMIXTURE

To estimate the different components of ancestry present in ancient Irish and Eurasian individuals in relation to diverse modern populations we used a model-based clustering approach implemented by the program ADMIXTURE v.1.23 (Alexander et al. 2009). Several filters were imposed on the dataset before analysis. SNPs were pruned for sites in strong linkage disequilibrium using PLINK v1.90, with a sliding window size of 1000 variants, a step size of 50 variants, and an r^2 threshold of 0.25 (--indep-pairwise 1000 50 0.25). Remaining variants with a MAF below 1% were subsequently removed. Related individuals (19 modern and 6 ancient) and those with genotype missingness over 55% were discarded. All modern individuals were chosen to have a missingness below 3.5%, with the exception of the Irish population (n=10) taken from Byrne et al. submitted, who only possessed genotype information for 51.35% of sites, but whose geographical relevance necessitated their inclusion. After the above filtering was completed a final pruning of variant sites with higher than 5% missingness across all 2692 individuals was carried out. This left a total of 145,098 SNPs for analysis on 2401 modern and 291 ancient individuals, including 83 ancient Irish.

ADMIXTURE was run with cross-validation enabled using the --cv flag for all ancestral population numbers from K=2 to K=16. This analysis was replicated 20 times over. CV error was observed to drop

with increasing values of K until K=12 at which point error values began to rise again. The results for the replicate with the highest achieved likelihood at K=12 are displayed in Fig. 1.2.

Principal Components Analysis

Principal component analysis was performed on present-day west Eurasian populations using smartpca version 10210 from EIGENSOFT (Patterson et al. 2006; Price et al. 2006), with 464 ancient genomes, including 93 Irish samples, projected onto the modern variation (lsqproject: YES option). The first two components of variation were then plotted against one another for a series of time transects using the ggplot2 package in R (Wickham 2009) (Fig. 1.3). Modern populations were grouped into larger geographical divisions and coloured appropriately. Ancient individuals were also classified using this same geographical colour key.

Haplotype-based Analysis with ChromoPainter and FineSTRUCTURE

Following a similar approach to Lazaridis et al. 2014, FineStructure v.2 (Lawson et al. 2012) was used to investigate haplotypic sharing between high coverage ancient and modern individuals. For each dataset, the genotype data was split by chromosome and phased genotypes with SHAPEIT v2.r778 (Delaneau et al. 2011). Haplotype files were converted to ChromoPainter format using “impute2chromopainter.pl” and a recombination map was created using “makeuniformrecfile.pl” both available at <http://www.paintmychromosomes.com/>. ChromoPainter analysis was then run in linked mode, first by estimating “mu” and “Ne”, followed by estimating “c” and using “Chromocombine” to create a genomewide ChromoPainter output for all individuals. Median chunk donation from each ancient individual to populations in the modern dataset was calculated in order to ascertain patterns of haplotypic affinity and linear regression analysis was performed using the “visreg” package (Breheny & Burchett 2012) built for the R programming language (Fig. 1.4A). To allow for a visual comparison of the geographical distribution of haplotype sharing between ancient samples and modern populations, the median of the chunks donated by each ancient individual to each separate population was first calculated. These values were then normalised by scaling between 0 and 1 and plotted as interpolated maps using the Geostatistical Analyst tool of the ArcMap component of the ArcGIS (v10.1) software suite (ESRI) (Fig. 1.4B).

Results

DNA extraction and sequencing

Sufficient endogenous DNA content was retrieved from 93 of the 140 individuals screened for this study to warrant further shotgun sequencing to higher coverages. A summary of the final dataset created, with respect to time depth and coverage achieved, is shown in Fig. 1.1A. The majority of samples were sequenced to between 1X and 2X coverage, with only 5 samples falling below 0.5X. Furthermore, a Neolithic and an Early Bronze Age individual were chosen for higher coverage sequencing to approximately 10X, to allow for direct diploid analysis. The final dataset spanned all periods of Irish prehistory, including two Mesolithic, 35 Neolithic, 29 Chalcolithic/Bronze Age, 25 Iron Age and two early modern individuals, as well as two individuals of unknown date from Late Bronze Age wetland contexts (See Appendices I and II for further information). Molecular sex determination was possible for all samples (Table II.2), with male to female ratio across time periods visualised in Fig. 1.1A. Data authenticity was confirmed using several approaches (Appendix II), with contamination estimates below 1% observed for all samples (Table II.2). The exception to this were the two lowest coverage samples, Derrymaquirk20 and Carrowkeel534 (contamination estimates between 2.91-4.02%), who were excluded from all further analyses, bar PCA and mitochondrial haplogroup assessment.

Uniparental analysis: Divergent patterns in male and female lineage history

Y chromosome haplogroup assignment was possible for all 53 male samples sequenced. Mesolithic and Neolithic individuals were found to place overwhelmingly within haplogroup I2a (Electronic Data Table S5-7), one of the dominant lineages of the European Mesolithic and commonly observed in other western Neolithic populations (Haak et al. 2015; Mathieson et al. 2015; Olalde et al. 2017). As expected, Mesolithic samples show more basal lineages with respect to the Neolithic cohort, a significant number of whom placed within the subclade I2a2a1-M284, found almost exclusively in Britain today (International Society of Genetic Genealogy 2017; FamilyTreeDNA 2017). A near complete turnover is then witnessed in the Chalcolithic and Early Bronze Age, with virtually all samples from this time onwards belonging to the R1b-L151 haplogroup, associated with the Atlantic Modal Haplotype. This lineage forms a downstream branch of R1b-M269, the most prevalent haplogroup in western Europe today (Myres et al. 2011). It was possible to place a further majority of samples (81.8%) into the subclade R1b-L21, a haplogroup whose distribution shows a steep and somewhat restricted peak in modern Ireland, where it accounts for almost half of male lineages (Myres et al. 2011). Only one Y chromosome examined fell outside R1b-M269 after the Neolithic period, belonging to an Early Bronze Age individual from the southwest, *Killuragh1*, who placed within I2a2a1-M284.

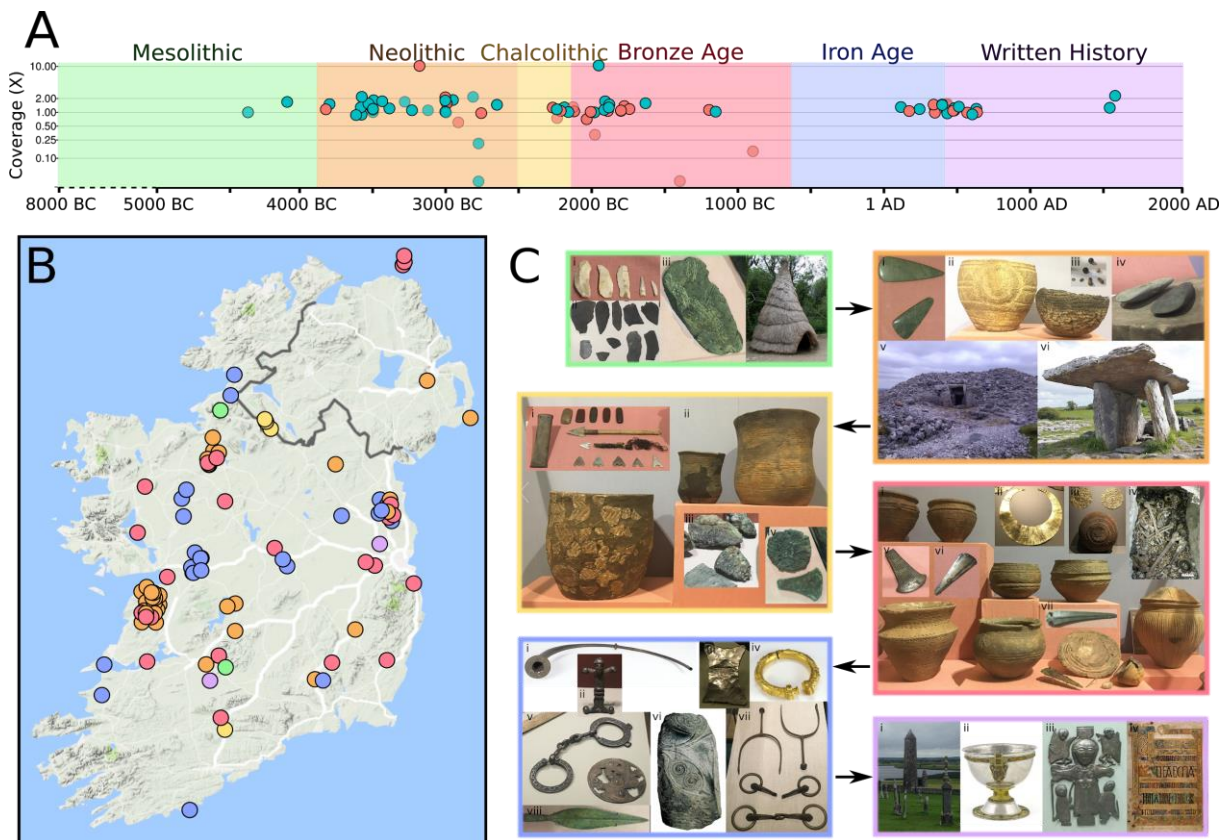


Figure 1.1. Summary of ancient Irish individuals studied alongside their geographical and archaeological contexts. **A)** A timeline of samples plotted with respect to genomic coverage (exponential axis). The key periods of Irish human history are highlighted. Samples are coloured by molecular sex (red: female, blue: male). Translucent points lack a direct radiocarbon date. **B)** Geographical locations of samples studied. Locations are approximate as multiple samplings from the same sites necessitated ‘jittering’ of plot points. Samples are coloured based on archaeological context, using the key outlined in the above timeline. **C)** Key cultural developments of the various periods sampled, outlined using the same colour key. The Mesolithic box shows microlithic (i) and later macrolithic (ii) technology, alongside recovered fish-traps (iii) and a reconstructed Mesolithic dwelling at Mount Sandel (iv). Neolithic artefacts include jadite (i) and stone axes (iv), ceramic-ware (ii), clay, stone and bone jewelry (iv). Megalithic passage (v) and portal (vi) tombs also appear at this horizon, represented by Carrowkeel and Poulnabrone monuments respectively. Chalcolithic developments are characterised by archery equipment (i), beaker vessels (ii), copper ore from mines in Cork (iii), a copper cake and axe (iv). The Bronze Age sees the emergence of food vessel pottery, a range of which is seen across (i). Examples of gold (ii and iii) and bronze (v, vi and vii) working are also shown, as well as single cist inhumation burial, typical of the earlier part of the period. In the Iron Age, bronze working persists over the period (i, ii, v, vii), with some evidence of iron metallurgy (viii). Many metal pieces are decorated in the Celtic La Tene style (i, ii, iv, v), also expressed through stone carving (v). Roman contacts (iii) are seen at the end of the period. Written history begins with the emergence of early Christian monastic sites (i), art (ii, iii) and manuscripts (iv). (Map Credit: Google. Photos taken by L.M.C at NMI, Kildare Street or sourced from NMI website and Wikimedia Commons)

Mitochondrial haplogroup analysis revealed more continuous trends in lineage distributions. The two Mesolithic individuals both belonged to haplogroup U5, currently the only haplogroup that has been identified in post-glacial Western Europe prior to agriculture. In contrast, the Neolithic of Ireland shows a range of maternal lineages, including H, J, K1a, T2 and X2, all associated with genetic influx from West Asia during the spread of farming (Richards et al. 2000; Bramanti et al. 2009; Haak et al. 2015). Several U5 lineages are also noted among the sample. This diversity prevails over the Neolithic to Bronze Age transition, with no obvious disruption in maternal lineage frequencies observed, though some new haplogroups are observed (I1, U2e, W4). The Iron age period also presents a similar distribution of the major haplogroups, with several more newly observed lineages, such as haplogroup V, appearing at this time.

ADMIXTURE

Three strands of ancestry

ADMIXTURE analysis provides an avenue to explore the major autosomal components of ancestry segregating across Irish prehistory, with reference to other ancient Eurasians and modern global populations, providing a more robust view of prehistoric variation. The model-based approach allows the decomposition of individual genomes into coefficients contributed by a set number (K) of ancestral populations. Fig. 1.2 displays plots of estimated coefficients ($K=12$) for ancient and modern populations, averaged across individuals, and panelled by time period. These partition similarly to previous analyses, with three major ancestral coefficients (red, orange and teal in colour) manifesting in west and central Europe across time (Allentoft et al. 2015; Haak et al. 2015). Moreover, three substantial turnovers in the genetic makeup of Ireland can be visualised, correlating well with the arrival of each of these components to the island, discussed in turn below.

European Hunter-Gatherer Ancestry

The first European component, colored red, forms the near totality of ancestry on the continent, prior to the arrival of agriculture (Fig. 1.2A). This includes both Irish Mesolithic samples, who are indistinguishable from the majority of other European hunter-gatherer individuals plotted. However, a notable exception is the Palaeolithic sample from El Miron in Spain, belonging to the Magdalenian culture, who, within a majority “red” background, possesses a mosaic of components that dominate in diverse modern populations, such as Papuans, East Asians and Indians. This noisy signal is likely to represent deep pre-glacial population structure between Iberian refugia populations and those further east in Italy and the Balkans (Fu et al. 2016). The signature is also seen to a lesser extent in the earliest Mesolithic sample of the region, but disappears in later Spanish samples. Importantly, Irish Mesolithic individuals do not show any detectable level of this ancestry, displaying profiles more similar to Palaeolithic samples from Switzerland and Italy.

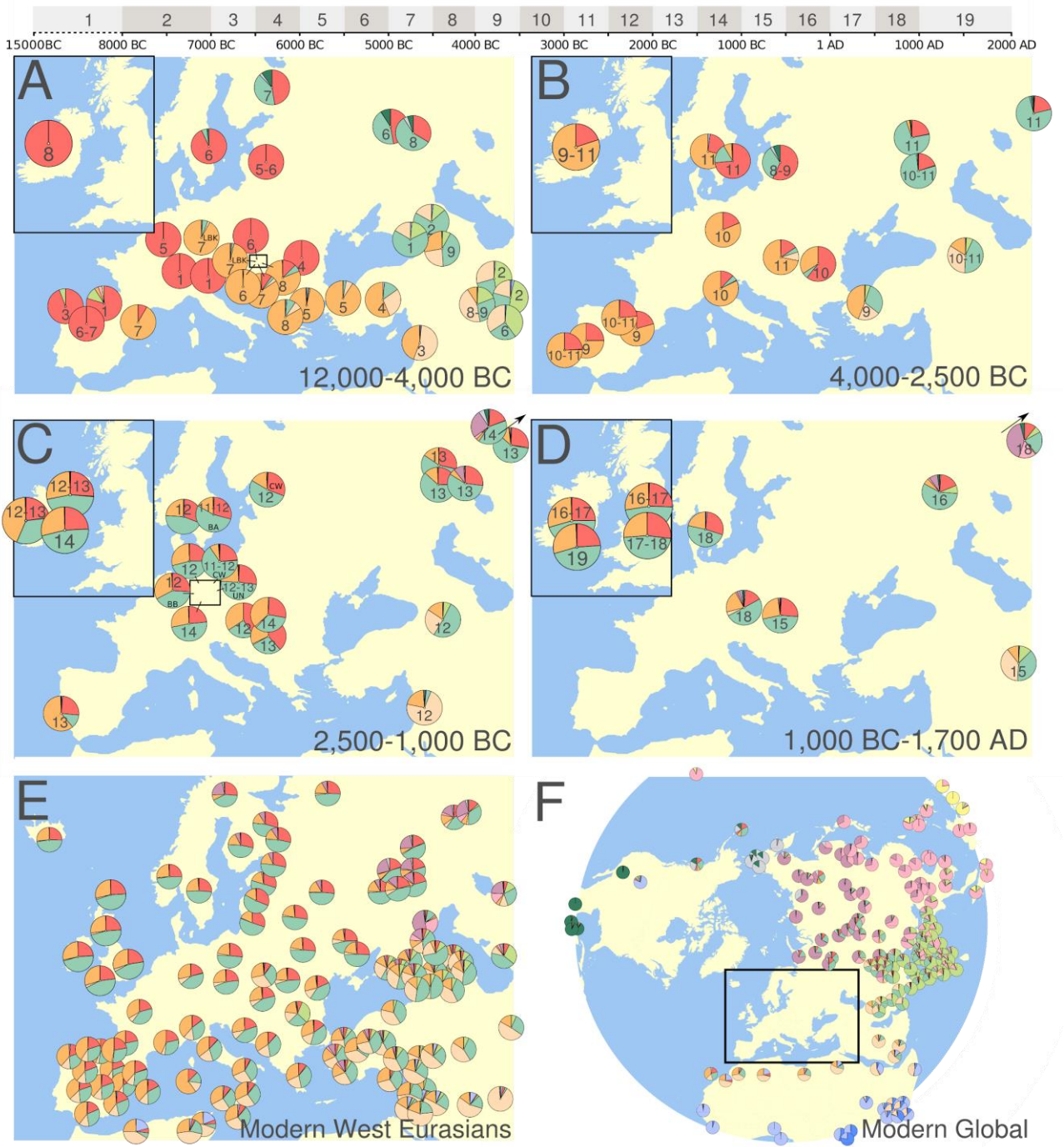


Figure 1.2. ADMIXTURE Analysis for 2401 ancient Eurasians and 291 diverse modern individuals at $K=12$. Individual admixture components are summarised in pie charts, averaged over larger populations, and located at the approximate geographical region of origin. Approximate time periods for populations are represented through the overhead numbered timeline **A-D**) Admixture components for ancient populations are panelled into four time divisions, with Irish populations magnified in each. Pie chart numbers represent further partitioning by date, represented by the overhead timeline. Four preglacial Palaeolithic individuals have been omitted from these plots. **E-F**) Admixture components for modern populations are shown, with the exception of 20 Southern hemisphere populations.

Early European Farmer Ancestry

The appearance of agriculture in the Balkans (Fig. 1.2A) clearly represents a major disruption in terms of autosomal ancestry, only explainable through large-scale population movements. This is visualised most strikingly in the earliest Greek and Hungarian Neolithic samples, who are composed almost entirely of the second major component of European ancestry (coloured orange), believed to have originated in northwest Anatolia (Hofmanová et al. 2016), and stands in sharp contrast to hunter-gatherer individuals of the region (Gamba et al. 2014). Strikingly, while a gradual increase in European hunter-gatherer ancestry (red) can be seen through time in the Hungarian Middle and Late Neolithic, this is not mirrored in Linearbandkeramik (LBK) populations who expanded further North via the Danubian route. Instead these culturally homogenous groups, show minor components of ancestry present in West Asia at the time (coloured teal and beige). These west Asian components also appear as larger minority contributions in the Greek Late Neolithic and to a lesser extent in Hungarians. In contrast, the earliest Spanish Neolithic individual, descended from populations who followed a Mediterranean route of expansion, is devoid of these smaller ancestries, presenting a majority early European farmer genome in orange, with marked European hunter-gatherer introgression. Overall, this heterogeneity suggest persistent contacts with West Asia after the initial colonisation of the Balkans, as well as differential interactions with local hunter-gatherer populations among Neolithic groups in Europe.

The spread of the Neolithic package to Europe's northern and western frontiers (Fig. 1.2B) is accompanied by a further increase in European hunter-gatherer admixture, though these populations are still dominated by the EEF component. The Neolithic population of Ireland follows this trend, with clear displacement of earlier hunter-gatherer ancestry apparent. This observation is only explainable through mass demographic upheaval. Indeed, in comparison to their contemporaries on the Southwest Atlantic coast and in Scandinavia, regions with dense and persistent hunter-gatherer populations (Cunliffe 2008), the Irish Neolithic population possesses, on average, less Mesolithic introgression. In this respect they are perhaps more comparable with contemporaneous groups from Germany and northeast Spain.

Importantly, the minor west Asian components (beige and teal) seen in LBK populations are not visible in the later populations of Northern Europe, suggesting that these groups, in their observed form, were not the main vehicle for agriculture into Ireland. These ancestries are however visible in Chalcolithic individuals from Italy and the Balkans (Fig. 1.2B), from the same approximate time depth as the Irish Neolithic, underscoring the dynamic zone of influence operating in the southeast Mediterranean, and also witnessed in the clear genetic turnover occurring in northwest Anatolia. A separate arena of demographic interplay is seen unfolding in the northeast of the continent, providing a contraposition to the northwest. Here, minimal penetrance of EEF ancestry is seen across the Neolithic transition (Jones et al. 2017). Importantly, another divergent group of hunter-gatherers existed in these northern regions and the Steppe, related not only to European hunter-gatherers further west (WHGs), but also to Caucasus (teal), Native American (green) and Eskimo groups (grey) (Fig. 1.2A). These eastern hunter-

gatherers (EHGs) appeared to have admixed with the earlier WHG population in Latvia by the Neolithic period (Jones et al. 2017). Some Scandinavian Mesolithic populations, who can be seen as a mixture of these two distinct hunter-gatherer groups (Lazaridis et al. 2014; Haak et al. 2015), also persisted across the Neolithic transition of the region with minor EEF influence, maintaining a largely non-agrarian lifestyle (Skoglund et al. 2014). No indication of such a phenomenon is apparent in Ireland in the current dataset.

Caucasus and Steppe-related Ancestry

During the Chalcolithic period, EHG-related populations in the Eurasian steppe see a significant inflation of the third major component of European ancestry, the west Asian coefficient coloured teal, identified in highest proportions in prehistoric and modern populations from the Caucasus (Fig. 1.2B). This has been linked to the spread of metallurgical technologies into the steppe from west Asia (Jones et al. 2015). This teal component is subsequently introduced on a major scale to the Late Neolithic and Bronze Age populations of Northern Europe, also accompanied by an increase in European hunter-gatherer ancestry (red), which is also present at substantial levels in steppe populations (Fig. 1.2C). As with the preceding Neolithic, this level of discontinuity implies widespread demographic replacement and has been explained by mass migration from the steppe into the region, through the vector of Corded Ware culture (Allentoft et al. 2015; Haak et al. 2015). Indo-European language introduction to the continent has also been linked to these movements.

Strikingly, this migration reverberated all the way to Ireland, visualised in the dramatic shift of ancestral coefficients corresponding with the arrival of metallurgy, and similar in profile to those observed in contemporary Northern Europeans. Notably, some heterogeneity in Irish Chalcolithic and Early Bronze Age populations is evident, when samples from the southwest are considered separately to those from the rest of the island (Fig. 1.2C, 12-13 interval); markedly reduced steppe-related ancestry is visible in this group (See Chapter Five for further exploration). However, despite this discrepancy, it is clear that large-scale migration impacted the entire island. A contrasting scenario is seen in the southwest of the continent, specifically in the Portuguese Bronze Age, which experiences significantly less penetrance of steppe-type ancestry relative to the northwest (Martiniano et al. 2017). In the southeast, Hungarian individuals show marked heterogeneity in ancestral coefficients over the course of the period, a testament to the demographic upheavals of the time and the region's position as a continental crossroads. Back migration from Europe to the Steppe is also apparent during the Bronze Age (Fig. 1.C), which also witnesses the earliest appearance of Siberian related ancestry (purple component) in the region.

Modern Distributions of the Three Ancestries

Genomic data from ancient individuals post-dating the Bronze Age is scarce in Europe, with few published continental Iron Age populations available with which to contrast the Irish sample set (Fig. 1.2D). Within the island itself, the overall interpretation is one of continuity, with the relative proportions of the three main components varying little from the Bronze Age through to modern populations (Fig.

1.2C-E). The Iron Age population of the neighbouring island of Britain shows a similar profile, with later Anglo-Saxon and Medieval samples giving a slight increase in Caucasus and European hunter-gatherer components (teal and red). The Nordic Iron Age shows even higher proportions of these ancestries relative to the orange EEF component. The few individuals sampled from the center of the continent show considerably more noise from West Asian and steppe sources, relative to those on the northern and western edges, highlighting their geographic isolation.

By modern times (Fig. 1.2E), a fourth component of ancestry (beige), found at highest levels in Bedouin populations, has left an impact on European populations, particularly those of Italy and the Balkans. This is most likely an effect of the Mediterranean civilisations that dominated the regions for many centuries, culminating in the Roman empire (27 BC – 395 AD), which saw much of the continent subsumed into this growing hegemony. Indeed, this component penetrates the more isolated northern regions to some degree, including France and Germany. Despite its absence in Medieval individuals, the component is also seen in Britain, where it decreases in scale from Southern England to Scotland. Ireland is bypassed almost entirely, as in Scandinavia. Intriguingly, in Iberia this component is also close to absent in the linguistically divergent Basque region, in sharp contrast to neighbouring populations.

Principal Component Analysis

Principal components analysis (PCA) allows for further dissection of the genetic affinities between ancient individuals and their modern counterparts, acting as a complementary analysis to model-based clustering. As with ADMIXTURE results, PCA projections of ancient variation have been panelled chronologically (Fig. 1.3), visualising the main demographic trends within each period. Fig. 1.3F presents the original PCA plot of modern populations, following the same geographical colour key as used for ancient samples. PC1 and PC2 are seen to be highly correlated with geography for modern populations, representing approximate north-south and east-west gradients respectively.

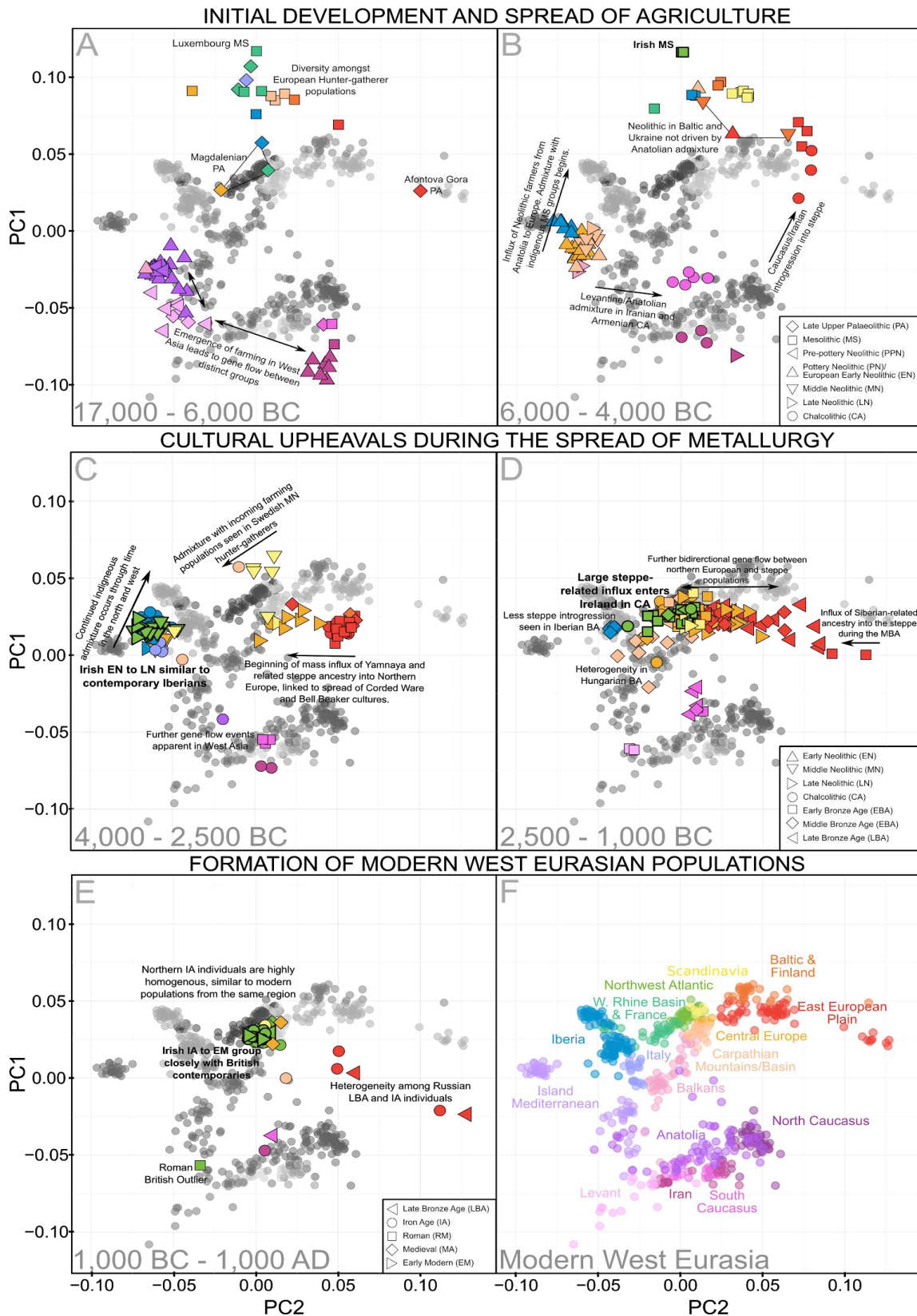


Figure 1.3. PCA plots for the first two components of modern West Eurasian genetic variation with ancient samples projected. Ancient samples are partitioned by date in panels A-E. The three shape keys in these plots correspond to approximate archaeological contexts for A-B, C-D, and E respectively. F) presents the original PCA plot, which has been set to grayscale in panels A-E. Modern individuals are coloured by geographical region, following the same key used for ancient individuals. Individuals from the current study are marked in bold, to distinguish them from other northwest Atlantic populations.

PCA projection plots of early ancient individuals (Fig. 1.3A-B) emphasise the genetic divergence between the Palaeolithic, Mesolithic and Early Farming populations of western Eurasia. Upper Palaeolithic and Mesolithic hunter-gatherer (HG) individuals from Europe place toward the top (or by analogy to moderns - the extreme north) of the PCA plot, showing a gradient of genetic variation that correlates well with an east-west divide between groups. More detailed relationships between these diverse hunter-gatherer groups can be discerned here. The east-west trajectory of variation shows HGs from western Europe placing at leftmost extreme, followed by those from the Balkans, Latvia, Scandinavia, Ukraine and finally EHGs from Russia at the other end of the gradient. Within the grouping of western European HGs, Spanish Mesolithic individuals group furthest south, towards Magdalenian Palaeolithic individuals, while those from Ireland and Luxembourg comprise the far north of the plot, forming a tight cluster with one another, indicative of high levels of genetic similarity. Nearby to this northwest Mesolithic grouping, Palaeolithic individuals from the Epigravettian of Italy and Azilian of Switzerland are seen.

At the bottom of the PCA plot, early populations from Anatolia and the Levant are seen at the southwestern edge of modern variation, while those from Iran and the Caucasus fall in the southeast (Fig. 1.3A). The emergence of agriculture in West Asia is accompanied by gene flow between these divergent groups, which, through time, are gradually pulled towards their modern-day counterparts, though the admixture events occurring over the course of the Neolithic and Metal Ages are clearly complex in nature (Fig. 1.3B-E). Importantly, together with the divergent groups of WHGs and EHGs in the north, early Levantine/Anatolian and Caucasus/Iranian groups plot at the approximate four corners of modern variation (Lazaridis et al. 2016), suggesting that modern day Europeans, including the Irish, may for the most part be a mixture of the genetic variation found in these early Holocene populations, brought about through the successive admixture events discussed in the previous section.

The Anatolian origin of Europe's first farmers is easily ascertained from the PCA plot, with early Neolithic individuals from the Balkans indistinguishable from the larger Anatolian cluster. Later Neolithic groups of Hungary, and Early Neolithic individuals from Spain and Germany show a slight degree of differentiation from both Anatolians and each other, most likely driven by the variation of minor ADMIXTURE components between these groups (Fig. 1.2A-B). Later Neolithic populations of Europe move further north along PC1, correlating with the increase in European hunter-gatherer ancestry seen in ADMIXTURE analysis (Fig. 1.2C). Strikingly, within this tight cluster of individuals some differentiation between regional groups is again observed, with German and Swedish Middle Neolithic individuals, as well as an Italian Chalcolithic group, falling further east compared to Irish and Iberian Neolithic populations, whom are notably indistinguishable from each other.

On the other side of the PCA plot a separate admixture event between EHGs and populations related to Iranian and Caucasus groups is seen to occur in the Russian Chalcolithic (Fig. 1.3B), culminating in the Bronze Age steppe cultures of Yamnaya, Afanasievo and Poltavka (Fig. 1.3C), who present as genetically homogenous despite their large geographical ranges. Fig. 1.3D visualises the dramatic

eastward shift of Northern European affinities that accompanies the mass movement of people related to these steppe populations. The result is a gradient of European Bronze Age populations that somewhat resemble present-day populations from the same regions. Irish Chalcolithic and Early Bronze Age individuals, particularly those from the southwest of the island, fall to the western edge of this large Northern European grouping, but east of the Iberian Bronze Age and heterogeneous Hungarian samples. By the Iron Age (Fig. 1.3E) Irish genetic diversity appears to have homogenised, forming a tight cluster indistinguishable from contemporary British individuals, and placing close to modern populations from both islands.

Haplotype-Based Resolution of Continuity

Uniparental, ADMIXTURE and principal components analysis all suggest some level of genetic and demographic continuity occurring in Ireland from the Chalcolithic onwards. To explore this possibility further, haplotypic analysis of the high-coverage Irish Early Bronze Age genome, Rathlin1, was carried out, alongside five other high-coverage ancient genomes to act as reference points, including an Irish Neolithic individual, Ballynahatty. Haplotype-based approaches are more powerful than those using unlinked genetic loci in identifying fine genetic structure, such as that displayed among Northern Europeans, and are relatively robust to bias from marker ascertainment (Lawson et al. 2012; Leslie et al. 2015). ChromoPainter in fineSTRUCTURE (Version 2) (Lawson et al. 2012) was used to decompose each ancient genome into a series of haplotypic chunks, and identify which modern individuals from a diverse set of Eurasian populations (Hellenthal et al. 2014) shared the same, or most similar, haplotype at each given chunk. Patterns of chunk donation between each ancient genome and modern populations were then considered.

Strikingly, the Irish Early Bronze Age genome revealed high median donation levels to Irish, Scottish, and Welsh populations (Fig. 1.4B). In regression with values from the other ancient genomes (Fig. 1.4A), these insular Celtic populations, and to a lesser degree the English, show an excess of sharing with Rathlin1, supporting strong continuity at the edge of Europe persisting over 4,000 years. As a contrast, the Hungarian Bronze Age genome examined shows more affinity with central European populations. Unsurprisingly, the pattern of haplotypic affinity of Ballynahatty among modern European populations is strongly correlated to that of the earlier Neolithic samples (Fig. 1.4B; $r > 0.74$, $P < 10^{-7}$), with Mediterranean samples in each analysis showing the highest levels of chunk copying. However, some differences are discernable; the NE1 and Stuttgart Neolithic genomes tend toward higher values in eastern Mediterranean (Sicilian, Italian, and Greek samples), while the Irish Neolithic has highest values in the west (Sardinian and Spanish). A further difference lies in the comparison of each to the affinities shown by the WHG, Loschbour, which shows no correlation in its modern affinities with the earlier continental Neolithics but does show a significant relationship ($P = 5.4 \times 10^{-4}$) with those of Ballynahatty, undoubtedly due to a greater WHG contribution to her ancestry.

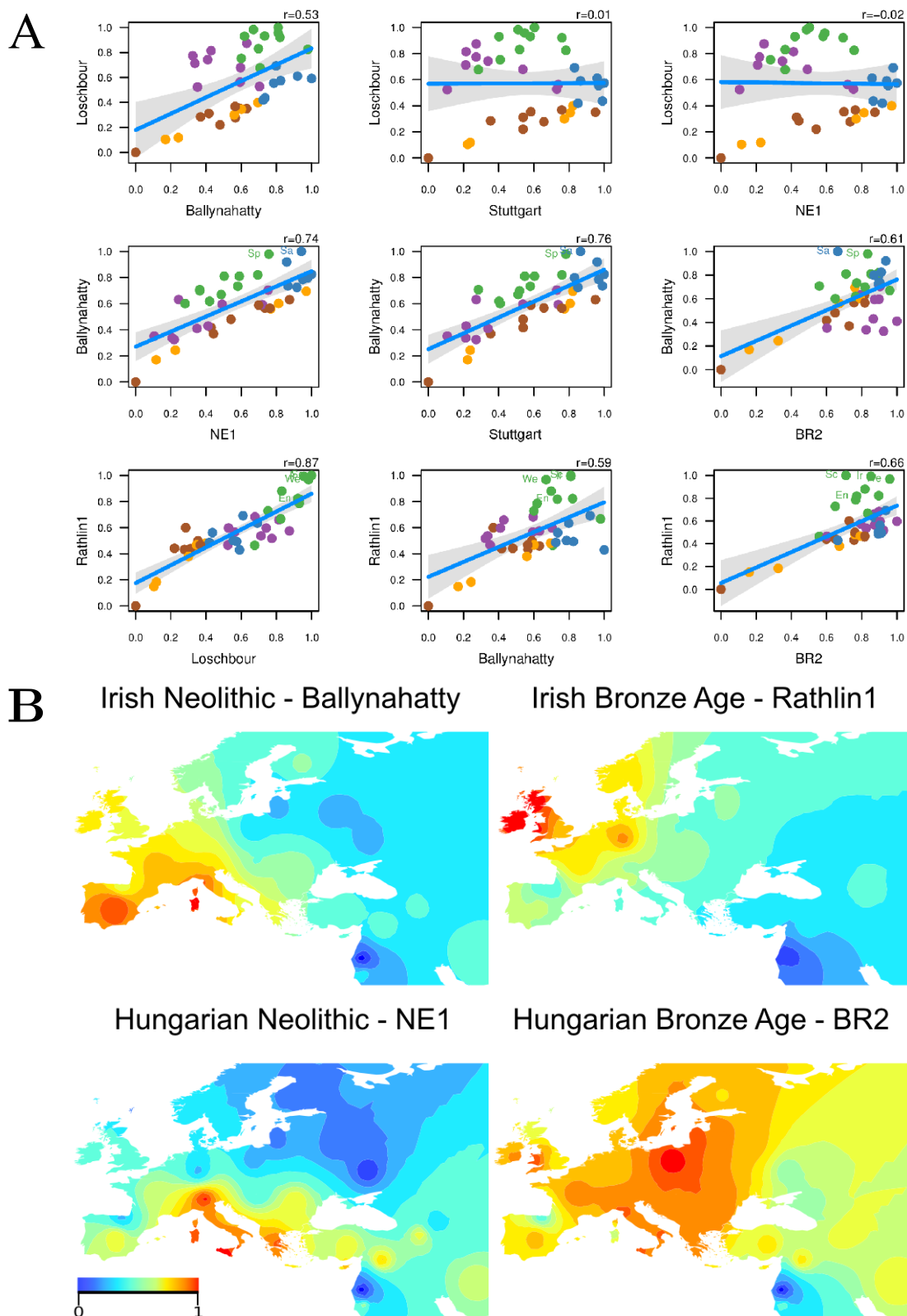


Figure 1.4. Comparison of Irish and Central European ancient genomes for haplotype-based affinity to modern populations. A) Selected pairwise regression plots of haplotype donation by ancient genomes to a set of modern genomes. Colour keys are West Asia (brown), South Middle East (orange), East European (purple) South Europe (blue) and NorthWest Europe (green) Specific samples are labelled: En - English; Ir - Ireland; Sa - Sardinian; Sc - Scottish; Sp - Spanish; We - Welsh. **B)** Interpolated heatmaps comparing relative haplotype donations by two Irish (Ballynahatty, Rathlin1) and two Hungarian (NE1, BR2) ancient genomes.

Conclusions

Flux and stasis on the Atlantic Edge

The oldest histories of Ireland, drawn from the written testaments of Christian scholars, outlined the origins of the Irish people as a series of invasions, a concept later seized upon in early linguistic and archaeological research. The theme has persisted in some form or another to the present day, most famously embodied by the proposed migration of Central European Celts to the island in the Iron Age. More recent opinion in both fields has emphasised the deeper roots of Atlantic populations, stressing the extensive maritime networks that delivered shared cultural phenomena along the seaboard for many millennia. Combined with genetic surveys of modern populations, these theories emphasised demographic continuity in Ireland, stretching back in some form or other to the Palaeolithic or Mesolithic. The palaeogenomic evidence presented here has, to a certain extent, proved both scenarios simultaneously true and untrue, with clear population replacement and long term continuity both evident on Europe's northwestern edge.

Multiple episodes of demographic replacement are implied by the existence of three genetically distinct populations, who occupied the island at different periods in time. This is attested to in both PCA and ADMIXTURE analysis, which clearly demonstrate that ancient Irish genetic affinities segregate within the major European archaeological horizons, rather than clustering geographically within the island through time. The exact points of demographic disruption can be readily narrowed to two major cultural transitions, that between the Mesolithic and Neolithic periods (~4,000-3,800 BC), and that between the Neolithic and Chalcolithic periods (~2,500 BC), the first of these is bookmarked by the youngest Mesolithic (4,224-3,950 cal BC) and the oldest Neolithic (3,942-3,702 cal BC) individuals, the second by the youngest Neolithic (2,834-2,469 cal BC) and oldest Chalcolithic (2,402-2,138 cal BC) individuals. The only reasonable explanation for these observations is extensive migration to the island at both horizons. Importantly, both Neolithic and Chalcolithic/Bronze Age cohorts are composed of over 26 individuals each, sampled from a diverse range of chronological, cultural and geographical contexts (Fig. 1.1), emphasising the widespread and persistent nature of these migrations.

Mesolithic Origins

The earliest population of the island, represented by two individuals from the end of the Mesolithic period, show closest similarity to other Mesolithic individuals from Western Europe, with an apparently heightened affinity to both each other and the Loschbour genome from Luxembourg in PCA. In terms of relationships to earlier hunter-gatherer groups, the Mesolithic Irish appear more similar to Italian and Swiss individuals from the Epigravettian and Azilian cultures, rather than the Iberian Magdalenian and later Mesolithic. This suggests the post-glacial colonisation of the island did not occur from a purely Iberian refugia, although these interpretations, based solely on PCA analysis, will be revisited in greater detail in Chapter Three. Importantly, no early Mesolithic samples from Ireland are available, and the possibility of demographic discontinuity within this period cannot be ruled out.

Neolithic Replacement

The second distinct population of the island existed from the onset of the Neolithic through to its conclusion. Over this period the Irish population presents as homogenous. They are composed in the majority of an ancestry dominant in West Asia, most probably Anatolian in origin, which is not detectable in any European population before the onset of agriculture. This demonstrates conclusively that the expansion of the European Neolithic into its northwestern frontier, was primarily mediated by mass migration of populations who possessed an unbroken chain of ancestry stretching back to the early farming communities of Anatolia.

A minority contribution of indigenous Mesolithic ancestry is also visualised in all Irish individuals, inflated relative to the Neolithic groups of Hungary and the Balkans, as well as early LBK (Germany) and Cardial (Spain) individuals from further west. This trend is seen in both PCA positioning and in the ratio of ADMIXTURE components. This could indicate some continuity across the Mesolithic to Neolithic transition on the island. However, as similar levels of Mesolithic ancestry are seen in contemporary populations from Spain, Scandinavia and Germany, there is no reason to suppose the indigenous hunter-gatherer ancestry observed in Irish Neolithic populations is in fact Irish in origin. On the other hand, even extensive admixture with local populations may not have detectably impacted the ancestral components of the Irish Neolithic population, given the likely large discrepancy in population size between the two groups (Woodman 2015), leaving the question of hunter-gatherer contribution to the Neolithic in Ireland largely unanswered. These issues will be returned to in Chapter Four.

Mitochondrial haplogroup frequencies mirror the autosomal patterns of population replacement with some indigenous input, showing a minority of Mesolithic U5 lineages observed among an abundance of west Asian haplogroups. However, in striking contrast, Y chromosomes show near complete continuity across the transition, with all but two male individuals belonging to the dominant Mesolithic lineage I2a. This may be indicative of unequal contribution by male and female hunter-gatherers to incoming farmer populations, though in a reverse of the usual trend witnessed across episodes of colonisation, that of female continuity and male replacement (Jobling & Tyler-Smith 2003). Alternatively the smaller effective population size and thus increased drift acting on the Y chromosome may be responsible for the differential patterns.

With respect to relationships between the Neolithic populations of western Europe, PCA shows Irish individuals to be indistinguishable from both each other and from Iberians of the same time depth, another region dominated by megalithic culture. However, despite this apparent uniformity, differing affinities among these individuals may be present, a possibility also addressed in Chapter Four, and a solely Iberian/Atlantic origin for Ireland's Neolithic communities cannot be confirmed. Moreover, despite the apparent homogeneity of the population, multiple colonisations of the island during the Neolithic, as has been theorised (Sheridan 2010), also remains a possibility, as these would likely originate from closely related source populations, followed by introgression between groups on the island. These

complex admixture scenarios, between relatively recently diverged populations, renders any differentiation between Irish samples undetectable at this level of autosomal analysis. Moreover, the lack of British and northwestern continental samples prevents access to key reference points with which to interpret results.

Bronze Age Continuity

The third and final population of Ireland that can be considered as a discrete genetic entity arrives in the Chalcolithic period (2,500-2,200 BC), likely alongside metallurgy and aspects of Beaker culture. It is at this horizon that a new component of ancestry in ADMIXTURE analysis, originating in pastoralist cultures of the Pontic Steppe, enters the island, also visualised in the dramatic eastward shift of Irish genetic variation in PCA. Importantly, no clear distinction can be made between the populations of Early Bronze Age, Iron Age and Modern Ireland, in either PCA or ADMIXTURE analysis. Also at this juncture, in striking contrast with the previous Mesolithic to Neolithic transition, male lineages witness a near complete turnover. R1b-L151 haplogroups, which segregate with the Atlantic Modal Haplotype (Y Chromosome Consortium 2002), dominate in the island's population from this point in time onwards, with the subclade R1b-L21, closely associated with the modern Irish population, occurring in the vast majority of Bronze and Iron Age samples. Conversely, the mitochondrial make-up of the population remains relatively similar to the preceding Neolithic, albeit with several steppe-related lineages appearing at this time (U2, I, W) (Haak et al. 2015).

Overall, the above observations are suggestive of substantial population continuity in Ireland since the Chalcolithic period. This possibility could be confirmed through more powerful haplotypic analysis of a high coverage Early Bronze Age individual from the north of the island, Rathlin1, who presented inflated sharing with the geographically closest modern populations, a trend not seen in other high coverage genomes analysed. This affinity with Irish, Scottish, and Welsh, emphasises demographic continuity stretching over 4,000 years at the insular Celtic edge of Europe. A weaker signal from modern English populations is likely due to the effects of Anglo-Saxon migrations; (Leslie et al. 2015; Martiniano et al. 2016; Schiffels et al. 2016). While further migration into the island has undoubtedly occurred, including recorded movements of Vikings, Normans and English planters, these appear to have produced a minimal genetic impact on the population, supporting studies of Y chromosomal diversity (Hill et al. 2000). Forthcoming research on the fine-scale population structure of modern Ireland, based on haplotypic analyses, also supports these conclusions. This demonstrates that, while some introgression of European and British haplotypes into Ireland is apparent at horizons of known historical migrations, it is more ancient Celtic population structure that remains the defining characteristic of the Irish population (Byrne et al. submitted).

However, it must be noted that substantial prehistoric migration to the island post-dating the Early Bronze Age from closely related populations of similar genetic makeup, particularly those of Celtic Britain or indeed historic Gaul, may not be detectable. Indeed, the presence of shared tribe names across

the various regions in Ptolemy's *Geography*, hint at such a possibility. Substantial demographic upheavals in Britain and the continent, such as those linked to the spread of Latin and Germanic languages, may have worked to reduce signals of haplotypic and Y chromosomal continuity with Early Bronze Age Ireland, that could have otherwise extended the affinities of Rathlin1 across a much wider region of western Europe. Indeed, for the population of England, as noted above, this is demonstrably the case. Such possibilities will be further addressed in Chapter Five.

Three Tongues

A demographic framework for the prehistory of Ireland can give insight into the question of language. What can be safely assumed, given the profound nature of the population turnovers witnessed at both the Neolithic and Chalcolithic transitions, is that at least three, probably highly divergent languages have been spoken on the island since the commencement of its human habitation, obviously excluding the later introduction of English. The earliest language, that of the Mesolithic peoples of Ireland, was in all likelihood a pre-Indo-European tongue. Furthermore, even if the popular Kurgan hypothesis is only partially accepted (i.e. sub-branches of the family entering the continent via steppe migration) we would expect that by the Early Bronze Age at least some occupants of the island were speaking an Indo-European tongue.

The language of the Neolithic inhabitants, and indeed the Chalcolithic and Early Bronze Age populations of southwest Ireland, are more difficult to define. If the Anatolian hypothesis is to be considered, then earlier substrates of Indo-European would have been present in the Neolithic communities of western Europe, and all pre-Indo-European languages, both extinct and living, that have been thus far documented, would have to be attributed to Mesolithic communities. This argument may find some support in the preponderance of pre-Indo-European in Iberia, a region with dense Mesolithic communities and extensive hunter-gatherer admixture apparent. However, harder to explain are the Tyrsenian languages of Italy, a region where substantially less Mesolithic introgression has so far been observed. What seems to link the two regions better together is the apparent lack of penetrance of Steppe ancestry, both in modern populations, and in the Iberian Bronze Age (Fig. 1.2), which would have allowed survival of pre-Indo-European Neolithic tongues, prior the expansion of Latin. In this scenario, the Neolithic language of Ireland would also have been pre-Indo-European, perhaps distantly related to the Basque language further south along the Atlantic. However, the possibility of Basque belonging to an even deeper Mesolithic linguistic layer cannot be excluded.

If it can be assumed that by the Early Bronze Age of Ireland some form of Indo-European language was present, then the question that must be asked is whether this could be considered a form of Celtic. The simplest scenario places the origin of Celtic in an early central European population, pre-dating the spread of steppe Ancestry both south- and northwest. Its subsequent dispersal, via Beaker networks, would then account for the deepest splits in the family. Alternatively, as the Atlantic region of Europe only saw ever increasingly contacts as the Bronze Age progressed, there are later windows in which

which Celtic languages could develop and spread. Moreover, a simple model of isolated linguistic evolution and subsequent dispersal may not be appropriate. Parallel evolution of mutually intelligible Indo-European dialects across a wide region could occur if sufficient connectivity was achieved, with later divergence only when contacts were cut, possibly at the end of the Bronze Age (Cunliffe and Koch 2012). In any case, whatever the temporal and geographic origins of the Celtic language family may be, the clear signal of continuity between Irish Early Bronze Age and modern populations, suggests that any introduction or evolution of Celtic language on the island after this point was mediated for the most part by the local population.

The only exception may be mass migration from a closely related group, as noted above. Indeed, the most recent expansions of Goidelic from Ireland to Scotland and Brittonic from Cornwall to Brittany, both involving some population migration, are a testament to how regularly such phenomena could occur. Future contributions of ancient genomics to linguistic models will involve the demonstration or disqualification of such events across the Bronze and Iron Ages, a task which demands both denser regional sampling and deeper genome sequencing of ancient individuals.

References

- Alexander, D.H., Novembre, J. & Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), pp.1655–1664.
- Allentoft, M.E. et al., 2015. Population genomics of Bronze Age Eurasia. *Nature*, 522(7555), pp.167–172.
- Ammerman, A.J. & Cavalli-Sforza, L.L., 1984. *The Neolithic transition and the population genetics of Europe*. New Jersey: Princeton University Press.
- Anthony, D.W., 2010. *The Horse, the Wheel, and Language: How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World*. New Jersey: Princeton University Press.
- Balaresque, P. et al., 2010. A predominantly neolithic origin for European paternal lineages. *PLoS Biology*, 8(1), p.e1000285.
- Bayliss, A. & Woodman, P., 2009. A New Bayesian Chronology for Mesolithic Occupation at Mount Sandel, Northern Ireland. *Proceedings of the Prehistoric Society*, 75, pp.101–123.
- Behar, D.M. et al., 2010. The genome-wide structure of the Jewish people. *Nature*, 466(7303), pp.238–242.
- Bramanti, B. et al., 2009. Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science*, 326(5949), pp.137–140.
- Breheny, P. & Burchett, W., 2012. visreg: Visualisation of regression models. *R package version*, 2.
- Broushaki, F. et al., 2016. Early Neolithic genomes from the eastern Fertile Crescent. *Science*, 353(6298), pp.499–503.
- Busby, G.B.J. et al., 2012. The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. *Proceedings of the Royal Society B: Biological Sciences*, 279(1730), pp.884–892.
- Byrne, R.P. et al., 2017 Celtic population structure and genomic footprints of migration. Manuscript submitted for publication.
- Carew, M., 2012. The Harvard Mission, Eugenics and the Celts. *Archaeology Ireland*, 26(4), pp.38–40.
- Carey, J. & Dumville, D.N., 1994. *The Irish national origin-legend: synthetic pseudohistory*. UK: Department of Anglo-Saxon, Norse, and Celtic, University of Cambridge.
- Cassidy, L.M. et al., 2016. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proceedings of the National Academy of Sciences of the United States of America*, 113(2), pp.368–373.
- Cavalli-Sforza, L.L., Menozzi, P. & Piazza, A., 1994. *The History and Geography of Human Genes*. New Jersey: Princeton University Press.
- Chang, C.C. et al., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4, p.7.
- Clark, J.G.D., 1965. Radiocarbon dating and the expansion of farming culture from the Near East over Europe. *Proceedings of the Prehistoric Society*, 31, pp.58–73.
- Coon, C.S., 1939. *The Races of Europe*. New York: The McMillan Company.
- Cooney, G., 1995. Theory and practice in Irish archaeology. In P. J. Ucko (ed.), *Theory in archaeology: A world perspective*. London: Routledge, pp.259–272.
- Corlett, C. & Potterton, M. (eds.), 2012. *Life and death in Iron Age Ireland: in the light of recent archaeological excavations. Research papers in Irish archaeology, no. 4*. Bray, Ireland: Wordwell.
- Cunliffe, B.W., 1997. *The Ancient Celts*. Oxford: Oxford University Press.
- Cunliffe, B.W., 2001. *Facing the ocean: the Atlantic and its peoples, 8000 BC-AD 1500*. Oxford: Oxford University Press.
- Cunliffe, B.W., 2008. *Europe between the Oceans: 9000 BC-AD 1000*. New Haven-London: Yale University Press.
- Cunliffe, B.W. & Koch, J.T., 2012. *Celtic from the West: Alternative Perspectives from Archaeology, Genetics, Language, and Literature*. Oxford: Oxbow Books.
- De Bernardo Stempel, P., 2007. Pre-Celtic, Old Celtic layers, Brittonic and Goidelic in ancient Ireland. In P. Cavill and G. Broderick (eds.), *Language contact in the place-names of Britain and Ireland*. Nottingham: English Place-Name Society, pp.137–163.
- Delaneau, O., Marchini, J. & Zagury, J.-F., 2011. A linear complexity phasing method for thousands of genomes. *Nature Methods*, 9(2), pp.179–181.
- Dowd, M. & Carden, R.F., 2016. First evidence of a Late Upper Palaeolithic human presence in Ireland. *Quaternary Science Reviews*, 139, pp.158–163.
- Esmonde-Cleary, A.S., 2002. *The Ending of Roman Britain*. London: Routledge.
- FamilyTreeDNA, 2017. I-M223 Y-Haplogroup Project. Available at: <https://www.familytreedna.com/groups/m223-y-clan/>.
- Flanagan, L., 1998. *Ancient Ireland: Life Before the Celts*. Dublin: Gill & Macmillan Ltd.
- Fu, Q. et al., 2016. The genetic history of Ice Age Europe. *Nature*, 534(7606), pp.200–205.

- Gamba, C. et al., 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications*, 5, p.5257.
- Gimbutas, M., 1956. *The prehistory of eastern Europe*. New Haven: Peabody Museum.
- González-Fortes, G. et al., 2017. Paleogenomic Evidence for Multi-generational Mixing between Neolithic Farmers and Mesolithic Hunter-Gatherers in the Lower Danube Basin. *Current Biology*, 27(12), pp.1801–1810.e10.
- Günther, T. et al., 2015. Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proceedings of the National Academy of Sciences*, 112(38), pp.11917–11922.
- Haak, W. et al., 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522(7555), pp.207–211.
- Harbison, P., 1975. The coming of the Indo-Europeans to Ireland: an archaeological assessment. *Journal of Indo-European Studies*, 3, pp.101–119.
- Hawkes, C., 1931. Hill-Forts. *Antiquity*, 5(17), pp.60–97.
- Hellenthal, G. et al., 2014. A genetic atlas of human admixture history. *Science*, 343(6172), pp.747–751.
- Henderson, J., 2007. *The Atlantic Iron Age: Settlement and Identity in the First Millennium BC*. London: Routledge.
- Henn, B.M. et al., 2011. Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proceedings of the National Academy of Sciences of the United States of America*, 108(13), pp.5154–5162.
- Herity, M. & Eogan, G., 1977. *Ireland in Prehistory*, London: Routledge and Kegan Paul.
- Hill, E.W., Jobling, M.A. & Bradley, D.G., 2000. Y-chromosome variation and Irish origins. *Nature*, 404(6776), pp.351–352.
- Hills, C., 1993. The Anglo-Saxon settlement of England. The State of research in Britain in the late 1980s. *Vorträge und Forschungen*, 41(1), pp.303–315.
- Hofmanová, Z. et al., 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proceedings of the National Academy of Sciences of the United States of America*, 113(25), pp.6886–6891.
- International Society of Genetic Genealogy, 2017. Y-DNA Haplogroup I and its Subclades. Available at: https://isogg.org/tree/ISOGG_HapgrpI.html.
- Jobling, M.A. & Tyler-Smith, C., 1995. Fathers and sons: the Y chromosome and human evolution. *Trends in Genetics*, 11(11), pp.449–456.
- Jobling, M.A. & Tyler-Smith, C., 2003. The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics*, 4(8), pp.598–612.
- Jones, E.R. et al., 2015. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications*, 6, p.8912.
- Jones, E.R. et al., 2017. The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Current Biology*, 27(4), pp.576–582.
- Keller, A. et al., 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nature Communications*, 3, p.698.
- Kılınc, G.M. et al., 2016. The Demographic Development of the First Farmers in Anatolia. *Current Biology*, 26(19), pp.2659–2666.
- Lawson, D.J. et al., 2012. Inference of population structure using dense haplotype data. *PLoS Genetics*, 8(1), p.e1002453.
- Lazaridis, I. et al., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 513(7518), pp.409–413.
- Lazaridis, I. et al., 2016. Genomic insights into the origin of farming in the ancient Near East. *Nature*, 536(7617), pp.419–424.
- Leslie, S. et al., 2015. The fine-scale genetic structure of the British population. *Nature*, 519(7543), pp.309–314.
- Lhuyd, E., 1707. *Archaeologia Britannica: Giving Some Account Additional to what Has Been Hitherto Publish'd, of the Languages, Histories and Customs of the Original Inhabitants of Great Britain: from Collections and Observations in Travels Through Wales, Cornwall, Bretagne, Ireland, and Scotland*, Printed at the Theater for the author, MDCCVII. And sold by Mr. Bateman in Pater-Noster-Row, London: and Jeremiah Pepyat bookseller at Dublin.
- Libby, W.F., 1954. Chicago Radiocarbon Dates V. *Science*, 120(3123), pp.733–742.
- Li, H. et al., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), pp.2078–2079.
- Li, H. & Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), pp.1754–1760.
- Li, J.Z. et al., 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science*, 319(5866), pp.1100–1104.
- Lozny, L.R. (ed.), 2011. Archaeology in the Age of Globalization: Local Meanings, Global Interest. In *Comparative Archaeologies*. New York: Springer, pp. 21–49.
- Macalister, R.A.S., 1949. *The Archaeology of Ireland*. London: Methuen.

- MacHugh, D.E. et al., 2000. The extraction and analysis of ancient DNA from bone and teeth: a survey of current methodologies. *Ancient Biomolecules*, 3(2), pp.81–103.
- Mahr, A., 1937. New Aspects and Problems in Irish Prehistory - Presidential Address for 1937. *Proceedings of the Prehistoric Society*, 3(1-2), pp.261–436.
- Mallory, J.P., 2013. *The origins of the Irish*. London: Thames & Hudson.
- Mallory, J.P., 2016. *In Search of the Irish Dreamtime: Archaeology and Early Irish Literature*. London: Thames & Hudson.
- Martiniano, R. et al., 2016. Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nature Communications*, 7, p.10326.
- Martiniano, R. et al., 2017. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genetics*, 13(7), p.e1006852.
- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), pp.10–12.
- Mathieson, I. et al., 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528(7583), pp.499–503.
- McEvoy, B. et al., 2004. The Longue Duree of genetic ancestry: multiple genetic marker systems and Celtic origins on the Atlantic facade of Europe. *American Journal of Human Genetics*, 75(4), pp.693–702.
- McEvoy, B. et al., 2006. The scale and nature of Viking settlement in Ireland from Y-chromosome admixture analysis. *European Journal of Human Genetics*, 14(12), pp.1288–1294.
- McEvoy, B., Simms, K. & Bradley, D.G., 2008. Genetic investigation of the patrilineal kinship structure of early medieval Ireland. *American Journal of Physical Anthropology*, 136(4), pp.415–422.
- McKenna, A. et al., 2010. The Genome Analysis Toolkit: a MapReduce framework for analysing next-generation DNA sequencing data. *Genome Research*, 20(9), pp.1297–1303.
- Meyer, M. & Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, 2010(6), p.db.prot5448.
- Moore, L.T. et al., 2006. A Y-chromosome signature of hegemony in Gaelic Ireland. *American Journal of Human Genetics*, 78(2), pp.334–338.
- Myres, N.M. et al., 2011. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *European Journal of Human Genetics: EJHG*, 19(1), pp.95–101.
- Novembre, J. et al., 2008. Genes mirror geography within Europe. *Nature*, 456(7218), pp.98–101.
- O'Curry, E., 1878. *Lectures on the Manuscript Materials of Ancient Irish History: Delivered at the Catholic University of Ireland, During the Sessions of 1855 and 1856*, Dublin: William A. Hinch
- O'Donnabháin, B., 2000. An appalling vista? The Celts and the archaeology of later prehistoric Ireland. In A. Desmond et al. (eds.), *New agendas in Irish prehistory*. Bray, Ireland: Wordwell, pp.189–196.
- O'Dushlaine, C.T. et al., 2010. Population structure and genome-wide patterns of variation in Ireland and Britain. *European Journal of Human Genetics*, 18(11), pp.1248–1254.
- Olalde, I. et al., 2014. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature*, 507(7491), pp.225–228.
- Olalde, I. et al., 2015. A Common Genetic Origin for Early Farmers from Mediterranean Cardial and Central European LBK Cultures. *Molecular Biology and Evolution*, 32(12), pp.3132–42. Olalde, I. et al., 2017. The Beaker phenomenon and the genomic transformation of northwest Europe. *bioRxiv*.
- Patterson, N., Price, A.L. & Reich, D., 2006. Population structure and eigenanalysis. *PLoS Genetics*, 2(12), p.e190.
- Piazza, A. et al., 1995. Genetics and the origin of European languages. *Proceedings of the National Academy of Sciences of the United States of America*, 92(13), pp.5836–5840.
- Piggott, S., 1954. *The Neolithic Cultures of the British Isles*. Cambridge: Cambridge University Press.
- Price, A.L. et al., 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8), pp.904–909.
- Prichard, J.C., 1831. *The Eastern Origin of the Celtic Nations: Proved by a Comparison of Their Dialects with the Sanskrit, Greek, Latin, and Teutonic Languages. Forming a Supplement to Researches Into the Physical History of Mankind*. S. Collingwood.
- Raftery, B., 1994. *Pagan Celtic Ireland: the enigma of the Irish Iron age*. London: Thames and Hudson.
- Renfrew, C., 1990. *Archaeology and Language: The Puzzle of Indo-European Origins*. Cambridge: CUP Archive.
- Richards, M. et al., 1996. Paleolithic and neolithic lineages in the European mitochondrial gene pool. *American Journal of Human Genetics*, 59(1), pp.185–203.
- Richards, M. et al., 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. *American Journal of Human Genetics*, 67(5), pp.1251–1276.
- Rosser, Z.H. et al., 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by

- language. *American Journal of Human Genetics*, 67(6), pp.1526–1543.
- Rowley-Conwy, P., 2007. *From genesis to prehistory: the archaeological three age system and its contested reception in Denmark, Britain, and Ireland*. Oxford: Oxford University Press.
- Schiffels, S. et al., 2016. Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nature Communications*, 7, p.10408.
- Sheridan, A. (ed.), 2010. The Neolithisation of Britain and Ireland: the big picture. In *Landscapes in transition*. Oxford: Oxbow, pp. 89–105.
- Simoni, L. et al., 2000. Reconstruction of prehistory on the basis of genetic data. *American Journal of Human Genetics*, 66(3), pp.1177–1179.
- Skoglund, P. et al., 2014. Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science*, 344(6185), pp.747–750.
- Smyth, J., 2014. *Settlement in the Irish Neolithic: New discoveries at the edge of Europe*. Oxford: Oxbow Books.
- Taylor, R.E. & Bar-Yosef, O., 2014. *Radiocarbon Dating, Second Edition: An Archaeological Perspective*. California: Left Coast Press.
- Torroni, A. et al., 1998. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *American Journal of Human Genetics*, 62(5), pp.1137–1152.
- Torroni, A. et al., 2001. A signal, from human mtDNA, of postglacial recolonisation in Europe. *American Journal of Human Genetics*, 69(4), pp.844–852.
- Waddell, J., 2005. *Foundation myths: the beginnings of Irish archaeology*. Bray, Ireland: Wordwell.
- Waddell, J., 2010. *The Prehistoric Archaeology of Ireland*. Bray, Ireland: Wordwell.
- Wickham, H., 2009. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
- Wilde, W.R., 1850. *The beauties of the Boyne, and its tributary, the Blackwater*. Dublin: J. McGlashan.
- Woodman, P., 2015. *Ireland's First Settlers: Time and the Mesolithic*. Oxford: Oxbow.
- Wood-Martin, W.G., 1886. The Rude Stone Monuments of Ireland. *The Journal of the Royal Historical and Archaeological Association of Ireland*, 7(67), pp.470–487.
- Wood-Martin, W.G., 1895. *Pagan Ireland; an archaeological sketch: a handbook of Irish pre-Christian antiquities*, London: Longmans, Green.
- Yang, D.Y. et al., 1998. Technical note: improved DNA extraction from ancient bones using silica-based spin columns. *American Journal of Physical Anthropology*, 105(4), pp.539–543.
- Y Chromosome Consortium, 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Research*, 12(2), pp.339–348.

3. The First Arrivals

A genetic insight into Ireland's Mesolithic inhabitants

Overview

This chapter concerns the genetic and demographic nature of the Irish Mesolithic, represented by two tooth samples from separate cave sites, Killuragh, Co. Limerick (4,660-4,500 cal BC) and Sramore, Co. Leitrim (4,224-3,950 cal BC), both detailed in Dowd 2015 dating to the end of the period (See Appendix I for further archaeological context). The individuals, sequenced to 1 and 1.7X coverage, were introduced in the previous chapter, where PCA and ADMIXTURE analysis showed them to be typical examples of European Mesolithic populations. Here, the genetic affinities of these samples are dissected in depth, using D- and f-statistics, alongside 71 genomes from Palaeolithic, Mesolithic and Neolithic hunter-gatherer cultures of Europe and West Asia (Table 2.1), as well as two early Neolithic populations from Iran and Anatolia. Moreover, statistical imputation of 18 hunter-gatherer genomes allowed novel analyses, such as identification of runs of homozygosity (ROH), polygenic estimation of height and pigmentation profiling.

To give proper context to this palaeogenomic dataset, a detailed picture of the genetic and archaeological landscape of the European Upper Palaeolithic (~50-12 kya) and Mesolithic (~12-6 kya) is first drawn in the introductory section. This gives an overview of evidence pertaining to the colonisation of Europe by modern humans and their movements within the continent during the last glacial period, culminating in the occupation of its most westerly isle *circa* 8,000 BC.

The Irish Late Mesolithic individuals are seen to share highest genetic affinities with contemporary populations in the north of France and Luxembourg, although ROH analysis suggest population bottlenecks occurred on the lineage specific to Irish hunter-gatherers. Strikingly, Irish and other northwestern hunter-gatherers share more affinity with the Late Palaeolithic of Italy, than they do with Spanish groups in the southwest, suggesting their ultimate origins lie in more eastern glacial refugia. Moreover, the Irish Mesolithic displays markedly decreased introgression from other more divergent hunter-gatherer groups, such as Scandinavian hunter-gatherers and the western Palaeolithic Magdalenian culture, relative to its continental counterparts. This suggests later admixture occurred on the continent to the exclusion of Irish populations. Finally, statistical imputation of genotypes allows for estimation of genomic height and pigmentation, giving a first glimpse of physical characteristics of Ireland's earliest inhabitants.

Introduction

Out of Africa and into Europe: Earliest Colonisations of the Continent

For the vast majority of its human habitation Europe has been populated by hunter-gatherer (HG) groups. As with all human ancestry, the origins of these early Europeans can be traced back to Africa, where anatomically modern humans (AMH) first emerged over 200 thousand years ago (kya). The earliest presence of AMHs in Eurasia, and potentially Europe, is placed at some time between 120 and 70 kya (Shea 2008; Armitage et al. 2011; Liu et al. 2015), the suggested result of multiple expansions driven by glacial-scale climatic changes (Timmermann & Friedrich 2016). However, genetic evidence suggests all non-Africans today are derived from a single later expansion event approximately 75-50 kya; the famous Out-of-Africa (OoA) migration, with little or no contribution from these earlier groups (Mallick et al. 2016), although possible exceptions in South Asia and Oceania, exist (Rasmussen et al. 2011; Pagani et al. 2016). Despite its proximity to the Levant and Africa, the northward colonisation of Europe appears to have proceeded at a slow pace relative to more southerly regions, possibly stalled by harsh climatic conditions. Although tool assemblages have been linked to modern human presence in southern Europe as early as 48-45 kya (Hoffecker 2009; Higham et al. 2012), the earliest direct fossil evidence (Kent's Cavern, UK) dates to ~44.2–41.5 kya (Higham et al. 2011). The earliest sequenced European genome, from Peștera cu Oase, Romania (Trinkaus et al. 2003) also belongs to this same period. This individual shows no specific contribution to the genetics of modern Europeans or later HG groups (Fu et al. 2015), which could be indicative of failed colonisation event, a fate that may have also been shared by migrants from older OoA expansions.

Approximately 40 kya, a new tool industry, the Aurignacian, likely originating in the Balkans (Hoffecker 2009), expanded through the continent via the Danubian corridor (Van Andel et al. 2003; Higham et al. 2012), providing Europe's oldest examples of figurine and cave art (Jöris et al. 2010). Two male samples associated with Aurignacian assemblages have been sequenced, Kostenki14 (37,000 BP) (Seguin-Orlando et al. 2014; Hublin 2015) from Eastern Europe and GoyetQ116-1 (35,000 BP) (Fu et al. 2016) from Belgium (Y haplogroups C1b and C1a respectively). Although divergent from each other, these two genomes, along with all other sampled ancient Europeans up until 14,000 years ago, are proposed to descend from a single founder population, which contributed directly to the ancestry of modern Europeans (Fu et al. 2016).

As the Aurignacian expanded, the previous inhabitants of Europe and West Asia, Neanderthals, disappeared completely from the prehistoric record (~39,000 BP; (Higham et al. 2014). These were the result of a previous *Homo* exodus from Africa that had split the ancestors of *Homo sapiens* from those of Neanderthals and their eastern sister group, Denisovans, approximately 700-400 kya (Meyer et al. 2016; Stringer 2016). It is now widely accepted that admixture occurred between these groups and AMHs. Aurignacian and later Gravettian (discussed below) genomes are observed to have relatively high levels of Neanderthal ancestry (3-4%) (Fu et al. 2016), while levels between 1-4% are seen in all modern non-

Africans (Green et al. 2010; Wall et al. 2013; Prüfer et al. 2014). Despite the geographical range of Neanderthals, it is East Asians today who possess the highest levels of related introgression, due to the widespread dispersal of 'Basal Eurasian' ancestry in modern Europe and West Asia, linked to the Neolithic and Bronze Age expansions (Lazaridis et al. 2014; Lazaridis et al. 2016). Basal Eurasians are a ghost population believed to have diverged from all other Eurasians early on in the OoA migration, with the main Neanderthal admixture event into modern humans proposed to have occurred after this split, but before the later groups split from each other. However, further episodes of admixture may also have occurred on lineages specific to Europe and other regions. The identification of a recent Neanderthal ancestor in the family tree of the individual from Oase, Romania, supports this possibility (Fu et al. 2015)

When the Ice Melts: The Genetic Legacy of the Last Glacial Maximum

Between approximately 30-20 kya, rapid climatic deterioration occurred in Europe, in advance of the Last Glacial Maximum (LGM), Aurignacian industries retreated south of the trans-European mountain barrier (Van Andel et al. 2003), replaced by another technocomplex, the Gravettian, which brought unprecedented cultural unification to the continent (Wojtal et al. 2015). Earliest sites are found in Russia (Kostenki), the Balkans and Danubian valley, indicating a possible eastern origin (Van Andel et al. 2003). The success of the Gravettian culture has been hypothesised to lie with more complex and mobile hunting skills, suited for the glacial conditions gripping the northern half of the continent where the majority of sites occur (Van Andel et al. 2003). Indeed, mammoth hunting played a clear central role in the economies of Eastern variants of the Gravettian culture.

Eight ancient genomes retrieved from such sites (Dolní Věstonice and Pavlov, Czech Republic; Krems-Wachtberg, Austria) form a discrete genetic grouping, with which 6 Belgian (Goyet Cave) and Italian Gravettian genomes also broadly cluster. The apparent shared genetic affinities of these individuals, despite their widespread geographical range, suggests the cultural complex expanded at least in part by population movement (Fu et al. 2016). Consistent with the distribution of early sites, the eastern Kostenki14 lineage is thought to be the predominant contributor to the Gravettian grouping, rather than the western Aurignacian lineage represented by GoyetQ116-1, which does not re-emerge until after the LGM in Iberia. Mitochondrial haplogroups in Gravettian genomes are similar to the preceding period, with both M and U (specifically U8c, U2, U5 and U6) represented. The earliest identification of Y haplogroup I, which came to dominate the European Mesolithic, is also first seen at this point in time.

As the ice sheets reached their peak expansion approximately 22 kya (Shakun & Carlson 2010), Gravettian cultures, similarly to the Aurignacian before them, were forced to retreat to southern refugia in the Mediterranean and Balkans, where they diversified and were replaced by regional groupings. In Italy and the Balkans, as well as possibly Central/Eastern Europe, characteristically Gravettian industries continued to develop after 20 kya, labelled as Epigravettian for their apparent continuity (Bietti 1990). In the west, the Gravettian was replaced by Solutrean (22-17 kya) and later Magdalenian (17-12 kya) industries, which flourished and expanded north after the LGM, most likely from Franco-Iberian refugia

(Straus 1991). While no Solutrean genomes have yet been sequenced, seven individuals from Spain, Belgium and Germany, associated with the Magdalenian complex have been seen to form a distinct genetic grouping (see PCA in Chapter 1), which possesses significant input from the western Aurignacian individual, GoyetQ116-1, indicating that this lineage did, in some form, survive the expansion of both the ice sheets and Gravettian cultures.

Only two autosomal genomes, both from Italy, are associated with the Epigravettian, one from Villabruna in the north (Fu et al. 2016), and the second from Sicily in the south (Mathieson et al. 2017). Both date to the very end of the period (~14kya). Notably, the Villabruna individual is the earliest Palaeolithic European genome to show specific genetic affinities to populations outside Europe, namely West Asians (Fu et al. 2016). This may be the result of populations movements and interactions across southern refugia, either during or after the LGM. Though some shared ancestry is apparent, the Epigravettian individuals are genetically distinct from the Magdalenian grouping, while sharp divergence is also noted between both groups and pre-LGM populations (Fu et al. 2016), suggestive of intense genetic bottlenecks and demographic upheavals during the LGM. Dramatic loss in mitochondrial diversity is also observed between pre- and post-LGM groups (Posth et al. 2016), with U5 emerging as the dominant lineage.

Hunter-gatherer groups in Holocene Europe

Prior to the Younger Dryas cold snap (~12.9-11.7 kya) Magdalenian culture had extended from southern Iberia all the way to Britain, where it diverged into the Creswellian, and eastwards into Poland. Magdalenian sites show clear evidence of cold-climate adaptation, with reindeer and horse hunting providing a major form of subsistence. Bone tools were pervasive and the period provides some of the final examples of cave art, a tradition that had carried down from the Aurignacian period and remained almost exclusive to western Europe. As the climate warmed, the Magdalenian and related cultures began to die off, linked to loss of large-game herd animals whose ecosystems had become compromised (Jochim 1980). It was replaced in western Europe by simpler Epipalaeolithic cultures, which provided the bridge between the Palaeolithic and Mesolithic periods. These included the Azilian in Northern Spain and Southern France, which gradually abandoned larger herd animals for smaller game, evidenced by the increase in microlithic tool sets, linked to archery. Notably, the earliest dates for the Azilian come from areas which witnessed early reforestation, such as the Swiss piedmont (Plisson et al. 2008). Fine art, including cave painting and bone engraving disappeared during this period. Subsequent related Mesolithic cultures, defined by in part by their microlithic toolkits, colonised the newly expanding temperate forests of Northern Europe, and fragmented into a variety of regional groups with complex subsistence strategies based on exploitation of aquatic, plant-based, and diverse faunal resources. Connectivity increased across the continent, while semi-sedentism became possible in some regions (Cunliffe 2008).

Some of the first published ancient human genomes, released in 2014, belonged to these Mesolithic populations of Europe. KO1 from Hungary, LaBrana1 from Spain and Loschbour from Luxembourg, all of similar time depth (~8.2-7.7 kya), were seen early on to cluster with one another at the extreme north of modern European variation on PCA (See Chapter 1). Their close affinities earned them the blanket term ‘Western Hunter Gatherers’ (WHGs), which has persisted in usage despite the identification of genetically similar individuals as far east as Latvia and Romania (González-Fortes et al. 2017; Jones et al. 2017). Strikingly, genome data from these later hunter-gatherer groups, as far removed as Spain, Croatia, Latvia and Luxembourg, all show strong affinity with the Epigravettian individuals from Italy (Fu et al. 2016; Mathieson et al. 2017). This includes individuals from regions previously associated with the genetically distinct Magdalenian, suggesting its disappearance represents not only a cultural, but demographic demise. However, several western Mesolithic samples show slightly increased affinity to Magdalenian genomes, relative to those further east, including the Epigravettian Villabruna individual (Fu et al. 2016). Notably, while the majority of individuals shown to cluster with the Italian Epigravettian samples belong to Mesolithic cultures from the Holocene era, two Epipalaeolithic individuals, from Bichon Switzerland (Azilian Culture) and Rochedane, France, also form part of this grouping, placing a lower date limit for this ancestry’s expansion into western Europe.

The mitochondrial haplogroup U8a, associated with the Magdalenian, has not been observed in later European hunter-gatherer samples, although interestingly, while rare in modern Europe, its deepest roots are found in the modern Basque country (González et al. 2006), a region associated with the Franco-Iberian refugia. In terms of Y chromosomes, haplogroup I, specifically I2, is found at high frequency across western Europe after the LGM, including in two Magdalenian samples, while R1 groups are seen in Latvia and Romania, and also intriguingly in the Villabruna individual himself. The rare C1a lineage, previously noted in the Aurignacian sample from Goyet, Belgium, is also seen in an individual from Mesolithic Iberia.

Taken together, it would appear the Villabruna individual belonged to a population that successfully dominated the post-glacial recolonisation of Europe, possibly through expansion from an Italian, Balkan or West Asian refugia. However, some regions of the continent, notably Britain and Ireland in the extreme northwest, have yet to be examined; although preliminary PCA and ADMIXTURE analysis from Chapter Two would suggest Irish Mesolithics samples potentially belong to the same grouping.

Other distinct hunter gatherer populations have also been identified in Holocene Europe. In Russia, eastern hunter gatherers (EHGs), represent a separate lineage (see Chapter Two, Fig. 1.2), albeit with some ancestry related to WHGs, as well as modern Caucasus, Native American and Eskimo-Aleut groups. Mitochondrial (U4a and C1) and Y haplogroups (R1a and J), absent in WHGs, are observed in these individuals. A separate group, Scandinavian hunter-gatherers (SHGs) were borne out of an admixture event, whereby EHG, entering the peninsula from the Northeast, followed the ice-free Norwegian coast downward where they met and hybridised with WHG groups who were expanding

into the region from the south (Günther et al. 2017). SHGs possess both U5 and U4 mitochondrial lineages, although their dominant Y haplogroup appears to be I2, associated with WHGs (Lazaridis et al. 2014; Günther et al. 2017). Hunter-gatherer populations from the Balkans, Hungary and Latvia, while all predominantly 'WHG' in ancestry (Fig. 1.3), also show varying degrees of introgression from sources related to EHG (González-Fortes et al. 2017; Jones et al. 2017; Mathieson et al. 2017). Conversely, Ukrainian HGs show a majority EHG ancestry, with significant WHG input. Moreover, hunter-gatherer individuals from the Balkans and Hungary possess a preponderance of mitochondrial haplogroups not observed in other WHG or EHG populations, such as K1 and H, possibly indicative of genetic contacts with west Asian populations (Mathieson et al. 2017).

In other regions of western Eurasia, hunter-gatherer populations resulting from deeper splits in the OoA migration have been identified, who possess the Basal Eurasian ancestry absent in their European counterparts (Lazaridis et al. 2016). In the Caucasus, continuity between two hunter-gatherer individuals (CHGs) is seen across the Later Upper Palaeolithic to Mesolithic boundary, which represents a lineage that split from WHG ~45 kya previously (Jones et al. 2015). This ancestry is also found at high levels in Iranian Mesolithic and Neolithic populations, as well as modern individuals from the South Caucasus (See Chapter Two, Fig. 1.3 for visualisation). To the southwest, in the Levant region and Anatolia, Epipaleolithic and early Neolithic populations also form distinct genetic groupings, which will be discussed in greater detail in Chapter 3. A notable difference in genetic height has been observed between these West Asian groups and their European counterparts, with WHGs, particularly those from the west of the continent, showing greatly increased estimates relative to individuals from the Caucasus and Anatolia (Martiniano et al. 2017).

Different populations also show variability in pigmentation. The first WHGs to be sequenced were seen to possess a unique configuration of dark skin and blue eyes (Lazaridis et al. 2014; Olalde et al. 2014). In contrast, SHGs and EHG, as well as hunter-gatherers from the Balkans, Latvia and the Ukraine all have intermediate to high frequencies of the derived allele at SLC24A5 associated with lighter skin pigmentation, which is fixed in modern European populations (Günther et al. 2017; Jones et al. 2017; Mathieson et al. 2017), suggesting that dark skin pigmentation was more of a Western European phenomenon. Moreover, SHGs and EHG also show significant frequencies of the derived allele at another skin colour locus, SLC45A2, which decreases in frequency on a North to South gradient in modern Europe (Norton et al. 2007). Alleles associated with blond hair in Europeans are also noted in Eastern and Scandinavian groups (Mathieson et al. 2017). Finally, blue eye colour, while common in Latvian and Scandinavian HGs, appears at lower frequencies in the Balkans, Ukraine and Russia (Mathieson et al. 2015; Günther et al. 2017; Mathieson et al. 2017).

Insight into the demographic history of HG populations in Europe has been limited due to the lack of available high coverage genomes. However, analysis of the Bichon genome, alongside two Caucasus HG, indicated these individuals had undergone population constrictions in the past, relative to early European

farmers (Jones et al. 2015). Furthermore, the historic effective population size for the Loschbour individual was estimated using PSMC, indicating a very small population contributed to its recent ancestry (Lazaridis et al. 2014). Many factors may influence these demographics, including the harsh climatic impact of the LGM and founder effects upon re-colonisations of northerly regions.

Ireland's Earliest Colonists

For the near entirety of the LGM Ireland and the majority of Britain were located deep underneath the expansive British-Irish Ice Sheet. As the climate began to warm again, the rapid release of meltwater into the Irish sea resulted in Ireland's emergence as an island from under the retreating glaciers no later than 15 kya (Edwards & Brooks 2008; Clark et al. 2012; Montgomery et al. 2014). Ireland has thus been isolated from the mainland continent for almost twice as long as the neighbouring island of Britain, which only lost its landbridge in the southeast *circa* 8 kya. This has profound implications for Ireland's recolonisation by floral and faunal species, as well as humans themselves. There is a clear decline in biodiversity from the southeast to the northwest of Europe, with Ireland, and a lesser extent Britain, possessing an extremely impoverished array of postglacial species. (Fløjgaard et al. 2011). Indeed, the rapid cycling of temperatures following the LGM in Europe would have pushed many populations, both cold- and warm-adapted, to extinction, and forced species, including humans, back into refugia. This not only includes established southern havens, e.g. Iberia and Northern Italy, but also higher latitude 'cryptic' refugia such as those proposed to the south and southwest of Ireland, as well as the Bay of Biscay (Provan & Bennett 2008; Montgomery et al. 2014).

While many subarctic and tundra species, such as arctic fox, lemming, stoat, reindeer, bear, hare and possibly wolf, were able to recolonise Ireland quickly after the LGM, the continued rapid rise in temperature may have compromised these species' ecosystems leading to the demise of some. Importantly, the initial post-LGM biome of Ireland was not completely restocked through Britain, but also from the aforementioned cryptic refugia of the Atlantic coast (Montgomery et al. 2014). This is evidenced by the western continental affinities and marked genetic divergence of several Irish mammals from their British counterparts, such as stoat (Martínková et al. 2007), hare (Hughes et al. 2006; Melo-Ferreira et al. 2007), otter (Finnegan & O'Neill 2010) and brown bear (Davison et al. 2011; Edwards et al. 2011). The increasing temperature also offered a potential window for warm-adapted species, such as deer, to enter Ireland before it was isolated by rising sea levels, however the Younger Dryas cold snap (13-12 kya) would have seen their complete disappearance, with any further introduction of species requiring a sea-crossing, most likely mediated by humans.

The impacts the above climatic and recolonisation events had on the movements of post-LGM human populations in the northwest of Europe is not fully understood. In Ireland, the earliest firm evidence of permanent human habitation dates to as recently as the Holocene, approximately 10,000 BP (Bayliss & Woodman 2009). However, recent evidence points to some level of human presence in the area during the Younger Dryas, in the form of several artificial cut markings on a bear bone dating to 12,810-12,590

cal BP) (Dowd & Carden 2016). The source of this possible Palaeolithic population, its permanence on the island across the Younger Dryas period, and its contribution, if any, to later Irish Mesolithic groups are all highly debatable points. While both the Magdalenian-related Creswellian Culture and the later, more temperate-adapted Federmesser, appear to have largely abandoned Britain by the onset of the Younger Dryas, specialised reindeer hunters (Ahrensburgian) from further north expanded into the Benlux and southern Britain during the period (Vermeersch 2011; Pettitt & White 2012). Although evidence of human occupation is sparse these groups may have contributed to seasonal occupation in other regions of Britain, and possibly Ireland.

As temperatures stabilised at the beginning of the Holocene, new Mesolithic cultures quickly repopulated northern-western Europe and Britain. These include the Sauveterrian of Northern France and Central Europe, related to the more southerly Azilian and characterised by minute geometric microliths, which expanded into the west of Britain (Wainwright 1961). Another microlithic culture, based on heavier axes, the Maglemosian (~11-8 kya), is seen to replace the Ahrensburgian of the Northern European Plain. Maglemosian sites are known across southern and eastern Britain, most famously at Starr Carr in Yorkshire. The Maglemosian is also identified as the earliest Mesolithic culture of Scandinavia, which was landlocked alongside Britain at the time, with extensive geographical connections between the regions via the Doggerland.

A 2,000 year delay occurred between the Mesolithic occupations of Britain and Ireland. Requirement of seafaring knowledge may have been the main hurdle for any new arrivals, though the island's dense forest cover and impoverished array of post-glacial flora and fauna may not have encouraged permanent settlement (Woodman 2015). When Ireland was eventually colonised, it was most probably from a British source, with the earliest Irish stone toolkits, geometric microliths (Fig. 1.1C), broadly resembling those that emerged in the Britain during its Later Mesolithic period. An accidental colonisation event is unlikely, given the seemingly deliberate introduction of wild boar to the island at the time (Mallory 2013). A suggested origin of these immigrants is the now submerged landscape surrounding the current Isle of Man, visible from Ireland, where rising sea levels would have forced population relocations. Continued communication between new colonies and their flooding homelands may have been compromised, which could provide explanation for the lack of parallel development between the Irish Early Mesolithic and any contemporary British population (Mallory 2013), although outside contacts cannot be ruled out.

The Late Mesolithic (~6,750-4,000 BC) in Ireland is defined by a major break in cultural development, with earlier microliths being replaced by less sophisticated macrolithic technology (Woodman 2015). Though some have suggested this represents a second colonisation event, no plausible outside source exists (Mallory 2013). What we may instead be witnessing is simply a loss of technological traditions due to isolation and low population size (Henrich 2004). Thus, there is no strong evidence to suggest that any outside contact occurred over the 4,000 year course of Ireland's Mesolithic. Furthermore, it is estimated that as little as 800 individuals, and no more than 8,000, inhabited the island during the period,

painting a picture of an extremely small and isolated island population (Mallory 2013; Woodman 2015). The genomic analysis of two Late Mesolithic Irish individuals in the following sections, offers an avenue to explore these demographic scenarios, and allows the placement of the Irish Mesolithic in the wider emerging narrative of Europe's post-glacial recolonisation, although the absence at present of British Mesolithic samples for comparison is noted.

Methods

Sample Processing

A full description of molecular methodology and bioinformatic data processing is provided in Appendix II and summarised in the Methods section of Chapter Two.

Variant Calling and Dataset Preparation

For f - and D -statistic analyses, to avoid the adverse effects of post-mortem damage, and to maximize the number of SNPs available, a set of 2,733,477 autosomal transversions from the 1000 Genomes Phase 3 release was chosen (1000 Genomes Project Consortium 2015), all of which possessed a minor allele frequency filter of $>1\%$ across 1000 Genomes European populations. Randomised pseudo-haploid genotypes for these sites were called in ancient samples using the Pileup tool from GATK (McKenna et al. 2010) (See Appendix II for further detail). In addition, these sites were extracted for four Mbuti individuals from the SGDP dataset (Mallick et al. 2016), to provide an outgroup population, and merged with the ancient genotype data. All dataset manipulation and merging was carried out in PLINK v1.90 (Chang et al. 2015).

Imputed genotypes of ancient samples were used for runs of homozygosity (ROH) and phenotypic analysis. Methods for imputation using Beagle software (Browning & Browning 2007) are detailed in Appendix II. For ROH analysis, transversion sites with a MAF above 5% across the 1000 Genomes dataset were used. Only sites called in all samples with a genotype probability above 99 were considered. Three samples from the SGDP Panel B were also included as reference points, B_Mbuti-4 (African), B_French-3 (European) and B_Karitiana-3 (Native American), as well as diploid calls for the Loschbour genome, released with the SGDP, to act as a measure of imputation accuracy.

D - and f -statistics

D - and f -statistics (Reich et al. 2009; Green et al. 2010) were calculated out using the AdmixTools package (Patterson et al. 2012) with significance assessed using a block jackknife of 5cM in size. Statistics were considered significant if they possessed a Z -score greater than 3 (Patterson et al. 2012), which corresponds to a p -value of less than 0.001. The Mbuti population of Central Africa were used as an outgroup, given their undetectable levels of Eurasian admixture (Gurdasani et al. 2015). For information on ancient individuals included in these tests and the population keys used see Table 2.1.

Outgroup f_3 -statistics of the form $f_3(\text{Mbuti}; X, Y)$ were used to measure shared genetic drift between pairs of individuals. A heatmap was created for the resulting values to explore genetic clustering among samples (heatmap.2 package from R), where all individuals included required over 7,500 overlapping variants with all other individuals (Fig. 2.2). To further investigate the specific affinities of Irish HGs to low coverage European HG genomes, particularly those from the northwest, further outgroup f_3 -statistics of the form $f_3(\text{Mbuti}; X, \text{Irish_HG})$ were carried out, where Irish_HG represents both Irish genomes grouped in the same population. This maximised the number of SNPs available for analysis. A cut-off of at least 5,000 overlapping variants was imposed for all tests. This allowed the inclusion of four northwestern samples not shown in the heatmap visualisation (BerryAuBac, Chaudardes, Falkenstein, Rochedane).

Runs of Homozygosity

Distributions of ROH were explored for the imputed genomes of 19 hunter-gatherers and 5 Early Neolithic Iranians and Anatolians (Broushaki et al. 2016; Hofmanova et al. 2016) as well as the high coverage genomes of three modern individuals. Diploid calls from the high coverage Loschbour genome were all included, for comparison with imputed calls from a downsampled (2X) version of the same genome. In total 643,143 SNPs were available for analysis (see above for selection process). ROH were calculated using PLINK, with the following parameters -

```
--homozyg --homozyg-density 50 --homozyg-gap 100 --homozyg-kb 500 --homozyg-snp 50 --homozyg-window-het 1 --homozyg-window-snp 50 --homozyg-window-threshold 0.05
```

Resulting individual ROH were then first divided into two classes, long ROH (≥ 1.6 Mb), indicative of recent endogamy, and short to intermediate (<1.6 Mb), a result of more ancient population bottlenecks (Pemberton et al. 2012). For both classes, the summed total length of ROH was calculated for every individual. These values were then plotted against each other, presented in Fig. 2.3.B, giving an overview of ROH length distribution for all genomes. To investigate the length distribution of ROH in more detail we placed our ROH into size bins (Kirin et al. 2010) and calculated the total length of ROH in each bin for each individual, plotted in Fig. 2.3A.

Genomic Height Estimation

Height is a strongly heritable trait, though environmental factors also play a role (Jelenkovic et al. 2016). The GIANT consortium has captured 60% of this heritability, based on a panel of common variants (Wood et al. 2014), and polygenic risk scores from this consortium have been provided for ~ 2.5 million SNPs. These sites were extracted for all imputed genomes and filtered for transversion sites called across all individuals at a genotype probability of 99% or higher, leaving 111,825 sites for analysis. Following the methods of (Martiniano et al. 2017), combined genetic effects were estimated using the “--q-score-

range” parameter in PLINK (Chang et al. 2015), unrestricted in terms of p-value. Results, consistent with those observed in (Martiniano et al. 2017), are shown in Fig. 2.4.

Pigmentation Profiling

Imputed genotypes for both individuals were assessed at the two major variants known to impact human skin pigmentation, rs16891982 (SLC45A2) and rs1426654 (SLC24A5) (Lamason et al. 2005; Soejima & Koda 2007), as well as along a 13 SNP haplotype found in 97% of individuals with blue eye colour (Eiberg et al. 2008).

Table 2.1 List of hunter-gatherer samples analysed in the current chapter. Sample identifiers and contexts are given. For samples from which both whole genome sequence (WGS) and 1240k SNP capture (Mathieson et al. 2015) data is available, the data type used in this study is noted in the sample ID. Furthermore, the inclusion or exclusion of each sample in the different analyses presented here is noted in the final four columns. No. of variants refers to the total number of secure calls for a set of 2,733,477 autosomal transversions. Anatolian Neolithic individuals used for *D*-statistics are not shown here, but were sourced from Mathieson et al. 2015; Broushaki et al. 2016, Hofmanova et al. 2016 and Kılınç et al. 2016.

Reference Code: ¹Skoglund et al. 2014; ²Jones et al. 2017; ³Haak et al. 2015; ⁴Mathieson et al. 2015; ⁵Gamba et al. 2014; ⁶Fu et al. 2016; ⁷Mathieson et al. 2017; ⁸Jones et al. 2015; ⁹Lazaridis et al. 2016; ¹⁰Current Study; ¹¹Lazaridis et al. 2014; ¹²González-Fortes et al. 2017; ¹³Olalde et al. 2014; ¹⁴Seguin-Orlando et al. 2014

Ref	Current Study ID	Other IDs	County	Lat	Long	Period	Arch. Culture	Date (BP) (*Not Direct)	Mito. Hg	Y Hg	Data Type	No. Variants	F ₃ -Stats	D-Stats	D-Stat Pop	ROH
1	PWC_Ajv52	Ajvide52	Sweden	57.29	18.21	MN	Pitted Ware	4,900-4,600*	V	..	WGS	199752	No	Yes	-	No
1	PWC_Ajv53	Ajvide53	Sweden	57.29	18.21	MN	Pitted Ware	4,900-4,600*	U4d	-	WGS	56596	No	Yes	-	No
1	PWC_Ajv58	Ajvide58	Sweden	57.29	18.21	MN	Pitted Ware	4,900-4,600*	U4d	I2a1	WGS	2180757	Yes	Yes	-	No
1	PWC_Ajv59	Ajvide59	Sweden	57.29	18.21	MN	Pitted Ware	4,900-4,600*	U	..	WGS	10636	No	Yes	-	No
1	PWC_Ajv70	Ajvide70	Sweden	57.29	18.21	MN	Pitted Ware	4,900-4,600*	U4d	..	WGS	344808	No	Yes	-	No
1	PWC_Ire8	Ire8	Sweden	57.29	18.21	MN	Pitted Ware	5100-4150*	U4d	..	WGS	95317	No	Yes	-	No
2	Latvia_HG3		Latvia	57.78	25.24	MS/EN	-	7,302-6,852	U5a2d	R1b1a1a	WGS	1289261	Yes	Yes	LVHG	Yes
3,4	Samara_HG	I0124	Russia	53.4	50.4	NE/CA	-	7,650-7,555	U5a1d	R1b1	Capture	145299	Yes	Yes	EHG	No
5,4	KO1_Koros_HG_WGS	KO1; I1507	Hungary	47.93	21.2	EN	Kőrös	7,780-7,640	R1b1	I2a1	WGS/Capture	1639877	Yes	Yes	SEHG	Yes
6,7	BerryAuBac		France	49.24	3.54	MS	-	7,320-7,170	U5b1a	I	Capture	10387	No	Yes	-	No
6	Chaudardes1		France	49.24	3.46	MS	-	8,360-8,050	U5b1b	I	Capture	19439	No	Yes	-	No
6	LesCloseaux13		France	48.52	2.11	MS	-	10,240-9,560	U5a2	-	Capture	1625	No	Yes	-	No
6	Ranchot88		France	47.91	5.43	MS	-	10,240-9,930	U5b1	-	Capture	83746	Yes	Yes	-	No
8	Kotias		Georgia	42.25	43.27	MS	-	9,895-9,529	H13c	J	WGS	2542214	Yes	Yes	CHG	Yes
6	Bockstein		Germany	48.33	10.09	MS	-	8,370-8,160*	U5b1d1	-	Capture	3476	No	Yes	-	No
6	Ofnet		Germany	48.49	10.27	MS	-	8,430-8,060*	U5b1d1	-	Capture	1135	No	Yes	-	No
6,7	Falkenstein		Germany	48.06	9.04	MS	-	9,410-8,990	U5a2c	I2a2a	Capture	15020	No	Yes	-	No
9	Iran_HotuIIIb	I1293	Iran	35.59	53.5	MS	-	11,120-10,650*	HV2	J	Capture	586128	Yes	Yes	-	No
10	Sramore62		Ireland	54.3	-8.33	MS	-	6,224-5,950	U5a2d	I2a2	WGS	2082759	Yes	Yes	Irish_HG	Yes
10	Killuragh6		Ireland	52.59	-8.34	MS	-	6,660-6,500	U5b2a	I2ab1	WGS	1651951	Yes	Yes	Irish_HG	Yes
6	Continenza		Italy	41.96	13.54	MS	-	11,200-10,510*	U5b1	-	Capture	3541	No	Yes	-	No
2	Latvia_HG1		Latvia	57.78	25.24	MS	Kunda	6,467-6,249	U5a1c	-	WGS	1571595	Yes	Yes	LVHG	Yes
2	Latvia_HG2		Latvia	57.78	25.24	MS	Narva	7,841-7,636	U2c1	R1b1a	WGS	2379461	Yes	Yes	LVHG	Yes
11	Loschbour		Luxembourg	49.81	6.4	MS	-	8,160-7,940	U5b1a	I2a1b	WGS	2656402	Yes	Yes	-	Yes
12	OC1_Meso		Romania	44.52	23.52	MS	-	8,973-8,435	K1 + 16362	R1b	WGS	2113936	Yes	Yes	SEHG	Yes
12	SC1_Meso		Romania	44.63	22.61	MS	-	9,075-8,553	U5b2c	R1	WGS	1642162	Yes	Yes	SEHG	Yes
12	SC2_Meso		Romania	44.63	22.61	MS	-	9,100-8,500*	U5a1c	R1	WGS	2374084	Yes	Yes	SEHG	Yes

4	Karelia_HG2	I0211	Russia	61.65	35.65	MS	-	8,800-7,950	U4a	J	Capture	37271	No	Yes	EHG	No
3,4	Karelia_HG1_WGS	I0061	Russia	61.65	35.65	MS	-	8,850-8,000*	C1	R1a1	Capture	1913693	Yes	Yes	EHG	Yes
12	Canes1_Meso		Spain	43.36	-4.72	MS	-	7,245-6,985	U5a2a	-	WGS	1615823	Yes	Yes	SPHG	Yes
13,4	LaBrana_WGS	LB1	Spain	42.91	-5.38	MS	-	7,990-7,740	U5b2c	C1a2	WGS/Capture	2216082	Yes	Yes	SPHG	Yes
12	Chan_Meso		Spain	42.73	-7.03	MS	-	9,255-9,007	U5b	-	WGS	2169843	Yes	Yes	SPHG	Yes
1	Sff11		Sweden	57.29	18	MS	-	7,500-7,250*	U5a1	..	WGS	205429	No	Yes	SHG	No
11,3,4	Motala2_SnpCap	I0012	Sweden	58.54	15.05	MS	-	7,714-7,575	U2e1	I2c	WGS/Capture	350263	Yes	Yes	SHG	No
11,3,4	Motala1_SnpCap	I0011	Sweden	58.54	15.05	MS	-	7,721-7,516	U5a1	-	WGS/Capture	238062	Yes	Yes	SHG	No
11,3,4	Motala12_WGS	I0017	Sweden	58.54	15.05	MS	-	7,721-7,631	U2e1	I2a1b	WGS/Capture	2245653	Yes	Yes	SHG	Yes
11,3,4	Motala4_SnpCap	I0014	Sweden	58.54	15.05	MS	-	7,877-7,629	U5a2d	-	WGS/Capture	277524	Yes	Yes	SHG	No
11,3,4	Motala6_SnpCap	I0015	Sweden	58.54	15.05	MS	-	7,964-7,629	U5a2d	I2a1	WGS/Capture	185382	No	Yes	SHG	No
11,3,4	Motala3_WGS	I0013	Sweden	58.54	15.05	MS	-	7,964-7,638	U5a1	I2a1b	WGS/Capture	1028253	Yes	Yes	SHG	No
2,7	Ukraine_HG1_WGS		Ukraine	47.44	35.25	MS	-	11,143-10,591	U5b2	R1a	WGS/Capture	365334	No	Yes	-	No
6	Ibousieres39		France	44.29	4.46	EP	-	12,040-11,410	U5b2b	..	Capture	42010	No	Yes	-	No
6,7	Rochedane		France	47.21	6.45	EP	-	13,090-12,830	U5b2b	I	Capture	48705	No	Yes	-	No
8	Bichon		Switzerland	47.1	6.58	EP	Azilian	13,770-13,560	U5b1h	I2a1	WGS	2589329	Yes	Yes	-	Yes
6	Villabruna		Italy	46.15	12.21	PA	Epigravettian	14,180-13,560	U5b2b	R1b1a	Capture	385662	Yes	Yes	-	No
6	GoyetQ-2		Belgium	50.26	4.28	PA	Magdalenian	15,230-14,780	U8a	HIJK	Capture	16358	No	Yes	MAG	No
6	Rigney1		France	47.23	6.1	PA	Magdalenian	15,690-15,240	U'2'3'4'7'8'9	-	Capture	7511	No	Yes	MAG	No
6	HohleFels79		Germany	48.22	9.45	PA	Magdalenian	15,070-14,270	U8a	..	Capture	2210	No	Yes	MAG	No
6	Burkhardtshohle		Germany	48.32	9.35	PA	Magdalenian	15,080-14,150	U8a	I	Capture	7082	No	Yes	MAG	No
6	Brillenhohle		Germany	48.24	9.46	PA	Magdalenian	15,120-14,780	U8a	..	Capture	2796	No	Yes	MAG	No
6	HohleFels49		Germany	48.22	9.45	PA	Magdalenian	16,000-14,260*	U8a	I	Capture	11813	No	Yes	MAG	No
6	ElMiron		Spain	43.26	-3.45	PA	Magdalenian	18,830-16,610	U5b	-	Capture	282054	Yes	Yes	MAG	No
8	Satsurblia		Georgia	42.22	42.36	PA	-	13,380-13,132	K3	J2a	WGS	1736218	Yes	Yes	CHG	Yes
6	AfontovaGora3		Russia	56.05	92.87	PA	-	16,930-14,490*	R1b	-	Capture	61571	No	Yes	-	No
6	KremsWA3		Austria	48.41	15.59	PA	Gravettian	31,250-30,690*	U5	..	Capture	10681	No	Yes	-	No
6	GoyetQ56-16		Belgium	50.26	4.28	PA	Gravettian	26,600-26,040	U2	-	Capture	5434	No	Yes	-	No
6	GoyetQ376-19		Belgium	50.26	4.28	PA	Gravettian	27,720-27,310	U2	-	Capture	1885	No	Yes	-	No
6	GoyetQ53-1		Belgium	50.26	4.28	PA	Gravettian	28,230-27,720	U2	-	Capture	228297	No	Yes	-	No
6	Vestonice16		Czech Republic	48.53	16.39	PA	Gravettian	30,710-29,310*	U5	IJK	Capture	376079	Yes	Yes	-	No
6	Vestonice43		Czech Republic	48.53	16.39	PA	Gravettian	30,710-29,310*	U	F	Capture	51641	No	Yes	-	No
6	Vestonice13		Czech Republic	48.53	16.39	PA	Gravettian	31,070-30,670*	U8c	CT	Capture	39955	No	Yes	-	No
6	Vestonice14		Czech Republic	48.53	16.39	PA	Gravettian	31,070-30,670*	U	C1a2	Capture	2344	No	Yes	-	No
6	Vestonice15		Czech Republic	48.53	16.39	PA	Gravettian	31,070-30,670*	U5	BT	Capture	9077	No	Yes	-	No
6	Pavlov1		Czech Republic	48.53	16.39	PA	Gravettian	31,110-29,410*	U5	C1a2	Capture	16606	No	Yes	-	No
6	Ostuni1		Italy	40.73	17.57	PA	Gravettian	27,810-27,430	M	-	Capture	102782	Yes	Yes	-	No

6	Paglicci108		Italy	41.65	15.61	PA	Gravettian	28,430-27,070*	U'2'3'4' 7'8'9	-	Capture	1397	No	Yes	-	No
6	Ostuni2		Italy	40.73	17.57	PA	Gravettian	29,310-28,640	U2	-	Capture	5159	No	Yes	-	No
6	Paglicci133		Italy	41.65	15.61	PA	Gravettian	34,580-31,210*	U8c	I	Capture	20486	No	Yes	-	No
6	Muierii2		Romania	45.11	23.46	PA	-	33,760-32,840	U6	-	Capture	27892	No	Yes	-	No
6	Cioclovina1		Romania	45.35	23.84	PA	-	33,090-31,780	U	CT	Capture	2506	No	Yes	-	No
6	GoyetQ116-1		Belgium	50.26	4.28	PA	Aurignacian	35,160-34,430	M	C1a2	Capture	2288	No	Yes	-	No
6	Kostenki12		Russia	51.23	39.3	PA	-	32,990-31,840*	U2	CT	Capture	14360	No	Yes	-	No
14,6	Kostenki14_SnpCap	K14	Russia	51.23	39.3	PA	-	38,680-36,260	U2	C1b	WGS/C apture	1111443	Yes	Yes	-	No

Results

Sequencing Results, Sex Determination and Uniparental Analysis of Irish Mesolithic samples

Methods and results for mitochondrial and Y chromosome haplogroup assignment were previously discussed in Chapter Two and are revisited in more detail here for the Mesolithic sample set. Killuragh6 was sequenced to 1X coverage, and Sramore62 to 1.7X. Each sample exhibited damage patterns typical of ancient DNA and extremely low levels of contamination (See Appendix II).

Y Chromosome Analysis

Both individuals were determined to be male (Electronic Data Table S3). Sramore62 was placed in Y haplogroup I2a2 and Killuragh6 in I2a1b. Neither individual could be placed in more specific subclades of these haplogroups, with ancestral alleles present for all possible downstream lineages. Moreover, Killuragh6 possessed the ancestral allele for one of the defining markers of I2a1b (CTS1802), and derived alleles at several other markers warranted his placement in the haplogroup. Both I2a1b and I2a2 lineages have been previously observed in Mesolithic individuals from western Europe, more specifically the Falkenstein (Germany) and Loschbour (Luxembourg) samples.

Mitochondrial Analysis

Mitochondrial coverages of 678X and 1,455X were retrieved for the Sramore62 and Killuragh6 tooth samples respectively. These were significantly higher than those obtained from petrous bone specimens sequenced to similar autosomal coverages, which fell between 61X and 290X (Extended Data Table X), fitting with previous observations of an inflated mtDNA to nuclear DNA ratio in teeth compared to petrous temporal bone (Hansen et al. 2017). Based on these high quality mitogenomes, Sramore62 was placed in haplogroup U5a2d and Killuragh6 in U5b2a. To place these results in a more detailed context, a map of mitochondrial haplogroup frequencies was constructed for the Pleistocene and Holocene HGs of post-glacial Europe (Fig. 2.1), based on published the whole genome displayed in Table 2.1 and mitochondrial data from (Bramanti et al. 2009; Posth et al. 2016; Mathieson et al. 2017).

The reported replacement of Magdalenian lineages, particularly U8a (Fu et al. 2016; Posth et al. 2016), is seen between the Pleistocene and Holocene in western Europe, with Epigravettian cultures in Italy and Epipalaeolithic individuals from France and Switzerland showing profiles more similar to later groups in the region. By the Mesolithic era U5a and U5b lineages are seen to be prevalent throughout the continent, with some geographical structure apparent. U5a1 is seen to be mostly restricted to EHGs and other eastern groups, while U5b and U5a2 lineages are found at highest frequencies in the west. Indeed, while only U5 haplogroups have been observed in the Mesolithic of western Europe so far, further east other U lineages belonging to the U2'3'4'7'8'9 clade are present in hunter-gatherers from Italy, the Balkans and the Baltic. Notably, the two haplogroups observed in Irish hunter-gatherers are

widely dispersed in Europe across the period, with U5a2d previously identified in Sweden, Latvia and Serbia, while U5b2a has also been noted in Latvia and Serbia, as well as France.

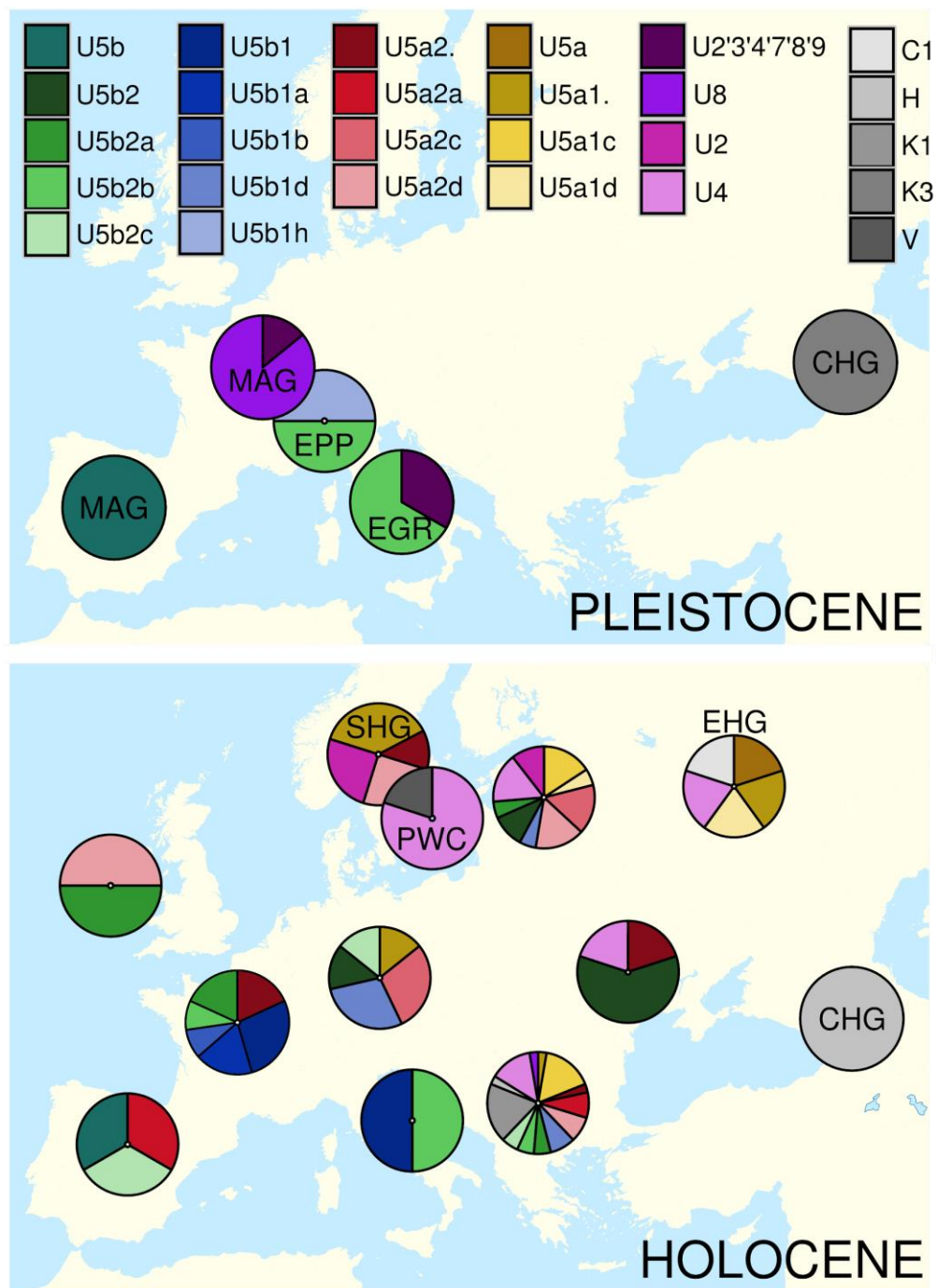


Figure 2.1. Mitochondrial Haplogroup distributions in postglacial hunter-gatherer groups of Europe. Frequencies are based on both published WGS and Mitochondrial sequence data. Individual data was partitioned into pie charts based on broader geographic regions. Samples predating the Younger Dryas event (~13-12kya) are presented in the top panel (Epigravettian - EGR, Magdalenian - MAG, and two Epipalaeolithic individuals - EPP), while hunter-gatherers from the Holocene era are shown in the bottom panel (Scandinavian -SHG, Caucasus - CHG, Pitted Ware Culture - PWC, Eastern -EHG).

Outgroup f_3 -statistics: Ireland's place within the genetic landscape of Mesolithic Europe.

To explore the patterns of affinity among the HG groups of Europe, and clarify the Late Mesolithic Irish population's placement within this structure, genetic clustering based on outgroup f_3 -statistics of the form $f_3(\text{Mbuti}; X, Y)$ was carried out. This revealed several distinct clusters of HG individuals, which correlated well with geography (Fig. 2.2). EHGs, CHGs and pre-LGM European individuals were clearly split from all postglacial European samples. The only exception to this observation was the postglacial Magdalenian ElMiron sample (Spain), which grouped together with the Aurignacian GoyetQ116-1 from Belgium, who predates ElMiron by almost 20,000 years, supporting previously observed continuity between these western samples (Fu et al. 2016). Unfortunately, other published Magdalenian samples were of too low coverage to be considered.

Within the larger grouping of post-glacial European hunter-gatherers, an approximate geographical split can be seen between eastern and western individuals, with further subgroupings apparent. Southeastern individuals from Hungary and Romania form a larger cluster with a group of Latvian HGs. Together these make up the approximately 'eastern' branch of the denodrogram, alongside Scandinavian HGs, who form their own discrete grouping. The other 'western' branching of the tree has one clear split, that seen between Spanish HGs, and individuals from the Northwest and Italy. The two Irish Mesolithic individuals are seen to fall securely within this latter group, which has samples stretching from northern Italy, along the Rhine basin to Ireland itself. Both Irish hunter-gatherers share extremely inflated genetic drift with each other, the highest of any pair of individuals tested. Following this, the Irish Mesolithic displays closest affinity to the Loschbour individual from Mesolithic Luxembourg, followed by the Azilian Epipalaeolithic individual, from Bichon, Switzerland.

The affinity of Irish HGs to other lower coverage northwestern continental samples was further explored using the statistic $f_3(\text{Mbuti}; X, \text{Irish_HG})$, where Irish_HG represents both Sramore62 and Killuragh6 combined as a population (Fig. 2.3). Similar values of shared drift are seen between Irish HGs and two northern French samples (BerryAuBac and Chaudardes), as that observed for Loschbour, in keeping with their geographic and temporal proximity to this individual.

Importantly, patterns of affinity between larger geographical groupings of HGs, identified in Fig. 2.2, are not always consistent for the individuals within them. For example, Latvia_HG2 and the KO1 individual from Hungary show increased shared drift with western groupings compared to other eastern samples, while Canes1 from Spain shows a slightly higher affinity to eastern groups, relative to earlier Spanish samples. However, some uniform patterns are apparent.

The first is the higher Magdalenian-type influence in western individuals, which can be visualised in the increased shared drift seen between these samples and the El Miron genome. Notably, two of the earliest western Mesolithic samples, Chan_Meso from Spain (9,255-9,007 cal BP) and Ranchot88 from France

(10,240-9,930 cal BP), show particularly inflated affinities to the Magdalenian sample, though the earlier Epipaleolithic individual from Switzerland does not (Bichon). A reverse trend in EHG affinity is seen, with Latvian and Southeastern HGs sharing more drift with the the Karelia and Samara HGs than their western counterparts. No significant inconsistencies can be seen between the patterns of shared drift displayed by Irish HGs and those seen in other northwestern samples.

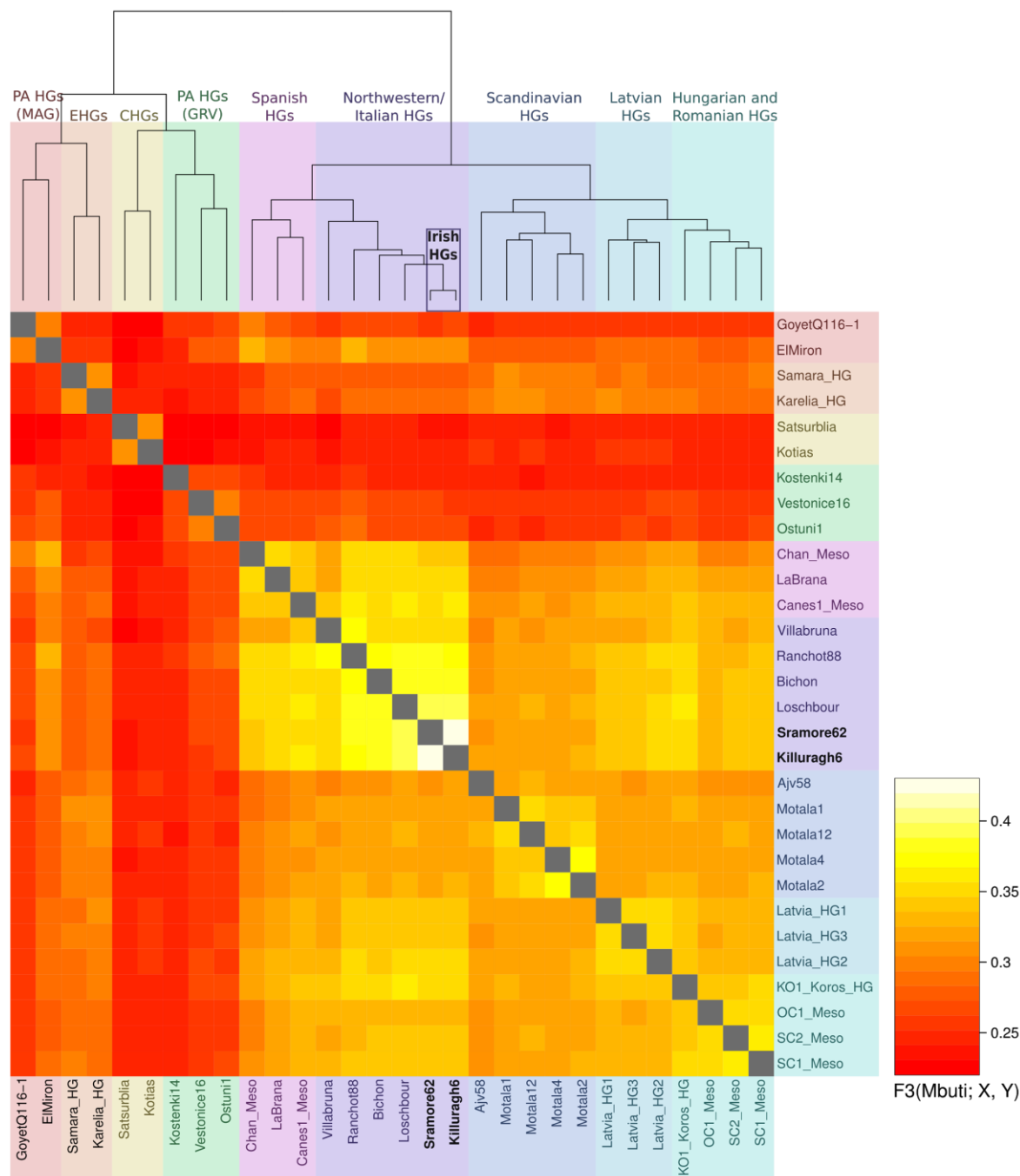


Figure 2.2. Heatmap of shared genetic drift for 30 hunter-gatherer individuals from pre- and postglacial Europe and West Asia. Drift is measured using the statistic $f_3(\text{Mbuti}; X, Y)$. All individuals share at least 7,500 overlapping SNP sites with all other individuals shown. Yellow indicates higher levels of shared drift.

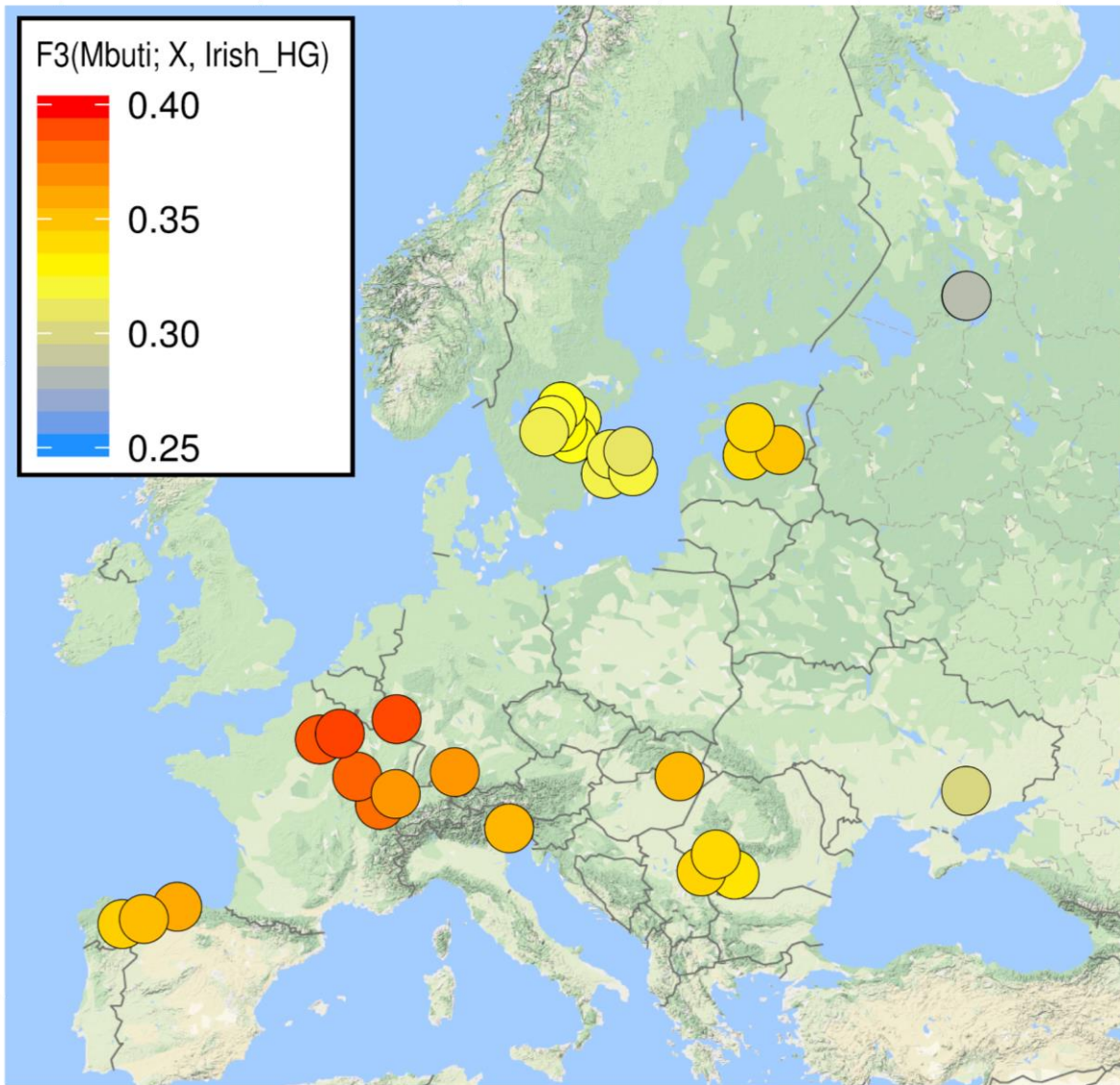


Figure 2.3. Levels of shared genetic drift between the Irish Mesolithic population and HG individuals from the continent. Only individuals dating to the Epipalaeolithic or later are considered. Drift is measured using the statistic $f_3(\text{Mbuti}; X, \text{Irish_HG})$. All individuals share at least 5,000 overlapping SNP sites with the Irish_HG population. Red indicates higher levels of shared drift. (Map Credit: Google)

D-statistics: Varying affinities within Mesolithic Europe.

Outgroup f_3 -statistics suggest that the Irish hunter-gatherer individuals can be viewed as an extension of the Northwestern Mesolithic of the continent. However, more subtle differences between the groups may be present, detectable through more direct methods of hypothesis testing achievable using D -statistics. Here, several four population test configurations are carried out to provide evidence for a number of statements, listed below.

i. Killuragh6 and Sramore62 form a distinct clade, separate from all other European hunter-gatherers

First, it was tested whether Irish HGs could be considered a clade distinct from all other HG groups in Europe and West Asia through the use of the statistic $D(\text{Mbuti}, X; \text{Sramore62}, \text{Killuragh6})$, where X is any other Palaeolithic or Mesolithic individual. No significant scores were obtained (Table 2.2). This was then retested using European HG groups, rather than individuals, and also HG and Neolithic groups from the Caucasus and Levant. Again, no significant scores were obtained (Table 2.2). Notably however, a handful of near significant scores were seen over these iterations ($Z > 2$; $p\text{-value} < 0.023$). These were for two SHG individuals from Motala, Sweden ($Z = 2.545\text{-}2.794$), as well as the SHG population as a whole ($Z = 2.028$), who showed some increased affinity to Killuragh6 relative to Sramore62.

The Anatolian population also displayed a similar aversion to Sramore62 ($Z = 2.067$), while the Gravettian individual Vestonice16 showed the opposite trend ($Z = 2.405$). It is unclear whether these signals are simply noise, suggested by their somewhat inconsistent nature, or if they are in fact indicative of outside contact during the Irish Mesolithic, or structure within the island over its duration. Perhaps it is worth remarking that several other Motala individuals also show slight affinity to Killuragh6 over Sramore62 ($1 < Z < 2$; $0.023 < p\text{-value} < 0.159$). It will be possible to explore these signals in a more robust manner when a denser Mesolithic sample set becomes available, specifically from Britain, which most likely mediated any prolonged incidence of contact between Ireland and the continent during the period.

To provide further evidence that Killuragh6 and Sramore62 can be considered as belonging to the same continuous population we asked whether the two samples formed a clade distinct from the higher coverage individual they share the largest amount of drift with, the Loschbour Mesolithic, as identified in outgroup f_3 -statistics. The test $D(\text{Mbuti}, \text{Irish_HG1}; \text{Loschbour}; \text{Irish_HG2})$ was constructed and in both possible formations majorly significant introgression was seen from one Irish_HG into another to the exclusion of Loschbour ($Z = 10.053$ and $Z = 12.599$). Given that the test $D(\text{Mbuti}, \text{Loschbour}; \text{Sramore62}, \text{Killuragh6})$ yields an insignificant score, the two Irish HGs are demonstrably part of the same population, distinct from northwestern continental groups.

ii. Loschbour shows higher amounts of introgression from varied sources, relative to Irish hunter-gatherers

The relationship between the Irish HG population (Irish_HG) and the northwestern Loschbour individual was further investigated through the use of the statistic $D(\text{Mbuti}, X; \text{Loschbour}, \text{Irish_HG})$, where X is rotated between different HG individuals and groupings, as well as Neolithic individuals from Anatolia. A clear pattern emerges, with a large number of values ($Z > 1$) suggestive of introgression into Loschbour from other European HGs to the exclusion of Irish HGs. In contrast, only four such values were indicative of gene flow in the opposite direction, i.e. introgression into Irish HGs to the exclusion of Loschbour, three from low coverage pre-glacial individuals, which may be treated as suspect due to lack of informative SNP sites (Table 2.2).

This suggests that while Loschbour forms a relatively secure grouping with the Irish hunter-gatherers, the Luxembourg individual has experienced admixture from continental sources that have not impacted the Irish population. This includes populations possibly related to the Magdalenian, evidenced both by affinity to samples sharing inflated drift with the Magdalenian cluster, such as the western Aurignacian individual GoyetQ116-1, Chan_Meso and the Spanish Mesolithic population as a whole ($Z=2.171-3.212$), as well as direct shared drift with the El Miron individual and the larger Magdalenian cluster itself ($Z=2.007-2.171$).

Swedish Motala samples and the larger SHG grouping are also seen to reach near-significant levels of introgression into Loschbour ($Z=2.529-2.884$), as they did in tests between Killuragh6 and Sramore62 from the previous section. Importantly, no such patterns are seen for EHG samples, suggesting it is not the EHG component of SHG ancestry that is driving this affinity. The recurrent uneven contributions of SHGs to clades composed of northwestern samples, perhaps is in part related to their own geographical origins, whereby they are acting as proxy for a yet to be sampled WHG population from the north or northwest Atlantic region of Europe. Good candidates may include Maglemosian groups of Britain, the Doggerland and Northern European Plain, which provided the first Mesolithic culture of Scandinavia.

iii. Irish hunter-gatherers form a more secure clade with the earlier Bichon individual, relative to Loschbour

The above tests were repeated using an older Epipalaeolithic individual from Bichon, Switzerland in place of the Mesolithic Loschbour, $D(\text{Mbuti}, X; \text{Bichon}; \text{Irish_HG})$. Despite its early date, Bichon possesses a high affinity to both Irish and Luxembourgian individuals, as seen in Fig. 2.2, suggesting substantial continuity in Western Europe from the Younger Dryas onwards. Unlike Loschbour, Bichon shows little evidence of introgression from continental sources to the exclusion of Irish samples. Moreover, the individuals who do produce near significant scores are low coverage in nature and do not follow any consistent geographic or temporal trend. Indeed, it is Irish HGs who show near significant introgression from several more northern sources, such as Mesolithic Latvians ($Z=1.492$), specifically Latvia_HG2 ($Z=2.338$) who showed more western affinities in outgroup f_3 -statistics; two SHG

individuals ($Z=1.49-2.103$); the French Mesolithic Ranchot88 ($Z=2.591$) and, to a lesser extent, Loschbour ($Z=1.357$).

Overall, it would appear that Irish HGs and Bichon have experienced little detectable admixture on their distinct lineages since they diverged from their shared ancestral population, though it would appear that Irish groups do have some slight increase in affinity to other Northern Mesolithic populations. Importantly, individuals who share highest drift with the Irish Mesolithic from the northwest, such as Loschbour, fail to significantly break the clade containing Bichon and Irish_HG, despite their much closer temporal and geographical proximity. This suggests that the Irish Late Mesolithic is not a recent offshoot of these or related groups, pointing to a deeper divergence for these northwestern populations.

Table 2.2 Testing clades of northwestern hunter-gatherers using D-statistics. For each test we ask whether individual or population X breaks the clade of Irish or Irish and northwestern hunter-gatherers. Red colours indicate higher shared drift between X and the right-hand individual or group (Killuragh6, Irish_HG) and blue scores with the left-hand (Sramore62, Loschbour, Bichon). Darker colours indicate higher significance.

Test	D(Mbuti, X; Sramore62, Killuragh6)			D(Mbuti, X; Loschbour, Irish_HG)			D(Mbuti, X; Bichon, Irish_HG)		
	Sample (X)	D Statistic	Z Score	Sites Used	D Statistic	Z Score	Sites Used	D Statistic	Z Score
Chan_Meso	0.0082	0.769	989463	-0.0262	-3.212	1715281	-0.0161	-1.524	1706522
Motala12_WGS	0.0122	1.618	1034581	-0.0176	-2.884	1797846	0.0078	0.884	204062
PWC_Ajv70	-0.0037	-0.445	164051	-0.0251	-2.658	280760	-0.0017	-0.141	278646
Motala3_WGS	0.0094	1.118	486143	-0.0171	-2.529	836998	0.003	0.383	830656
AfontovaGora3	0.026	1.307	33204	-0.0219	-2.449	55390	-0.0112	-1.237	55072
Vestonice14	-0.0668	-0.831	1119	-0.1167	-2.428	1952	-0.0845	-1.319	1947
GoyetQ116-1	0.0029	0.297	120408	-0.023	-2.399	204141	-0.0278	-1.382	14606
ElMiron	-0.0022	-0.241	144867	-0.0142	-2.007	249208	-0.0043	-0.443	247991
PWC_Ajv58	0.0102	1.761	1004858	-0.0149	-1.693	1742445	0.0082	1.064	1727779
Latvia_HG3	0.006	0.854	617544	-0.0121	-1.548	1063439	-0.0031	-0.281	1054280
Kostenki14	0.0081	1.252	532550	-0.011	-1.534	927506	-0.0109	-1.465	922459
PWC_Ajv52	0.0045	0.376	94928	-0.0176	-1.475	163064	0.0185	1.518	161702
LaBrana	-0.0014	-0.158	1023435	-0.0098	-1.466	1775247	-0.0023	-0.273	1764855
BerryAuBac	-0.0328	-0.69	5618	-0.0389	-1.453	9445	0.0043	0.136	9387
GoyetQ56-16	0.004	0.039	1043	-0.0945	-1.4	1730	-0.1123	-1.712	1717
GoyetQ-2	0.0319	0.987	8817	-0.0272	-1.374	14704	-0.0059	-0.739	203151
KO1_Koros_HG_WGS	0.0025	0.24	773294	-0.0119	-1.317	1334275	0.0003	0.023	1323643
Burkhardtshohle	-0.0484	-0.94	3817	-0.0292	-1.261	6410	-0.0653	-1.55	6370
Falkenstein	0.0176	0.564	7973	-0.0314	-1.227	13348	0.008	0.318	13253
Motala6_SnpCap	0.0278	2.545	97159	-0.0142	-1.21	163993	0.019	2.103	163181

A Genomic Compendium of an Island

Rigney1	-0.0306	-0.643	4075	-0.0329	-1.205	6778	-0.0514	-1.572	6743
Satsurblia	0.012	1.747	811932	-0.0078	-1.2	1403348	-0.0114	-1.45	1390465
Latvia_HG2	-0.0006	-0.066	1107775	-0.0083	-1.153	1926213	0.0151	2.338	1908824
Canes1_Meso	0.0135	1.612	756438	-0.0096	-1.09	1306019	0.0073	0.956	1299589
Bockstein	-0.1227	-1.941	1935	-0.0572	-1.065	3164	-0.0483	-1.233	3149
Karelia_HG2	0.0045	0.596	925294	0.0044	0.332	33499	-0.0335	-2.458	33329
Iran_HotuIIIb	0.0255	1.444	22164	-0.0096	-0.827	37152	-0.0148	-0.944	36981
Ranchot88	-0.0129	-0.858	45305	-0.0078	-0.811	76314	0.0312	2.591	75941
Bichon	-0.0075	-0.726	1156230	-0.0092	-0.794	2022121	-	-	-
SC2_Meso	-0.0023	-0.269	1071040	-0.0058	-0.784	1864803	0.0074	1.12	1850302
Vestonice16	-0.0207	-2.405	187506	-0.0058	-0.726	325687	-0.0004	-0.071	324120
PWC_Ajv53	0.003	0.128	27498	-0.012	-0.655	46699	0.0051	0.359	46483
OC1_Meso	0	0	960470	-0.0053	-0.631	1666898	-0.0034	-0.428	1654279
GoyetQ53-1	-0.076	-0.863	1249	-0.0282	-0.626	2047	0.0013	0.025	2040
Latvia_HG1	-0.0131	-1.483	750029	-0.0036	-0.58	1290260	0.0072	0.899	1279033
Motala2_SnpCap	0.0265	2.794	173786	-0.0049	-0.577	298537	0.0137	1.496	297111
SC1_Meso	0.01	1.241	756613	-0.0048	-0.52	1306397	0.0073	1.057	1297051
Pavlov1	0.0119	0.418	8424	-0.0126	-0.504	14569	0.0048	0.168	14509
Vestonice13	-0.016	-0.795	21229	-0.0073	-0.442	35940	-0.0342	-2.247	35750
Cioclovina1	0.0309	0.41	1322	-0.0241	-0.44	2263	0.0173	0.393	2257
HohleFels79	-0.0231	-0.355	1232	-0.0222	-0.405	1994	-0.0091	-0.155	1981
PWC_Ire8	-0.0172	-1.341	42752	-0.0048	-0.351	72536	0.0041	0.323	71888
Motala1_SnpCap	0.0121	1.37	120054	-0.0028	-0.283	205026	0.001	0.162	1782789
Ofnet	0.031	0.196	590	-0.0183	-0.187	1016	0.0266	0.421	1012
Motala4_SnpCap	0.0041	0.363	142016	-0.001	-0.131	241904	0.0027	0.222	240741
SfF11	0.0023	0.213	97171	-0.0008	-0.1	167240	0.0056	0.532	166119
GoyetQ376-19	-0.0093	-0.157	2780	-0.0045	-0.097	4812	0.0173	0.343	4778
Kotias	0.0054	0.772	1145749	0.0001	0.024	2003842	0.0023	0.35	1985440
Vestonice15	0.0154	0.319	4596	0.0014	0.045	7899	0.023	0.827	7871
Continenza	0.029	0.355	1713	0.0037	0.081	3080	0.0093	0.229	3058
Brillenhohle	-0.0058	-0.085	1466	0.0084	0.198	2463	0.0197	0.478	2449
Paglicci108	-0.0679	-0.573	696	0.0235	0.267	1230	0.0029	0.029	1222
PWC_Ajv59	-0.0281	-0.561	5083	0.0068	0.271	8643	0.0061	0.21	8603
HohleFels49	0.0587	1.386	6394	0.0079	0.303	10715	-0.0045	-0.202	10639
Karelia_HG1_WGS	0.0121	0.451	19863	-0.0063	-0.885	1596372	0.0018	0.242	1585246
KremsWA3	0.0025	0.06	5623	0.0116	0.425	9460	0.0197	0.942	9407
LesCloseaux13	-0.1628	-1.596	873	0.031	0.45	1482	0.0483	0.777	1466
Vestonice43	0.0006	0.03	27495	0.0057	0.454	46466	-0.003	-0.33	46218

Villabruna	-0.0042	-0.56	206404	0.0042	0.478	347163	0.01	0.912	345512
Ostuni1	0.006	0.49	54255	0.0056	0.59	91963	-0.0104	-1.011	91543
Ukraine_HG1	0.0053	0.406	169382	0.0056	0.67	289760	0.0066	0.805	286808
Ibousseries39	-0.083	-0.799	840	0.0481	0.837	1371	0.0141	0.236	1364
Rochedane	0.0112	0.458	26071	0.0098	0.839	43829	-0.0295	-1.745	43595
Muierii2	-0.0312	-1.154	15035	0.0152	0.907	25206	0.0032	0.166	25074
Chaudardes1	-0.0031	-0.079	10504	0.0205	0.958	17540	0.0162	0.726	17457
Samara_HG	-0.0047	-0.301	74899	0.0086	1.027	127103	0.0119	1.321	126431
Kostenki12	-0.013	-0.281	7311	0.0367	1.397	12434	0.004	0.19	12384
Ostuni2	0.0068	0.102	2557	0.0846	2.044	4484	0.034	0.832	4465
Paglicci133	-0.0125	-0.401	10907	0.037	2.634	18237	0.0063	0.376	18152
Loschbour	-0.002	-0.245	1165297	-	-	-	0.0169	1.357	2022121
Population (X)	D Statistic	Z Score	Sites Used	D Statistic	Z Score	Sites Used	D Statistic	Z Score	Sites Used
LVHG	-0.0005	-0.068	1154300	-0.0085	-1.429	2013324	0.0098	1.492	1994833
MAG	-0.0005	-0.057	153754	-0.0147	-2.171	264177	-0.0061	-0.665	262853
EHG	0.0036	0.494	970829	-0.0048	-0.732	1677630	0.0022	0.33	1665052
SEHG	0.0029	0.515	1167035	-0.0069	-1.025	2042037	0.0032	0.475	2023797
SPHG	0.0044	0.62	1109141	-0.0162	-2.581	1931913	-0.006	-0.747	1919681
CHG	0.0086	1.42	1163208	-0.0029	-0.547	2034834	-0.003	-0.479	2015937
SHG	0.0134	2.028	1095974	-0.0156	-2.746	1909719	0.0029	0.49	1893747
Anatolian	0.0094	2.067	1171182	-0.008	-1.827	2051136	-0.0028	-0.571	2032073

iv. Irish Hunter-Gatherers possess minimal Magdalenian influence in comparison to other western Mesolithic individuals

Given the apparent discontinuity between Irish hunter-gatherer population and their close relative, the Loschbour individual, with respect to Magdalenian-type ancestry, we decided to compare introgression from Magdalenian samples across a wider range of Mesolithic Europeans (Table 2.3). Increase in Magdalenian-type ancestry in certain Mesolithic samples was previously noted in (Fu et al. 2016), but can be expanded on here through the inclusion of newly published Latvian, Romanian and Spanish data. This was carried out using tests of the form $D(\text{Mbuti}, \text{MAG}; X, Y)$, where X and Y are any pair of European hunter-gatherers, and MAG represents the cluster of Magdalenian genomes identified in (Fu et al. 2016).

A clear division can be seen between eastern and western samples, mirroring that observed in outgroup f_3 -statistic analysis, with western samples, including the Irish Mesolithic, possessing inflated levels of Magdalenian-type ancestry relative to eastern samples from the Baltic, Hungary and the Balkans. Interestingly, among these more eastern samples, Latvians, particularly Latvia_HG2, show some higher affinity to Magdalenian samples relative to those in the Balkans, perhaps indicative of western contacts across the Northern European Plain. In the west, Irish samples show some of the lowest levels of

Magdalenian-type introgression. Indeed, only comparisons to the early Epigravettian Villabruna individual from Palaeolithic Italy produce a slight signal of increased drift between the Magdalenian group and Irish hunter-gatherers. Patterns most similar to the Irish Mesolithic are observed for the Epipalaeolithic Bichon individual, as well as the latest Spanish Mesolithic sample, Canes1_Meso, which shows marked discontinuity and eastern affinities relative to earlier Spanish samples. Magdalenian-type introgression into Loschbour relative to both Irish HGs is confirmed, but eclipsed by values seen for the earlier Ranchot88 individual of the same region (10,240-9,930 cal BP).

D(Mbuti, MAG; X, Y)	Y															
	Chan_Meso	LaBrana	Canes1_Meso	Villabruna	Ranchot88	Bichon	Loschbour	Killuragh6	Sramore62	KO1_Koros_HG	OC1_Meso	SC1_Meso	SC2_Meso	Latvia_HG1	Latvia_HG2	Latvia_HG3
Chan_Meso	0	-3.977	-5.303	-5.175	-1.101	-4.096	-5.213	-4.246	-5.638	-7.736	-9.393	-8.619	-10.437	-10.5	-8.909	-8.129
LaBrana	3.977	0	-2.408	-3.974	1.296	-1.872	-1.073	-2.01	-2.577	-5.655	-7.92	-7.343	-10.932	-7.267	-6.237	-5.403
Canes1_Meso	5.303	2.408	0	-1.217	4.955	0.827	1.482	-0.013	0.03	-4.095	-5.566	-5.107	-6.055	-4.672	-5.468	-5.022
Villabruna	5.175	3.974	1.217	0	4.666	2.312	3.137	1.202	1.922	-3.302	-4.022	-3.286	-4.29	-2.559	-2.255	-2.703
Ranchot88	1.101	-1.296	-4.955	-4.666	0	-2.107	-2.902	-3.514	-3.308	-6.517	-6.294	-6.692	-8.024	-6.936	-7.804	-4.655
Bichon	4.096	1.872	-0.827	-2.312	2.107	0	0.911	-0.276	-0.764	-5.72	-7.332	-6.612	-10.039	-4.662	-4.175	-4.567
Loschbour	5.213	1.073	-1.482	-3.137	2.902	-0.911	0	-1.784	-2.017	-6.333	-7.327	-7.378	-8.614	-5.17	-5.976	-5.73
X Killuragh6	4.246	2.01	0.013	-1.202	3.514	0.276	1.784	0	0.057	-5.329	-5.821	-6.528	-6.892	-3.916	-3.078	-4.018
Sramore62	5.638	2.577	-0.03	-1.922	3.308	0.764	2.017	-0.057	0	-4.154	-5.714	-7.441	-7.9	-7.13	-4.007	-4.42
KO1_Koros_HG_WGS	7.736	5.655	4.095	3.302	6.517	5.72	6.333	5.329	4.154	0	-0.876	-1.477	-2.701	-1.256	0.259	0.429
OC1_Meso	9.393	7.92	5.566	4.022	6.294	7.332	7.327	5.821	5.714	0.876	0	1.159	0.168	1.987	2.858	1.897
SC1_Meso	8.619	7.343	5.107	3.286	6.692	6.612	7.378	6.528	7.441	1.477	-1.159	0	-1.11	0.699	3.212	1.388
SC2_Meso	10.437	10.932	6.055	4.29	8.024	10.039	8.614	6.892	7.9	2.701	-0.168	1.11	0	1.858	3.194	2.935
Latvia_HG1	10.5	7.267	4.672	2.559	6.936	4.662	5.17	3.916	7.13	1.256	-1.987	-0.699	-1.858	0	0.691	0.42
Latvia_HG2	8.909	6.237	5.468	2.255	7.804	4.175	5.976	3.078	4.007	-0.259	-2.858	-3.212	-3.194	-0.691	0	-2.335
Latvia_HG3	8.129	5.403	5.022	2.703	4.655	4.567	5.73	4.018	4.42	-0.429	-1.897	-1.388	-2.935	-0.42	2.335	0

Table 2.3. Z-scores for D(Mbuti, Magdalenian; X, Y). Red colours indicate higher shared drift between the Magdalenian cluster and the Y individual and blue scores with the X. Darker colours indicate higher significance.

V. Irish hunter-gatherers display similar levels of EHG ancestry relative to other western Mesolithics.

Following a similar logic to that described above, we applied the test D(Mbuti, EHG; X, Y) across all pairs of European Mesolithic samples to investigate differential patterns of EHG introgression. Again, confirmation of patterns seen in outgroup f_3 -statistics (Fig. 2.2) was obtained, with a relative increase in EHG ancestry apparent in eastern populations from Hungary, Romania and the Baltic, as previously reported in (González-Fortes et al. 2017; Jones et al. 2017) and fitting with the geographical range of EHG. Furthermore, the early Spanish Mesolithic Chan_Meso displays significantly lower levels of EHG relative to all other individuals, including other Spanish HGs, suggesting this individual derives a substantial portion of their ancestry from relic populations of southwestern refugia, with inflated

Magdalenian-type ancestry and little outside influence. The decrease in Magdalenian influence in later Spanish individuals coupled with the increase of EHG is most likely linked to population movements from the east to west across Europe and into Spain. Interestingly, the Early Mesolithic Ranchot88 individual, whom also displayed inflated Magdalenian affinities, does not show the same aversion to EHG as seen in Chan_Meso, perhaps related to increased connectivity between hunter-gatherer groups further north, via the European plain. Indeed, the majority of northwestern hunter-gatherers, including the Irish Mesolithic, show no differences in EHG ancestry, though all appear to have somewhat higher affinities to this group relative to the earlier Italian Villabruna individual. This suggests that any EHG-related admixture in these individuals occurred after population expansions from southern refugia.

D(Mbuti, EHG; X, Y)	Y															
	Chan_Meso	LaBrana	Canes1_Meso	Villabruna	Ranchot88	Bichon	Loschbour	Killuragh6	Sramore62	KOI_Koros_HG	OCI_Meso	SCI_Meso	SC2_Meso	Latvia_HG1	Latvia_HG2	Latvia_HG3
Chan_Meso	0	4.225	6.616	3.783	2.408	5.435	8.099	8.036	4.898	8.895	7.804	8.965	8.286	10.837	11.27	11.656
LaBrana	-4.225	0	2.929	-0.4	0.526	0.953	2.221	2.107	1.179	4.317	4.135	4.535	4.608	8.699	7.033	7.72
Canes1_Meso	-6.616	-2.929	0	-3.187	0.639	-1.44	-0.459	-0.288	-1.457	1.963	1.468	2.489	1.184	7.077	5.667	4.948
Villabruna	-3.783	0.4	3.187	0	1.729	1.041	2.261	2.085	1.493	3.606	3.044	3.969	2.413	7.669	5.624	3.598
Ranchot88	-2.408	-0.526	-0.639	-1.729	0	-0.463	-0.175	0.36	-1.625	2.4	2.394	1.412	1.676	4.035	3.329	3.915
Bichon	-5.435	-0.953	1.44	-1.041	0.463	0	0.922	0.694	0.12	3.345	2.882	3.167	3.138	8.204	7.302	5.06
Loschbour	-8.099	-2.221	0.459	-2.261	0.175	-0.922	0	-0.623	-0.621	2.727	2.274	3.127	2.439	5.34	6.544	4.595
X Killuragh6	-8.036	-2.107	0.288	-2.085	-0.36	-0.694	0.623	0	-0.494	3.206	2.968	3.931	3.007	6.856	5.798	5.608
Sramore62	-4.898	-1.179	1.457	-1.493	1.625	-0.12	0.621	0.494	0	3.148	3.1	3.527	3.286	13.994	5.833	4.914
KOI_Koros_HG	-8.895	-4.317	-1.963	-3.606	-2.4	-3.345	-2.727	-3.206	-3.148	0	-0.854	0.078	-0.4	3.427	1.743	3.232
OCI_Meso	-7.804	-4.135	-1.468	-3.044	-2.394	-2.882	-2.274	-2.968	-3.1	0.854	0	0.542	0.03	3.827	2.98	3.025
SCI_Meso	-8.965	-4.535	-2.489	-3.969	-1.412	-3.167	-3.127	-3.931	-3.527	-0.078	-0.542	0	-0.333	3.495	2.902	2.381
SC2_Meso	-8.286	-4.608	-1.184	-2.413	-1.676	-3.138	-2.439	-3.007	-3.286	0.4	-0.03	0.333	0	3.892	2.434	2.483
Latvia_HG1	-10.837	-8.699	-7.077	-7.669	-4.035	-8.204	-5.34	-6.856	-13.994	-3.427	-3.827	-3.495	-3.892	0	-1.554	-1.065
Latvia_HG2	-11.27	-7.033	-5.667	-5.624	-3.329	-7.302	-6.544	-5.798	-5.833	-1.743	-2.98	-2.902	-2.434	1.554	0	-0.029
Latvia_HG3	-11.656	-7.72	-4.948	-3.598	-3.915	-5.06	-4.595	-5.608	-4.914	-3.232	-3.025	-2.381	-2.483	1.065	0.029	0

Table 2.4 Z-scores for D(Mbuti, EHG; X, Y). Red colours indicate higher shared drift between EHG and the Y individual and blue scores with the X. Darker colours indicate higher significance.

vi. Loschbour possesses inflated SHG ancestry relative to other European Mesolithics

Finally, we tested differences in SHG ancestry among hunter-gatherer individuals from Europe, using the statistic D(Mbuti, SHG; X, Y). SHGs showed a less obvious pattern of affinities between eastern and western groups, with no definite split between the geographical regions evident (Table 2.5). The clearest signal was the lack of SHG introgression into Spanish Mesolithic individuals, relative to all other samples. In contrast, the Loschbour Mesolithic was seen to possess the relatively highest affinities to

SHGs, though this signal was more pronounced in comparisons with other western samples than with eastern groups. As noted above, this observation does not seem to be driven by EHG admixture, which shows a differing range of distributions (Table 2.4). The Killuragh6 hunter-gatherer also shows some near-significant values of SHG introgression relative to other individuals, including Sramore62, as noted above. Comparisons between other western samples (Villabrana, Bichon and Ranchot88) reveal little differences between the individuals.

		Y															
D(Mbuti, SHG; X, Y)		Chan_Meso	LaBrama	Canes1_Meso	Villabrana	Ranchot88	Bichon	Loschbour	Killuragh6	Sramore62	KO1_Koros_HG	OC1_Meso	SC1_Meso	SC2_Meso	Latvia_HG1	Latvia_HG2	Latvia_HG3
X	Chan_Meso	0	4.825	5.989	8.329	5.645	9.263	12.925	10.773	9.108	10.962	8.455	8.71	12.095	8.602	10.588	8.327
	LaBrama	-4.825	0	4.881	3.144	3.005	3.501	6.293	4.945	3.695	5.339	2.5	3.124	5.063	3.852	6.102	4.875
	Canes1_Meso	-5.989	-4.881	0	1.27	1.211	0.66	2.551	2.025	0.063	2.453	0.047	0.861	1.387	1.684	2.642	1.321
	Villabrana	-8.329	-3.144	-1.27	0	-1.046	-0.411	3.399	1.741	-0.255	0.984	-0.624	1.338	0.411	0.716	0.886	-0.426
	Ranchot88	-5.645	-3.005	-1.211	1.046	0	-0.035	1.892	0.52	-0.251	1.338	0.225	1.349	1.426	0.782	0.636	1.699
	Bichon	-9.263	-3.501	-0.66	0.411	0.035	0	2.761	1.473	-0.464	1.102	-0.712	0.355	1.364	0.486	1.259	-0.117
	Loschbour	-12.925	-6.293	-2.551	-3.399	-1.892	-2.761	0	-1.014	-3.484	-0.816	-2.987	-1.04	-1.02	-1.026	-1.025	-1.575
	Killuragh6	-10.773	-4.945	-2.025	-1.741	-0.52	-1.473	1.014	0	-2.028	-0.15	-2.029	-0.561	-0.46	-0.587	-0.297	-1.054
	Sramore62	-9.108	-3.695	-0.063	0.255	0.251	0.464	3.484	2.028	0	1.617	-0.343	1.027	2.114	0.916	1.463	0.475
	KO1_Koros_HG	-10.962	-5.339	-2.453	-0.984	-1.338	-1.102	0.816	0.15	-1.617	0	-2.037	-0.343	0.172	-0.209	-0.48	-1.198
	OC1_Meso	-8.455	-2.5	-0.047	0.624	-0.225	0.712	2.987	2.029	0.343	2.037	0	1.3	2.169	1.304	1.841	0.557
	SC1_Meso	-8.71	-3.124	-0.861	-1.338	-1.349	-0.355	1.04	0.561	-1.027	0.343	-1.3	0	0.599	-0.045	0.199	-0.655
	SC2_Meso	-12.095	-5.063	-1.387	-0.411	-1.426	-1.364	1.02	0.46	-2.114	-0.172	-2.169	-0.599	0	-0.54	-0.264	-1.383
	Latvia_HG1	-8.602	-3.852	-1.684	-0.716	-0.782	-0.486	1.026	0.587	-0.916	0.209	-1.304	0.045	0.54	0	0.322	-0.616
	Latvia_HG2	-10.588	-6.102	-2.642	-0.886	-0.636	-1.259	1.025	0.297	-1.463	0.48	-1.841	-0.199	0.264	-0.322	0	-1.196
	Latvia_HG3	-8.327	-4.875	-1.321	0.426	-1.699	0.117	1.575	1.054	-0.475	1.198	-0.557	0.655	1.383	0.616	1.196	0

Table 2.5 Z-scores for D(Mbuti, SHG; X, Y). Red colours indicate higher shared drift between SHG and the Y individual and blue scores with the X. Darker colours indicate higher significance.

Runs of Homozygosity: Constricted population history for northwestern European hunter-gatherers

Genome-wide distributions of ROH are influenced by population history, both ancient and recent, as well as cultural practices, such as endogamy (Kirin et al. 2010; Pemberton et al. 2012). Previous studies have identified three approximate classes of ROH - short, medium and long. The first two of these categories encompass ROH under approximately 1.6 Mb in size, the total genomic length of which correlates strongly with geographic distance from East Africa in modern populations. In contrast, long ROH (> 1.6 Mb) lack this correlation, and tend to be more indicative of recent inbreeding. Here, such distributions are explored for the first time across a diverse set of European hunter-gatherer individuals, made possible through the use of imputed genotypes (Fig. 2.4). Divergent west Asian populations of a similar time depth are also analysed, to allow for relative comparisons of the demographic histories of the different regions.

In terms of short to medium ROH, a clear distinction can be seen between European and West Asian populations, implying more restricted historical populations sizes in the former, visualised through their inflated genomic lengths of short runs. The most parsimonious explanation for such an observation is the harsher climatic conditions which engulfed higher latitudes during the last glacial period, culminating in the LGM, although other factors such as serial migratory founder effects and persistent low population densities should also be considered. When longer ROH are considered, west Asian samples show more of a diversity of genomic lengths. Anatolian Neolithics possess the lowest levels of ROH across the entire ancient dataset and not dissimilar to that observed in the modern French population, demonstrative of a large outbreeding population. In contrast, both CHGs, but particularly the Palaeolithic sample Satsurblia, show some low level of inbreeding (Jones et al. 2015).

Within Europe a clear distinction can be made between individuals from the south and east, and those in the northwest, with the latter group displaying the highest levels of ROH seen across samples. EHGs and groups noted to have experienced EHG introgression, including Scandinavian, Baltic and southeastern hunter-gatherers, present relatively diminished levels of short and medium ROH, compared to their more western counterparts. This may be the result of increased incidences of admixture between diverse populations in these regions, driven by proximity to Asia. Indeed, SHGs are known to be the product of mixing between divergent western and eastern groups (Günther et al. 2017), while hunter-gatherers groups in the southeast have noted West Asian influence (Mathieson et al. 2017). Later Spanish Mesolithic individuals also display similar distributions of short and medium ROH to those from further east, though the earlier Chan_Meso sample is somewhat inflated for both shorter and higher runs. This concurs with the apparent change in genomic affinities seen in Spain over time through *D*- and *f*-statistics, which may have been mediated in part by migration, which works to increase diversity and lower ROH.

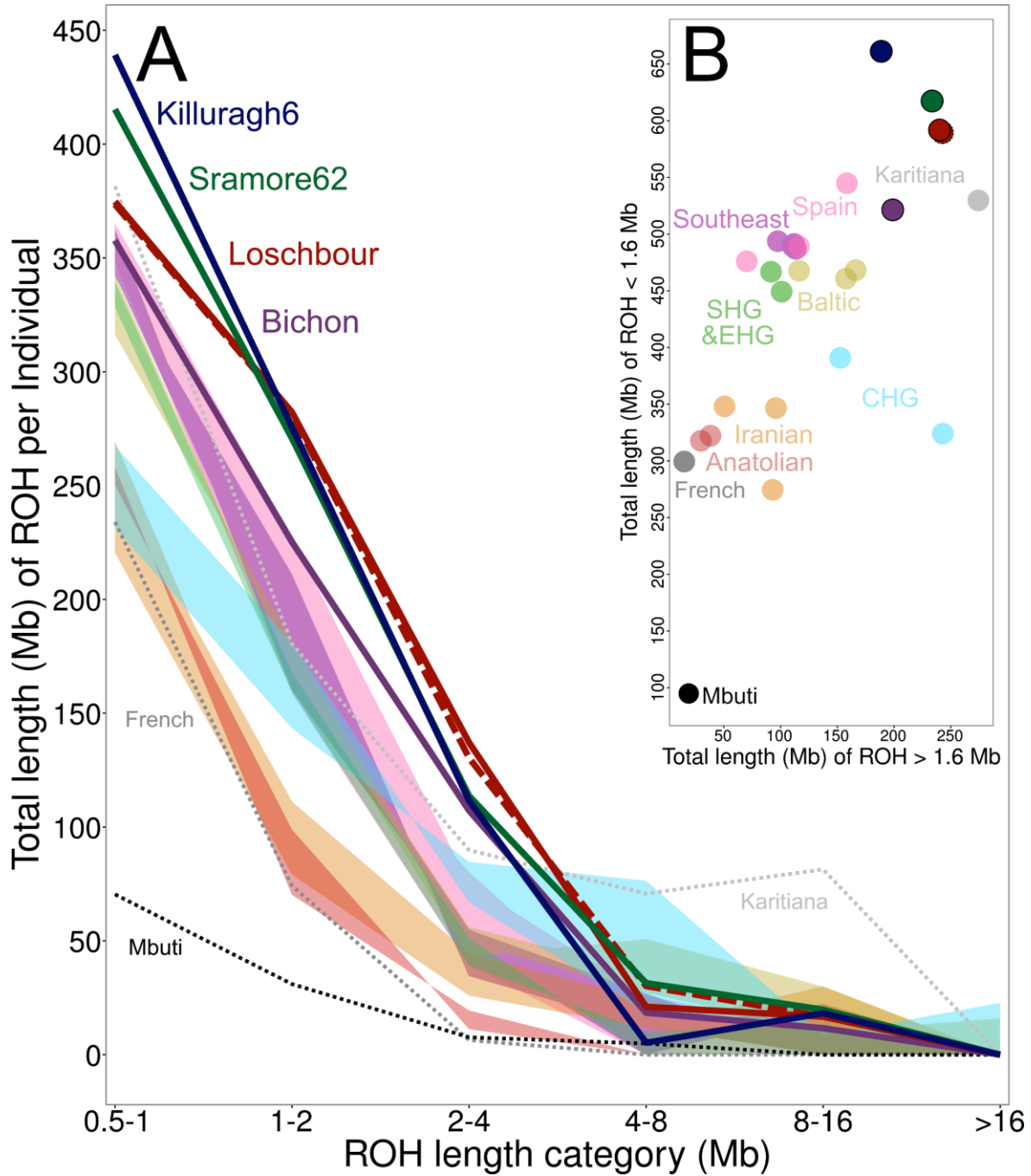


Figure 2.4. Estimated distribution of ROH for imputed ancient samples, placed in the context of values from modern individuals. A) Total length of ROH for a series of length categories in 23 ancient and 3 modern (dotted line) individuals. Distributions for ancient samples possessing similar ancestry and/or ROH profiles have been condensed into population-level silhouettes. Northwestern hunter-gatherer individuals are represented by solid bold lines. For Loschbour, both high-coverage diploid genotypes (dashed line), as well as imputed genotypes from downsampled data (solid line) are displayed. **B)** The total length of long ROH (> 1.6 Mb) against short ROH (< 1.6 Mb) plotted for the same set of individuals. Colour keys are identical throughout.

Strikingly, the two Irish Mesolithic samples display unprecedented lengths of short to medium ROH, particularly those of the shortest category (< 1 Mb), eclipsing all other samples in the dataset. This includes a modern Karitiana individual from an isolated Native American population in the Amazon, which forms the geographical extreme of the OoA migration. Indeed, the Irish samples, particularly Killuragh6, represent a clear step up in levels of short genomic ROH, relative to the closely related continental Loschbour Mesolithic, who is in turn elevated with respect to the earlier Epipalaeolithic Bichon sample from Switzerland, whose distribution is more similar to the southerly Spanish Chan. This may be in some part the result of serial founder effects occurring during the recolonisation of Northern Europe, as well as the demographically destabilising climatic effects of the Younger Dryas, which Bichon predates. Further bottlenecks can be expected to have occurred during the colonisation of a relatively under-provisioned postglacial island, which may help explain the large discrepancy between Loschbour and the Irish HGs.

Notably, all four northwestern samples also share inflated levels of longer ROH relative to southern and eastern samples, though this does not extend to ROH past 4 Mb in length. Indeed, all European HG individuals considered share remarkably similar low distributions of very long ROH, indicating that none of them are the result of recent inbreeding events. Thus, while the Irish Mesolithic clearly underwent severe population constrictions in its past, the samples investigated here appear to belong to breeding populations with enough complexity to prevent consanguineous pairings.

Finally the reduction of short ROH in Sramore62, relative to Killuragh6, is notable given the individual's later date (4,224-3,950 cal BC), which straddles the Neolithic boundary in Ireland. During this period severe demographic upheavals were occurring on the continent and Britain, which may have worked to displace HG populations. Any resultant migrations across the region would have worked to increase diversity and lower levels of ROH among HGs. Furthermore, if such events impacted Ireland they may have contributed the subtle differences in genetic affinities seen between the Killuragh6 and Sramore62 samples (Table 2.2), as well as in short ROH levels, though full investigation of such possibilities will require denser sampling.

Appearance of the Irish Mesolithic

Pigmentation

Both Killuragh6 and Sramore62 displayed pigmentation profiles identical to that reported for other WHGs (Gamba et al. 2014; Lazaridis et al. 2014; Olalde et al. 2014). Ancestral alleles were found at both rs1426654 and rs16891982, indicative of dark skin and hair pigmentation. The two individuals were also homozygotes for the 13-SNP haplotype found in 97% of Caucasians with blue eye colour, including the likely causative mutation (rs12913832; Eiberg et al, 2008).

Height

Previous work showed inflated genomic height in European HGs, relative to contemporaneous west Asian samples (Martiniano et al. 2017). This is expanded here with the addition of newly imputed individuals, including Russian, Latvian and Romanian HGs (Fig. 2.5). The same general trends were observed, with HGs from across Europe estimated to be genetically taller than contemporaneous Iranian, Anatolian and Georgian individuals. Clear variability in genomic height exists across European HG groups, possibly with some geographical trends, although samples sizes are too low to ascertain these reliably. Northwestern samples from the Rhine Basin (Bichon and Loschbour) show the highest values, while Irish HGs display somewhat reduced scores, despite their close relation to these individuals. Swedish, Spanish and Russian individuals display the lowest scores observed for European HGs, though the risk scores estimated were still net positive.

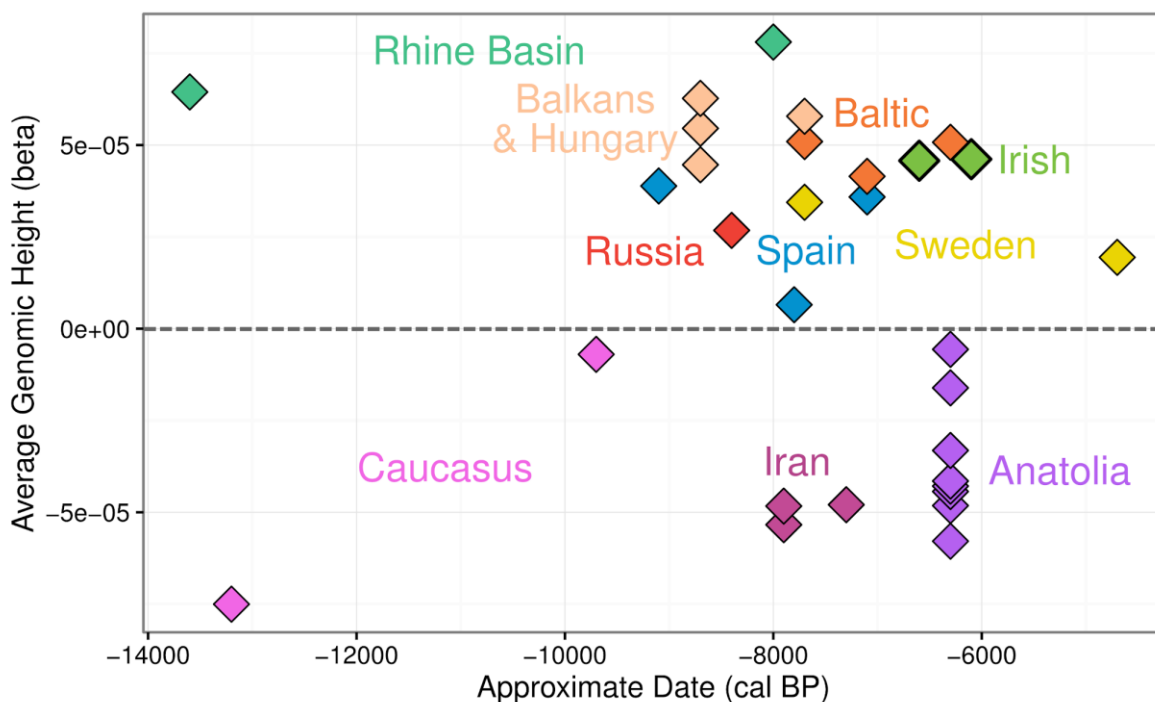


Figure 2.5 Average genomic height for imputed European and West Asian ancient samples. Individual polygenic scores are plotted against date and data points are coloured by geographical region.

Conclusions

The arrival of *Homo sapiens* on the island of Ireland is a recent phenomenon in the species' history, with the earliest firm evidence of permanent occupation dating back approximately only 10,000 years (Bayliss & Woodman 2009). To put a global perspective on this, the majority of Europe, including the neighbouring island of Britain, had seen intermittent human settlement from at least 42,000 years ago onwards (Higham et al. 2011), while more intrepid migrants had covered thousands of kilometers to reach the southern end of South America by at least 14,500 BP (Dillehay et al. 2015). Notably, the Mesolithic period accounts for more than a third of the island's brief human occupation and yet relatively little archaeological evidence remains of Ireland's earliest inhabitants. Less than a dozen sites have produced Mesolithic human remains, and the majority of those recovered are not amenable to ancient DNA analysis (Meiklejohn & Woodman 2012). Thus, the successful retrieval of whole genome sequence data (1-1.7X) from two separate cave sites on the island, presented here, offers an unparalleled insight into Ireland's first, and relatively late, Mesolithic arrivals.

Importantly, the dataset of published European HG genomes (Table 2.1), used here to provide a reference for these Irish samples, has not been analysed previously as a whole. Their inclusion not only provides interpretative fodder for the origins, demographics and phenotypic characteristics of the Irish Mesolithic, but also contributes to a more holistic synthesis of the wider population dynamics of Palaeolithic and Mesolithic Europe. The next sections consider the questions pertaining to the origin of the Irish Mesolithic, specifically the latter part of the period from which both samples date (6,660-5,950 cal BP), at three different ancestral layers.

Postglacial Expansions

The first, and deepest of these layers lies in the hunter-gatherer populations who emerged from the LGM to dominate the recolonisation of Western Europe. Two distinct genetic clusters successively occupied the post-glacial landscape of the region, indicating that some level of demographic discontinuity occurred across the Younger Dryas and into the Holocene (Fu et al. 2016; Posth et al. 2017). The genomic legacy of both groups are relevant to the question of Irish Mesolithic origins. The earlier cluster, sampled from Spain to Germany (18,830-14,780 cal BP), is associated with the Magdalenian Culture, which depended on large-game hunting that became increasingly unsustainable as temperate forests reclaimed much of the continent.

While the Younger Dryas (12,900-11,700 BP) cold snap offered a brief respite, the Magdalenian and its daughter cultures in the north, including Britain and potentially Ireland (Dowd & Carden 2016; Pettitt & White 2012), eventually disappeared, replaced by diverse Epipalaeolithic and Mesolithic microlith-users, who exploited woodland and aquatic resources. Strikingly, the earliest Epipalaeolithic individuals from Western Europe (13,770-12,830 cal BP) (Jones et al. 2015; Fu et al. 2016), alongside all later Northwestern hunter-gatherers, including the Irish Late Mesolithic, are divergent from prior

Magdalenian samples, clustering instead with a Palaeolithic individual from the Epigravettian site of Villabruna in Northeastern Italy (14,180-13,560 cal BP) (Fig. 2.2).

Magdalenian type ancestry did not disappear entirely, with substantial input seen into the earliest sampled Mesolithic individuals of Spain and France (10,240-9,007 cal BP). Its persistence in these regions, particularly Spain, is consistent with the proposed origin of the Magdalenian in Franco-Iberian refugia. Importantly, of the western samples considered Irish HGs exhibit some of the lowest levels of Magdalenian introgression (Table 2.3). This indicates the majority of their ancestry derives not from such western Atlantic refugia, but further east, in groups related to the North Italian Villabruna individual, the only western sample that displays lower Magdalenian-type introgression relative to the Irish HGs.

The genetic composition of the Villabruna individual may itself have contributions from further east still, in Balkans or Near Eastern refugia, evident by its inflated affinity to modern West Asian populations compared to all earlier European HGs so far sampled (Fu et al. 2016). However, the analysis of newly published Mesolithic Latvian and Romanian genomes here (González-Fortes et al. 2017; Jones et al. 2017), alongside the Villabruna genome, allows us to demonstrate the individual's higher affinity to HGs in the west, compared to those further east. This suggests other post-glacial populations contributed to eastern groups, fitting with the mitochondrial diversity seen in these regions (Mathieson et al. 2017; Fig. 2.1), though they also exhibit substantial allele sharing with the Italian individual (Fig. 2.2). Thus the exact geographical origin of the LGM population that gave rise to prevailing 'Villabruna-type' ancestry of the European Mesolithic, both east and west, including Ireland, remains uncertain.

An Early Isolation?

The second layer that can be considered in the assessment of Irish Mesolithic origins is embodied by the Epipalaeolithic and Mesolithic cultures that diversified north across western Europe with rising temperatures. In terms of shared genetic drift, the Irish Late Mesolithic appears to be an extension of such groups (Fig. 2.2), with highest allele sharing seen for its close continental contemporaries in Luxembourg (Loschbour; 8,160-7,940 cal BP), as well as Northern France (BerryAuBac and Chaudardes1; 8,360-7,170 ca; BP) (Fig. 2.3). These regions are situated just south of the once existent land-bridge that gave entry into Britain, the western coast of which provides the most likely springboard into Ireland. Genetic clustering suggests the Irish and northwestern individuals form a cohesive group with older Epipalaeolithic samples from Jura region of France and Switzerland (13,770-12,830 cal BP), who date prior to the Younger Dryas cold snap (Fu et al. 2016; Fig. 2.2). This includes the Bichon individual belonging to Azilian Culture, a suggested progenitor industry of the microlithic traditions that later spread north, which now can be more securely interpreted as a demographic expansion. Somewhat congruently, the earliest sites of Azilian industries are at the southern feet of the Alps in Switzerland and Northwest Italy, which saw early reforestation (Plisson et al. 2008). Thus, the Italian peninsula can be considered both a leading geographic and genetic candidate region for much of the 'Villabruna-type' ancestry present in the Mesolithic populations of northwestern Europe.

Comparison of Irish HGs to their closely related contemporaries in Northern France and Luxembourg, using older Epipaleolithic samples from the south as reference points, can allow inferences to be made about the ancestral population of the Irish Late Mesolithic, based on differences between the groups. As noted above, Irish HGs exhibit substantially less Magdalenian-type introgression (Table 2.3) relative to other northwestern Mesolithics, more on par with that observed in Bichon, and only slightly above that seen in Villabruna. The large inflation of Magdalenian-type ancestry in closely related early samples, such as Ranchott88, suggests the progenitor population of the Irish Late Mesolithic was isolated early on from such groups and did not reside in an area where prolonged admixture with Magdalenian-type populations could occur. Britain or Ireland provide good candidates for such regions, as, given their near total abandonment during the Younger Dryas, any Mesolithic colonists were likely entering uninhabited territory. Iberia presents quite the reverse situation, which may explain the persistence of such ancestry on the peninsula, embodied most strikingly in the Chan_Meso individual (González-Fortes et al. 2017), though its decrease over time on the northern coast suggests continued dilution from outside sources.

Such dilution may also provide an alternative explanation for the limited observance of Magdalenian-type ancestry in Late Mesolithic Ireland. In this scenario, any Magdalenian-type introgression accumulated by the progenitors of the Irish Mesolithic, post-divergence from the ancestral population shared with Bichon, may have been subsequently displaced through admixture with other unsampled groups. The decrease in Magdalenian-type introgression between Ranchot88 and the later Loschbour individual from a nearby region may provide evidence for such a dilution over time. However, the lack of introgression in Irish HGs from any other tested HG source to the exclusion of Loschbour (Table 2.2) argues against such a scenario. Indeed, while Irish HGs show some slightly increased affinity to Northern populations relative to Bichon, such as SHGs and Baltic HGs, which may provide the diluting factor necessary via the Northern European Plain, Loschbour shares more alleles with both these groups, while also retaining more Magdalenian-type introgression relative to the Irish.

Thus, the more parsimonious explanation remains the demographic isolation of the ancestral population of Irish HGs in comparison to related continental groups. Substantially inflated levels of short ROH runs (<1 Mb) in Irish HGs (Fig. 2.4), relative to the continental Loschbour provides further evidence for such a scenario. Notably, both Loschbour and Irish HGs also show evidence of ancestral bottlenecks relative to Bichon. This may be related to the latter's more southern location and its pre- Younger Dryas date, an event which would have encouraged population constriction and isolation, particularly in groups further north.

Origins of the Irish Late Mesolithic

The final and most recent layer of ancestry pertaining to Irish Mesolithic origins, regards the exact progenitor group that gave rise to the Late Mesolithic population of Ireland sampled here, and whether differences in ancestry exist between the two individuals that compose this population. Without a

comparable sample set from Britain, or indeed the Irish Early Mesolithic, such questions are by far the hardest to address, and this final section of the conclusion is thus more speculative in nature.

The archaeological record provides no firm evidence of any outside contact or invasive cultures throughout the course of the Irish Mesolithic, though a clear shift in lithic technology is apparent from approximately 6,500 BC, marking the start of the later phase of the period and providing the most viable window for new population entry. The genomic data presented here mimics the archaeology, with no outside migration across the period required to interpret the patterns of allele sharing and ROH distributions observed in Mesolithic Irish and continental samples, which as noted above, indicates relatively early isolation between the two groups. Indeed, short ROH in Irish HGs are substantially higher than those witnessed in some of the most isolated human societies in the Amazon today, placed at the end of the OoA migration. This extreme signature is exactly what would be expected from a population that has undergone a narrow ancient bottleneck, such as an island colonisation, that was not alleviated by recurrent migration and contact. Though not inconceivable, such an inflation of short ROH is harder to explain if the progenitor population of the Irish Late Mesolithic was in Britain, which could be simply considered another region of continental Europe up until 6,000 BC, with land connections both south into Belgium and Northern France, and east across the Doggerland and Northern European Plain. Moreover, if migration did occur from an outside source into Ireland at onset of the Late Mesolithic, ROH in the resulting population would be reduced due to likely admixture between indigenous and colonising groups, while stable sea levels would have encouraged maintained contact with source populations, further working to erode these tracts.

Arguments against prolonged isolation within Ireland may find some support in the several slight inconsistencies in allele sharing between Killuragh6 and Sramore62, particularly in relation to SHGs, which may be the result of subtle structure within the island, or alternatively differential contacts from outside the island. Given that the earliest Scandinavian Mesolithic industries (Maglemosian) were also present in Britain, this is not an improbable scenario, as British populations themselves may have presented diversity in their affinities to the Scandinavian Mesolithic. The inflated allele sharing also seen between Loschbour and SHGs highlights the potential complexity of such structure across northern Europe, which can only be resolved through dense sampling, though this is unfortunately impossible for many submerged regions. Slight deflation of short ROH in Sramore62 may also point to outside interaction. However, it must be noted that as this sample dates to the very end of the Mesolithic, such speculated contacts, if correct, may not have been typical across the Irish Mesolithic period, but induced by unprecedented environmental and demographic change brought about by the arrival of the Neolithic in northwest Europe.

Finally, the Irish Mesolithic population seem typical physical examples of western Europeans of the time (Lazaridis et al. 2014; Olalde et al. 2014). They were dark-skinned, black-haired and blue-eyed. Moreover, they exhibited polygenic risk scores associated with increased height similar to other Mesolithic

Europeans. However, the impact of environment on the physical manifestation of such a complex trait must be noted. Indeed, the Loschbour skeleton measures only 160cm in physical height (Toussaint et al. 2009) and, while not included in analysis, the skeleton of the Villabruna measures no more than 169cm (Vercellotti et al. 2008). Notably, while no complete Mesolithic skeleton from Ireland has been recovered, the single such find in Britain (Cheddar Man) though only 166cm tall, possessed relatively long arms and legs, and showed several signs of growth stress (Cunliffe 2013).

References

- 1000 Genomes Project Consortium, 2015. A global reference for human genetic variation. *Nature*, 526(7571), pp.68–74.
- Armitage, S.J. et al., 2011. The southern route “out of Africa”: evidence for an early expansion of modern humans into Arabia. *Science*, 331(6016), pp.453–456.
- Bayliss, A. & Woodman, P., 2009. A New Bayesian Chronology for Mesolithic Occupation at Mount Sandel, Northern Ireland. *Proceedings of the Prehistoric Society*, 75, pp.101–123.
- Bietti, A., 1990. The late Upper Paleolithic in Italy: An overview. *Journal of World Prehistory*, 4(1), pp.95–155.
- Bramanti, B. et al., 2009. Genetic discontinuity between local hunter-gatherers and central Europe’s first farmers. *Science*, 326(5949), pp.137–140.
- Broushaki, F. et al., 2016. Early Neolithic genomes from the eastern Fertile Crescent. *Science*, 353(6298), pp.499–503.
- Browning, S.R. & Browning, B.L., 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localised haplotype clustering. *American Journal of Human Genetics*, 81(5), pp.1084–1097.
- Chang, C.C. et al., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4, p.7.
- Clark, J. et al., 2012. Response of the Irish Ice Sheet to abrupt climate change during the last deglaciation. *Quaternary Science Reviews*, 35(Supplement C), pp.100–115.
- Cunliffe, B.W., 2008. *Europe between the Oceans: 9000 BC-AD 1000*. New Haven-London: Yale University Press.
- Cunliffe, B., 2013. *Britain Begins*. Oxford: Oxford University Press.
- Davison, J. et al., 2011. Late-Quaternary biogeographic scenarios for the brown bear (*Ursus arctos*), a wild mammal model species. *Quaternary Science Reviews*, 30(3), pp.418–430.
- Dillehay, T.D. et al., 2015. New Archaeological Evidence for an Early Human Presence at Monte Verde, Chile. *PLoS One*, 10(11), p.e0141923.
- Dowd, M., 2015. *The Archaeology of Caves in Ireland*. Oxford: Oxbow Books.
- Dowd, M. & Carden, R.F., 2016. First evidence of a Late Upper Palaeolithic human presence in Ireland. *Quaternary Science Reviews*, 139, pp.158–163.
- Edwards, C.J. et al., 2011. Ancient hybridisation and an Irish origin for the modern polar bear matriline. *Current Biology*, 21(15), pp.1251–1258.
- Edwards, R. & Brooks, A., 2008. The island of Ireland: Drowning the myth of an Irish land-bridge? *The Irish Naturalists’ Journal*, 29, pp.19–34.
- Eiberg, H. et al., 2008. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the *HERC2* gene inhibiting *OCA2* expression. *Human Genetics*, 123(2), pp.177–87.
- Finnegan, L.A. & Ó’Neill, L., 2010. Mitochondrial DNA diversity of the Irish otter, *Lutra lutra*, population. *Conservation Genetics*, 11(4), pp.1573–1577.
- Flojgaard, C. et al., 2011. Deconstructing the mammal species richness pattern in Europe—towards an understanding of the relative importance of climate, biogeographic history, habitat heterogeneity and humans. *Global Ecology and Biogeography: a Journal of Macroecology*, 20(2), pp.218–230.
- Fu, Q. et al., 2015. An early modern human from Romania with a recent Neanderthal ancestor. *Nature*, 524(7564), pp.216–219.
- Fu, Q. et al., 2016. The genetic history of Ice Age Europe. *Nature*, 534(7606), pp.200–205.
- Gamba, C. et al., 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications*, 5, p.5257.
- González, A.M. et al., 2006. The mitochondrial lineage U8a reveals a Paleolithic settlement in the Basque country. *BMC Genomics*, 7, p.124.
- González-Fortes, G. et al., 2017. Paleogenomic Evidence for Multi-generational Mixing between Neolithic Farmers and Mesolithic Hunter-Gatherers in the Lower Danube Basin. *Current Biology*, 27(12), pp.1801–1810.e10.
- Green, R.E. et al., 2010. A draft sequence of the Neandertal genome. *Science*, 328(5979), pp.710–722.
- Günther, T. et al., 2017. Genomics of Mesolithic Scandinavia reveal colonisation routes and high-latitude adaptation. *bioRxiv*, p.164400.
- Gurdasani, D. et al., 2015. The African Genome Variation Project shapes medical genetics in Africa. *Nature*, 517(7534), pp.327–332.
- Hansen, H.B. et al., 2017. Comparing Ancient DNA Preservation in Petrous Bone and Tooth Cementum. *PLoS One*, 12(1), p.e0170940.
- Henrich, J., 2004. Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case. *American Antiquity*, 69(2), pp.197–214.

- Higham, T. et al., 2011. The earliest evidence for anatomically modern humans in northwestern Europe. *Nature*, 479(7374), pp.521–524.
- Higham, T. et al., 2012. Testing models for the beginnings of the Aurignacian and the advent of figurative art and music: the radiocarbon chronology of Geißenklösterle. *Journal of Human Evolution*, 62(6), pp.664–676.
- Higham, T. et al., 2014. The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature*, 512(7514), pp.306–309.
- Hoffecker, J.F., 2009. The spread of modern humans in Europe. *Proceedings of the National Academy of Sciences*, 106(38), pp.16040–16045.
- Hofmanová, Z. et al., 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proceedings of the National Academy of Sciences of the United States of America*, 113(25), pp.6886–6891.
- Hublin, J.-J., 2015. The modern human colonisation of western Eurasia: when and where? *Quaternary Science Reviews*, 118, pp.194–210.
- Hughes, M., Montgomery, W.I. & Prodöhl, P., 2006. Population genetic structure and systematics of the Irish Hare. *Report prepared by Quercus for the Environment and Heritage Service (DOE, NI), UK*.
- Jelenkovic, A. et al., 2016. Genetic and environmental influences on height from infancy to early adulthood: An individual-based pooled analysis of 45 twin cohorts. *Scientific Reports*, 6, p.28496.
- Jochim, M.A., 1980. Postglacial Adaptations. In *Proceedings of the Conference on Northeastern Archaeology. University of Massachusetts Department of Anthropology Research Reports*. pp. 49–70.
- Jones, E.R. et al., 2015. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications*, 6, p.8912.
- Jones, E.R. et al., 2017. The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Current Biology*, 27(4), pp.576–582.
- Jöris, O. et al., 2010. The radiocarbon chronology of the Aurignacian to Mid-Upper Palaeolithic transition along the Upper and Middle Danube. In Ch. Neugebauer Maresch and L. Owen (eds.), *New aspects of the Central and Eastern European Upper Palaeolithic - methods, chronology, technology and subsistence*. Vienna: OAW, pp.101-137.
- Kılınc, G.M. et al., 2016. The Demographic Development of the First Farmers in Anatolia. *Current Biology*, 26(19), pp.2659–2666.
- Kirin, M. et al., 2010. Genomic runs of homozygosity record population history and consanguinity. *PLoS One*, 5(11), p.e13996.
- Lamason, R.L., 2005. SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science*, 310(5755), pp.1782–6.
- Lazaridis, I. et al., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 513(7518), pp.409–413.
- Lazaridis, I. et al., 2016. Genomic insights into the origin of farming in the ancient Near East. *Nature*, 536(7617), pp.419–424.
- Liu, W. et al., 2015. The earliest unequivocally modern humans in southern China. *Nature*, 526(7575), pp.696–699.
- Mallick, S. et al., 2016. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature*, 538(7624), pp.201–206.
- Mallory, J.P., 2013. *The origins of the Irish*. London: Thames & Hudson.
- Martiniano, R. et al., 2017. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genetics*, 13(7), p.e1006852.
- Martínková, N., McDonald, R.A. & Searle, J.B., 2007. Stoats (*Mustela erminea*) provide evidence of natural overland colonisation of Ireland. *Proceedings of the Royal Society B: Biological Sciences*, 274(1616), pp.1387–1393.
- Mathieson, I. et al., 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528(7583), pp.499–503.
- Mathieson, I. et al. 2017. The Genomic History Of Southeastern Europe. *bioRxiv*.
- McKenna, A. et al., 2010. The Genome Analysis Toolkit: a MapReduce framework for analysing next-generation DNA sequencing data. *Genome Research*, 20(9), pp.1297–1303.
- Meiklejohn, C. & Woodman, P.C., 2012. Radiocarbon dating of Mesolithic human remains in Ireland. *Mesolithic Miscellany*, 22(1), pp.22–41.
- Melo-Ferreira, J. et al., 2007. The rise and fall of the mountain hare (*Lepus timidus*) during Pleistocene glaciations: expansion and retreat with hybridisation in the Iberian Peninsula. *Molecular Ecology*, 16(3), pp.605–618.
- Meyer, M. et al., 2016. Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature*, 531(7595), pp.504–507.
- Montgomery, W.I. et al., 2014. Origin of British and Irish mammals: disparate post-glacial colonisation and species introductions. *Quaternary Science Reviews*, 98(Supplement C), pp.144–165.
- Norton, H.L. et al., 2007. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Molecular Biology and Evolution*, 24(3), pp.710–722.

- Olalde, I. et al., 2014. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature*, 507(7491), pp.225–228.
- Pagani, L. et al., 2016. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature*, 538(7624), pp.238–242.
- Patterson, N. et al., 2012. Ancient admixture in human history. *Genetics*, 192(3), pp.1065–1093.
- Pemberton, T.J. et al., 2012. Genomic patterns of homozygosity in worldwide human populations. *American Journal of Human Genetics*, 91(2), pp.275–292.
- Pettitt, P. & White, M., 2012. *The British Palaeolithic: Human Societies at the Edge of the Pleistocene World*. London: Routledge.
- Plisson, H., Dubreuil, L. & Guilbert, R., 2008. The functional significance of Sauveterrian microlithic assemblages: broadening the focus of investigation. *Prehistoric Technology*, 40, pp.147–157.
- Posth, C. et al., 2016. Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Current Biology*, 26(6), pp.827–833.
- Provan, J. & Bennett, K.D., 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, 23(10), pp.564–571.
- Prüfer, K. et al., 2014. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*, 505(7481), pp.43–49.
- Rasmussen, M. et al., 2011. An Aboriginal Australian Genome Reveals Separate Human Dispersals into Asia. *Science*, 334(6052), pp.94–98.
- Reich, D. et al., 2009. Reconstructing Indian population history. *Nature*, 461(7263), pp.489–494.
- Seguin-Orlando, A. et al., 2014. Paleogenomics. Genomic structure in Europeans dating back at least 36,200 years. *Science*, 346(6213), pp.1113–1118.
- Shakun, J.D. & Carlson, A.E., 2010. A global perspective on Last Glacial Maximum to Holocene climate change. *Quaternary Science Reviews*, 29(15), pp.1801–1816.
- Shea, J., 2008. Transitions or turnovers? Climatically-forced extinctions of Homo sapiens and Neanderthals in the east Mediterranean Levant. *Quaternary Science Reviews*, 27(23-24), pp.2253–2270.
- Soejima, M. & Koda, Y., 2007. Population differences of two coding SNPs in pigmentation-related genes SLC24A5 and SLC45A2. *International Journal of Legal Medicine*, 121(1), pp.36-9.
- Straus, L.G., 1991. Human Geography of the Late Upper Paleolithic in Western Europe: Present State of the Question. *Journal of Anthropological Research*, 47(2), pp.259–278.
- Stringer, C., 2016. The origin and evolution of Homo sapiens. *Philosophical transactions of the Royal Society B: Biological Sciences*, 371(1698).
- Timmermann, A. & Friedrich, T., 2016. Late Pleistocene climate drivers of early human migration. *Nature*, 538(7623), pp.92–95.
- Toussaint, M. et al., 2009. The Mesolithic site of Heffingen-Loschbour (Grand Duchy of Luxembourg). A yet undescribed human cremation possibly from the Rhine-Meuse-Schelde culture: anthropological, radiometric and archaeological implications. In P. Crombé et al. (eds.), *Chronology and Evolution within the Mesolithic of north-west Europe*. Cambridge: Cambridge Scholars Publishing, pp.239–260.
- Trinkaus, E. et al., 2003. An early modern human from the Peștera cu Oase, Romania. *Proceedings of the National Academy of Sciences of the United States of America*, 100(20), pp.11231–11236.
- Van Andel, T.H., Davies, W. & Weninger, B., 2003. The human presence in Europe during the last glacial period I: human migrations and the changing climate. In T.H Van Andel and W. Davies (eds.), *Neanderthals and modern humans in the European landscape during the last glaciation: archaeological results of the Stage, 3*. Cambridge: McDonald Institute for Archaeological Research, p.31e56.
- Vercellotti, G. et al., 2008. The Late Upper Paleolithic skeleton Villabruna 1 (Italy): a source of data on biology and behavior of a 14,000 year-old hunter. *Journal of Anthropological Sciences = Rivista di antropologia: JASS / Istituto italiano di antropologia*, 86, pp.143–163.
- Vermeersch, P.M., 2011. The human occupation of the Benelux during the Younger Dryas. *Quaternary International: the Journal of the International Union for Quaternary Research*, 242(2), pp.267–276.
- Wainwright, G.J., 1961. *The Mesolithic period in South and Western Britain*. UK: Doctoral Thesis, University of London.
- Wall, J.D. et al., 2013. Higher levels of neanderthal ancestry in East Asians than in Europeans. *Genetics*, 194(1), pp.199–209.
- Wojtal, P., Wilczynski, J. & Haynes, G., 2015. World of Gravettian Hunters. *Quaternary International: the Journal of the International Union for Quaternary Research*, 359(360), p.1e2.
- Wood, A.R. et al., 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, 46(11), pp.1173–1186.

4. The Genomics of Megaliths

Origins and structure of Irish Neolithic societies

Overview

In this chapter the genomes of 34 individuals, retrieved from a range of megalithic structures constructed across the Irish Neolithic (See Appendix I for further archaeological context), are analysed alongside continental farming populations (Table 3.1). Approximate affinities of the Irish Neolithic dataset could be deduced from ADMIXTURE and PCA analysis in Chapter Two, which indicated the island's population during the period was relatively homogenous, despite the dataset's large temporal and geographical range, with a similar ancestral profile to that observed in other contemporary Neolithic populations of western Europe. These groups all share a similar majority ancestry from Neolithic Anatolia, though with significant minor input from European Mesolithic groups. The archaeological and genomic evidence pertaining to the distribution and spread of such ancestries across Neolithic Europe is the subject of the introductory section in this chapter.

Utilisation of both allelic and haplotypic-sharing methods allows for further exploration here of Neolithic ancestral affinities, and suggests a substantial contribution from Mediterranean farming groups and northwestern hunter-gatherers to the Irish population. Moreover, an individual from the Parknabinnia court tomb is seen to share an excess of long haplotypes with an Irish HG, providing evidence for local Mesolithic survival and introgression in Ireland long after the commencement of the Neolithic. These methods also allow for the identification of potential regional outliers, emphasising the extensive connectivity of Neolithic Europe. Societal complexity is suggested in patterns of Y chromosome and autosomal structure in Ireland, which may indicate differential community usage of megalithic tombs. Most strikingly, the son of a likely sibling pairing is identified through ROH analysis, retrieved from the heart of the mega-passage grave at Newgrange, a clear anomaly amongst the observably large and outbreeding Irish population. Ethnographic comparisons and the individual's burial context strongly imply his belonging to a ruling elite, and suggest the elaboration and expansion of megalithic monuments over the course of the Neolithic was accompanied in some regions by dynastic social hierarchies.

Introduction

The Neolithisation of West Asia: A societal transformation across genetically diverse groups

The development of agriculture is arguably the most profound shift that has occurred in the course of human history. Through the domestication, cultivation and exploitation of animal and plant species, humans were able for the first time to create significant food surpluses, paving the way for large-scale sedentism, population growth and ultimately complex civilisation. Agriculture has emerged independently in several parts of the globe at different points in prehistory. The oldest of these, and the one to have a direct impact on the genetic makeup of Europe and Ireland, occurred in the Fertile Crescent, a region stretching from the southern Levant, through eastern Anatolia and into the Zagros Mountains region of West Asia. Here, in the wake of the Last Glacial Maximum (LGM), favourable environmental conditions encouraged sedentism among hunter-gatherer communities. Subsequent population growth necessitated more complex social organisation, while food management and storage became a communal requirement (Cunliffe 2008). By the Late Epipalaeolithic (14-11.6 kya) permanent settlements had sprung up across the region, between which exchange networks were in operation. Following the climatic downturn of the Younger Dryas (12.9-11.7 kya), the exploitation of wild cereals and animals intensified, and domestic forms of these species soon emerged. This period (9,600-6,900 BC), known as the Aceramic or pre-pottery Neolithic (PPN) for its lack of ceramics, marked the appearance of the first fully fledged agricultural societies.

Strikingly, genetic continuity has been witnessed across the transition from hunting and gathering to farming in both the Levant and the southern Caucasus–Iran highlands (Broushaki et al. 2016; Lazaridis et al. 2016). Importantly these Levantine and Iranian groups appear genetically distinct from one another, visualised in Chapter Two (Fig. 1.2; Fig. 1.3). In the Levant, sedentary hunter-gatherer (HG) groups belonging to the Natufian culture contribute to approximately two-thirds of the ancestry of the region's later PPN farmers, with a further third coming from a source related to Neolithic Anatolians. In Anatolia itself, the early adoption of agriculture seems to have been a relatively indigenous affair, with individuals belonging to early Aceramic Neolithic sites displaying levels of genetic diversity and ROH profiles similar to that of hunter-gatherer groups in Europe (Kılınç et al. 2016). This suggests little admixture accompanied the agricultural transition in these small populations and is in agreement with archaeological evidence proposing local forager adoption of cultivation and herding on the peninsula (Baird 2012). Neolithic Iranians appear equally divergent from both Levantine and Anatolian groups (Broushaki et al. 2016; Lazaridis et al. 2016), highlighting their relative isolation. Moreover, these individuals, and a preceding Iranian HG, show the highest levels of Basal Eurasian ancestry of all ancient groups so far sampled (~65%), almost twice that seen in related Caucasus HG individuals, who share unique ancestry with European HG groups (Broushaki et al. 2016; Lazaridis et al. 2016). The Levantine and Anatolian Neolithic also shows substantial Basal Eurasian input (~20-40%), despite their deep divergence from Iranians, indicating the progenitors of these West Asian farming groups derived much of their ancestry from the earliest known split in Eurasian population history.

Taken together, these recent studies paint a picture whereby the early spread of farming across west Asia was achieved for the most part through movement of ideas and technology rather than populations, with convergent cultural evolution occurring in deeply genetically diverged groups (Broushaki et al. 2016). However, intensified trade networks and demographic expansions would work over the succeeding millennia to erode pre-farming structure and increase genetic diversity across the region (Lazaridis et al. 2016). Economic collapses in the 7th millennium BC, brought on in part by overexploitation of the land, would have further encouraged population relocations (Cunliffe 2008). These twin pressures of exponential demographic growth and diminishing returns on food resources, likely created the impetus for migration that would ultimately push farming communities across the Aegean into Europe. By 6,600 BC farming groups had expanded into western Anatolia, and it was this region that served as the launchpad of agriculture into the Balkans, where it arrived via large-scale migration into the Thessalian Plain circa 6,500 BC (Gamba et al. 2014; Hofmanová et al. 2016). Genomic surveys have demonstrated this movement to be derived from populations related to the Pottery Neolithic of Northwestern Anatolia, whose ancestry went on to contribute majorly to the gene-pool of all modern Europeans (Mathieson et al. 2015; Hofmanová et al. 2016; Omrak et al. 2016). Notably, earlier Aceramic Neolithic sites (circa 7,000 BC) have also been identified in Southern Greece, though these early pioneers are thought to derive from a separate, possibly more southern, genetic source, linked to the PPN of Cyprus and the Levant (Mathieson et al. 2017).

The Neolithisation of Europe: Colonisation and admixture along two pioneering trails

After a brief period of consolidation, farming communities continued to expand northwards and eastwards into the Balkans from Greece (Fig. 3.1A). Early settlements clustered around river valleys which gave eventual access to the Danube, allowing farmers to reach the Hungarian plain by 5,550 BC. As they did so, regional groupings began to diversify, (e.g. Starčevo–Kőrös–Criş), though substantial similarities remained between the Early Neolithic cultures of the region. Clear selection and rejection of certain aspects of the Neolithic package in different ecological niches, as well as continued exploitation of wild resources, highlights the complexity by which agriculture spread (Cunliffe 2008). Importantly, several coastal and river regions of the Balkans maintained dense hunter-gatherer populations (Roksandic 2012; Gurova & Bonsall 2014), such as the Iron Gates gorge on the Danube, and interaction between incoming farmers and such groups is evident in the genomic data. While the majority of Early Neolithic individuals from the Balkans and Hungary show no difference to Western Anatolians in genetic composition, a farming community close to the Danube in Bulgaria possess a clear increase in European hunter-gatherer introgression (~15%), as does a Starčevo individual from Hungary (~8%) (Lipson et al. 2017; Mathieson et al. 2017). Moreover, the discovery of genetically Mesolithic individuals buried in Neolithic contexts and *vice versa* provide direct snapshots of such early crossovers (Gamba et al. 2014; Mathieson et al. 2017). Evidence suggest HG ancestry continued to increase in both regions over time), reaching approximate levels of almost 15% by the Chalcolithic (Lipson et al. 2017).

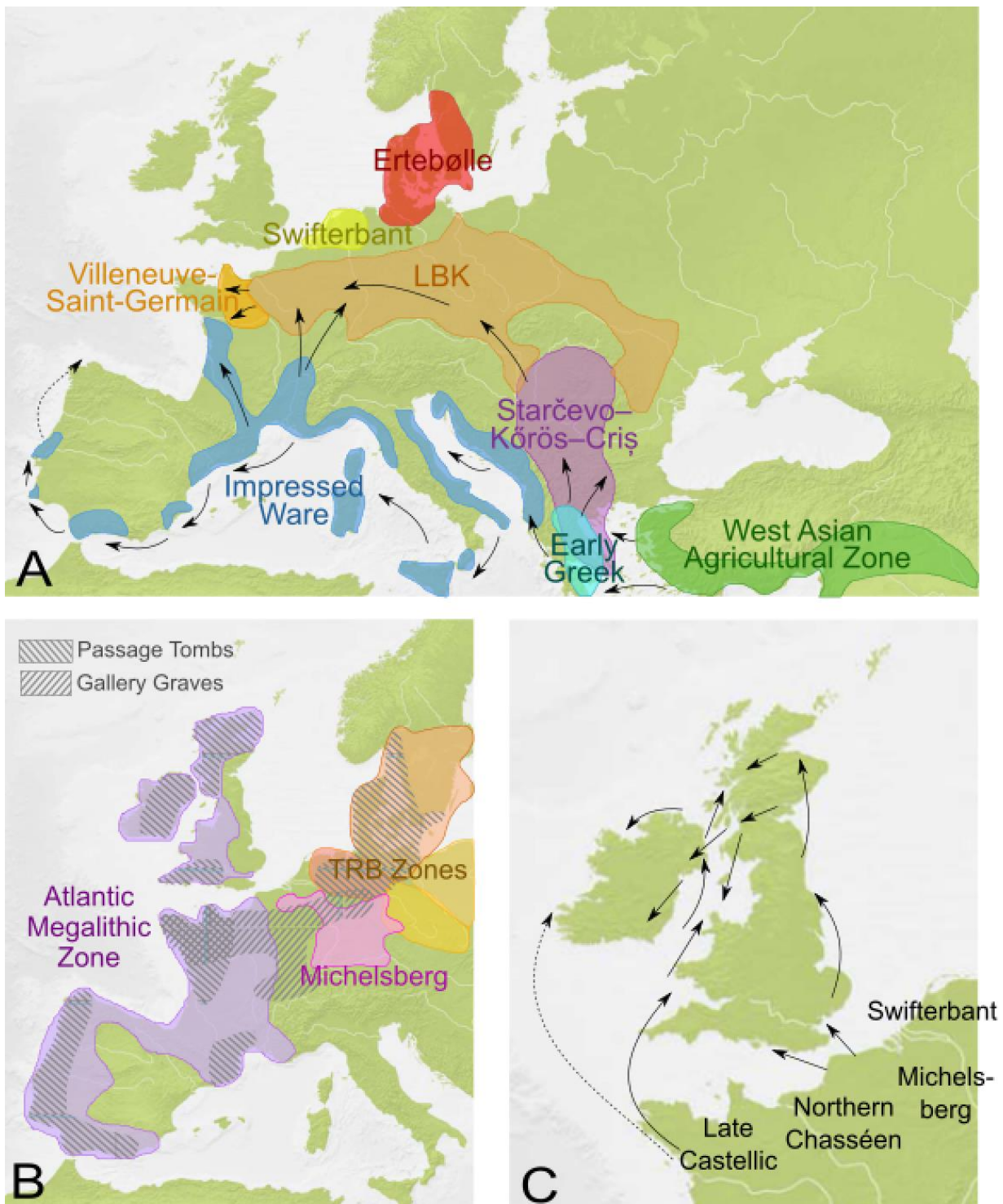


Figure 3.1. A map of archaeological cultures discussed in the chapter. A) Migratory routes followed by Neolithic communities from Anatolia to northwest Europe. **B)** Cultural spheres of influence in operation during the Irish Neolithic. **C)** Possible routes of entry taken by Irish Neolithic settlers.

Cultural distributions sourced from Cunliffe 2008, Müller 2014 and Cummings 2017. (Map Credit: freeworldmaps.net)

It was from this southeastern hotbed that the Neolithic revolution spread quickly out across the European Peninsula, following two key migratory routes, one travelling overland into Central Europe, the other following the Mediterranean seaboard westwards (Fig. 3.1A). Importantly, pioneers along both these thoroughfares remained closely related to the Early Neolithic populations of the Balkans, and all

descend primarily from the same Anatolian source (Olalde et al. 2015; Hofmanová et al. 2016; Lipson et al. 2017; Mathieson et al. 2017), though, debatably, some minor unique input from groups linked to the Aceramic Neolithic of Southern Greece may have contributed to Mediterranean populations (Hofmanová et al. 2016; Mathieson et al. 2017).

The overland route continued the initial expansion northward from Hungary, following the Danubian corridor into the heart of the continent and reaching the Paris Basin by 5,100 BC. This entire trajectory was dominated by a strikingly uniform cultural complex that emerged in Central Europe *circa* 5,500 BC, the Linearbandkeramik (LBK), who produced distinctive pottery and housing forms. LBK pioneers found themselves expanding into heavily forested territory with seemingly little indigenous presence (Cunliffe 2008). No uptake in hunting practices is seen and little of the preceding Mesolithic culture can be recognised in LBK assemblages. The farmers confined themselves to a distinct ecological zone, defined by fertile loess soils, which pushed expansion along a westward corridor, hemmed by the Alps to the south and the less fertile Northern European Plain to the north (Cunliffe 2008). Genetic data from German LBK individuals complements their uniform archaeology, revealing a largely homogenous group with very limited European HG input (~4-5%) (Haak et al. 2015; Lipson et al. 2017; Mathieson et al. 2017). Moreover, virtually no HG ancestry is seen in these groups' likely progenitor, the Transdanubian LBK (LBKT) of Western Hungary (Szécsényi-Nagy et al. 2015). This is in sharp contrast to derivatives of LBK culture (ALP or Bükk) in the eastern regions of Hungary (the Alföld or Great Hungarian Plain), who show substantially inflated hunter-gatherer ancestry in comparison (~8-10%), supporting archaeological evidence of increased Mesolithic settlement in these regions (Bánffy 2006; Lipson et al. 2017). Interestingly, these populations, like their progenitors in the Balkans, were highly lactose intolerant (Gamba et al. 2014; Mathieson et al. 2015), despite ample evidence for cattle husbandry and dairying (Gillis et al. 2017).

The second route followed by early European farmers leapfrogged across the Mediterranean coastline, reaching the Atlantic seaboard of Portugal as early as 5,600 BC, at a time when the LBK settlements of Central Europe were still in their earliest phases. The pattern suggests a mode of rapid maritime enclave colonisation (Zilhão 2001), with disparate sites linked by characteristic 'Cardial' or 'Impressed Ware' pottery. These groups had spread from the west coast of Greece northwards to Croatia and across to the Italian Peninsula from whence Neolithic colonies could be planted along the French and Spanish coasts through long distance sea travel. Notably, one of the islands settled along the way, Sardinia, retains the largest proportion of ancestry from these early farmers seen in Europe today (Keller et al. 2012; Gamba et al. 2014; Lazaridis et al. 2014; Skoglund et al. 2014), suggesting it experienced relative isolation after the Neolithic period. Mesolithic sites along the northern Mediterranean coastline are sparse, though some regions, such as Sicily, produce evidence of agricultural uptake by foraging communities (Cunliffe 2008). In other areas new Neolithic settlements show complete discontinuity from previous Mesolithic cultures (e.g La Draga), and suggest pioneering groups chose largely unoccupied regions in which to establish themselves.

While the exact interactions with indigenous groups must have varied along the seaboard, it is apparent that by the time Early Cardial groups arrived in Iberia they had procured some minority Mesolithic ancestry (~10%), greater than that observed in LBK individuals (Olalde et al. 2015; Lipson et al. 2017; Mathieson et al. 2017). Later Neolithic and Chalcolithic Iberians, from both the southwest and northeast of the peninsula, display substantially higher proportions of Mesolithic ancestry (~25%) (Günther et al. 2015; Haak et al. 2015; Mathieson et al. 2015; Lipson et al. 2017; Martiniano et al. 2017), suggesting further admixture took place with the Mesolithic populations of Iberia. Contrarily, a recent study has suggested, based on an admixture graph framework, that this large increase in Mesolithic ancestry has a non-local origin, due to some shared ancestry of Middle Neolithic Iberians from the northeast with western hunter-gatherer individuals other than the Spanish LaBranã1, such as the Luxembourgian Loschbour (Lipson et al. 2017). However, this does not take into account the likely heterogeneity of Mesolithic Iberian populations, as evidenced in González-Fortes et al. 2017, and discussed in Chapter Three.

Hunter-gatherer groups are believed to have been somewhat unevenly distributed in Iberia, with the dry interior largely abandoned at the onset of the Holocene, in favour of more bountiful coastal zones (Zilhão 2000). In particular, the Atlantic littoral provided rich foraging resources, which could maintain large sedentary hunter-gatherer communities. In Portugal these persisted for many centuries after the arrival of Cardial Ware culture (Cunliffe 2008) and genomic surveys suggest extensive Mesolithic admixture from western sources in Portuguese Middle Neolithic individuals, including LaBranã1 (Martiniano et al. 2017). Thus, the dominant theme across the Neolithic in Iberia appears to be one of hunter-gatherer introgression, seen at a magnitude never reached further east in the Balkans and Hungary (Haak et al. 2015; Lipson et al. 2017), although the exact focal points of such admixture events remain debatable.

A similar pattern of Mesolithic persistence and resurgence is also seen in the more northern Atlantic regions of Europe, which provided the entry point of agriculture into Britain and Ireland. Indeed, the ease with which early Neolithic migrants traversed the Mediterranean and Central European forests stands in sharp contrast with the delay seen in the delivery of agriculture to the Northern extremes of the continent. Many of these regions possessed large and well-established hunter-gatherer societies, which persisted long after the establishment of farming communities in neighbouring areas. Moreover, the more northern latitudes presented substantial ecological barriers to cultivation. Some level of climatic or environmental change may have been required to lure pioneer farmers into these more hostile climes, with further encouragement possibly provided by demographic and economic stresses from within existing communities (Cunliffe 2008; Cunliffe 2013). The genetic and cultural interplay between incoming farming groups from both the Danubian and Mediterranean migratory streams and the indigenous populations of Scandinavia, Germany, the Low Countries and France, gives crucial context to the Irish and British Neolithic, and the Neolithisation of each of these regions is discussed in turn below.

The Neolithisation of the North: The enduring legacy of the European Mesolithic

In the Northern European Plain and southern Scandinavia, Mesolithic groups belonging to Ertebølle and other Post-Maglemosian cultures (Fig. 3.1A), had continued their sedentary foraging lifestyles for many centuries, despite clear interaction with the LBK zone further south, from whom they adopted pottery and toolkits (Cunliffe 2008). After almost a millennia of persistent rejection of agricultural uptake, a drastic shift then occurred across Northern Europe and Scandinavia *circa* 4,000 BC, with the spread of the Neolithic Funnel Beaker or Trichterrandbecher Culture (TRB), named for its characteristic pottery forms (Fig. 3.1B). While the ultimate geographical source(s) of TRB is debated, the earliest evidence suggests the ceramic style originated near the western regions of the Baltic Sea, from where it radiated quickly outwards, reaching the Netherlands, Bohemia, Sweden and southeastern Poland (Czekaj-Zastawny et al. 2013). This was followed slightly later by the emergence of agricultural practices across the region, with farmers utilising sparsely inhabited inland regions through forest clearance.

There is considerable heterogeneity seen across the TRB cultures of northern Europe, with a number of regional groupings distinguished on the basis of pottery styles and mortuary practices (Midgley 2008; Czekaj-Zastawny et al. 2013). Burials in flat graves and monumental non-megalithic long barrows are widespread and form part of the earliest horizon, appearing *circa* 4,400 BC on the North European Plain, and 3,900 BC in Scandinavia. Similar earthen long mound structures appear in southern Britain during the same time period, though limited evidence of accompanying TRB material culture has been found (Müller 2014). Non-megalithic long barrows (Passy type) are also a prominent feature of the post-LBK or Cerny tradition of the Paris Basin. Popular archaeological opinion suggests these burial structures find their origins in the timber longhouses of LBK culture and its later derivatives, acting as symbolic ‘houses of the dead’, although parallels with elongated Mesolithic shell middens of coastal regions have also been drawn (Cunliffe 2008; Müller 2014). Megalithic burials (structures based on large stones), including dolmens and passage graves, are also an important feature of some TRB societies. However, these structures do not appear in TRB regions until nearly 500 years after the commencement of agricultural practices, and are then restricted mainly to the west, including the Netherlands, Scandinavia and some parts of Germany (Midgley 2008).

Genomes retrieved from a TRB megalithic passage grave in Sweden (*circa* 3,000 BC) indicate that by this time large scale migration from Anatolian-related populations had occurred into Scandinavia, although substantial HG admixture is present in these individuals (Skoglund et al. 2014), similar to that seen in Middle Neolithic Iberians (Haak et al. 2015). Contemporaneous individuals from various German TRB cultures, such as the Baalberge, Salzmünde and Bernberg groups, have also been sequenced and demonstrate broadly similar affinities to the Swedish farmers (Haak et al. 2015), with inflated Mesolithic ancestry relative to earlier LBK groups of the same region (~17%) (Lipson et al. 2017). Combined archaeological and genetic evidence also highlights the survival of populations of predominantly Mesolithic ancestry across the Scandinavian Neolithic. Hunter-gatherer individuals belonging to the Pitted-Ware Culture, which persisted to the very end of the era, show a reverse trend in ancestries relative

to their agriculturalist TRB neighbours, possessing minor Anatolian admixture within a larger European Mesolithic background (Skoglund et al. 2014), suggesting cultural barriers to gene flow were maintained over many centuries.

Further west, at a cave site near the Dutch-German border (Blätterhöhle), just outside the western fringe of TRB culture, isotopic and genetic analysis has identified another temporally Neolithic individual (4,000-3,500 BC) of majority Mesolithic ancestry (~75%) living on a predominantly fish-based diet (Bollongino et al. 2013; Lipson et al. 2017). Isotopically 'farming' individuals have been recovered from the same time depth at the site, and possess substantially less Mesolithic ancestry, though still relatively inflated compared to contemporaneous TRB individuals (~40%). This is remarkable given the regions encircling the Blätterhöhle site had been occupied from an early date (~5,300 BC) by LBK culture and its later derivatives, including Rössen and Michelsberg groups. While no palaeogenomic data is available for these cultures, mitochondrial analyses support continuity of the Rössen and Michelsberg with the LBK of Central Europe, though some HG input is apparent (Rivollat et al. 2016; Beau et al. 2017). Notably the Blätterhöhle site postdates the Neolithic expansion into Britain and Ireland (~4,000 BC), a time when Michelsberg agricultural populations, known for their large ditched enclosures, were widespread across northeastern France and the Low Countries (Fig. 3.1B). While such late survival of genetically distinct HG groups so far inland at Blätterhöhle is surprising, the persistence of coastal foragers and fishers across the Rhine delta may provide a reservoir for such ancestry. Indeed, the uptake of farming practices by Mesolithic groups of the Northern Netherlands (Swifterbant) was an extremely gradual process of apparent acculturation (Cummings 2017), lasting well over a millenia (5,000-3,400 BC). It is worth noting how sharply this contrasts with the spread of farming further east in Scandinavia and Northern Germany, which occurred in little more than a century (~3,900 BC), after outright rejection by Mesolithic Ertebølle groups for almost a millenia (Cunliffe 2008). Neolithic expansion into Britain and Ireland also occurred at this time and at an equally rapid pace.

The Neolithisation of the Northwest: Megaliths and Maritime Migrations

The northern regions of France are perhaps the most crucial in understanding the mechanisms by which agriculture was delivered to the offshore islands of Britain and Ireland. It was here that the two main migratory arms of the Neolithic, which had diverged almost a millennium earlier in the Balkans, were reunited and together gradually subsumed the established foraging communities. Unfortunately, no palaeogenomic data from the Neolithic has yet been obtained for this dynamic region and archaeological debate continues over the relative influence of the two migratory strands, as well as the role played by local Mesolithic groups, whose cultures persisted for many centuries, particularly in coastal regions such as Brittany (Thorpe 2003; Cunliffe 2013).

Impressed Ware groups, settled in the Mediterranean Golfe du Lion, had made early use of overland river routes to reach both the Atlantic coast of France and delve inland toward the Rhine, where they met with LBK culture, expanding from the east (Fig. 3.1A). Indeed, the influence of Mediterranean

Cardial Ware is seen in the Limburg and La Hoguette pottery styles, found inland along the western fringes of LBK's early distribution (Thorpe 2003; Hofmann 2016). These ceramic types, thought to be the result of indigenous interaction with more southern Impressed Ware groups, may have predated the arrival of LBK in some regions, and indeed are clearly replaced by LBK ware in many areas at a later date, including the southern Netherlands (~5,300 BC) (Cunliffe 2008). On the Atlantic coast, Neolithic communities using Cardial and Epicardial pottery were known to be settled as far north as the Loire estuary between 5,500-4,500 BC, from where farming practices and people could have easily spread to Brittany via maritime networks. However, these sites lack clear shared markers of identity, possibly the result of indigenous influence or adoption, which creates difficulties in tracing the spread of such cultures northward (Thomas 2013).

The expansion of traditional LBK culture into Northern France is more easily followed, through identifiers such as the Danubian longhouse and characteristic ceramics (Thomas 2013). It reached its western limit in the Paris Basin *circa* 5,300 BC, where it stalled for several centuries. Rössen culture subsequently evolved across western LBK territories, before more regional complexes, such as TRB and Michelsberg, discussed above, emerged. Another LBK derivative, Villeneuve-Saint-Germain culture, expanded south and westwards between 5,000-4,700 BC, gradually encroaching on the Mesolithic societies of Normandy and Brittany (Cummings 2017). Very rapid culture change is subsequently apparent across the maritime regions of Brittany, with the sudden adoption of agricultural practices, the erection of standing stone menhirs and the construction of long mound burials, a tradition also seen in the Paris Basin and further east on the Northern European Plain.

During the same period, burials in stone cists and early passage graves begin appearing on the Brittany peninsula, set within small circular mounds, or later within the long mounds themselves (Cunliffe 2008), representing some of the earliest forms of collective megalithic burial. Notably, the polygonal forms of these structures are similar to Epicardial and Cerny burial traditions further south (Müller 2014). Initially passage graves appear restricted to the Northwest tip of Brittany (~4,700-4,500 BC) but later disperse across the coastal regions of the peninsula (~4,000-3,800 BC), replacing or incorporating the previous long mounds and menhirs, indicating multiple spheres of influence were acting in the region. Broadly contemporaneous construction of passage tombs and other collective megalithic burial structures, such as simple dolmens (single chamber flanked by multiple monoliths), is also seen in Portugal (~4,500 BC) (Cunliffe 2008). Passage tombs subsequently disperse across Atlantic Europe, showing a distribution typically restricted to coastal regions (Fig. 3.1B), suggestive of extensive maritime contacts stretching from Iberia to Orkney throughout the course of the Neolithic. However, some Atlantic regions, including northeastern France, the Lowlands and Eastern England, reject passage graves completely, though other megalithic structures do appear. Most notably are gallery graves, which show a more inland distribution relative to passage tombs, both on a continental level (Fig. 3.1B), and in more local regions such as Brittany (Cunliffe 2008). Gallery graves lack a distinctive entrance passage and tend to be found under rectangular long barrows, rather than the rounded mounds of passage graves.

The Neolithisation of Britain and Ireland: The Final Frontier

After roughly a millennium of delay in Northern France, Neolithic communities finally crossed the English Channel, breaching the final frontier of agricultural expansion in Europe. Despite its relatively late arrival, the spread of the Neolithic across Britain and Ireland was a very rapid affair, so rapid in fact that there has been some difficulty constructing a secure chronological framework for its expansion. Nevertheless, a combination of Bayesian statistics and refined calibration methods have given some order to the sequence of events (Whittle et al. 2011). It is now believed that farming communities arrived in the southeast edges of England *circa* 4,050 BC, and, after several centuries of consolidation in the south of Britain, spread into most other regions approximately 3,800 BC. Notably, this suggests little to no delay occurred between the Neolithization of Britain and Ireland, as might have been expected (Mallory 2013). Moreover, several curiously early dates exist in Ireland, which could push Neolithic contacts on the island back several centuries (~4,000 BC), including the remains of a cow at Ferriter's Cove in Co. Kerry (Woodman & McCarthy 2003), and an interrupted ditched enclosure in Co. Sligo, a structure primarily seen in Southern Britain. It is then important to consider that Ireland may not have derived the entirety of its Neolithic culture directly from Britain, though the substantial similarities between the two islands strongly suggest a common origin. Communication between the islands was also clearly sustained throughout the Neolithic period, evidenced by continued material exchange networks (e.g. porcellanite axes).

With the exception of the earliest sites in the southeast, Neolithic culture appears to have been delivered across Britain and Ireland as a complete package, including domesticated plants and animals, ceramics, toolkits and in many regions megaliths (Cummings 2017), with little to no evidence of preceding Mesolithic culture after its arrival. Chapter Two and the publication on which it is based (Cassidy et al. 2016), as well as a more recent genomic survey of Neolithic Britain (Olalde et al. 2017), demonstrates that this package arrived through large scale migration, a feat which must have required extensive maritime knowledge. The Atlantic coasts of France and the Low Countries are the most obvious staging areas for such a seaborne colonisation, with two broad zones of influence considered as potential sources, although the two are by no means mutually exclusive (Fig. 3.1C). The first is the extensive cultural interface linking northeastern France and the Lowlands with southern and eastern Britain (Cunliffe 2008). The second centers on more western Atlantic seaways, which show some evidence of dispersed movement from Brittany along the coastlines of the Irish Sea towards Northern Scotland, echoing the maritime mode of enclave colonisation seen in the Mediterranean.

Northeastern France (Calais-Picardy) provides the most popular candidate homeland for Ireland and Britain's first farmers (Cunliffe 2013; Mallory 2013). Strong parallels are seen between the Neolithic of southeastern Britain and the continental Michelsberg and Chasséen cultures across the channel (Fig. 3.1C), which share similar round-bottom Carinated Bowl pottery, domesticated faunal assemblages and flint leaf-shaped arrowheads (Whittle 2007; Cunliffe 2013). Carinated Bowls are most abundant along the eastern seaboard of Britain, though they are ubiquitous across the Neolithic of both islands, as are

the characteristic arrowheads. Another component of the Neolithic package in Britain and Ireland that has clear connections with northern continental traditions are the large rectangular timber houses that first appear a few generations after pottery and domesticates in southeast England (Sheridan 2013; Cummings 2017). In Ireland, this so-called ‘housing horizon’ begins no earlier than 3,720–3,680 cal BC (Whitehouse et al. 2014; McLaughlin et al. 2016), accompanied by the appearance of cereals. Together, such material culture has been grouped under the the blanket term ‘Carinated Bowl Neolithic’ (Sheridan 2010), and is proposed to have spread along the east coast of Britain, before penetrating into Scotland and crossing over to Northern Ireland, where it travelled southwards down the east and west coasts.

Funerary traditions featuring non-megalithic monuments, such as earthen long mounds and mortuary enclosures, largely restricted to the eastern side of Britain, have also been linked to this Carinated Bowl Neolithic (Sheridan 2010). In Ireland, the initial construction of megalithic court tombs broadly corresponds to the same horizon (3,700-3,570 cal BC) (Schulting et al. 2012; McLaughlin et al. 2016). These are a type of chambered cairn with a characteristic forecourt at the entrance. The majority of these are found in the northern third of the island, though atypical groupings further south exist. Strong resemblance is seen between these and Clyde-Carlingford cairns in the southwest of Scotland, suggesting cultural links between the regions (Cunliffe 2013). The Severn-Cotswold tombs of Wales and southwestern England are also of a similar type. The origins of such structures are unknown, though they fall into the same general family as continental gallery graves, and may represent a merging of megalithic and long mound traditions, which could have occurred more than once in different regions. A population from the southwestern Parknabinnia court tomb, located in the Burren, Co. Clare, is sampled here, as well as an individual from a more typical northern site in Cohaw, Co. Cavan.

A second strand of Neolithization has been suggested, based on the confinement of Breton-style megalithic structures to Ireland and western Britain. This includes simple passage graves and closed polygonal cists, the latter suggested as being related to portal dolmens (Sheridan 2010; Cunliffe 2013). The theory proposes an earlier spread of Neolithic communities to the islands from the Morbihan region of Brittany, primarily through maritime networks, and indeed both tomb types show somewhat coastal distributions. Within the islands, simple passage tombs and closed chambers are seen in Wales, western Scotland and the north coasts of Ireland. Sheridan argues that these mark the beginning of a long and complex tradition of passage tomb construction in Scotland and Ireland, culminating in the massive cemeteries seen in Orkney, the Boyne valley and Co. Sligo. In contrast, passage tombs appear to die out in Wales, reintroduced towards the end of the fourth millennium from eastern Ireland (Sheridan 2004). Portal dolmens show slightly more southern affinities, absent from Scotland and present in Cornwall (Cunliffe 2013). Their chronology is not well understood, though an early appearance pre-dating the ‘house horizon’ is likely (Kytmanow 2009; Lynch 2014; McLaughlin et al. 2016), supported by early burials from the portal dolmen at Poulnabrone, Co. Clare (Lynch 2014), several of which are sampled here.

In general, radiocarbon evidence indicating an earlier colonisation of Ireland and Western Britain is rare (for obvious reasons megaliths themselves cannot be dated), though support is drawn also from early dates at the enigmatic Altanagh megalith in Co. Tyrone (Murphy et al. 2010), the Carrowmore passage cemetery in Co. Sligo (4,200-3,800 BC), and possible pre-4,000 BC Breton-style pottery from a simple passage tomb in Scotland, though these have been contentious. More generally, the aforementioned Ferriter's cove dates hint at some oversea contact between Ireland and the continent prior to 4,000 BC (possibly a failed colonisation attempt). Another line of evidence supporting long distance sea travel between Atlantic regions is the presence of distinctive rodent species in Orkney and Ireland, unknown in mainland Britain, whose closest relatives are found in Western France and Spain (Cunliffe 2013). If such early maritime networks and enclave farming communities did indeed exist, they may have helped hasten the transmission of the 'Carinated Bowl Neolithic' when it arrived, possibly contributing to the near-simultaneous Neolithisation of many regions of the islands (Mallory 2013). Moreover, continued use or later activity at megalithic sites may have obscured any evidence of pre-Carinated Bowl activity (Sheridan 2004).

Regardless of their exact mode of arrival, megaliths remained an enduring feature of the Irish Neolithic landscape across the era. Most impressive is perhaps the monumental passage grave cemetery of the Boyne Valley, which bears a close resemblance to similar sites across the Irish Sea in Anglesey and Orkney, all erected towards the end of the fourth millennium BC (Cunliffe 2013). The tombs here reach massive proportions, with satellite structures built around the larger mounds, drawing comparison to the pyramids of Egypt (Hensey 2015). Close cultural and spiritual links between the disparate cemeteries are captured in their shared architecture and art, with a focus on the solar calendar and solstice alignments, hinting at underlying astronomical knowledge that one would associate with sea travel (Mallory 2013). Passage graves such as these also show clear links to continental groups, particularly in Iberia and Brittany. In Ireland, the construction and expansion of these mega-passage grave cemeteries was largely restricted to the north and east, with other regions, such as the Burren in the southwest possibly remaining more insulated from these Late Neolithic cultural developments (Jones forthcoming). The end of the Irish Neolithic is also marked by the spread of Grooved Ware ceramics, believed to have origins in Orkney, alongside henge structures (Mallory 2013), both associated with the passage tomb tradition. Two individuals from the famous Newgrange site in the Boyne Valley are analysed here, one from the centre of the main passage grave, the other from a satellite tomb. Another population from the equally imposing Carrowkeel passage cemetery in Co. Sligo is also sampled, which provides some of the latest Neolithic dates of the presented dataset (~2,800-2,500 BC).

Methods

Sample Processing

A full description of molecular methodology and bioinformatic data processing is provided in Appendix II and summarised in the Methods section of Chapter Two.

Variant Calling and Dataset Preparation

A set of 2,733,477 autosomal transversions from the 1000 Genomes Phase 3 release (European MAF > 1%) were used for f - and D -statistic analyses, as described in Chapter Three. Randomised pseudo-haploid genotypes for these sites were called in ancient samples using the Pileup tool from GATK (McKenna et al. 2010) (See Appendix II for further detail). The Mbuti population from the SGDP dataset (Mallick et al. 2016) again provided an outgroup population. All dataset manipulation and merging was carried out in PLINK v1.90 (Chang et al. 2015).

For lcMLkin maximum likelihood estimation of identity-by-descent (IBD), genotype likelihoods for the above set of variants were called in 90 Neolithic and hunter-gatherer samples of sufficient coverage (See Table 3.1). This was carried out using the SNPbam2vcf.py tool, recommended for lcMLkin - (https://github.com/COMBINE-lab/maximum-likelihood-relatedness-estimation/tree/master/src_python/SNPbam2vcf).

These sites were further filtered to include only positions also present in the 1240k SNP capture described in Mathieson et al. 2015, to maximize number of usable SNPs shared across all individuals. An additional minor allele frequency (MAF) filter of 5% was then imposed on the Neolithic dataset and sites were thinned to be at least 10 kb apart.

Imputed genotypes of ancient samples were used for ChromoPainter, runs of homozygosity (ROH) and phenotypic analysis. Methods for imputation using Beagle software (Browning & Browning 2007) are detailed in Appendix II. For ROH and ChromoPainter analyses, only transversion sites with a MAF above 5% across the 1000 Genomes dataset were used. Furthermore, both analyses only considered sites for which all samples included possessed a genotype probability (GP) above 99% (i.e no missingness was allowed). For ROH this dataset consisted of 56 Neolithic individuals (See Table 3.1), three diverse reference genomes from the SGDP (Mallick et al. 2016) also included in Chapter Three, and diploid calls for the Stuttgart genome (Lazaridis et al. 2014), downloaded with the SGDP, to act as a measure of imputation accuracy.

For ChromoPainter, 178 imputed ancient individuals with sufficiently low levels of missingness were selected to preserve the number of usable variant sites (353,005 SNPs in total). As ancient data from two separate rounds of imputation were to be included, the dataset was first rephased using SHAPEIT V2

(Delaneau et al. 2011) with the 1000 Genomes Phase 3 reference panel (1000 Genomes Project Consortium 2015). Samples were phased by chromosome using default settings and the GRCh37 build genetic map to estimate linkage disequilibrium.

***D*- and *f*-Statistics**

D- and *f*-statistics (Reich et al. 2009; Green et al. 2010) were calculated using the AdmixTools package (Patterson et al. 2012) with significance assessed using a block jackknife of 5cM in size. The Mbuti population of Central Africa were used as an outgroup, given their undetectable levels of Eurasian admixture (Gurdasani et al. 2015). For information on ancient individuals included in these tests and the population keys used see Tables 2.1 and 3.1.

Outgroup *f*₃-statistics of the form $f_3(\text{Mbuti}; X, Y)$ were used to measure shared genetic drift between pairs of Irish Neolithic individuals. A heatmap was created for the resulting values to explore genetic clustering among samples (heatmap.2 package from R) (Fig. 3.4). Parknabinnia675 was excluded from this analysis due to recent relatedness with another individual, Parknabinnia357.

Test scores for different *D*-statistic configurations were regressed against one another for Neolithic individuals (Fig. 3.7). Outliers were identified based on studentised residuals (degrees freedom=69, CI=95%). All samples considered singularly were required to have at least 170,000 variants called from the set of 2,733,477 autosomal transversions used.

lcMLkin estimation of IBD

IBD sharing was estimated using lcMLkin (Lipatov et al. 2015), summing over all genotypes (-g all), given their raw likelihoods (-l raw). Pi-hat scores were extracted for pairs of Irish Neolithic individuals (Parknabinnia675 excluded) and visualised in a heatmap (heatmap.2 package from R).

ChromoPainter analysis

ChromoPainter was run under default settings to paint each ancient individual using all other individuals (-a 0 0), as detailed in Chapter Two. Raw chunk donations and average chunk lengths for Mesolithic and Neolithic individuals are displayed in Fig. 3.5 and Fig. 3.6.

Runs of Homozygosity analysis

A total of 596,569 transversion sites called across 56 ancient and three modern samples were used for ROH (see above for selection criteria). Distributions of runs of homozygosity (ROH) were calculated using PLINK, with the same parameters and visualisation techniques described in Chapter Three (Fig.

3.8). Given the large number of long homozygous tracts seen in Newgrange10, further analysis was carried out on this individual alone. There were 2,047,204 transversions (1000G MAF > 5%) imputed in this genome with a GP above 99%. For these sites heterozygote and homozygote positions were plotted against a physical map of the autosomal genome, to visualise larger tracts of homozygosity (Fig. 3.9A).

In addition, the inbreeding pedigree that may have led to this distribution was explored, following methods similar to Prüfer et al. 2014, whereby breeding simulations were carried out on six unrelated Neolithic genomes from the Parknabinnia site. These individuals were chosen as they represent an Irish Neolithic community of similar ancestral background to the Newgrange10 individual. The Parknabinnia and Newgrange10 samples had 1,649,640 transversion sites (1000G MAF > 5%) called securely between them (GP > 99%), which were used for these simulations.

Given the extent of the ROH tracts observed (Fig. 3.8A), two possible scenarios of inbreeding were considered for Newgrange10, offspring of siblings (inbreeding coefficient : 25%) or offspring of half-siblings (inbreeding coefficient : 12.5%). Other pedigrees leading to a 12.5% inbreeding coefficient exist (double first cousins; aunt and nephew; grandfather and granddaughter etc.), but these were not simulated given the difficulty in distinguishing such scenarios from one another (Prüfer et al. 2014). Simulations were carried out on the Parknabinnia dataset, whereby haploid genomes were created based on the phased imputation output for each sample, from which offspring between pairs of samples could be produced. For the six samples, 20 offspring were produced from each possible pairing, with ten F2 offspring generated by pairing these 20. This resulted in a total of 150 offspring from unique sibling pairings for analysis. Furthermore, 8 offspring from half-sibling pairings were created from each possible triple of the Parknabinnia dataset, resulting in a total of 160 unique offspring of half-siblings. Crossover rates of 41.1 in females and 26.4 in males were assumed based on estimates from Chowdhury et al. 2009 and sample sex was considered mutable depending on the pairing. No mutation rate or sequencing error was considered and so estimated tracts from these simulations represent an upper limit of homozygosity.

ROH distributions for both Newgrange10, and all simulated data were then estimated in PLINK using the same parameters described in Chapter Three. Six different measures of ROH were considered as indicators of inbreeding, based on both overall distributions of homozygosity, as well as runs over 7.5 Mb in length alone. The ranges for these in simulated half-sibling and sibling offspring are plotted in Fig. 3.9B, alongside the value obtained for Newgrange10. ROH distributions were also calculated for Newgrange10 based on a more relaxed heterozygote allowance of 2 sites within the sliding window, and values for this output are also presented.

Table 3.1 List of Neolithic and Chalcolithic samples analysed in the current chapter. Sample identifiers and archaeological contexts are given. For samples from which both whole genome sequence data (WGS) and 1240k SNP capture data from the panel released in Mathieson et al. 2015 (Cap.) is available, the data type used in this study is noted in the sample ID. Furthermore, the inclusion or exclusion of each sample in the different analyses presented here is noted in the final three columns. Individuals considered only as part of a population for D-statistic analyses are marked as such (Pop). Parknabinnia samples used for inbreeding simulations are noted with an asterisk (*) in the last column. Y haplogroup placements that may be confounded by post-mortem damage are bracketed. ‘Number of Variants’ refers to the total number of secure calls for a set of 2,733,477 autosomal transversions from the 1000 Genomes Phase 3 release. The European country of origin for samples is denoted by its first three letters. Anatolian Neolithic individuals used for D-statistics are not shown here, but were sourced from Mathieson et al. 2015; Broushaki et al. 2016, Hofmanova et al. 2016 and Kılınç et al. 2016.

Reference Code: ¹Haak et al. 2015; ²Mathieson et al. 2015; ³Lazaridis et al. 2014; ⁴Hofmanova et al. 2016; ⁵Gamba et al. 2014; ⁶Current Study; ⁷Keller et al. 2012; ⁸Allentoft et al. 2015 ; ⁹Martiniano et al. 2017; ¹⁰Olalde et al. 2015; ¹¹Günther et al. 2015; ¹²Skoglund et al. 2014.

Reference	Current Study ID	Other IDs	Country	Latitude	Longitude	Period	Archaeological Context	Date (cal BC) *not direct	Mitochondrial Haplogroup	Chromosome Y Haplogroup	Data Type	Number of Variants	D-Statistics	D-Statistic Population	ROH & ChromoPainter
1,2	Baalberge_MN2	I0560	Ger	51.79	11.14	MN	TRB	3640-3376	T2c1	-	Cap.	173377	Yes	TRB	No
1,2	Esperstedt_MN	I0172	Ger	51.42	11.68	MN	TRB	3360-3086	T2b	I2a1b1a	Cap.	1399295	Yes	TRB	Yes
1,2	LBK_EN1	I0025	Ger	48.78	9.18	EN		5500-4850*	T2b	-	Cap.	361918	Yes	LBK_EN	No
1,2	LBK_EN10	I0659	Ger	51.89	11.04	EN		5211-4963	N1a1a1a2	G2a2a1	Cap.	360435	Yes	LBK_EN	No
1,2	LBK_EN11	I0795	Ger	51.27	11.66	EN		5217-5041	H1	T1a	Cap.	51186	Pop	LBK_EN	No
1,2	LBK_EN12	I0821	Ger	51.90	11.05	EN		5201-4850	X2d1	G2a2a1	Cap.	82649	Pop	LBK_EN	No
2	LBK_EN13	I0797	Ger	51.28	11.65	EN		5500-4850*	H46b	T1a	Cap.	39951	Pop	LBK_EN	No
2	LBK_EN14	I1550	Ger	51.89	11.04	EN		5500-4850*	K1a2	-	Cap.	152839	Pop	LBK_EN	No
1,2	LBK_EN2	I0026	Ger	48.78	9.18	EN		5500-4850*	T2b	-	Cap.	344339	Yes	LBK_EN	No
1,2	LBK_EN3	I0046	Ger	51.89	11.04	EN		5211-4991	T2c1	-	Cap.	527581	Yes	LBK_EN	No
1,2	LBK_EN4	I0054	Ger	51.66	11.53	EN		5223-5021	J1c17	-	Cap.	949586	Yes	LBK_EN	Yes
1,2	LBK_EN5	I0100	Ger	51.89	11.04	EN		5202-4852	N1a1a1a	-	Cap.	617958	Yes	LBK_EN	No
1,2	LBK_EN6	I0022	Ger	48.78	9.18	EN		5500-4850*	T2c	-	Cap.	96693	Pop	LBK_EN	No
1,2	LBK_EN7	I0048	Ger	51.89	11.04	EN		5210-5002	K1a	G2a2a	Cap.	154407	Pop	LBK_EN	No
1,2	LBK_EN8	I0056	Ger	51.89	11.04	EN		5212-5006	T2b	G2a2a	Cap.	93842	Pop	LBK_EN	No

1,2	LBK_EN9	I0057	Ger	51.89	11.04	EN		5219-5021	N1a1a1	-	Cap.	37940	Pop	LBK_EN	No
3	Stuttgart_LBK		Ger	48.78	9.18	EN		5500-4850*	T2c1d1	-	WGS	2623405	Yes	LBK_EN	Yes
4	Klei10		Gre	40.43	21.86	LN		4230-3995	K1a2	G2a2a1b	WGS	2158312	Yes		Yes
4	Pal7		Gre	40.51	22.50	LN		4452-4350	J1c1	-	WGS	1756951	Yes		Yes
4	Rev5		Gre	40.33	22.57	EN		6438-6264	X2b	-	WGS	1476345	Yes		Yes
2,5	CO1_Badden_WGS		Hun	47.17	19.83	CA		2903-2806	H	-	WGS /Cap.	1422986	Yes		Yes
2,5	NE1_ALP_WGS		Hun	47.88	21.19	MN		5306-5229	U5b2c	-	WGS /Cap.	2589902	Yes		Yes
2,5	NE2_ALP_WGS		Hun	48.52	21.17	MN		5291-5144	H	-	WGS	359528	Yes		No
2,5	NE3_Bukk_SnpCap		Hun	47.88	21.19	MN		5281-5132	X2b-T226C	-	Cap.	286115	Yes		No
2,5	NE4_Tis-Bukk_SnpCap		Hun	47.17	20.83	MN		5211-5116	J1c5	-	Cap.	248341	Yes		No
2,5	NE5_Late_ALP_WGS		Hun	47.17	20.83	MN		5295-5116	J1c1	C1a2	WGS /Cap.	1360375	Yes		No
2,5	NE6_LBK_WGS		Hun	47.17	19.83	MN		5206-5002	K1a3a3	C1a2	WGS /Cap.	1543616	Yes		Yes
2,5	NE7_Lengyel_WGS		Hun	47.17	19.83	LN		4491-4416	N1a1a1a	I2a1	WGS /Cap.	1503771	Yes		Yes
6	Annagh1		Ire	52.68	-8.46	MN	Link.	3638-3137	K1a+195	I2a2a1	WGS	1789701	Yes	Irish_NE	Yes
6	Ardcroyne2		Ire	52.93	-8.15	MN	Link.	3625-3366	J2b1a	I2a1b1a1	WGS	2168879	Yes	Irish_NE	Yes
6	Ashleypark1		Ire	52.93	-8.19	MN	Link.	3635-3522	K2a9	I2a2a1a1a2	WGS	1475206	Yes	Irish_NE	Yes
6	Ashleypark3		Ire	52.93	-8.19	NE	Link.	3635-2500*	T2c1d1	I2a1a(2)	WGS	1892923	Yes	Irish_NE	Yes
6	Ballynahatty		Ire	54.54	-5.95	MN	Passage-like	3343-3020	HV0+195	-	WGS	2586002	Yes	Irish_NE	Yes
6	Baunogenasraid72		Ire	52.81	-6.82	MN	Link.	3631-3384	K1a4a1	H2a	WGS	1811726	Yes	Irish_NE	Yes
6	Carrowkeel530		Ire	54.06	-8.38	LN	Passage	2883-2635	W5b	I2a2(a1a1a)	WGS	1577911	Yes	Irish_NE	Yes
6	Carrowkeel531		Ire	54.06	-8.38	NE	Passage	3600-2500*	H1	I2a2(a1a1a)	WGS	507576	Yes	Irish_NE	No
6	Carrowkeel532		Ire	54.06	-8.38	LN	Passage	3015-2891	J1c3	I2a2a1a1a	WGS	2162806	Yes	Irish_NE	Yes
6	Carrowkeel533		Ire	54.06	-8.38	LN	Passage	3086-2903	H	-	WGS	1984873	Yes	Irish_NE	Yes
6	Carrowkeel68		Ire	54.06	-8.38	LN	Passage	2834-2469	H	I2a2a1	WGS	1974579	Yes	Irish_NE	Yes
6	Cohaw448		Ire	54.05	-7.03	EN	Court	3652-3514	H1	I2a1b1a1	WGS	1829661	Yes	Irish_NE	Yes
6	Jerpoint14		Ire	52.52	-7.17	MN	Link.	3694-3369	T2c1d1	H2a	WGS	2020974	Yes	Irish_NE	Yes
6	MillinBay6		Ire	54.36	-5.49	MN	Passage-like	3376-3090	J1c3	I2a2a1a1a	WGS	1706873	Yes	Irish_NE	Yes

A Genomic Compendium of an Island

6	Newgrange10	Ire	53.70	-6.48	LN	Passage	3339-3028	U5b1+16189+16192	I2a2a1(a1a)	WGS	1729559	Yes	Irish_NE	Yes	
6	NewgrangeZ1	Ire	53.70	-6.48	LN	Passage	3332-2921	X2b+226	-	WGS	1065442	Yes	Irish_NE	Yes	
6	Parknabinnia1239	Ire	52.99	-9.11	NE	Court	3700-2800*	K1a1a	I2a1(a)	WGS	2036600	Yes	Irish_NE	Yes*	
6	Parknabinnia186	Ire	52.99	-9.11	MN	Court	3518-3355	X2b+226	I2a1(a2)	WGS	2130178	Yes	Irish_NE	Yes*	
6	Parknabinnia2031	Ire	52.99	-9.11	NE	Court	3700-2800*	K1a2b	-	WGS	1593364	Yes	Irish_NE	Yes*	
6	Parknabinnia357	Ire	52.99	-9.11	MN	Court	3175-2835*	U8b1b	-	WGS	2177030	Yes	Irish_NE	No	
6	Parknabinnia443	Ire	52.99	-9.11	NE	Court	3700-2600*	K1b1a1	I2a1b1a1	WGS	2276485	Yes	Irish_NE	Yes*	
6	Parknabinnia581	Ire	52.99	-9.11	NE	Court	3635-2930*	T2b	I2a2a1a1a2	WGS	2101407	Yes	Irish_NE	Yes*	
6	Parknabinnia672	Ire	52.99	-9.11	EN	Court	3639-3514	T2c1d+152	I2a1(a)	WGS	2270509	Yes	Irish_NE	Yes*	
6	Parknabinnia675	Ire	52.99	-9.11	MN	Court	3100-2910	H1	I2a1(a)	WGS	2123099	Yes	Irish_NE	Yes	
6	Parknabinnia768	Ire	52.99	-9.11	NE	Court	3700-2600*	H4a1a1	I2(a1a)	WGS	1624097	Yes	Irish_NE	Yes	
6	Poulnbrone02	Ire	53.05	-9.14	NE	Portal	3900-3200*	U5b1c1	I2a2a1	WGS	1750820	Yes	Irish_NE	Yes	
6	Poulnbrone03	Ire	53.05	-9.14	NE	Portal	3900-3200*	K1a1	I2a2a1	WGS	1820474	Yes	Irish_NE	Yes	
6	Poulnbrone04	Ire	53.05	-9.14	NE	Portal	3900-3200*	H1+16189	I2a2a1	WGS	1793419	Yes	Irish_NE	Yes	
6	Poulnbrone05	Ire	53.05	-9.14	NE	Portal	3900-3200*	K1a+195	I2a2a1a1a2	WGS	1620350	Yes	Irish_NE	Yes	
6	Poulnbrone06	Ire	53.05	-9.14	NE	Portal	3900-3200*	H1	I2a(xI2a2)	WGS	1763091	Yes	Irish_NE	Yes	
6	Poulnbrone07	Ire	53.05	-9.14	NE	Portal	3629-3571	U5b1c	I2a2a(1a1a)	WGS	1806768	Yes	Irish_NE	Yes	
6	Poulnbrone107	Ire	53.05	-9.14	EN	Portal	3928-3666	U4a2f	I2a2a1(a1a)	WGS	1980161	Yes	Irish_NE	Yes	
6	Poulnbrone112	Ire	53.05	-9.14	EN	Portal	3696-3535	U5b2b	I2a2a1a1	WGS	1474113	Yes	Irish_NE	Yes	
6	Poulnbrone113	Ire	53.05	-9.14	EN	Portal	3942-3702	H4a1a1a	-	WGS	1693502	Yes	Irish_NE	Yes	
7	Oetzi	Ita	46.78	10.83	CA		3359-3105	K1	G2a2a1b1a	WGS	2166146	Yes		No	
8	Remedello2	RISE 487	Ita	45.26	10.38	CA	3483-3107	H2a	I	WGS	350349	Yes		No	
8	Remedello3	RISE 489	Ita	45.26	10.38	CA	2908-2578	X2c1	I	WGS	860886	Yes		No	
9	CabecoArruda117B	Por	39.05	-9.33	LN	Megalith	3500-2500*	J1c1b	I2a1b	WGS	769641	Yes		No	
9	CabecoArruda122A	Por	39.05	-9.33	LN	Megalith	3500-2500*	H1e1a	G2a2a1	WGS	2057369	Yes		Yes	
10	Cardial_EN4	F19	Por	39.30	-8.36	EN	Cave	5325-5209	H4a1a1a	-	WGS	32604	Pop	Spain_EN	No
10	Cardial_EN5	G21	Por	39.30	-8.36	EN	Cave	5345-5223	H3	..	WGS	9152	Pop	Spain_EN	No

9	CovaMoura364		Por	39.08	-9.24	LN	Cave	3700-2100*	H	I2a1b	WGS	1362690	Yes	Yes
9	CovaMoura9B		Por	39.08	-9.24	LN	Cave	3700-2100*	J1c1	-	WGS	2354329	Yes	Yes
9	DolmenAnsiao96B		Por	40.00	-8.50	LN	Dolmen	3640-3100*	K1b1a	I2a1a1	WGS	2078256	Yes	Yes
9	LugarCanto41		Por	39.44	-8.85	MN	Cave	4230-3700*	U5b1		WGS	1567822	Yes	Yes
9	LugarCanto42		Por	39.44	-8.85	MN	Cave	4230-3700*	H3	-	WGS	2378298	Yes	Yes
9	LugarCanto44		Por	39.44	-8.85	MN	Cave	4230-3700*	H1c1a2	-	WGS	2137552	Yes	Yes
11	A1P12-1420		Spa	42.21	-3.31	CA	Cave	2960-2829	H3c3	I2a2a	WGS	1840514	Yes	Spain_CA
11	A1P16		Spa	42.21	-3.31	CA	Cave	3211-2866	X2c	-	WGS	1765396	Yes	Spain_CA
11	A1P2		Spa	42.21	-3.31	CA	Cave	2849-2628	U5b3	H2	WGS	2445704	Yes	Spain_CA
2	CA_MIR22	I1300	Spa	42.33	-3.50	CA	Cave	2900-2346*	K1a2a	-	Cap.	260696	Yes	Spain_CA
2	CA_MIR5	I1303	Spa	42.33	-3.50	CA	Cave	2900-2346*	X2b	I2a2a1	Cap.	786832	Yes	Spain_CA
10	Cardial_EN1	CB13	Spa	41.22	1.53	EN	Cave	5469-5327	K1a2a	-	WGS	1499849	Yes	Yes
10	Cardial_EN2	CB14	Spa	41.22	1.53	EN	Cave	5500-5000*	X2c	-	WGS	664	Pop	Spain_EN
10	Cardial_EN3	CS7675	Spa	39.30	-8.36	EN	Cave	5363-5302	K1a4a1	..	WGS	2125	Pop	Spain_EN
10	Cardial_EN6	H3C6	Spa	38.50	-0.21	EN	Cave	5400-5280	H4a1a1a	..	WGS	1944	Pop	Spain_EN
1,2	Spain_EN1	I0410	Spa	42.50	0.50	EN	Epi-cardial	5294-5066	T2c1d	R1b1	Cap.	242130	Yes	Spain_EN
1,2	Spain_EN2	I0412	Spa	42.50	0.50	EN	Epi-cardial	5310-5078	N1a1a1	I2a1b1	Cap.	932042	Yes	Spain_EN
1,2	Spain_EN3	I0413	Spa	42.50	0.50	EN	Epi-cardial	5303-5075	V	-	Cap.	224056	Yes	Spain_EN
1,2	Spain_EN4	I0409	Spa	42.50	0.50	EN	Epi-cardial	5311-5218	J1c3	-	Cap.	218351	Yes	Spain_EN
1,2	Spain_MN1	I0406	Spa	41.25	-2.33	MN	Passage	3900-3600*	H1	I2a2a1b2	Cap.	291636	Yes	No
1,2	Spain_MN2	I0407	Spa	41.25	-2.33	MN	Passage	3900-3600*	K1b1a1	-	Cap.	189408	Yes	No
1,2	Spain_MN3	I0408	Spa	41.25	-2.33	MN	Passage	3893-3661	U5b1	-	Cap.	686357	Yes	Yes
12	TRB_Gok2		Swe	58.18	13.40	MN	TRB	5050-4750*	H1c	-	WGS	1371428	Yes	TRB

Results

Sequencing Statistics

Detailed sequencing statistics and data authenticity results can be found in Appendix II for all Irish Neolithic samples. Whole genome data was retrieved for a total of 35 individuals from an assortment of diverse megalithic sites located across the island (See Appendix I for further archaeological context). Samples ranged in date from 3,942-3,702 to 2,834-2,469 cal BC, effectively covering the entire Neolithic period of the island (~4,000-2,500 BC). While all 35 individuals were included in PCA (Chapter Two), one individual, Carrowkeel534, has been excluded from the autosomal analysis presented in this chapter, due to low genomic coverage. The remainder of samples included here fall between 0.20X and 2.18X (See Fig. 1.1; Electronic Data Table S2), with a single sample, Ballynahatty, sequenced to higher 10.27X coverage.

Uniparental Analysis of Irish Neolithic Samples

No autosomal data from British Neolithic samples are yet available, however Y chromosome and mitochondrial haplogroups have been reported (Olalde et al. 2017), and provide at present the only avenue by which the Irish dataset can be compared with contemporary individuals from the neighbouring island. Y haplogroup frequencies taken from the above study and other published and unpublished datasets (Keller et al. 2012; Gamba et al. 2014; Allentoft et al. 2015; Günther et al. 2015; Mathieson et al. 2015; Hofmanová et al. 2016; Kılınç et al. 2016; Lazaridis et al. 2016; Lipson et al. 2017; Martiniano et al. 2017; Mathieson et al. 2017; Olalde et al. 2017) are plotted alongside those observed in Neolithic Ireland in Fig. 3.2. To explore mtDNA frequencies, the data utilised in Beau et al. 2017 and from Olalde et al. 2017 provided a reference for Irish samples (Fig. 3.3).

Y chromosome

The most common Y chromosome lineages of the European Neolithic are G2, also seen at high frequencies in Neolithic Anatolia, and I2, the dominant lineage of the European Mesolithic. Frequencies of non-I haplogroups are displayed in Fig. 3.2A, which demonstrates the influx of new lineages into Europe with the onset of agriculture. G2a2 is ubiquitous in the Balkans and Hungary across the Neolithic and Chalcolithic periods. It is also seen at lesser frequencies along the Mediterranean seaboard towards Portugal, though it is perhaps noteworthy that none of the four Early Cardial Neolithic individuals so far sampled (Croatia and Spain) belong to G2. Indeed, two possess haplogroups rare or absent in Balkans and Hungarian Neolithic (R and E1b1).

I2 haplogroups increase in frequency on an east-west gradient and also through time in Central Europe and Spain, indicative of Mesolithic admixture. Remarkably, not a single I2 lineage is observed in a large population of LBK individuals, in agreement with the low levels of autosomal European HG ancestry

seen in these samples (See Fig. 1.3 for visualisation). After G2 the most common haplogroup seen in the LBK group is T1a, the only other observation of which is in contemporaneous Bulgarian individuals. The following German Middle Neolithic shows a very different Y chromosomal profile, with marked increase in I2 lineages (specifically I2a1), but also R haplogroups, associated with HG populations from further east. This is perhaps indicative of the widespread distribution of TRB culture across the Northern European Plain towards the Baltic Sea. Notably, R1b haplogroups are observed as far west as the Blätterhöhle site near the Dutch Border, close to the fringes of Michelsberg Culture. Discontinuity is again seen in Northern Europe with the Globular Amphora Culture (3,400-2,800 BC), which exhibits total fixation of haplogroup I (predominantly I2a2a1b), despite the similar levels of autosomal WHG ancestry seen in these individuals relative to TRB groups (Mathieson et al. 2017).

In Britain and Ireland virtually all individuals from across the Neolithic period belong to I2 lineages, with the exception of two samples from southeastern Ireland, whom possess the rare haplogroup H2a. Intriguingly, both these samples are from proximate Linkardstown cists, distinctive singular burials, typically male, which possess a somewhat restricted temporal and geographical distribution (Brindley & Lanting 1989). I2a haplogroups also dominate in the Iberian Middle to Late Neolithic and Chalcolithic, though not to the same extent as in Britain and Ireland, with Anatolian-related G2a2 and H lineages also present. Notably, a group of four individuals from the Middle Neolithic of southern France also display only I2a haplotypes.

Within Britain and Ireland three main lineages are present, I2a1a (CTS595), I2a1b (M423) and I2a2a (M223), and some structure is apparent between sites. Strikingly, this seems to correspond not only with geography, but with burial type. I2a1b dominates Northern Scottish groups in Orkney and Caithness (stalled cairns and Maeshowe-type passage tombs; 3,700-2,500 BC), while the majority of the southwestern Argyll group (cave sites; 3,800-3,200 BC) belongs to I2a2a. This is also the sole lineage observed in the Poul nabrone Portal Tomb (n=8) and the Late Neolithic Passage/Passage-like tombs from the northwest and northeast coasts of Ireland (n=5). Notably, the only two I2a2 lineages seen in Northern Scotland (n=10) date to this same Late Neolithic period (3,000-2,800 BC). Moreover, the majority of samples were of sufficient coverage to further place them into haplogroup I2a2a1a1 (M284), a lineage whose distribution is generally restricted to modern British and Irish populations (International Society of Genetic Genealogy 2017; FamilyTreeDNA 2017a), though incidences in Portugal, France, Germany and Norway have been noted. This may suggest some male line continuity from the Neolithic period in Britain and Ireland.

The Parknabinnia court tomb (n=7), despite being located within 10km of Poul nabrone and used across the same chronological period, shows a radically different profile of Y chromosome frequencies, with highest frequencies for I2a1a (CTS595). A simple Z-test for two proportions suggests this is not likely to be the result of sampling error (p=0.0007). The few continental Neolithic samples belonging to I2a1a place within the I2a1a1 (M26) subclade, common in modern Europe (peaking in Sardinians at 40%, and

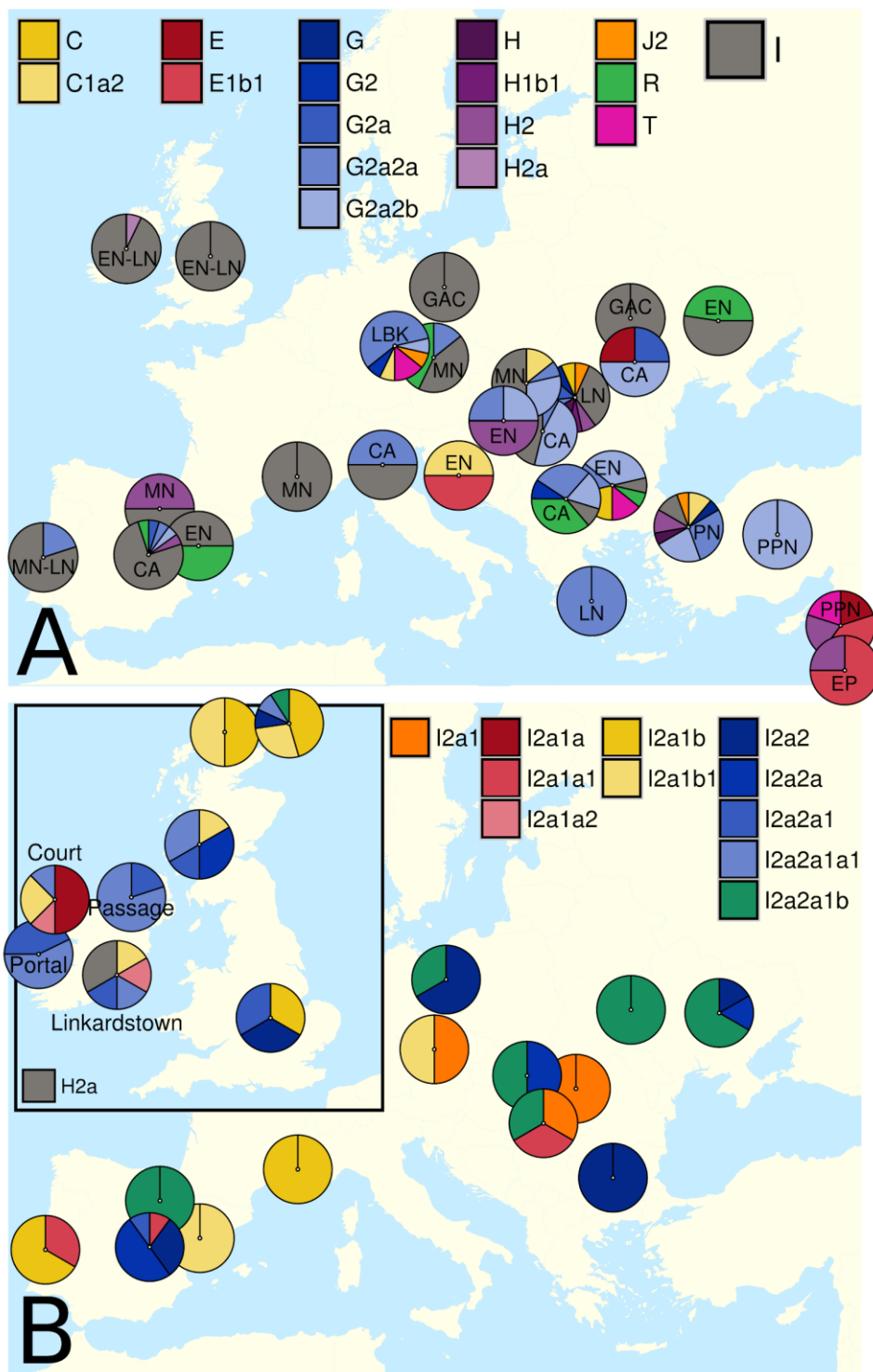


Figure 3.2. Y chromosome haplogroup frequencies for farming populations across Europe and West Asia. Haplogroups are defined with the ISOGG Y-DNA Haplogroup Tree (June 2017 Version) **A**) Population frequencies for all haplogroups, with I lineages greyed out. Importantly, low coverage data may prevent samples being placed in specific subclades, leading to high frequencies of basal lineages. Population Key: EP - Epipalaeolithic, Pre-Pottery Neolithic - PPN, Pottery Neolithic - PN Early Neolithic - EN, Middle Neolithic - MN, Late Neolithic - LN, Chalcolithic - CA, Globular Amphora Culture - GAC). **B**) Frequencies of I2a haplogroups within relevant populations. Irish and British samples have been split into regional subgroupings. Samples without sufficient coverage to place downstream of I2a were excluded. (Note: Passage also refers to the undefined “Passage-like” megaliths of Ballynahatty and Millin Bay)

Basques at 5%), while all Irish samples are ancestral for this marker. Furthermore, the two Irish samples with sufficient coverage could be further placed into the rare I2a1a2 clade (S21825). Downstream lineages of S21825 are scarce but have been noted in northern and northwestern Europe today (Britain, France, Germany) (FamilyTreeDNA 2017b). Interestingly, I2a1a is not observed in the British sample set.

Taken together, these patterns suggest that denser surveys of Y chromosomal variation, both existing and novel, across the Neolithic of Ireland and Britain may have the power to detect small-scale local structure, not present on an autosomal or mitochondrial level, due to the more fluid movement of women. Indeed, patrilocality, higher mobility of women, and an increase in monogamy have all been suggested to accompany the transition to agriculture and sedentism in Europe (Rasteiro & Chikhi 2013), phenomena which may also have biased reproductive success to male HG lineages in regions densely populated by Mesolithic communities. The relative lack of maternal structure seen in European Neolithic populations (Fig. 3.3), discussed below, and the more subtle resurgence of HG mtDNA haplogroups, relative to that seen for the Y chromosome, supports this view.

Mitochondrial DNA

Neolithic populations across Europe show similar frequencies of mtDNA haplogroups, with three major clades dominating, H, K and JT, all associated with genetic influx from West Asia during the spread of farming (Richards et al. 2000; Bramanti et al. 2009; Haak et al. 2015). However, some differentiation is apparent, such as the reduction in JT lineages in the Spanish Early Neolithic, relative to that of Central Europe, as well as an elevation of N(xN1a) haplogroups. Haplogroup X is also present in Spain and France, while virtually absent in LBK populations. The earliest appearance of the lineage in the north of Europe is seen in Rössen and related cultures, possibly the result of admixture with populations expanding into the region from the south. Interestingly, haplogroup X is also noted in the Irish and British Neolithic, as well as several TRB cultures further east. Interestingly, TRB groups also possess the highest incidences of haplogroup W, which is also present within the Irish and British cohorts, as well as the Michelsberg group.

In terms of HG introgression, mtDNA haplogroups show a similar trend to the Y chromosome, with U haplogroups seen at higher frequencies both in the west and in later Neolithic populations (Fig. 3.3). However, the scale of this introgression is much reduced compared to the increase seen in I2a Y chromosome lineages, again supporting a male HG bias during these admixture events. Notably, there is a large increase in U lineages between the Rössen culture and the succeeding Michelsberg culture in the northwest (Beau et al. 2017). Moreover, the Rössen culture itself sees some inflation of HG lineages in comparison to its predecessor, the LBK culture. This may be due to admixture with HG groups in the northwest, or with other farming populations who possess higher levels of HG ancestry. Notably, Irish and British groups display U haplogroup frequencies similar to that seen in the Michelsberg culture

and the French MN. In contrast, the different TRB cultures display substantial variance in U haplogroup frequencies, suggesting diverse interactions with HG populations were occurring in the region.

Finally, very little difference is seen in haplogroup frequencies between the Irish and British populations and the two groups also show similar profiles to northwestern continental groups, including one of their suggested progenitors, the Michelsberg culture. The discontinuity seen between LBK and these later northwestern populations is suggestive of the suspected complex admixture events that occurred in the region between diverse Neolithic cultures, as well as indigenous groups.

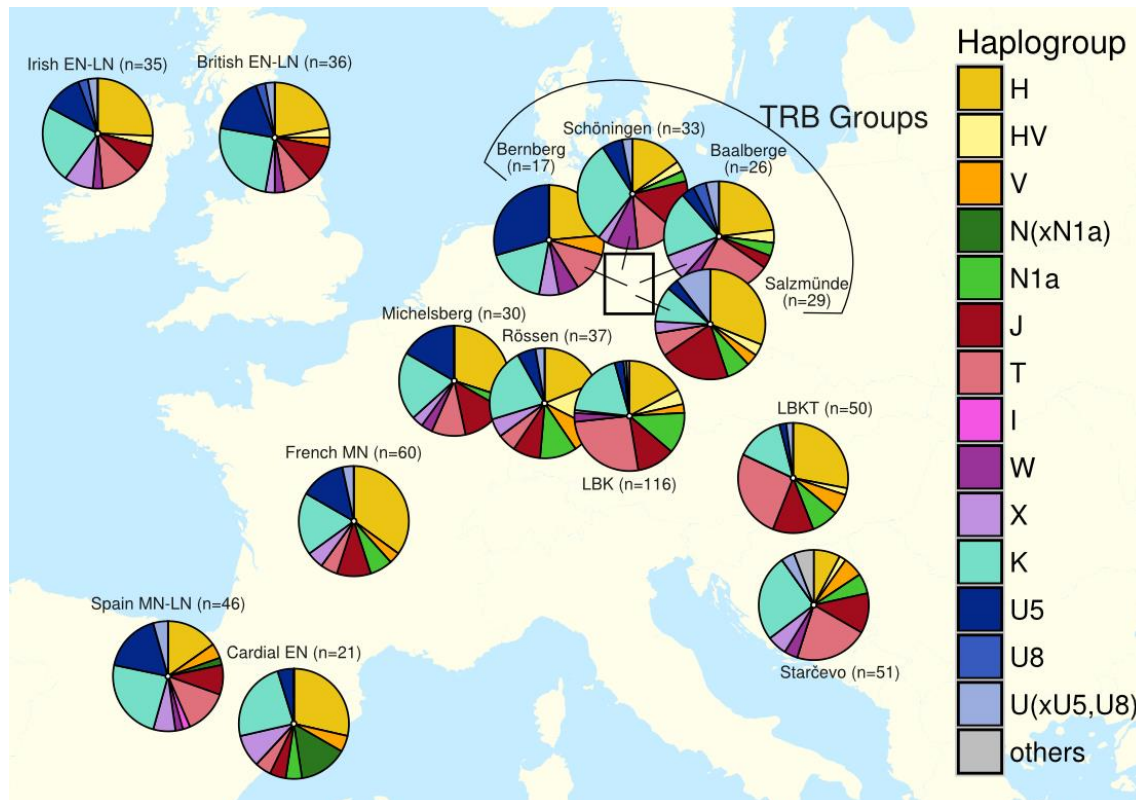


Figure 3.3. mtDNA haplogroup frequencies for farming populations in Europe. Population frequencies are represented in pie-charts, coloured by haplogroup. Rössen culture also refers to Grossgartach and Planig-Friedberg complexes, which intervene between LBK and Rössen.

Identifying Autosomal Structure among Irish Neolithic Individuals

Distributions of several Y chromosome haplogroups hint at possible genetic structure between Irish Neolithic sites. To investigate such a possibility further the autosomal affinities of Irish Neolithic samples were explored using two distinct measures of relatedness. Firstly, the *lcMLkin* program (Lipatov et al. 2015), specifically designed for low coverage data, was used to estimate IBD sharing between pairs of Irish individuals based on genotype likelihoods and population allele frequencies. This analysis revealed two individuals, Parknabinnia675 and Parknabinnia357 to be third degree relatives (Appendix II; Fig. SII.9) and Parknabinnia675 was subsequently excluded from downstream clustering analysis. Secondly, outgroup f_3 -statistics of the form $f_3(\text{Mbuti}; X, Y)$ were used to measure shared genetic drift between pairs of individuals. Two heatmaps and dendrograms were created based on the resulting matrices to explore genetic clustering among samples (Fig. 3.4).

Sample placement varied considerably between the two graphs, emphasising the homogeneity of the dataset. However, both heatmaps identified a cluster composed primarily of individuals from the Carrowkeel passage tomb cemetery, dating roughly between 3,000-2,500 BC. Moreover, both the MillinBay6 (3,376-3,090 cal BC) and Newgrange10 (~3,300-2,900 BC) individuals placed consistently within this cluster, with Newgrange10 showing definite inflated affinity to Carrowkeel532 (3,015-2,891 cal BC). The clear temporal overlap across the cluster is remarkable, as well as the shared archaeological context of Carrowkeel and Newgrange samples, retrieved from two of the largest passage grave cemeteries in Ireland. Notably, NewgrangeZ1, an individual from a satellite structure off the main Newgrange mound, showed no marked affinity to this grouping. In both dendrograms, this cluster was placed within wider branchings composed chiefly of Poul nabrone samples.

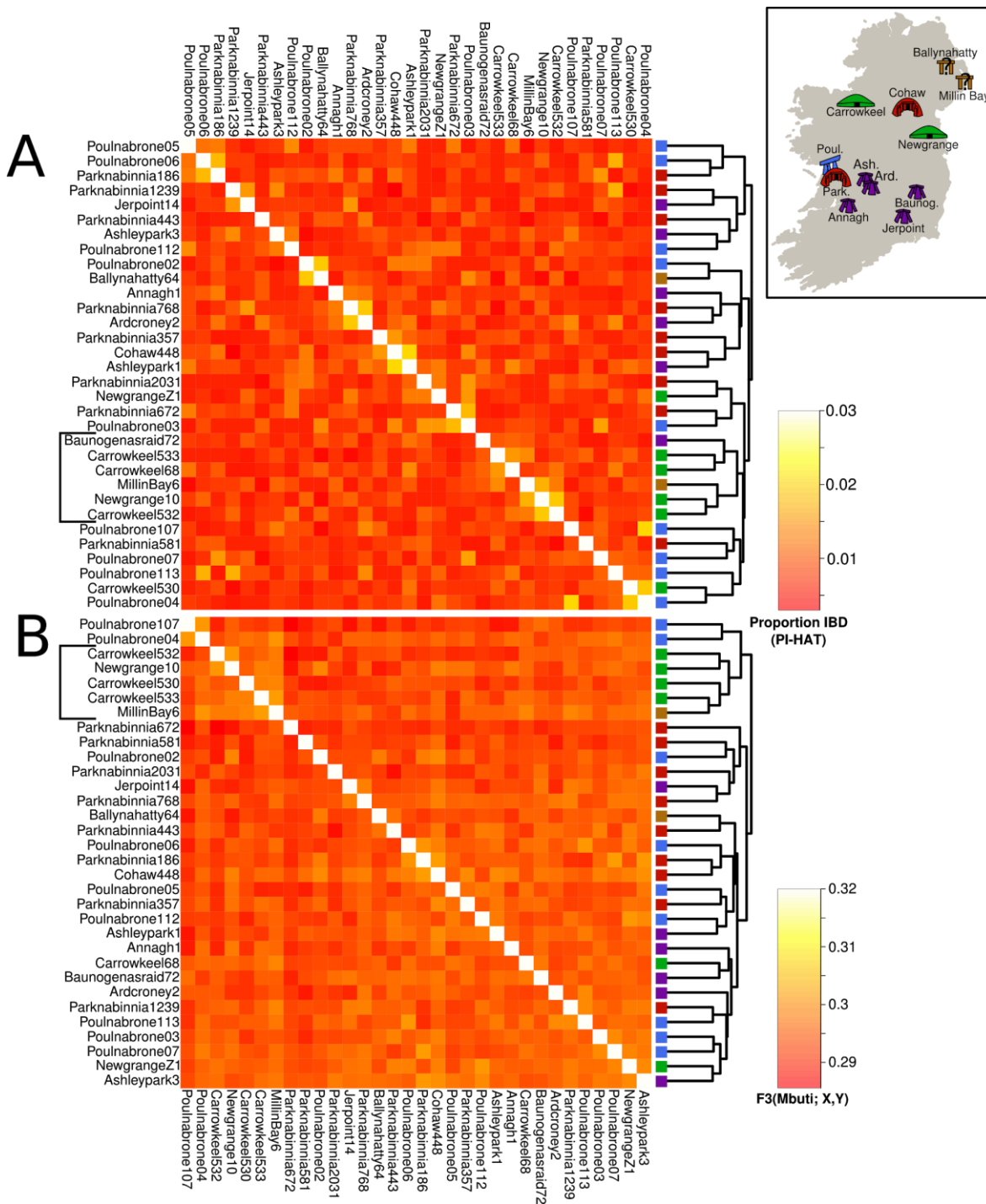


Figure 3.4. Genetic clustering of Irish Neolithic samples. A) IBD sharing heatmap based on PI-HAT values **B)** Shared genetic drift between samples, calculated with the test $f_3(\text{Mbuti}; X, Y)$. Yellow indicates higher relatedness/shared drift. A group of individuals from Carrowkeel, Newgrange and Millin Bay (bracketed) are seen to cluster consistently in both plots.

Dissecting Irish Neolithic Affinities through Linked and Unlinked Methods

ADMIXTURE analysis in Chapter Two (Fig. 1.2) revealed Irish Neolithic individuals to share a similar ancestral profile to contemporaneous individuals from Iberia and Northern Europe, possessing a minor European HG component (17-22%, with outlying value of 26% for Parknabinnia675) within a majority Anatolian background. In projection PCA, Irish samples are indistinguishable from Middle Neolithic to Chalcolithic Iberians, but on examination of the second component of variation, which separates present-day eastern and western Europeans, are seen to place slightly ‘west’ of Scandinavian, German and Italian individuals of the same time depth. The overall pattern is one of low genetic differentiation between Irish and European Neolithic individuals, relative to that seen among their Mesolithic predecessors, demonstrated in Lazaridis et al. 2016; an expected signature from a recent population expansion. This lack of structure, coupled with the likely complex admixture and migratory events that accompanied agricultural spread in northwest Europe, can be expected to have a confounding effect when establishing relationships between Irish individuals and continental Neolithic groups. To tackle this issue two distinct approaches were taken in exploring Irish Neolithic genetic affinities. The first is based on sensitive haplotype sharing methods, implemented using ChromoPainter (Lawson et al. 2012), and the second on more commonly used measures of allele sharing, summarised with D - and f -statistics (Patterson et al. 2006). The contributions of both Mesolithic and Early Neolithic groups to Irish individuals were considered and contrasted to patterns seen in continental samples.

Differential haplotypic donation into Neolithic individuals.

The level of haplotypic chunk donation seen from HG individuals to Neolithic Europeans is visualised in Fig. 3.5, and from Early Neolithic individuals to Later Neolithic individuals in Fig. 3.6. Two Anatolian individuals from Hofmanova et al. are included in Fig. 3.6, but not displayed in Table 1. Absolute chunk counts and the average length of chunks (cM) are both presented. While the former measure is likely to provide a better estimation of overall affinities, the latter can aid in the detection of recent admixture events. The main findings are presented below.

i. Local and increasing HG ancestry in continental samples

Greek, LBK and Early Spanish Neolithic individuals display the lowest levels of HG chunk donation (Fig. 3.5). In Greek samples highest levels haplotypic sharing are seen with HGs from Eastern Europe, including Romania, Hungary, Latvia and the EHG individual from Karelia. Increased sharing is also observed with SHG individual, likely due to EHG affinity. LBK samples show similar profiles, though some increased donation from HGs of more western provenance is apparent, fitting with the geographic distribution of the culture. In later Hungarian samples higher HG chunk counts are seen, with largest numbers provided by Hungarian and Romanian groups.

Surprisingly, Spanish EN samples show lower numbers of HG chunk donations relative to LBK, though unlinked analysis (PCA, ADMIXTURE and D -statistics) reveal them to possess on average higher

overall levels of HG admixture. Such a discrepancy may be reflective of more recent HG ancestry in the Spanish EN, in comparison to central Europeans, leading to a lower number of longer chunks. With this in mind, it is interesting to observe the clearly inflated levels of Spanish HG chunk donation in Spanish EN samples, coupled with an increase in average chunk lengths. The Cardial_EN1 individual in particular shows a signature of recent introgression from a source related to the Canes_Meso HG. Elevated haplotype sharing with Spanish HGs is also highly evident in later Neolithic samples from both Spain and Portugal, coupled with an increase in overall HG chunk counts, providing strong evidence for continued indigenous admixture within the peninsula. Moreover, Spain_MN3 exhibits a clear signal of a recent introgression event.

Several outliers are visible among the Iberian samples. The Chalcolithic sample ATP2 shows increased EHG chunk donation and also an inflation of sharing with Northwestern HGs, such as Bichon and Sramore62, possibly linked to increased interaction with the north and east at the dawn of the Beaker period. More intriguing is the signal of a recent admixture event involving a source related to SHG in DolmenAnsiao96B. The Portuguese sample isn't directly dated but is believed to belong to a Late Neolithic/Chalcolithic context, a period when Iberian had trade networks stretching the the Baltic and North Africa. Such outliers serve as a potential reminder that Neolithic societies in Europe, particularly in the maritime regions, were dynamic and well-connected (Cunliffe 2008).

ii. Local Mesolithic ancestry in a recently admixed Irish individual

Irish Neolithic samples display broadly similar profiles to their Iberian contemporaries, but without the inflation of Spanish HG chunk donation (Fig. 3.5). Instead, highest chunk numbers and average chunk lengths for Irish samples appear to come from Northwestern (Loschbour and Bichon) and Irish (Sramore62 and Killuragh6) hunter-gatherers. That said, a handful of samples show slightly elevated Spanish HG chunk counts and lengths, including Millinbay6 and Annagh1.

Most remarkably, the individual Parknabinnia675 displays a major inflation of chunk donation from the Irish HG Killuragh6, the highest of all pairings considered across the dataset. Moreover, these chunks appear to be large in size. The other Irish HG, Sramore62, is the second highest contributor to this individual. ADMIXTURE proportions (not shown) and the f_4 -ratio test displayed in Fig. 3.7F, suggest Parknabinnia675 is also an outlier in terms of overall HG ancestry. The most parsimonious explanation is a recent introgression event in the ancestry of Parknabinnia675 from a population related to Killuragh6. A number of other samples from Parknabinnia show inflated HG chunk numbers and lengths relative to other Irish samples, including the early Parknabinnia672, though these are northwestern (Loschbour and Bichon) rather than Irish in source.

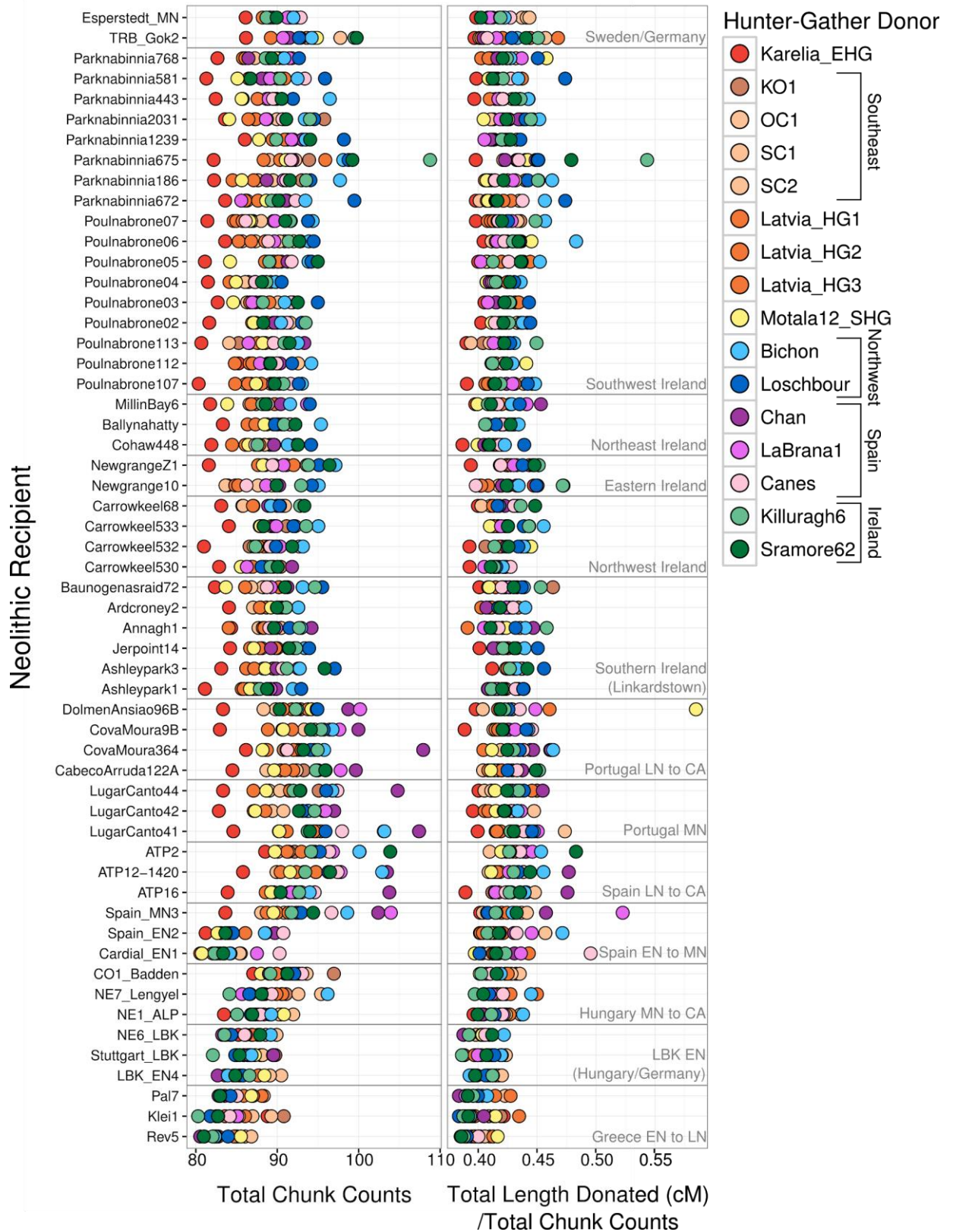


Figure 3.5. Patterns of haplotypic donation from Hunter-Gatherer to Neolithic samples. Estimates for both absolute chunk counts and average chunk lengths are displayed. Hunter-gather samples are coloured by geographical region. Neolithic samples are ordered by both region and archaeological context.

iii. Contemporaneous TRB samples show inflated eastern affinities relative to Irish Neolithic

Haplotypic sharing with HGs from further east appears reduced across the Irish dataset relative to continental Neolithic samples, although a number of exceptions exist, including Parknabinnia1239, Poulnabrone112 and several Linkardstown individuals. In contrast, two TRB individuals from Sweden and Germany show clear inflation of EHG chunk donation (Fig. 3.5). However, perhaps surprisingly, the Swedish sample receives highest donations from the two Irish HGs and a Romanian, though the Motala SHG individual also shows some increased sharing. Indeed the two Irish HGs contribute more to TRB_Gok2 than they do to any Irish sample, bar Parknabinnia675. Archaeological connections between Scandinavian and British Mesolithic groups were highlighted in Chapter Three and may be a factor in this affinity, though the megalithic context of TRB_Gok2 could also point to more recent Atlantic interactions between Sweden, Britain and Ireland.

iv. Similar Early Neolithic contributions to Irish and Iberian groups, with outlying individuals in both groups

As observed for HG chunk donation, the Irish and Iberian cohorts were seen to display broadly similar profiles of EN haplotypic sharing relative to TRB and Hungarian samples (Fig. 3.6). Specifically, Irish and Iberian samples had reduced donations from Earlier LBK, Hungarian and Aegean groups and inflated sharing with Early Neolithic samples from Spain. However, Spanish EN chunk donation was seen to be substantially higher in Iberian samples compared to Irish. In terms of average chunk length, a number of samples from Poulnabrone, two individuals from Ashleypark and NewgrangeZ1 showed distributions more similar to Iberian Individuals.

Interestingly, regional outliers with respect to HG chunk donation also tended to be outliers with respect to EN groups. In Iberia, Spanish EN individuals were the top contributors in both overall chunk counts and average lengths across all samples bar DolmenAnsiao96B, who showed inflated haplotype sharing from a source related to Neolithic Anatolians, and ATP2, for whom a Hungarian Neolithic was the second highest donator. This again emphasises eastern ancestry within these individuals, possibly via Mediterranean or more northern networks. In Ireland, while the major contributors were also Spanish EN individuals, more variability was seen in the top donators. Parknabinnia1239, who showed relatively increased EHG affinity, received highest chunk counts from LBK and Greek individuals, and possessed the smallest average lengths of Spanish EN haplotypes seen across Irish samples. Carrowkeel530 also shows unusually low levels of Spanish EN ancestry, as well as the lowest amount of HG chunk donation across Irish samples. Finally, the Newgrange10 individual, for reasons unsurmountable, appears to receive an inflated number of long haplotypes from a Late Neolithic Greek. This again emphasises the potential for long distance mobility during the course of the European Neolithic. Denser population sampling, combined with sensitive haplotypic methods, allows not only for the identification of outlying individuals, but also avoids the possible confounding effects such samples might have in the establishment of regional relationships.

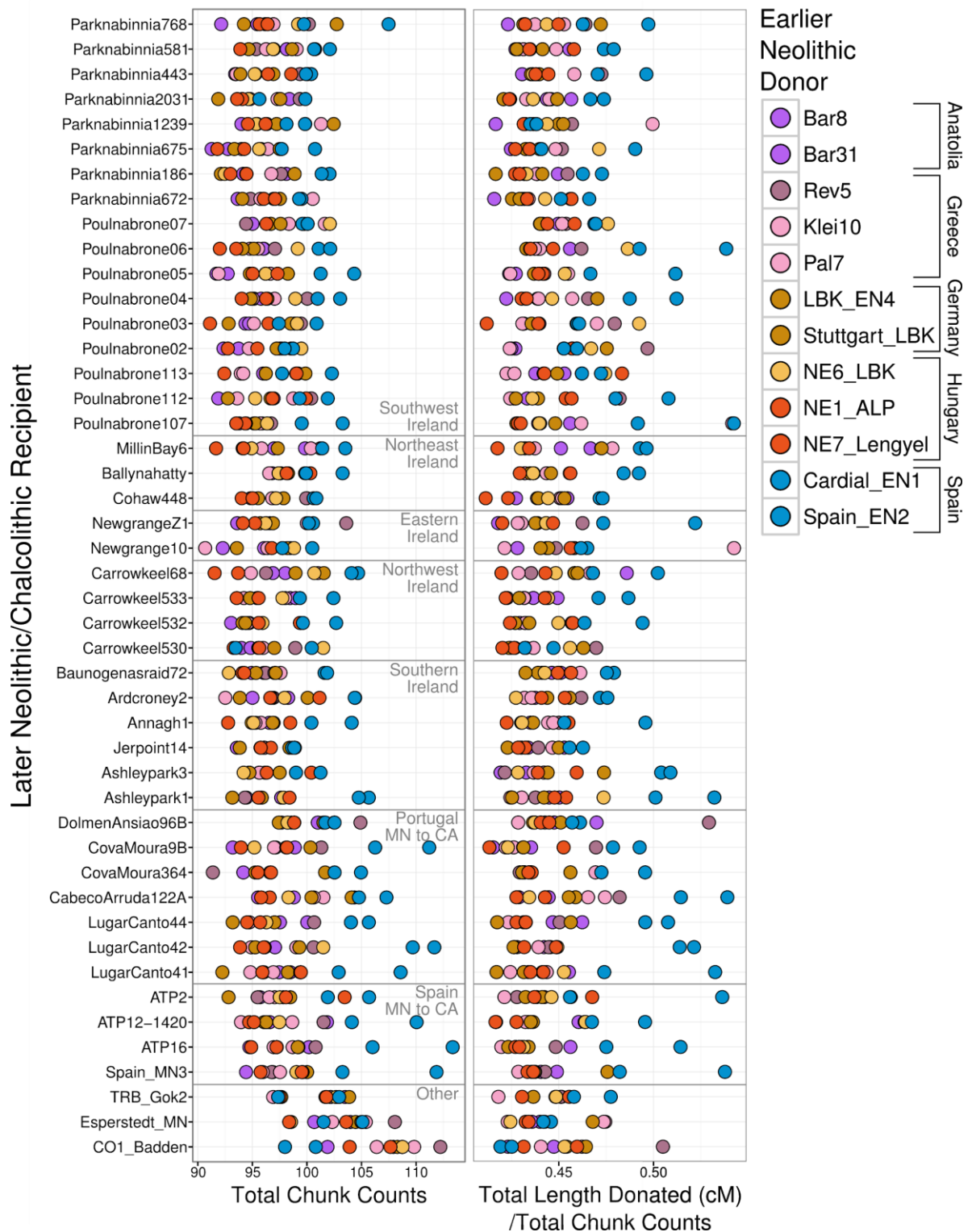


Figure 3.6. Patterns of haplotypic donation from Early Neolithic to Later Neolithic and Chalcolithic samples. Estimates for both absolute chunk counts and average chunk lengths are displayed. Early Neolithic samples are coloured by geographical region or culture. Later samples are ordered by both region and archaeological context.

Confirmation of haplotypic patterns using D-Statistics

D-statistics were used mainly to compare the contributions of larger HG populations (characterised in Chapter Three; See Table 2.1 for population key) to Neolithic individuals (Fig. 3.7). Affinities of later Neolithics to LBK and Early Spanish (Cardial and Epicardial) populations was also assessed (See Table 3.1 for population key). Notably, these tests could utilise larger numbers of samples than the above ChromoPainter analysis, given the latter's requirement for diploid genotypes.

i. Differential contributions of diverse HG groups to European Neolithic individuals

As Mesolithic Europeans possess relatively high levels of genetic differentiation, the minor components of diverse HG ancestry in farming individuals may provide insight into the geographical regions their ancestors occupied if admixture with indigenous groups took place. However, the size of this overall HG component tends to differ between Neolithic individuals, which can confound comparisons of affinities. To circumvent this, we first estimated approximate levels of general WHG introgression into Neolithic samples, relative to a population of Neolithic Anatolians, and then regressed these values against estimates from other HG groups (EHG, SHG, MAG). This was done using the statistic $D(\text{Mbuti, HG Group; Anatolian, X})$. The WHG population used was composed of Mesolithic individuals from across the continent (SEHG, SPHG, LVHG, NWHG, Irish_HG) to avoid regional bias as far as possible. The results are displayed in Fig. 3.7A-C. WHG introgression was seen to be a broadly good predictor of other HG ancestries, as expected given their shared demographic histories. However, outliers possessing more or less EHG, SHG or MAG introgression than expected were apparent, identified using studentized residuals.

In terms of WHG ancestry, lowest estimates of introgression were seen in LBK and Greek individuals ($Z=1.214-3.455$), and the highest in Swedish, Irish and Iberian samples ($Z=8.524-17.987$), with Early Neolithic Spanish, Hungarians, Italian Chalcolithic and German TRB samples forming the middle ground. EHG introgression shows a somewhat different pattern (Fig. 3.7A) in that Hungarian and German TRB individuals possess levels on par with Irish and Iberians. This is expected given their closer geographic proximity to EHG occupied regions and supports haplotypic analysis. In fact, it is possibly noteworthy that the TRB individuals do not possess more EHG introgression, given the preponderance of R1b lineages observed in this group, though these may have originated from Baltic WHG sources. The earlier German LBK population show substantial variability in EHG-related ancestry when compared to its tighter distribution of WHG introgression, with outlying individuals identified on both ends of the spectrum.

A wide distribution of EHG introgression is also seen in Atlantic populations with respect to WHG ancestry. In Iberia, two general clusters of Portuguese samples can be seen above and below the regression line. No chronological differentiation is apparent, though these samples are dated by context alone. Direct radiocarbon dates may shed some light on the observed distribution, as they appear to have in Northeast Spain, where Chalcolithic samples tend to have slightly larger positive residuals than the

preceding Neolithic. Most strikingly, a tentative Middle Neolithic sample from Portugal (LugarCanto41) possesses the highest level of EHG introgression across all samples, and is a clear outlier relative to other Iberians, hinting at a later, possibly Bell Beaker date, or, alternatively, earlier eastern connections in the region via networks unknown. Haplotypic analysis did not reveal any strong signal of EHG ancestry for the sample, though some inflation in Hungarian and Romanian HG chunk donations is seen. Less surprisingly, the Swedish TRB sample, Gok2, falls just below this Portuguese outlier on the regression plot, and when SHG introgression is considered (Fig. 3.7C) their positions reverse, suggestive of local Scandinavian HG admixture in this individual.

In Ireland, the majority of samples fall below the regression line, indicating no particularly strong eastern contribution to the island and supporting haplotypic results. The intriguing exception to this is the individual from Cohaw, Co. Cavan, the only individual sampled from a northern court tomb, who is an outlier when both SHG and EHG introgression are considered, although this is not captured in haplotypic analysis. Several samples from Poulnabrone, Carrowkeel and Ballynahatty show a reverse trend, with substantially less EHG or SHG than expected.

When Magdalenian-related introgression is regressed against that from WHG, Iberian samples are observed as the main outliers (Fig. 3.7B). This is particularly evident from the Middle Neolithic onwards, though several Early Cardial samples also possess substantial positive residuals. Given the persistence of Magdalenian-related HG groups in Iberia well into the Mesolithic period (Chapter Three) this again suggests local admixture on the peninsula from an early point in time. Magdalenian type ancestry also provides a measure by which Irish and Iberian samples, so far indistinguishable in other haploid analyses, can be pulled apart.

ii. Less localised introgression in Ireland relative to Iberia

To investigate differences in WHG ancestry between Neolithic groups, we estimated the level of introgression seen by Neolithic individuals into a clade containing the SEHG population (Hungarians and Romanians) and a more western HG population (NWHG, Irish_HG or IHG, and SPHG) - D(Mbuti, X; SEHG, Other HG). As expected, later Neolithic individuals from Ireland and Iberia showed less introgression into SEHG groups (Fig. 3.7D-E) relative to samples from further east. While Iberian and Irish samples showed indistinguishable levels of NWHG ancestry, Iberian samples could again be separated from Irish individuals on the basis of introgression into SPHG (Fig. 3.7D). It is possibly noteworthy that northwestern Neolithics do not possess more northwestern Mesolithic ancestry than their Iberian counterparts, a trend also seen in haplotypic analysis, and may indicate the extension of NWHG-related populations in or close to Iberia. Indeed, the discontinuity seen between the different Spanish HG samples in terms of Magdalenian and EHG related ancestry (Chapter Three) suggests the Mesolithic populations of Iberia were dynamic and possibly structured. Alternatively, the later Neolithic populations of Iberia may have sequestered their NWHG ancestry from a source outside the peninsula, as suggested in Lipson et al. 2017. It is also notable that two TRB individuals have seemingly less

continental NWHG ancestry than the Irish/Iberian cohort, suggesting that the contributing population did not extend into the TRB zone of influence.

Following a similar logic, we regressed values retrieved for IHG against NWHG to explore whether the Irish Neolithic population possess more IHG than expected relative to continental counterparts. However, no strong signal was seen and Irish and Iberian individuals were indistinguishable. It may be that IHG and NWHG are too closely related for this to be a meaningful test, though some interesting outliers were seen, including the Swedish TRB sample, reflecting haplotypic donation analysis. One of the earliest samples from the Irish dataset, Poul nabrone107 (3,928-3,666 cal BC), also had an inflated positive residual. However, the overall implication is that Irish Mesolithic groups played a reduced role in the island's transition to agriculture, relative to the situation in the Iberian Peninsula.

iii. Confirmation of Irish Affinities to the Cardial Neolithic

Finally, we looked at the contribution of LBK and Spanish Early Neolithic ancestry to later Neolithic samples in Europe using D(Mbuti, X; LBK_EN, Spain_EN). As levels of HG ancestry might be a confounding factor in this we plotted these values against the percentage WHG ancestry, calculated using an F4 ratio. Two Palaeolithic HG individuals were used for the ratio in an attempt to avoid the regional biases more likely among Mesolithic groups and estimates corresponded closely to ADMIXTURE values for HG ancestral components, although with slightly more variance (~15-25%, with an outlying value of 30% for Parknabinnia675).

No strong correlation was seen between percentage WHG ancestry and shared drift with the Spanish EN. However, the positive scores retrieved across samples caution that there may be a systematic shift among later Neolithic groups towards the Spanish EN due its increased HG ancestry, though relative comparisons are still possible. The TRB and Hungarian Chalcolithic samples show the least introgression into the Spanish EN, while Italian Chalcolithic individuals show a range of values similar to the Irish dataset, possibly reflective of their location on the Mediterranean route. Later Iberian samples are seen in general to have the most Early Spanish ancestry, though substantial overlap with Irish contemporaries exists. Interestingly, Parknabinnia1239, an outlier in haplotypic analysis, is also an outlier here, showing noticeably deflated allele sharing with the Spanish EN. More generally, later northern and eastern Irish samples (passage graves and Linkardstown cists) showed slightly decreased introgression into the Spanish EN, relative to individuals from the Burren, though exceptions exist, such as MillinBay6 and Parknabinnia1239.

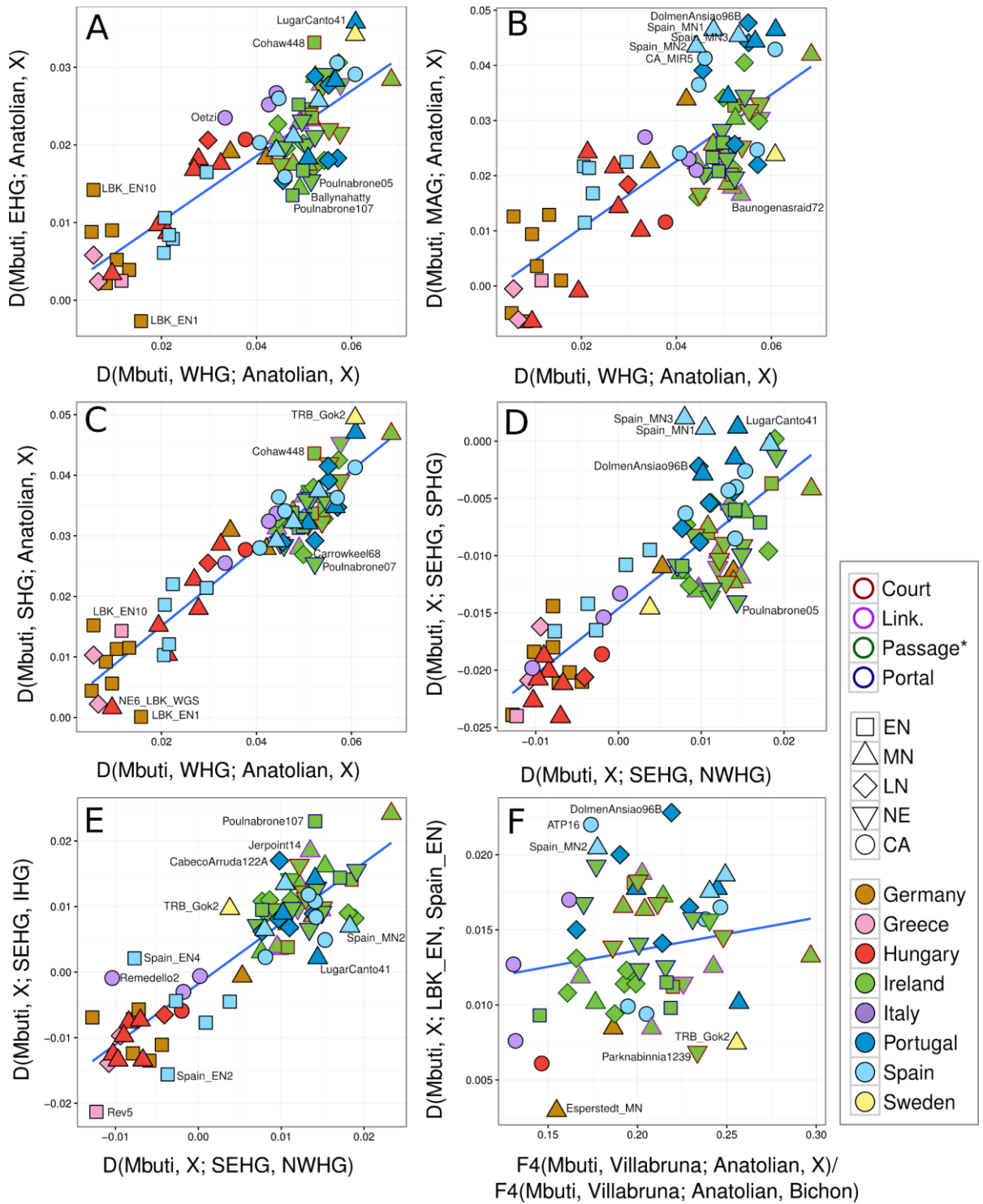


Figure 3.7. Patterns of allelic-sharing between hunter-gatherer and Neolithic groups. *D*-statistic values on the Y axes are regressed against those on the X. Outliers are identified and labelled based on studentised residuals (CI=95%). *Passage tombs also refer to the passage-like megaliths at Ballynahatty and Millin Bay.

Exploring Distributions of Homozygosity among European Neolithic Individuals

Through the use of imputed genotypes, runs of homozygosity could be explored for the first time among a large dataset of European Neolithic individuals, following the same approach outlined in Chapter Three. As for Mesolithic individuals, ROH were divided into different length categories and plotted to give insight into both recent and ancient patterns of inbreeding (Kirin et al. 2010; Pemberton et al. 2012) (Fig. 3.8). Three modern individuals were included as reference points, as well as for comparison with Fig. 2.4.

All European Neolithic individuals were seen to have similar levels of short ROH, a testament to their shared demographic origins. Values fall just above that seen in a modern French individual, a substantial reduction relative to the European Mesolithic period (Fig. 2.4), and similar to Neolithic Iranians and Anatolians, suggestive of a large ancestral population size. However, when considering ROH < 1.6 MB (Fig. 3.8A), one possible outlier can be identified, a Greek individual from the Early Neolithic (Rev5). It is perhaps noteworthy that both this sample, and an LBK individual, who also shows some inflation of short runs (Fig. 3.8A-B), have little to no European HG admixture, a process which would work to reduce ROH across length categories.

Low levels of long ROH across samples are also indicative of large outbreeding populations, the major exception being Newgrange10, who shows clear evidence of recent inbreeding, discussed further below. That said, several samples show slight inflations of long ROH, suggestive of a recent population constriction or inbreeding episode. Four individuals fall away from the main cluster seen in Fig. 3.8A, when ROH longer than 1.6 Mb are considered. These include the earliest Cardial Neolithic individual from Spain (Cardial_EN1) and Parknabinnia672, the earliest dated sample from the court tomb site. Founder effects would not be unexpected during initial colonisation events, especially via long-distance maritime routes, though any resultant inflation in homozygosity would likely be alleviated by further migration and also possibly admixture with indigenous communities. With this in mind, it is interesting to observe the longest ROH tracts (> 6Mb) occur in genomes from earlier Neolithic individuals, including the aforementioned Cardial_EN1, two LBK individuals (Stuttgart_LBK and NE6_LBK_WGS) and the Early Greek Neolithic, Rev5. However, several later Neolithic individuals also possess a number of long homozygous tracts, namely a Later Hungarian individual (NE7_Lengyel_WGS), the Swedish TRB sample (Gok2_TRB) and one Portuguese later Neolithic (CovaMoura364).

Among the 31 imputed Irish Neolithic samples (Newgrange10 excluded) only five possessed a tract longer than 6 Mb, suggesting that inbreeding was not common on the island during the period. Notably, three of these individuals are from passage or passage-like structures, contemporaneous with Newgrange10, including two individuals from anomalous megaliths on the northeast coast (Ballynahatty and MillinBay6), and an individual from Carrowkeel (Carrowkeel530), with the latter possessing three

such tracts. The other two samples were the aforementioned Parknabinnia672, and another undated individual from the same site (Parknabinnia443).

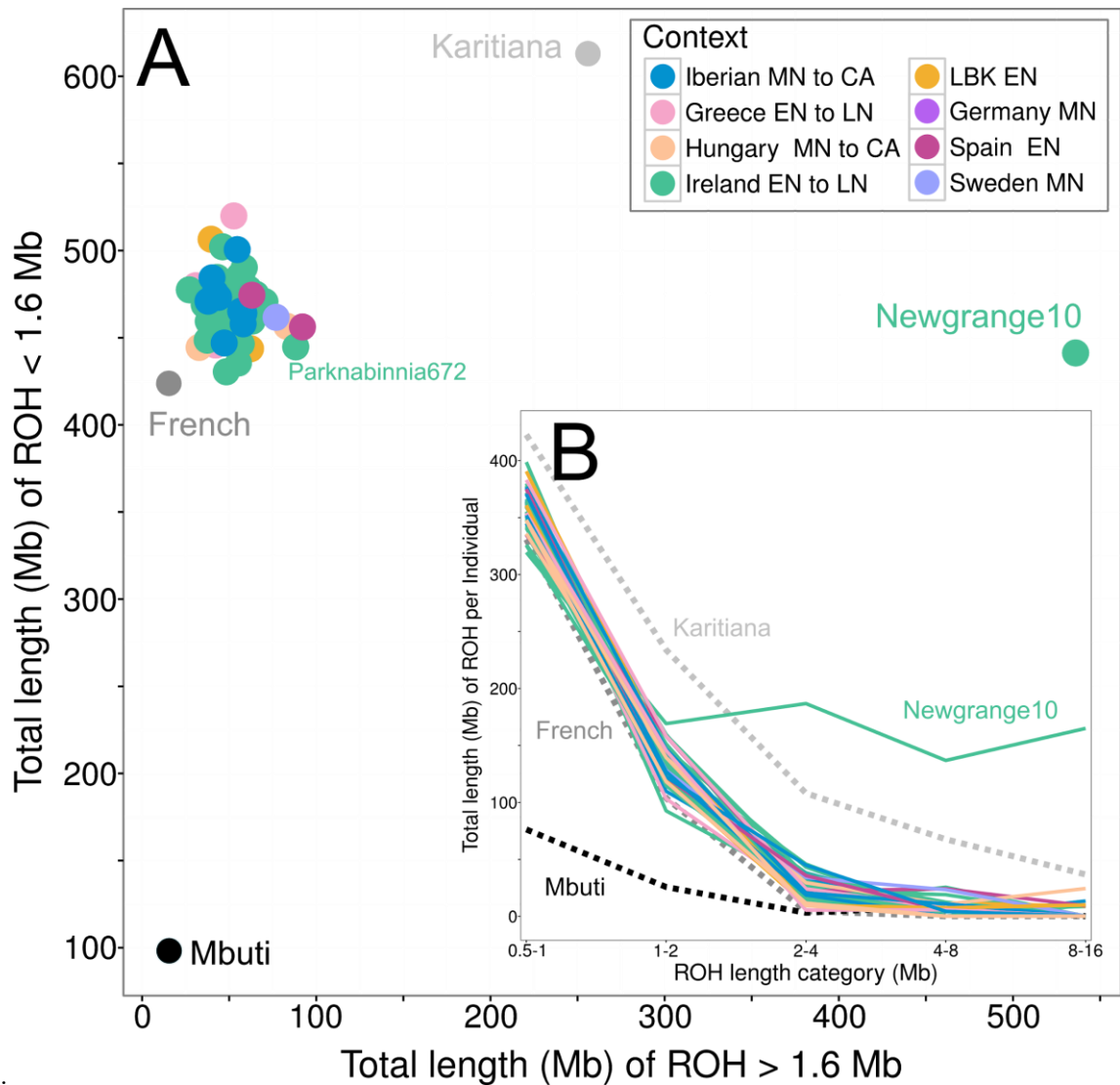


Figure 3.8. Estimated distribution of ROH for imputed Neolithic samples, placed in the context of values from modern individuals. Colour keys are identical throughout and samples noted in the text are labelled. **A)** The total length of long ROH (> 1.6 Mb) against short ROH (< 1.6 Mb) plotted for 56 Neolithic and 3 modern individuals (grey). **B)** Total length of ROH for a series of length categories for the same set of individuals.

Inbreeding within the immediate family of Newgrange10

ROH analysis of imputed Neolithic individuals showed Newgrange10 to possess massively increased numbers of long ROH (>1.6 Mb) (Fig. 3.8). Longer tracts of ROH can be broken by false heterozygote calls, which can result from a number of potential confounders in aDNA data processing pipelines, including post-mortem damage, sequencing error, alignment error and imputation error. To circumvent this issue, homozygote and heterozygote calls were plotted along a physical map of the genome for Newgrange10, to identify by eye the distribution of homozygous tracts (Fig. 3.9A). A number of extremely long ROH were observed, including several roughly 40-50 Mb in length, indicating a high degree of relatedness between the individual's parents.

To establish their exact coefficient of relatedness, different inbreeding scenarios were simulated on a population of Neolithic individuals from the Parknabinnia court tomb and the resultant ROH distributions were compared to that seen in Newgrange10, following similar methods to those described in Prüfer et al. 2014. The size and number of tracts in the sample allowed us to restrict simulations to inbreeding coefficients of 12.5% and 25%. A coefficient of 12.5% was simulated through half-sibling pairings and 25% through sibling pairings (See Methods for details). While other inbreeding pedigrees can lead to these coefficients (12.5% - grandparent and grandchild, uncle and niece/aunt and nephew, double first cousins; 25% - parent and child), these could not be differentiated in Prüfer et al. 2014 and so were not considered.

When total length of the genome under ROH is considered, Newgrange10 shows a value close to the average of sibling offspring (Fig. 3.9B). However, Newgrange10 has an inflated number of ROH tracts compared to both sibling and half-sibling offspring and thus the average length of ROH is somewhat diminished relative to sibling offspring. That said, the values seen for Newgrange10 are securely in the range of that seen for sibling offspring and among the highest values achieved for half-sibling offspring. No allowance for mutation or genotyping error was made during inbreeding simulations, which work to break up longer tracts, and is the likely reason for the comparatively inflated numbers of ROH segments seen in Newgrange10 and their diminished average length.

When the total number of long ROH (>7.5 Mb or ~10 cM) is considered, Newgrange10 again shows distributions most similar to sibling offspring, though when the total length of these tracts is considered Newgrange10 falls slightly below the interquartile range observed for these simulations. The length of the longest single ROH segment in Newgrange10 (~42 Mb with a relaxed heterozygote allowance) is also outside the interquartile range for sibling offspring. However, visualisation of ROH in Fig. 3.9A reveals a clear tract over 50 Mb on chromosome 6, the average seen for sibling offspring. Taken together, an inbreeding coefficient of 25% is most likely for Newgrange10.

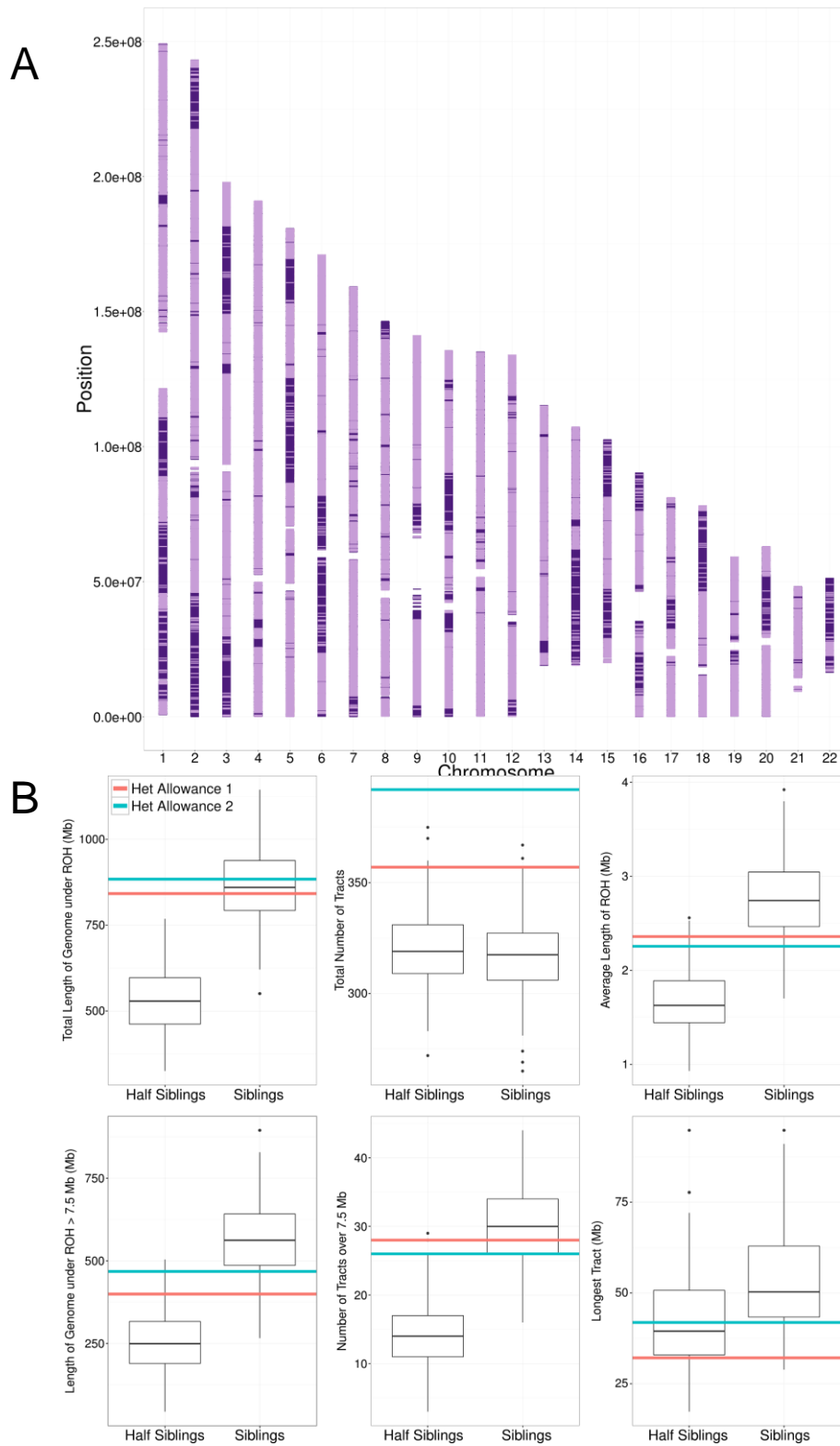


Figure 3.9. Evidence of recent inbreeding in the pedigree of Newgrange10. A) A physical map of the autosomal genome with heterozygote (light purple) and homozygote (dark purple) genotypes for Newgrange10 plotted. Large homozygous tracts are evident on over half of the chromosomes. **B)** Different measures of genomic ROH for simulated offspring of half-sibling and sibling pairings. Coloured horizontal lines represent values seen for Newgrange10 when two different heterozygote allowances are considered.

Conclusions

Ireland, on the northwestern fringe of the Atlantic, was the final frontier reached by the European Neolithic after nearly three millennia of expansion from their Anatolian source. Indeed, agricultural communities had existed in northwestern Europe for over a thousand years before the English Channel was at last crossed *circa* 4,000 BC. This period also saw the final assimilation of Scandinavia, the Northern Netherlands and the Cantabrian coast, pulling the entire continent into a larger sphere of interaction, evidenced by long distance trade networks and widely shared cultural phenomena, unseen in the preceding Mesolithic era. The Atlantic megalithic zone, of which Ireland formed a clear focal point, is perhaps the most impressive manifestation of such extensive connectivity and social complexity.

A new window is offered here into both the genetic, and potentially societal structure of Neolithic communities in Ireland, their relationships to those on the continent and to preceding Mesolithic groups. Perhaps most importantly, this dataset exhibits the future potential of large-scale genomic surveys to identify outlying individuals and subtle patterns of differentiation hidden within largely homogenous ancient groups, a key requirement for addressing more regional and site-specific questions in archaeology. Indeed, the hints of Y chromosome and autosomal structure seen here between sites, and anomalous samples such as Parknabinnia675 and Newgrange10, could be easily contextualised with denser sampling of Neolithic individuals across Ireland and Western Europe. Fortunately, this is becoming an achievable goal, as evidenced by the excellent rates of aDNA survival seen in the majority of megalithic sites surveyed here (Appendix II), though obstacles still exist. While some regions, such as the karstic landscape of the Burren, displayed unprecedentedly high levels of aDNA preservation, other sites, most notably the northern court tombs of Audleystown and Ballyalton, yielded virtually no endogenous molecules. Differential DNA survival based on soil type or burial rite, including cremation, may bias sampling to particular groups, and should be kept in mind as genomic datasets continue to grow.

In this spirit, the following sections present a base foundation on which future genomic research can be built, with three central questions surrounding the Irish Neolithic considered.

1. Where did they come from and how did they arrive?

Archaeological evidence strongly suggests an ultimate origin for Irish and British Neolithic populations along the northern coast of France, with both Atlantic and Northern European influences evident in the ensuing material cultures (Cunliffe 2013; Mallory 2013). Importantly, ROH analysis revealed all Irish Neolithic individuals, the anomalous Newgrange10, excluded, to belong to a large outbreeding population, with even the earliest samples from Poul nabrone (3,942-3,702 and 3,928-3,666 cal BC) showing no trace of a recent founding bottleneck (Fig. 3.8). This suggests the seaborne colonisation of Ireland was not carried out by a few isolated pioneer groups, but by a population of large effective size,

maintained by mechanisms such as continuous migration into regions or persistent back-and-forth interaction among a large network of communities.

The best evidence seen for any elevated background inbreeding (Newgrange10 excluded) comes from the sample Parknabinnia672 (3,639-3,514 cal BC), which gave the earliest radiocarbon determination from this court tomb, initially constructed *circa* 3,700 – 3,570 BC (Jones forthcoming). The individual is a clear outlier relative to Irish and other Neolithic samples, showing a distribution most similar to the earliest Cardial Neolithic from Spain (Cardial_EN1), with both possessing comparatively inflated levels of long ROH (>1.6 Mb). This could suggest some constriction of their ancestries. It is perhaps notable that long distance maritime enclave colonisation was the main process by which the Cardial Neolithic spread, while the outlying and isolated nature of north Munster court tombs is suggestive, though earlier Neolithic communities were already present in the Burren upon their appearance (Jones forthcoming). However, any resultant population constrictions that may have occurred upon colonisation were not prolonged and the signal seems to have been quickly erased in both regions.

The continental communities that gave rise to the Irish and British Neolithic belonged to a region which saw substantial overlap between the two main trajectories of Neolithic spread (Fig. 3.1). However, with no available palaeogenomic evidence, the relative genetic input of LBK and Impressed Ware groups into the different cultures of northern France remains unknown. That said, trends in both haplotypic (Fig. 3.6) and allelic sharing (Fig. 3.7F) suggest a substantial contribution from the Mediterranean route of expansion into the Irish Neolithic, relative to the Danubian corridor, also reported in British data (Olalde et al. 2017). Irish samples show inflated sharing with Spanish Early Neolithic individuals over Hungarian and LBK groups, in a pattern similar to contemporary Italians and Iberians. Moreover, clear differentiation is seen between the Irish Neolithic and TRB culture, which possesses not only a larger Central European Neolithic input, but also higher affinities to eastern HG groups. Nevertheless, Irish outliers in terms of genetic affinities are identifiable, most notably two court tomb samples, Cohaw448 (inflated EHG allele-sharing), and Parknabinnia1239 (inflated haplotype and allele-sharing with Central European Neolithics).

Importantly, later Iberian samples show increased haplotypic and allelic input from the Spanish EN, relative to Irish individuals. This, together with the clear lack of Spanish HG ancestry in the Irish Neolithic (though again potential outliers, such as MillinBay6 exist) argues against any large-scale direct contribution from Iberian populations to the island, via the Atlantic. From an archaeological perspective it is more likely that the Mediterranean-type ancestry within the Irish Neolithic diverged from Spanish groups at the Golfe du Lion, where it spread both inland via the Rhône and westwards toward the coast. Furthermore, admixture with LBK and derivative groups in the north most probably took place, suggested by the increased haplotype and allele sharing seen with LBK individuals for some Irish samples. The extent of such admixture cannot at present be estimated, but it may be partially responsible for the overall dilution of Cardial ancestry in Irish samples relative to Iberians.

Thus, the progenitor population of the Irish Neolithic was likely a group with majority Mediterranean and minority Danubian ancestral input, a description that can be potentially fitted to a large number of cultures living along the Northern French coast. Impressed Ware influence on the earliest LBK communities in the Paris Basin, prior to their expansion west, is suggested in the Limburg and La Hoguette ceramic styles, while coastal Epicardial sites have been identified not far south of Brittany and the Loire valley. In this respect, when considering the two major proposed routes of Neolithic spread across the channel (Fig. 3.1C), the findings here neither invalidate nor demonstrate Armorican or Upper French source points. This question can only be addressed by dense genomic sampling of the French Neolithic itself. The best hint of evidence for multiple colonisations of the islands comes from uniparental markers, namely the differentiation in Y chromosome lineages seen between passage and portal tomb sites in Ireland, linked to Atlantic coastal communities, and the Parknabinnia and Cohaw court tombs. A similar differentiation is visible between western and northern Scottish groups. However, the temporal component of such patterns remains to be established. Finally, it is worth remembering the extensive connectivity that existed across the Neolithic of Atlantic Europe, which may have aided in the long-distance dispersal of ancestries, and produced the several outlying individuals seen in Portugal and Ireland.

2. What happened to the Mesolithic population?

Roughly one fifth of Irish Neolithic ancestry is derived from European HG populations, estimated using both ADMIXTURE analysis and an f_4 -ratio (Fig. 1.2; Fig. 3.7F). This can be assumed to have accumulated along the various migratory paths of farming communities, discussed above. A key question is whether any small component of this HG ancestry was contributed by the indigenous population of Ireland itself. However, if such admixture is to be detected there are two major confounding issues to overcome.

Firstly, Chapter Three highlighted inflated levels of shared drift between Irish HGs and contemporaries from the northwest of the continent. Moreover, as both archaeological and genetic evidence suggests admixture occurred between these continental populations and incoming farming communities, who would eventually give rise to the Irish Neolithic. Identifying Irish-specific Mesolithic ancestry within such a background of closely related NWHG introgression is expectedly a difficult task. The general inability to differentiate between NWHG and Irish HG affinities in both allelic and haplotypic analysis highlights this problem, though one of the earliest Irish Neolithic samples (Poulnabrone107; 3,928-3,666 cal BC) may show some slightly increased local affinity (Fig. 3.7E). Moreover, yet unsampled British HGs are likely to form part of the same ancestral continuum, and distinguishing between HG groups from the two islands may prove even more arduous. Secondly, the likely scale of the population movement into the island over the course of the period, indicated here by ROH analysis (Fig. 3.8), suggests that even the total assimilation of Irish HG communities, estimated as only a few thousand individuals in number (Mallory 2013; Woodman 2015), would have produced only a marginal

contribution to the genomic pool. Indeed, it may have been completely erased by continued migration over the generations.

For these reasons, the identification of an outlying individual from the Parknabinnia Court Tomb in southwest Ireland is a hugely significant finding. Both ADMIXTURE analysis and *D*-statistics revealed Parknabinnia675 to have substantially inflated levels of Mesolithic ancestry (26-30%), higher than that seen for any other Neolithic farming individual so far published. This alone is suggestive of a recent admixture event, which could be confirmed through ChromoPainter analysis. Haplotypic chunk donations were inflated in Parknabinnia675 for all HG samples, but a substantial excess was seen from the Irish Mesolithic individual Killuragh6, the highest observed across the entire dataset. Moreover, these haplotypes were not only numerous but long, a typical signal of recent introgression.

Taken together, this is strongly suggestive of a HG ancestor (or farming individual of majority HG ancestry) closely related to Killuragh6 within the extended pedigree of Parknabinnia675. While the population to which this individual belonged may not have been situated in Ireland, the close geographic proximity between the Killuragh and Parknabinnia sites is noteworthy. Moreover, the strong preference seen for Killuragh6 haplotypes, over those from the highly-related Sramore62 HG from the north of the island, implies localised affinities. While the recent HG introgression seen in Parknabinnia675 may have come from a British or continental source with differential affinities to Sramore62 and Killuragh6, due to structure within the island or outside contacts as discussed in Chapter Three, the far more parsimonious explanation for the haplotypic imbalance is that the extended HG haplotypes in Parknabinnia675 are indeed Irish in origin.

Asides from his inflated HG ancestry, Parknabinnia675 shows genetic affinities similar to other Irish Neolithics and possesses a third degree relative (Parknabinnia357) from within the same burial, suggesting the individual was an integrated member of the community. Perhaps most puzzling, the court tomb had been used for some centuries before the inhumation of Parknabinnia675, who dates to the boundary of the Middle and Late Neolithic (3,100-2,910 cal BC). Such a late radiocarbon determination implies the survival of a community of majority Mesolithic ancestry well into the Neolithic era. This is slightly problematic as scant archaeological evidence for such a group exists in Ireland, or for that matter Britain. No definite trace of Irish Mesolithic culture can be seen in the succeeding Neolithic period, although admittedly, archaeological sites from the Mesolithic period itself are rare (Woodman 1977; Mallory 2013; Woodman 2015). However, possible exceptions exist, most notably a wetland site dating to *circa* 3,700-3,400 BC, which yielded hazelnut shells, wild pig, swan and cattle bone, a reed basket and human skull fragment, taken as evidence of foraging activity on the marshes (O'Sullivan 2001). Remarkably, this site is located only 30km from the Killuragh site, on the Shannon estuary, just south of the Neolithic Burren landscape to which Parknabinnia belongs.

Thus, it is not improbable that small HG bands continued their way of life in environmental niches unoccupied by farming groups, and this is perhaps the first genetic evidence for some such survival. However, if these communities did exist, their impact on the larger Neolithic population appears minimal, both in terms of archaeological culture and genetic ancestry. Future genomic surveys of megalithic sites may unearth further evidence of these potential invisible communities, as well as continued archaeological exploration of caves (Dowd 2015), which recently pushed possible HG occupation of the island back several thousand years (Dowd & Carden 2016), and, when combined with aDNA analysis, may also have the potential to push it forward.

3. How was society in Neolithic Ireland structured?

Perhaps the most profound implications for how we view Neolithic societies in Ireland, and more broadly across the Atlantic seaboard, come from the petrous bone (3,339-3,028 cal BC) sampled from the heart of the main passage grave at Newgrange, in use between 3,300 and 2,900 BC (O'Kelly 1982). ROH analysis of this individual revealed substantial tracts of homozygosity within his genome (Fig. 3.9A), which can only be the result of a recent inbreeding event in his ancestry. Simulations of inbreeding pedigrees in Irish Neolithic samples suggest that by far the most probable scenario is a sibling-sibling pairing (or possibly parent-child), while even the most conservative estimates rule out anything less than an inbreeding coefficient of 12.5% (Fig. 3.9B). The large and outbreeding nature of Irish Neolithic communities (Fig. 3.8) suggests this was not an unavoidable or inadvertent pairing brought on by demographic stress or population isolation. Moreover, the burial of this individual within one of the most imposing megaliths of the larger Brú na Bóinne monumental complex, suggests the relationship between his parents did not break a taboo or result in social ostracization. In fact, from an ethnographic perspective, it seems highly likely that the individual's ancestral status was a key determinant in his receiving a burial of such perceived distinction.

Incest within extended families has been commonly used throughout history to confine power to a particular kin group, including many European royal houses. However, marriage between immediate family members is relatively rare. Indeed, sibling couplings have been held in near universal abhorrence across all recorded human cultures, with only a few known exemptions from this taboo (Bixler 1982). The most famous cases of sibling marriages are provided by the pharaohs of Egypt, including Tutankhamun's parents (Hawass et al. 2010), though other examples exist, most well documented being the royal lines of Hawaii and the Incan Empire. Outside of ruling classes, societal acceptance of sibling marriage is vanishingly rare, with Roman Egypt providing possibly the only exception, where the royal practice, for reasons unknown, seems to have extended to the general populace (Hopkins 1980). This does not appear to be the case for Neolithic Ireland, with no evidence of background inbreeding seen in any other sample, supporting an elite status for Newgrange10.

Notably, very similar ideologies were shared amongst the disparate cultures of Peru, Egypt and Hawaii. Monarchs did not rule simply by divine right but as gods themselves, with no differentiation between

socio-political and religious roles. By mating within the immediate family, this divinity remains undiluted, alongside the power and riches that accompany it. Incest can also legitimise succession in situations where a ruler takes many wives and concubines, a common occurrence in all three cultures, with royal blood taking preference over lesser peers. Given the archaeological context from which the sample was retrieved, it is likely that a similar set of social dynamics were motivating factors in the pairing that gave rise to Newgrange10. Moreover, the extensive engineering and astronomical knowledge, not to mention sheer human expenditure required to construct the monuments of Brú na Bóinne, strongly suggests the presence of such an underlying elite. Indeed, social hierarchies and monumental architecture have long been argued to be inextricably linked (Abrams 1989; Renfrew 1973). Intriguingly, both Egyptian and Hawaiian societies also invested heavily in monument building, which served to provide combined symbols of religious and political authority (Kolb et al. 1994). The clear spiritual undertones evident in the artwork and solstice alignment of Newgrange may hint at a similar role for the structure (Hensey 2015). Further sampling from the other mega-passage tombs in the Brú na Bóinne complex, as well as clearly connected sites in Orkney, Brittany and Iberia, will be required to establish whether incestual relationships were a common occurrence among Atlantic Neolithic elites.

The dynastic implications of the Newgrange10 genome stand in sharp contrast to the lack of kinship seen in other Irish megalithic sites. Across the population samples that were taken from the Parknabinnia, Poul nabrone and Carrowkeel sites, only one pair of third degree relatives was identified (Parknabinnia357 and Parknabinnia675; Appendix I). Individuals from a Spanish Middle Neolithic passage grave also showed no evidence of recent relatedness (Haak et al. 2015). While long term use and reuse of megaliths may obscure kinship dynamics, the implication is that these tombs were not restricted to single familial groups. Notably, a later wedge tomb from the Chalcolithic period (discussed in Chapter Five) does not follow this trend. Here, while DNA was only retrieved from two individuals, these were seen to be sisters (Appendix I). Overall, this emphasises the variety of roles megalithic structures likely served across time in different communities or sections of society.

Some evidence of differential use of megaliths comes from the two sites sampled in the Burren, the Parknabinnia court tomb (n=9) and Poul nabrone portal tomb (n=9), separated by a mere 10km, which show clear discontinuity in Y chromosome lineages (Fig. 3.2B). Both were constructed in the Early Neolithic period, with Poul nabrone pre-dating Parknabinnia by a century or two, and continued in use, possibly intermittently, for many centuries (Lynch 2014; Jones forthcoming). Neolithic societies in the Burren and north Munster were likely of smaller-scale, with lower and more stable demographics relative to populations further east and north, appearing somewhat isolated from the increasingly elaborate developments in megalith construction and probable social hierarchies that accompanied them in the Middle and Late Neolithic (Hensey 2015; Jones forthcoming). The differentiation in Y chromosome lineages seen between Poul nabrone and Parknabinnia across a wide chronological period, may be indicative of two separate and somewhat unchanging communities operating within the region for a prolonged time, or alternatively, societal structure maintained within a single group (see Beckett 2015

for discussion of differential burial practices in the Burren). The atypical, outlying and unchanging nature of the Parknabinnia tomb, isolated from the main distribution of court tombs in the north, and its continued use after court tombs were likely abandoned elsewhere (Jones forthcoming), may be suggestive of such an inward looking community, and it is perhaps notable that this tomb yielded the only pair of related individuals from the Neolithic sample. Importantly, any genetic structure present between Poul nabrone and Parknabinnia appears to be restricted to the male line, with no autosomal or mitochondrial differentiation detectable at this resolution. This is in keeping with previous evidence for patrilocality and female migration among sedentary Neolithic populations (Rasteiro & Chikhi 2013). With this in mind, it is perhaps remarkable that the majority of Neolithic individuals randomly selected from both sites were male (8/9 and 8/12), a bias also possibly present in the Carrowkeel sample (4/6), though this has not been noted in osteological analysis (Beckett & Robb 2006).

Another example that hints at emphasis on male line ancestry is seen in two broadly contemporaneous burials from Linkardstown cists at Jerpoint West and Baunogenasraid. This distinctive burial type is a Middle Neolithic phenomenon, most common in the southeast of the island, typically containing the remains of a single adult male and argued to represent the emergence of a social elite in the region (Brindley & Lanting 1989). Both Jerpoint14 and Baunogenasraid72 belong to the rare Y haplogroup H2a, unobserved in other British and Irish samples, which could be indicative of shared male line ancestry between the similar sites. Alternatively, this haplogroup may simply be present at higher frequency in the southeast of the island. Notably, other Linkardstown cists further west in Munster do not belong to the same lineage.

Lastly, it is worth mentioning the slightly inflated autosomal affinity that Newgrange10 exhibits to samples from other passage (or passage-like) tomb sites in Ireland, most particularly a contemporaneous individual from the Carrowkeel cemetery (Carrowkeel532), evidenced in both outgroup f_3 -statistics and IBD sharing (Fig. 3.4). Such signals hint at possible connectivity between the sites, although they may be the result of temporal rather than cultural links. Moreover, male individuals from these sites all belong to the same Y chromosome lineage (I-M284), which is the sole lineage of the Poul nabrone portal tomb. In IBD and f_3 -statistics, passage tomb samples also appear more likely to cluster with individuals from Poul nabrone than Parknabinnia. It remains to be seen whether such perceived patterns hold on denser sampling of the island, but their emphasis here may help provide lines of inquiry with which to inform future research.

References

- 1000 Genomes Project Consortium et al., 2015. A global reference for human genetic variation. *Nature*, 526(7571), pp.68–74.
- Abrams, E.M., 1989. Architecture and Energy: An Evolutionary Perspective. *Archaeological Method and Theory*, 1, pp.47–87.
- Allentoft, M.E. et al., 2015. Population genomics of Bronze Age Eurasia. *Nature*, 522(7555), pp.167–172.
- Baird, D., 2012. The Late Epipaleolithic, Neolithic, and Chalcolithic of the Anatolian Plateau, 13,000–4000 BC. *A companion to the archaeology of the Ancient Near East*, pp.431–465.
- Bánffy, E., 2006. Eastern, central and western Hungary-variations of Neolithisation models. *Documenta Praehistorica*, 33, p.125.
- Beau, A. et al., 2017. Multi-scale ancient DNA analyses confirm the western origin of Michelsberg farmers and document probable practices of human sacrifice. *PLoS One*, 12(7), p.e0179742.
- Beckett, J.F., 2006. *Death and burial on the Burren : a taphonomic study of three megalithic monuments in County Clare, Ireland*. UK: Doctoral Thesis, University of Cambridge.
- Beckett J.F. & Robb, J., 2006. Neolithic burial taphonomy, ritual and interpretation in Britain and Ireland: a review. In R. Gowland & C. Knüsel (eds.), *The Social Archaeology of Funerary Remains*. Oxford: Oxbow, pp.57–80.
- Bixler, R.H., 1982. Sibling incest in the royal families of Egypt, Peru, and Hawaii. *Journal of Sex Research*, 18(3), pp.264–281.
- Bollongino, R. et al., 2013. 2000 years of parallel societies in Stone Age Central Europe. *Science*, 342(6157), pp.479–481.
- Brindley, A.L. & Lanting, J.N., 1989. Radiocarbon Dates for Neolithic Single Burials. *The Journal of Irish Archaeology*, 5, pp.1–7.
- Broushaki, F. et al., 2016. Early Neolithic genomes from the eastern Fertile Crescent. *Science*, 353(6298), pp.499–503.
- Browning, S.R. & Browning, B.L., 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localised haplotype clustering. *American Journal of Human Genetics*, 81(5), pp.1084–1097.
- Cassidy, L.M. et al., 2016. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proceedings of the National Academy of Sciences of the United States of America*, 113(2), pp.368–373.
- Chang, C.C. et al., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4, p.7.
- Chowdhury, R. et al., 2009. Genetic analysis of variation in human meiotic recombination. *PLoS Genetics*, 5(9), p.e1000648.
- Cummings, V., 2017. *The Neolithic of Britain and Ireland*. London: Routledge.
- Cunliffe, B.W., 2008. *Europe between the Oceans: 9000 BC-AD 1000*. New Haven-London: Yale University Press.
- Cunliffe, B.W., 2013. *Britain Begins*. Oxford: Oxford University Press.
- Czekaj-Zastawny, A., Kabaciński, J. & Terberger, T., 2013. The Origin of the Funnel Beaker Culture from a southern Baltic coast perspective. In S. Kadrow & P. Włodarczak (eds.), *Environment and subsistence—forty years after Janusz Kruk's "Settlement studies"*. Poland: Institute of Archaeology Rzeszów University, pp.409–428.
- Delaneau, O., Marchini, J. & Zagury, J.-F., 2011. A linear complexity phasing method for thousands of genomes. *Nature Methods*, 9(2), pp.179–181.
- Dowd, M., 2015. *The Archaeology of Caves in Ireland*. Oxford: Oxbow Books.
- Dowd, M. & Carden, R.F., 2016. First evidence of a Late Upper Palaeolithic human presence in Ireland. *Quaternary Science Reviews*, 139, pp.158–163.
- FamilyTreeDNA, 2017a. I-M223 Y-Haplogroup Project. Available at: <https://www.familytreedna.com/groups/m223-y-clan/>
- FamilyTreeDNA, 2017b. I2a Y-Haplogroup. Available at: <https://www.familytreedna.com/groups/i-2a-hap-group/>
- Gamba, C. et al., 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nature communications*, 5, p.5257.
- Gillis, R.E. et al., 2017. The evolution of dual meat and milk cattle husbandry in Linearbandkeramik societies. *Proceedings of the Royal Society B: Biological Sciences*, 284(1860).
- González-Forbes, G. et al., 2017. Paleogenomic Evidence for Multi-generational Mixing between Neolithic Farmers and Mesolithic Hunter-Gatherers in the Lower Danube Basin. *Current Biology*, 27(12), pp.1801–1810.e10.
- Green, R.E. et al., 2010. A draft sequence of the Neandertal genome. *Science*, 328(5979), pp.710–722.
- Günther, T. et al., 2015. Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proceedings of the National Academy of Sciences*, 112(38), pp.11917–11922.
- Gurdasani, D. et al., 2015. The African Genome Variation Project shapes medical genetics in Africa. *Nature*, 517(7534), pp.327–332.
- Gurova, M. & Bonsall, C., 2014. “Pre-Neolithic” in Southeast Europe: a Bulgarian perspective. *Documenta Praehistorica*, 41(0), p.95.
- Haak, W. et al., 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522(7555), pp.207–211.

- Hawass, Z. et al., 2010. Ancestry and pathology in King Tutankhamun's family. *The Journal of the American Medical Association*, 303(7), pp.638–647.
- Hensey, R. et al., 2015. *First Light: The Origins of Newgrange*. Oxford: Oxbow Books.
- Hofmann, D., 2016. The changing role of La Hoguette pottery in an LBK context. In L. Amkreutz, F. Haack & D. Hofmann (eds.), *Something Out of the Ordinary? Interpreting Diversity in the Early Neolithic Linearbandkeramik and Beyond*. Cambridge: Cambridge Scholars Publishing, p.191.
- Hofmanová, Z. et al., 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proceedings of the National Academy of Sciences of the United States of America*, 113(25), pp.6886–6891.
- Hopkins, K., 1980. Brother-Sister Marriage in Roman Egypt. *Comparative Studies in Society and History*, 22(3), pp.303–354.
- International Society of Genetic Genealogy, 2017. Y-DNA Haplogroup I and its Subclades. Available at: https://isogg.org/tree/ISOGG_HapgrpI.html.
- Jones, C. forthcoming. Regional traditions and distant events - Parknabinnia and other atypical court tombs in North Munster, western Ireland. In J. Müller & M. Hinz (eds.), *Megaliths, Societies, Landscapes - Early Monumentality and Social Differentiation in Neolithic Europe*.
- Keller, A. et al., 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nature Communications*, 3, p.698.
- Kılınc, G.M. et al., 2016. The Demographic Development of the First Farmers in Anatolia. *Current Biology*, 26(19), pp.2659–2666.
- Kirin, M. et al., 2010. Genomic runs of homozygosity record population history and consanguinity. *PLoS One*, 5(11), p.e13996.
- Kolb, M.J. et al., 1994. Monumentality and the Rise of Religious Authority in Precontact Hawai'i [and Comments and Reply]. *Current Anthropology*, 35(5), pp.521–547.
- Kytmanow, T., 2009. *Portal tombs in the landscape: the chronology, morphology and landscape setting of the portal tombs of Ireland, Wales and Cornwall*. UK: Doctoral Thesis, Queen's University Belfast.
- Lawson, D.J. et al., 2012. Inference of population structure using dense haplotype data. *PLoS Genetics*, 8(1), p.e1002453.
- Lazaridis, I. et al., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 513(7518), pp.409–413.
- Lazaridis, I. et al., 2016. Genomic insights into the origin of farming in the ancient Near East. *Nature*, 536(7617), pp.419–424.
- Lipatov, M. et al., 2015. Maximum Likelihood Estimation of Biological Relatedness from Low Coverage Sequencing Data. *bioRxiv*, p.023374.
- Lipson, M. et al., 2017. Parallel ancient genomic transects reveal complex population history of early European farmers. *bioRxiv*, p.114488.
- Lynch, A., 2014. *Poulnabrone: An early Neolithic portal tomb in Ireland*. Bray, Ireland: Wordwell.
- Mallick, S. et al., 2016. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature*, 538(7624), pp.201–206.
- Mallory, J.P., 2013. *The origins of the Irish*. London: Thames & Hudson.
- Martiniano, R. et al., 2017. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genetics*, 13(7), p.e1006852.
- Mathieson, I. et al., 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528(7583), pp.499–503.
- Mathieson, I. et al., 2017. The Genomic History Of Southeastern Europe. *bioRxiv*.
- McKenna, A. et al., 2010. The Genome Analysis Toolkit: a MapReduce framework for analysing next-generation DNA sequencing data. *Genome Research*, 20(9), pp.1297–1303.
- McLaughlin, T.R. et al., 2016. The Changing Face of Neolithic and Bronze Age Ireland: A Big Data Approach to the Settlement and Burial Records. *Journal of World Prehistory*, 29(2), pp.117–153.
- Midgley, M., 2008. *The Megaliths of Northern Europe*. London: Routledge.
- Müller, J., 2014. Monuments and Ideologies in the Neolithic Landscape. In J.F. Osborne (ed.), *Approaching Monumentality in Archaeology*. New York: SUNY Press, p.181.
- Murphy, E. et al., 2010. INSTAR—The People of Prehistoric Ireland: Phase 1. *Archaeology Ireland*, 24(1), pp.23–25.
- O'Kelly, M.J., 1982. *Newgrange: archaeology, art and legend*. London: Thames & Hudson.
- Olalde, I. et al., 2015. A Common Genetic Origin for Early Farmers from Mediterranean Cardial and Central European LBK Cultures. *Molecular Biology and Evolution*, 32(12), pp.3132–42.
- Olalde, I. et al., 2017. The Beaker phenomenon and the genomic transformation of northwest Europe. *bioRxiv*, p.28.
- Omrak, A. et al., 2016. Genomic Evidence Establishes Anatolia as the Source of the European Neolithic Gene Pool. *Current Biology*, 26(2), pp.270–275.

- O'Sullivan, A., 2001. *Foragers, farmers and fishers in a coastal landscape: an intertidal archaeological survey of the Shannon estuary*. Dublin: Royal Irish Academy.
- Patterson, N. et al., 2012. Ancient admixture in human history. *Genetics*, 192(3), pp.1065–1093.
- Patterson, N., Price, A.L. & Reich, D., 2006. Population structure and eigenanalysis. *PLoS Genetics*, 2(12), p.e190.
- Pemberton, T.J. et al., 2012. Genomic patterns of homozygosity in worldwide human populations. *American Journal of Human Genetics*, 91(2), pp.275–292.
- Prüfer, K. et al., 2014. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*, 505(7481), pp.43–49.
- Rasteiro, R. & Chikhi, L., 2013. Female and male perspectives on the neolithic transition in Europe: clues from ancient and modern genetic data. *PLoS One*, 8(4), p.e60944.
- Reich, D. et al., 2009. Reconstructing Indian population history. *Nature*, 461(7263), pp.489–494.
- Renfrew, C., 1973. *Monuments, mobilization and social organisation in Neolithic Wessex*. London: Duckworth.
- Rivollat, M. et al., 2016. Ancient mitochondrial DNA from the middle neolithic necropolis of Obernai extends the genetic influence of the LBK to west of the Rhine. *American Journal of Physical Anthropology*, 161(3), pp.522–529.
- Roksandic, M., 2012. Mobile and Terrestrial but Firmly Rooted on the River Banks: Biological Anthropology of Lepenski Vir and the Iron Gates Gorge Mesolithic. *AA: l'architecture d'aujourd'hui*, 02(03), pp.117–124.
- Schulting, R.J. et al., 2012. New dates from the north and a proposed chronology for Irish court tombs. *Proceedings of the Royal Irish Academy. Section C. Archaeology, Celtic studies, history, linguistics and literature*, 112C, pp.1–60.
- Sheridan, A., 2013. Early Neolithic Habitation Structures in Britain and Ireland: a Matter of Circumstance and Context. In D. Hofmann & J. Smyth (eds.), *Tracking the Neolithic House in Europe*. New York: Springer, pp. 283–300.
- Sheridan, A., 2004. Neolithic connections along and across the Irish Sea. In V. Cummings & C. Fowler (eds.), *The Neolithic of the Irish Sea*, pp.9–21.
- Sheridan, A., 2010. The Neolithisation of Britain and Ireland: the big picture. In B. Finlayson & G. Warren (eds.), *Landscapes in transition*. Oxford: Oxbow, pp. 89–105.
- Skoglund, P. et al., 2014. Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science*, 344(6185), pp.747–750.
- Szécsényi-Nagy, A. et al., 2015. Tracing the genetic origin of Europe's first farmers reveals insights into their social organisation. *Proceedings of the Royal Society B: Biological Sciences*, 282(1805).
- Thomas, J., 2013. *The Birth of Neolithic Britain: An Interpretive Account*. Oxford: Oxford University Press.
- Thorpe, I.J., 2003. *The Origins of Agriculture in Europe*. London: Routledge.
- Whitehouse, N.J. et al., 2014. Neolithic agriculture on the European western frontier: the boom and bust of early farming in Ireland. *Journal of Archaeological Science*, 51(0), pp.181–205.
- Whittle, A., 2007. The temporality of transformation: dating the early development of the southern British Neolithic In A. Whittle & V. Cummings (eds.), *Going Over: The Mesolithic-Neolithic Transition in North-West Europe*. Oxford: Oxford University Press.
- Whittle, A., Healy, F. & Bayliss, A., 2011. *Gathering Time: Dating the Early Neolithic Enclosures of Southern Britain and Ireland*, Oxford: Oxbow Books.
- Woodman, P.C., 1977. Problems of Identification of Mesolithic Survivals in Ireland. *Irish Archaeological Research Forum*, 4(2), pp.17–28.
- Woodman, P. & McCarthy, M., 2003. Contemplating some awful (ly interesting) vistas: importing cattle and red deer into prehistoric Ireland. In I. Armit et al. (eds.), *Neolithic settlement in Ireland and western Britain*. Oxford: Oxbow Books, pp.31–39.
- Woodman, P., 2015. *Ireland's First Settlers: Time and the Mesolithic*. Oxford: Oxbow.
- Zilhão, J., 2000. From the Mesolithic to the Neolithic in the Iberian peninsula. In T.D. Price (ed.), *Europe's First Farmers*. Cambridge: Cambridge University Press, pp. 144–182.
- Zilhão, J., 2001. Radiocarbon evidence for maritime pioneer colonisation at the origins of farming in west Mediterranean Europe. *Proceedings of the National Academy of Sciences of the United States of America*, 98(24), pp.14180–14185.

5. Bronze Age Beginnings

Signals of continuity across the Irish Metal Ages and the establishment of the Insular Atlantic Genome

Overview

Chapter Two demonstrated a mass influx of steppe-related ancestry into Ireland during the transition to metallurgy, with potentially reduced impact in the southwest of the island, after which point in time a large degree of haplotypic continuity is evident between a high coverage Early Bronze Age genome and the modern Irish population. These signals are more fully explored here using two populations sampled from the beginning (~2,500-1,600 BC) and end (~100 BC-600 AD) of the Irish Metal Ages. The introduction first discusses the archaeological and genetic evidence pertaining to the formation of the three great cultural complexes of the 3rd millennium BC - Yamnaya in the steppe, Corded Ware in the north and Bell Beaker in the west - which together set the foundations of European and Irish metallurgical societies, and by proxy the foundations of the modern Irish genome.

Using both haplotype and allele sharing methods, we show the transition to metallurgy in Ireland was complex, with a secondary influx of steppe-related ancestry from Britain likely during the widespread introduction of individualised inhumation burial to the north and east of the island, which roughly coincides with the start of the Early Bronze Age. Preceding Chalcolithic individuals are haplotypically distinct from these samples, and highly heterogeneous in terms of steppe-related introgression. Early Bronze Age samples from the southwest also show differing haplotypic affinities to those from the north and east. Most strikingly, individuals from megalithic sites in the Burren, dating to the end of the Early Bronze Age, possess substantially inflated levels of Neolithic ancestry, suggesting the persistence of megalithic culture in the western peripheries of Europe during the Beaker period was associated with some level of population continuity across the transition. Excesses of allele sharing with the Irish Neolithic population are seen for the majority of samples, suggesting some of this surviving Neolithic ancestry was local in origin. Signals of genetic continuity and change after this initial establishment of the Irish population are also explored, with haplotypic diversification evident between both the Bronze Age and Iron Age, and the Iron Age and present day. Across these intervals selection pressures related to nutrition appear to have acted, with variants involved in lactase persistence and skin depigmentation showing steady increases in frequency through time.

Introduction

After Agriculture: The first metallurgical civilisations of Europe

The Neolithic revolution can be viewed as the first domino in a chain of ever more complex technological and demographic developments, which are still impacting on human societies and genetics today (Harari 2014). Food surpluses allowed for the growth of densely populated settlements, within which division of labour and specialisation could occur, providing the skilled craftspeople needed for technological advance. To keep such large populations in continued co-operation, hierarchical ideologies were required, bringing with them centralised political and religious structures, as well as uniformity of culture. Sedentism also put an increased emphasis on property and ownership, and allowed for the accumulation of wealth among emerging social elites. Art and architecture flourished, alongside invention and discovery, as increased impetus was put on craftspeople to develop objects of worth. This new culture of materialism heavily encouraged trade, not only of required resources, but also of rare and luxury goods. Vast regions of Eurasia, such as the Atlantic megalithic zone, were gradually subsumed into highly connected exchange networks, which not only aided the quick dispersal of goods and people, but also new technical knowledge and ideologies.

The pace at which fledgling societies advanced along this post-agricultural trajectory varied from region to region, dependent on a myriad of environmental and social factors. The stop-and-start nature by which the Neolithic radiated out from its points of origin added to these discrepancies, with city states blossoming in West Asia at a time when the northern and western extremities of Europe, including Ireland, were still populated by hunter-gatherers (HGs). It is then no surprise that the earliest European civilisations were found in the Balkans, the home of the continent's first farmers. These societies were also some of the first in the world to make the technological leap to metallurgy, with evidence of copper working and mining in Serbia and Bulgaria *circa* 5,500-5,000 BC (Cunliffe 2008; O'Brien 2014). This is proposed to be a largely indigenous development, although some argue that the technical knowledge was introduced to the region from Anatolia and the Fertile Crescent (Roberts et al. 2009). However, it was in the Balkans where production reached industrial scales. The huge quantities of copper objects produced, originally decorative items but later tools and weaponry, were subsequently distributed over vast territories, extending towards the foothills of the Alps in the west and eastwards over the northern shores of the Black Sea. Such extensive contacts also brought with them an increased level of interdependency and the entire region during the period is routinely referred to as the "Carpatho-Balkan metallurgical province".

Several of the major cultures of this province have been the subject of ancient genomic analyses, which demonstrate continuity between the Anatolian-derived Neolithic and later Chalcolithic of both Hungary and the Balkans (Gamba et al. 2014; Lipson et al. 2017; Mathieson et al. 2017), albeit with some resurgence of indigenous European HG ancestry through time (discussed in Chapter

Four). These include the Varna culture of Bulgaria (4,400-4,100 BC), known for its early and elaborate usage of gold metallurgy, emphasising deepening societal inequalities at the time; and the Cucuteni-Trypillian culture (5,200-3,500 BC), centring on modern-day Moldova, which expanded north-eastwards into the Dniester and Southern Bug river valleys of the Ukraine, forming densely populated settlements in the region (Anthony 2007). These two rivers formed a major cultural and genetic boundary between the Anatolian-derived farming populations of Europe and steppe groups further east, which would remain intact for roughly two millennia.

Herding and Hierarchy: The transition to pastoralism on the steppe

The postglacial hunter-gatherer groups of the Pontic-Caspian steppe had begun to exhibit some aspects of Neolithic culture long before contact with more southern farmers. Specifically, ceramic technology appears to have arisen independently in these populations, emerging in the Samara region of the Volga River valley between 7,000-6,500 BC (Vasilieva 2011). Agriculture and animal husbandry were first delivered to the western fringes of the steppe by Early European farmers of the Criş culture, who moved eastwards from the Carpathian mountains *circa* 5,800-5,500 BC (Cunliffe 2015). However, Criş farmers never extended past the Dniester river valley, which was already densely occupied by foraging populations (Bug-Dniester culture). Through prolonged contact, these foragers gradually adopted some elements of herding and cultivation from Criş groups, although their culture remained distinct. The later penetrance of the aforementioned Cucuteni-Trypilian complex across the Dniester into the Bug Valley appears to have been a watershed moment for steppe populations (Anthony 2007). Bug-Dniester culture disappeared completely, while forager groups further east in the Dnieper valley began to transition to animal herding *circa* 5,200 BC, although cereal cultivation was not adopted. From there, cattle, sheep and goats spread quickly across the steppe, reaching the Volga-Ural region between 4,700-4,600 BC.

Animal husbandry triggered a radical transformation of societal structure in the steppe, with increasing cultural divergence apparent between these populations and those of the temperate forest zone further north, who continued to maintain a hunting and gathering lifestyle for at least another several millennia (Anthony 2007). Foraging cultures also persisted on the Kazakh steppe, with the Ural River valley, which had been submerged under postglacial meltwater between 16-11 kya, forming a major cultural and likely genetic boundary. The adoption of pastoralism on the Pontic-Caspian steppe coincides with the widespread appearance of Balkan copper in the region, heightened social stratification in the form of elaborate male burials, new feasting and sacrifice rituals, as well as apparent population growth indicated by the increase in cemeteries. Despite clear rise in trade and interaction between the societies of the Carpatho-Balkan metallurgical province and the pastoralist groups of the steppe, they remained on divergent cultural trajectories. The highly productive agricultural economies of the Balkans and Danube valley supported large numbers of densely populated settlements and the technical and aesthetic output of these civilisations dwarfed any comparable activity seen on the steppe (Anthony 2007).

This divide is also reflected genomically. Substantial continuity is clear across the Mesolithic to Neolithic transition in the Ukraine (Jones et al. 2017; Mathieson et al. 2017), with individuals from both periods possessing majority EHG ancestry with some minor WHG input (See Fig. 1.3 for visualisation). Notably, in some Ukrainian Neolithic samples the proportion of WHG is seen to be relatively increased (Mathieson et al. 2017), perhaps the result of Mesolithic European populations being pushed eastwards by incoming farmers. Importantly, no strong evidence of Anatolian-related ancestry is seen in the Ukrainian Neolithic populations so far sampled (Jones et al. 2017; Mathieson et al. 2017), including an individual from the Dnieper-Donets culture directly dated to 4,469-4,293 cal BC (Jones et al. 2017), despite their clear participation in reciprocal exchange networks with European Neolithic and metallurgical societies further west. That said, steppe-type introgression has been noted in several European Chalcolithic individuals from the Varna and Cucuteni-Trypillian cultures (Mathieson et al. 2017), indicating that the Southern Bug boundary was somewhat porous.

The Caucasus Connection: Genetic discontinuity in steppe populations

Strikingly, the main genetic influx into the Pontic-Caspian steppe over the course of the Chalcolithic did not come from the southwest, but from the southeast. This is indicated by the large increase in Caucasus/Iranian-type HG ancestry seen in later populations of the region (E. R. Jones et al. 2015), which would eventually reverberate through the European continent (See Figs. 1.2 and 1.3 for visualisation). The earliest signals of Caucasus-type introgression so far identified in the steppe come from the Samara area of the Volga river. Here, discontinuity is seen between a Mesolithic EHG (5,657-5,541 cal BC) and Chalcolithic samples (*circa* 5,200-4,000 BC) (Haak et al. 2015). Later pastoralist groups, most notably the Early Bronze Age Yamnaya culture (*circa* 3,500-2,000 BC; Fig. 4.1A), exhibit even larger amounts of Caucasus-type ancestry, suggestive of population movement either from or across the isthmus northwards into the steppe (Allentoft et al. 2015; E. R. Jones et al. 2015).

The North Caucasian piedmont was occupied by pastoralist populations by 5,000 BC. These were culturally, and likely also ethnically, distinct from steppe groups further north, though contacts between the regions are apparent (Anthony 2007). Indeed, they most probably received domesticated animals and copper through trade with steppe groups, as opposed to West Asian societies south of the somewhat impervious North Caucasus Mountain range. The first significant links to West Asia are seen later with the Maikop culture (3,700-3,000 BC), which introduced the earliest kurgan (barrow) burial mounds to the region, alongside new metals such as arsenic bronze, potter's wheels and Mesopotamian symbolism, possibly via the Black Sea coast. While the precise origins of Maikop culture are debated, a homeland somewhere within the Caucasian isthmus seems likely (Anthony 2007), and the culture has been identified as a prime candidate vector for the spread of CHG ancestry into the steppe (E. R. Jones et al. 2015).

Steppe populations embraced many Maikop innovations, though the groups remained culturally distinguishable. The Maikop period also coincides with the first substantial evidence of horse

domestication and riding on the steppe, which together with wagons, possibly Maikop in origin, paved the way for a highly mobile form of pastoralism (Anthony 2007; Chernykh 2008). The aforementioned Yamnaya culture, whose origins lie just north of Maikop territories on the lower Volga and Don River valleys, expertly exploited the new niche, and succeeded in rapidly spreading out across the Pontic-Caspian steppe, creating the first unified and fully pastoralist culture of this vast region. Expanding Yamnaya groups brought with them individualised kurgan burials, further social stratification and wealth inequality, improved metal-working technology, CHG ancestry, and likely a form of Indo-European language (Anthony 2007; Allentoft et al. 2015; Haak et al. 2015). A testament to Yamnaya mobility is seen in the long-distance trans-continental migration to the Altai mountains undertaken by some eastern groups. This movement gave rise to Afanasievo culture, which shows striking genetic homogeneity with Yamnaya individuals despite the thousands of kilometres between the two populations (Allentoft et al. 2015).

Stepping into the West: Demographic upheaval in Europe during the third millennium BC

Yamnaya and related steppe populations also expanded west, breaking through the Cucuteni-Trypillian border and migrating *en masse* southwards along the Black Sea and into the Carpathian Basin (*circa* 3,100-2,800 BC; Fig. 4.1A). There they settled in pockets along the Danubian corridor and the steppe land of the Great Hungarian Plain, erecting large numbers of distinctive kurgan mounds in the process (Anthony 2007; Harrison & Heyd 2007; Cunliffe 2008). The majority of the cultural-economic systems of the Carpatho-Balkan metallurgical province had by this time disintegrated (Chernykh 1992), most likely due to unprecedentedly cold climatic conditions between 4,200-3,800 BC, though other factors, such as conflict with invasive steppe groups or unsustainable agricultural and mining may have hastened their demise. The larger province was replaced by simpler, more regional culture complexes, which went on to catalyse the region's transition to the Bronze Age. Notably, genetic discontinuity is observed between the Chalcolithic and Bronze Age of both the Balkans and Hungary, with a clear increase in steppe ancestry observed, defined here as a mixture of EHG and CHG components (Gamba et al. 2014; Mathieson et al. 2017). In the Balkans this ancestry is seen to increase over the course of the Bronze Age (*circa* 3,400-1,000 BC), averaging at approximately one third across sampled individuals (Mathieson et al. 2017). Similar levels of steppe introgression are seen in the Hungarian Bronze Age, although high variability is apparent between individuals (See Fig. 1.3D), highlighting the region's position as a cultural crossroads.

Most strikingly, during the same approximate period, northern Europe underwent an even more dramatic genomic transformation, closely associated with the Single Grave or Corded Ware (CW) horizon (Fig. 4.1A), which emerged *circa* 2,900 BC and spread to cover vast swathes of territory from the Rhine Valley to Russia (Cunliffe 2008). The CW populations so far surveyed have ancestry predominantly derived from the steppe (75%) (Haak et al. 2015) and the culture exhibits a large number of steppe-derived traits, such as single burials beneath barrows and a highly mobile pastoralist economy, though elements of preceding TRB Neolithic groups are apparent in their increased emphasis on

ceramics and in the persistence of some agricultural practices. A minor portion of their ancestry is clearly derived from the Neolithic farming groups of Europe, reflecting this cultural hybridisation (Allentoft et al. 2015, Haak et al. 2015; See Fig. 1.2C for visualisation). Introgression from indigenous farming populations is suggested to have been largely mediated by women, whereby migrant steppe groups, dominated by men, would adopt local wives (Kristiansen et al. 2017). This scenario is supported by isotopic analysis of CW cemeteries, which demonstrate discontinuity between childhood and adulthood diet for female burials (Sjögren et al. 2016), and aDNA studies, which indicate a near complete replacement of common Neolithic Y chromosome lineages by eastern R1a and R1b haplogroups (Allentoft et al. 2015; Haak et al. 2015). These events have left a profound impact on modern European Y chromosomal diversity, which is also dominated by R1a and R1b subclades (Underhill et al. 2010; Myres et al. 2011). Notably, the two lineages show an east-west dichotomy, with R1a found at highest frequencies in modern-day Poland and Russia, and R1b peaking in the Irish and Basque populations.

Bayesian analysis of radiocarbon determinations from CW sites suggests the earliest occurrence of this culture was centered in central Poland and the Baltic region, from where it radiated out north, south, east and west (Wencel 2015). However, despite these new chronologies, and the clearly massive contribution of steppe groups to CW Culture, a definite model for its demographic formation remains elusive (Heyd 2017). This is in part due to the near-simultaneous emergence of CW Culture across relatively large regions of Europe (Fig. 4.1A), likely due to its mobility, and the sparse temporal and spatial genomic sampling that has so far been undertaken across the transition. Indeed, while at present Yamnaya samples act as the best proxy for the steppe ancestry present in CW individuals, it has been argued that this culture is unlikely to be the direct progenitor of CW (*ibid.*), given their overlapping chronologies and non-overlapping occupation of distinct ecozones (steppe vs. temperate forest).

The areas occupied by CW groups were also home to a number of other wide-reaching cultural complexes, which both preceded and overlapped with the CW horizon. This includes the Globular Amphora and TRB cultures (See Chapter Four), both of which derive their ancestry predominantly from Anatolian and WHG sources (Skoglund et al. 2014; Haak et al. 2015; Lipson et al. 2017; Mathieson et al. 2017), and the Pit-Comb culture, which represents an earlier intrusion of EHG ancestry into the Baltic region (Jones et al. 2017; Mathieson et al. 2017). Across these vast territories a large number of regional variants of CW Culture emerged, some influenced by neighbouring cultures or the available environmental niches. These included the Battle-Axe group in Scandinavia and the Protruding Foot Beaker Culture of the Lower Rhine, the latter forming the western boundary of CW Culture. Importantly, the Rhine valley provided key access to the extensive trade and exchange networks of Atlantic Europe.

In many ways CW Culture mirrored the Atlantic megalithic zone which it faced, representing a common belief and social value system shared across a vast region (Cunliffe 2008). However, unlike megalith builders, CW groups appear to have put strong focus on the individual and small kin groups, evidenced by individualised burials within single graves or cists. These were typically accompanied by grave goods,

including eponymous cord-decorated beakers, which have been linked to drinking culture, and the aforementioned stone axes. Differentiation between male and female funerary rites is clear, with male burials dominating, indicative of distinct gender roles within the society. In contrast, the collective megalithic burials of Atlantic Europe clearly emphasize the ancestral group shared by a larger community. However, over the course of the third millennium BC, burial rites similar to CW culture began to emerge in western Europe, associated with the spread of new social identities, metallurgy and another distinctive ceramic type, the Bell Beaker (Fig. 4.1A).

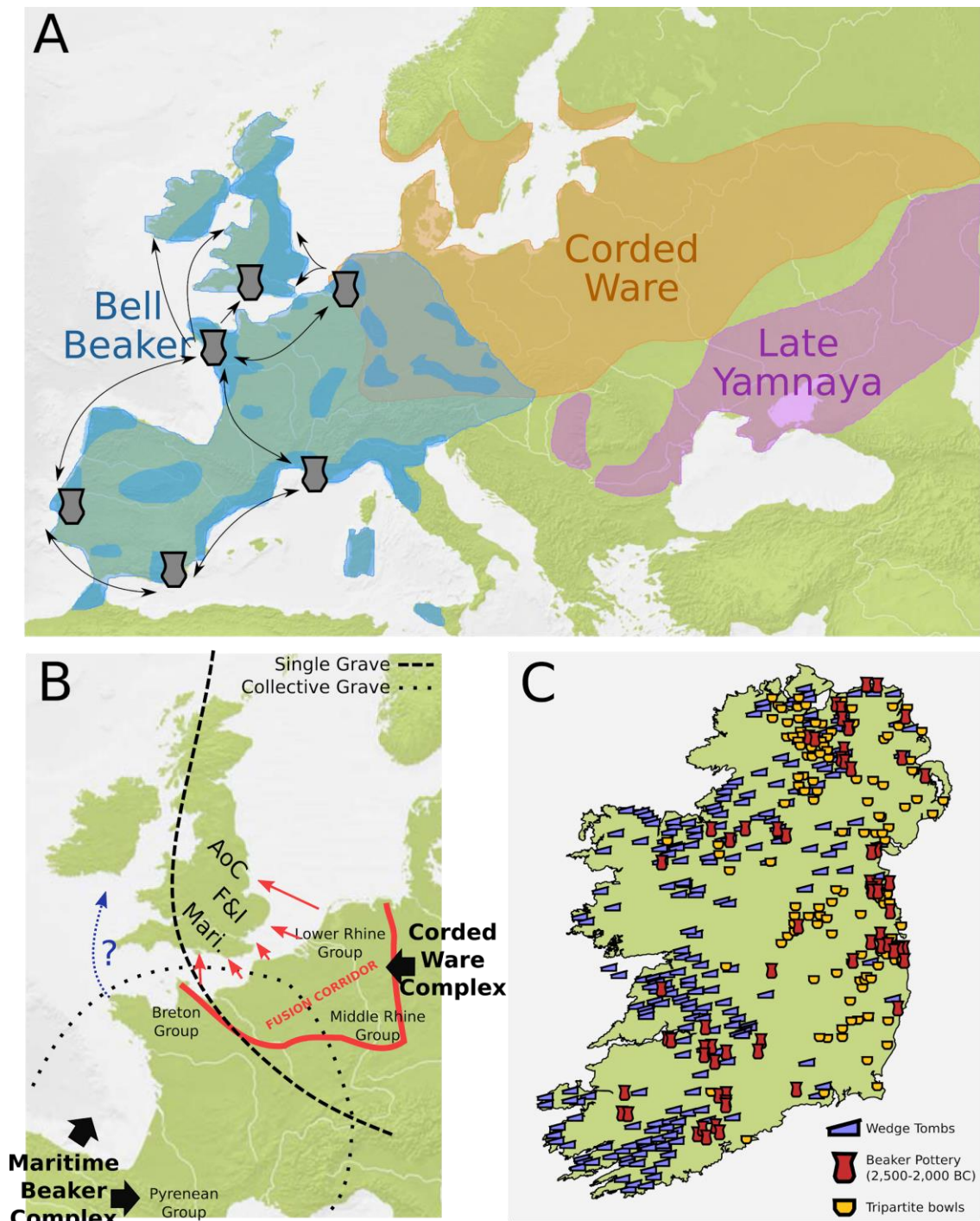


Figure 4.1 Cultural upheavals of the third millennium BC in Europe. A) Distributions of Yamnaya, Corded Ware and Bell Beaker cultural phenomena. Maritime Bell Beaker networks are also shown. Key hubs are marked with symbols. **B)** Spheres of cultural influence on the eve of Britain and Ireland's 'Beakerisation'. Regional early beaker groups are marked alongside the immediate source zone (Fusion Corridor) for British beakers (Mari. - Maritime-Derived; F&I - Fingernail and Incised; AoC - All-Over-Cord). Prevalent burial practices in Britain and the continent are also shown. **C)** Distribution of beaker and bowl pottery in Ireland, alongside megalithic wedge tombs.

Cultural distributions sourced from Cunliffe 2013 (A & B), Fitzpatrick 2013b (B), Flanagan 1998 (C) and O'Brien 2012 (C). Distributions in C are based in part on earlier surveys from (Ó Nualláin 1989; Waddell & Ó Ríordáin 1993). (Map Credit: freeworldmaps.net)

The Bell Beaker Phenomenon: The expansion of steppe ancestry into Atlantic Europe

Copper metallurgy had arrived in Atlantic Iberia by the end of the fourth millennium BC, likely as a result of east-west transmission along the coastlines and southern river basins of the Mediterranean (Ottaway & Roberts 2008), although some argue for independent invention (Ruiz-Taboada & Montero-Ruiz 1999). Regardless of the exact origins of the Iberian Chalcolithic, it is clear that the period involved extensive and long-distance trade, with networks stretching from the Baltic to North Africa (Murillo-Barroso & Martín-Torres 2012; Cunliffe 2013). A new degree of social complexity and hierarchy also accompanied the transition, seen most clearly in the urban developments which arose in primarily in the south and southwest of the Peninsula. Strong genetic continuity is evident across the Iberian Neolithic to Chalcolithic periods (Günther et al. 2015; Mathieson et al. 2015), and burials in the west remained primarily megalithic in nature.

Against this backdrop, the Maritime Bell Beaker, which served both domestic and funerary purposes, appears to have emerged from indigenous ceramic styles in the Tagus region of Portugal (2,800-2,700 BC, although its exact origin(s) are still under debate and some inspiration may have been indirectly drawn from both Corded Ware beakers in the north and African styles in the south through trade (Cunliffe 2008; Czebreszuk 2014; Jeunesse 2015). In the succeeding centuries, Maritime Beakers rapidly spread across Western Europe along the Atlantic and Mediterranean seaways and inland river corridors (Fig. 4.1A), forming part of a broader 'Beaker package', which included metal and flint working technologies and various artifacts, such as copper daggers, archery equipment and v-perforated buttons. The phenomenon is thought to represent a new ideology, emphasising individual knowledge, power and status. As well as its origins, the mode of spread of the Beaker package also remains contentious. Diffusion of ideas and of people along well-established networks have both been invoked to explain the ubiquity of Bell Beaker material in western Europe and are not by any means mutually exclusive (Cunliffe 2008). When migration is considered, theories range from small bands of mobile elites or craftspeople to large-scale folk migration. Notably, Maritime Beakers are seen to cluster in regions with an abundance of copper ore, which may have driven such movements (Cunliffe 2013). The debate is further confounded by the plethora of regional varieties underlying the more unifying aspects of the package, which casts doubt on the definition of the Bell Beaker phenomenon as a single archaeological culture.

In northern and central Europe the Beaker package overlapped substantially with CW territory (Fig. 4.1), replacing it in some regions, and variable cultural interactions can be seen between the two groups across the interface (Cunliffe 2008). In the core contact zone, centering on the Rhine Valley, a distinctive vessel type emerged: the All-over Ornamented Beaker, which typically made use of cord-impressions (All-over-Cord) (Cunliffe 2013; Fitzpatrick 2013b). This Beaker type was commonly placed in single graves with flint objects sourced from western France, highlighting the extensive overland connectivity between the regions via the Loire valley and Paris Basin. These networks helped mediate the spread of single grave burial customs beyond the boundaries of CW culture, into the Low Countries and northern France (Fig. 4.1B). It is in this area, and further east in Central Europe, that the 'classic' type of Beaker burial is

typically found, which shows many parallels to steppe and CW practices. Typically, these burials consisted of a crouched individual inhumation, placed within a cist or grave alongside a characteristic set of grave-goods (vessels, daggers and wristbands), with a barrow constructed overhead. Gender differentiation is observed in body orientation and the associated artefacts, with female burials exhibiting lower levels of wealth and no weaponry. It is likely such burials served only a small fraction of the population and social structure is clearly visible (Salanova 2016).

These funerary rites stand in sharp contrast to those observed in Atlantic coastal regions, such as Portugal and Brittany (Fig. 4.1B), where the Neolithic tradition of collective burial in megaliths and artificial caves persists across the Beaker period (Salanova 2003; Salanova 2004). Moreover, while beaker vessels are common in these regions, other elements of the Beaker grave-good package are rare, warranting their consideration as a potentially distinct cultural sphere. In northern France this dichotomy is clearly visible in the concentration of single grave burials eastwards of the Seine, with the Armorican Peninsula remaining primarily megalithic, although deposits within these collective tombs can be individualised (Salanova 2016). No singular Beaker burial has yet been identified in southern Portugal (Salanova 2004), while in southern France 87% of Beaker burials are found in collective contexts (Guilaine 2001).

Recent aDNA studies have contributed to, rather than reduced, the lively debate surrounding Bell Beaker origins, demonstrating clear signals of migration in some regions, and continuity in others (Allentoft et al. 2015; Haak et al. 2015; Olalde et al. 2017). In Germany, Hungary and the Czech Republic, individuals from Bell Beaker sites derive a substantial proportion of their ancestry from steppe-related populations (~50%), although they possess higher levels of European Neolithic introgression relative to German CW groups and heterogeneity in ancestry is apparent (Allentoft et al. 2015; Haak et al. 2015). Moreover, isotopic analyses have revealed high mobility of Bell Beaker individuals in these regions (Price et al. 2004), suggesting this interaction zone was a genetic, as well as cultural, melting pot. It is still undecided whether CW culture represents the primary vector of steppe ancestry into Beaker groups, with denser sampling of Central Europe required.

In contrast, Iberian individuals from Beaker sites exhibit little to no steppe-related introgression (Olalde et al. 2017), resonating with the continued use of Neolithic settlement and burial sites in the Peninsula. No Beaker sites from the French coast have yet been sampled, and so it is unclear whether this survival of Neolithic ancestry is part of a wider Atlantic phenomenon, where collective burial remained common. However, French Bell Beaker individuals from the Rhône and Rhine Valleys further east show levels of steppe-related ancestry similar to that seen in central Europe, although again heterogeneity in ancestry is apparent (Olalde et al. 2017). Despite this variability in steppe introgression seen in Beaker populations, they are somewhat unified by the predominance of the Y chromosome haplogroup R1b-P312, which is common in individuals across central and western Europe, although it has not yet been observed in Iberia until the later Bronze Age period (Martiniano et al. 2017; Olalde et al. 2017)

Facing West, Facing East: Beaker culture in Ireland and Britain

As with the preceding introduction of farming, the French and Benelux coasts are of key importance when considering the spread of the Beaker package to Ireland and Britain (Cunliffe 2013). During this period both islands were quickly subsumed into highly mobile maritime networks, reaching new heights of connectivity with the continent unseen in Neolithic. Just as the islands showed differential continental affinities in their uptake of Neolithic burial practices, so too do they sharply diverge in their adoption of the Bell Beaker phenomenon, with Ireland showing stronger affinity to Atlantic communities, relative to the neighbouring island of Britain, which derived much of its Beaker culture from the Rhineland region. Such disparities have been argued to represent different continental sources of Irish and British Beaker traditions, with adoption of the more northern elements of the package only occurring later in Ireland via Britain (Burgess 1979; Case 1995), although continued debate surrounds these theories (O'Brien 2012; Mallory 2013; C. Jones et al. 2015), which form a main subject of inquiry in the current chapter.

In Britain, the wholesale adoption of the classic northern Beaker package occurs rapidly between 2,500 and 2,250 BC. Sites are concentrated along the eastern seaboard, the Thames Valley and Wessex, all areas easily accessible from the North Sea and Eastern Channel (Cunliffe 2013), with more muted infiltration of the Beaker package seen in the west of the island. A 'fusion corridor' of Atlantic and eastern influences stretching from Northern France to the Low Countries is believed to have been the main source of these traditions (Fig. 4.1B), with little contribution thought to have occurred from Brittany (Fitzpatrick 2013b). Regional British Beaker designs reflect the continental styles to which they are most geographically proximate, with Maritime-derived Beakers common in central southern Britain and All-over-Cord Beakers typically found further northeast (Cunliffe 2013). The infiltration of the Beaker package has been argued to represent population movement into the island during the period and is supported by isotopic identification of recent migrants in Beaker burial contexts, such as the famous Amesbury Archer (of supposed Alpine stock (Fitzpatrick 2013a)). A recent aDNA survey of English Beaker burials has confirmed these suspicions, with substantial steppe introgression apparent in the east of the country (Olalde et al. 2017). Strikingly, this is the largest amount of steppe-related ancestry seen in any European Bell Beaker population so far surveyed, with a level most similar to a later Bronze Age population from the Netherlands. However, some resurgence of Neolithic ancestry is seen by the island's Bronze Age period (Olalde et al. 2017).

The Beaker period in Ireland commenced at approximately the same time as in Britain, *circa* 2,450 BC, with exclusive use of Beaker ceramics seen until 2,150 BC (Brindley 2007; O'Brien 2012), completely replacing previous Neolithic styles (grooved ware) and showing both Atlantic and eastern affinities (Case 1995). Other Beaker artefacts, such as flint arrowheads, v-perforated buttons and wristbands, also make an appearance at this horizon. However, in Ireland, not a single classic northern Beaker burial is known. Instead, Irish Beaker material is mostly found within settlement contexts or megalithic tombs (Carlin 2011), with substantial reuse of older megalithic structures apparent (Carlin 2011; Schulting et al. 2012; Bayliss & O'Sullivan 2013; McLaughlin et al. 2016). A small number of simpler pit and cist single graves

with Beaker pottery have also been identified, but vast majority of these contain burned rather than inhumed remains (Mount 2011), following the typical Neolithic tradition of cremation.

The clearest break with previous Irish funerary practices is seen with the emergence of a new megalithic burial structure, the wedge tomb, the most common type of megalithic monument known on the island (McLaughlin et al. 2016). These show a primary phase of construction during the Beaker period and continue in usage well into the Early Bronze Age, centuries after the decline of the Beaker pottery tradition (Schulting et al. 2008; O'Brien 2012). Wedge tombs are generally restricted to the western half of the island (Fig. 4.1C), with particular concentrations seen in the Burren, Cork and Sligo (O'Brien 2012; C. Jones et al. 2015), and are thought to represent small segmentary societies (O'Brien 1999; O'Brien 2000). Individuals from two such tombs at Labbacallee, Co. Cork (Leask et al. 1935) and Roughan Hill in the Burren (Ó Maoldúin 2015) are sampled here. While wedge tombs have been argued to represent a largely indigenous development (Waddell 2010), the influx of steppe-related ancestry in the Irish Beaker period, demonstrated in Chapter Two, suggests migration occurred into the island during the time of their initial construction. The most cited potential external inspiration for wedge tomb architecture are the *allées couvertes* tombs of Brittany, a region which shows persistent contact with Ireland from the very beginning of the island's megalithic tradition (O'Brien 2012; C. Jones et al. 2015).

Brittany is also the likely source for what was arguably the most transformative innovation of the Irish Beaker period, copper metallurgy (O'Brien 2012; Cunliffe 2013). The introduction of this fully established technology, which had previously spread from Iberia through maritime networks, would have required the migration of skilled craftspeople, possibly drawn to the island by its abundance of copper ore. Highly productive mines were established in the southwest at Ross Island, the earliest seen in northwestern Europe, and these served to produce the majority of arsenical copper items in circulation in Ireland, with large quantities also exported to Britain. Later copper mines were established in Cornwall and Wales, while gold deposits in Wicklow mountains were also soon exploited, and served as the most likely primary source of the metal for both islands. These new technologies brought an unprecedented level of craftsmanship to Irish societies, and the island soon became a key hub in Atlantic exchange networks. As Ireland transformed into a source of innovation, rather than a sink, new metallurgical technologies were experimented with, resulting in the creation of novel bronze alloys, based on tin sourced predominantly from Cornwall. The transition to full bronze usage commenced in Britain and Ireland between 2,200-2,000 BC, predating other regions of western and central Europe by at least two centuries (Cunliffe 2013). Indeed, in Scandinavia the Late Neolithic period continued until as late as 1,700 BC, after which time gold, copper and tin entered the region in large quantities.

Early Bronze Age Consolidation: Convergence of Irish and British burial traditions

The Early Bronze Age transition is marked by a number of cultural discontinuities in Ireland, most strikingly the widespread uptake of single burial practices in the east and north of the island from 2,150 BC onwards, with wedge tomb usage persisting in western regions (Flanagan 1998; O'Brien 2012;

Mallory 2013). These new traditions have strong parallels with the classic Beaker burials seen across the sea in Britain, from where they were possibly introduced. Typically they consist of a cist or pit grave, within which the remains of likely high status individuals, usually in the form of a crouched inhumation, were placed with grave-goods, including v-bored buttons and the earliest examples of tin-bronze in Ireland in the form of small daggers (O'Brien 2012). Ceremonial ceramic vessels were also placed within these graves. However, somewhat non-intuitively these were not beakers, but instead a new type of bowl food vessel (Fig. 4.1C), which remained in use until 1,920 BC (Brindley 2007). Beaker pottery is still seen in domestic and megalithic contexts in Ireland, but was rejected entirely from this new funerary custom, despite its clear Beaker influences. These Irish developments occurred in parallel with a more widespread adoption of Beaker burial rites seen across Britain in the same period (2,250-1,950 BC), as well as a clear diversification in both British material culture and ceramic styles, including the emergence of food vessels (Needham 1996; Cunliffe 2013). In particular, influence from the Low Countries is apparent with the appearance of stone battle-axes and new beaker pot designs.

Further societal change is apparent in both islands in the succeeding centuries. In Ireland, bowl pottery is supplanted by a new form of vase food vessel between 2,000-1,920 BC, which served similar funerary purposes, although with a greatly increased emphasis on cremation (Brindley 2007). This period also saw a more obvious uptake in tin-bronze technology. Unlike the bowl tradition, vase pottery is seen to successfully penetrate some western regions of the island, although highest concentrations are still seen in the north and east, with the southwest remaining relatively untouched (Flanagan 1998). Between 1,900-1,800 BC cremation became the near-universal burial rite in individualised graves and container food vessels were gradually replaced by urn pottery, including a clearly invasive tradition from Britain, the collared urn (again confined to the east), which appears alongside other British-type artefacts, such as the battle axe. This time marks the beginning of substantial cultural convergence between the islands after several centuries of parallel evolution (Brindley 2007).

This same period sees the final disappearance of Bell Beaker culture across the European continent, where it is replaced by a variety of regional Bronze Ages. In Central Europe, the widespread Únětice culture (2,300-1,700 BC) emerged, which was built on far-reaching trade networks that included Britain and Ireland (Kristiansen & Larsson 2005). Únětice individuals show similar genetic affinities to the Bell Beaker culture, with some suggested resurgence of European HG ancestry (Haak et al. 2015). In some Atlantic regions, such as Brittany and Iberia, the Bronze Age marks the appearance of wealthy individual graves and the abandonment of megaliths. In Portugal, single cist burials appear *circa* 1,900 BC, typically accompanied by Bronze daggers and other warrior equipment (Senna Martínez 2013). These individuals have been demonstrated to possess some small amount of steppe-related introgression relative to preceding Chalcolithic populations and the first identification of R1b-P312 Y chromosomes on the Peninsula is seen at this horizon, suggested to have arrived through predominantly male-mediated migrations (Martiniano et al. 2017).

The majority of Irish Early Bronze Age individuals analysed in this chapter are sourced from singular burials, typically with associated food vessels. This includes four samples from cist burials in the west of the island, associated with vase pottery (Annaghkeen84, Moyveela381, Stonepark3409 and Treanmacmurtagh116), as well as a rare southwestern cist burial from the Burren (Coolnatullagh1005), dating to the end of the Early Bronze Age. Of the remaining single grave samples, all but one (Grange10) come from the east coast and are typically associated with bowl ceramics. The ubiquity of cremation from 1,800 BC onwards severely restricts any attempted genomic survey of the succeeding Middle (1,600-1,200 BC) and Late (1,200-600 BC) Bronze Age periods (McGarry 2010). Only one potential inhumation is known, at Cherrywood, Co. Dublin, indirectly dated to 1,400-1,000 BC and is sampled here. A few unburnt skull deposits from wetland contexts have also been identified as Late Bronze Age in date, one of which from the southwest has been included in this study (Inchagreenoge134), with a direct date of 1,270-1,040 cal BC.

On the Cusp of Written History: Ideological changes in Late Iron Age Ireland

Very few ancient genomes from the later Bronze and Iron Age cultures of Europe have been sampled thus far, and, given the lack of unburnt Irish remains from the same periods available for analysis in this thesis, these are not reviewed here. However, some discussion of the Late Atlantic Bronze Age and Early Iron Age in Britain and Ireland is given in the introductory section of Chapter Two, alongside a thorough summary of the debate surrounding Celtic language origins. The majority of the Late Iron Age population presented here dates to the very end of Irish prehistory, *circa* 400 AD, at a time when inhumation rites became common on the island. The ubiquity of cremation on the island for nearly two millennia prior to this suggests a strong social taboo had surrounded unburnt burial, and the sudden emergence of extended and supine inhumations is argued to represent a drastic shift in belief systems, specifically the introduction of Christianity (McGarry 2010).

Only a handful of Iron Age inhumation burials can be securely dated to before 400 AD. The earliest of these is a crouched burial associated with the Knowth passage tomb, dating to 195-50 cal BC and sampled here (Knowth10) alongside a slightly later individual from the same context (Knowth4; 86-252 AD). The Knowth burials form part of a small group of crouched inhumations, centered around the Boyne Valley and eastern coast, which date prior to 250 AD. It is suggested that the group represents migrant communities from Britain, where inhumation was the main form of burial from 700 BC onwards, with crouched positions clearly predominating after 400 BC, until the continental rite of cremation was introduced in the first century BC. It has been argued that the appearance of continental funerary rites in Britain and the broadly contemporaneous introduction of British rites to Ireland may be the result of population displacements (McGarry 2010), and it is notable that the Roman invasions of Gaul (121 BC) and Britain (55 BC) were initiated during the same period.

The later practice of extended inhumation was also most probably introduced from Britain (O'Brien 2003), with clear communication and bidirectional movement between the islands during and after the

Roman period. This likely occurred during the fourth century AD, although several anomalous radiocarbon dates exist. Two sites with earlier than expected date ranges, Ballyglass Middle (80-420 AD) and Claristown (60-420 AD) are sampled here, although the latter may be the result of a high error margin in the absolute radiocarbon determination. Notably, the body positionings of the burials at Ballyglass Middle are unknown, while the Claristown individual was extended (See Appendix One for further detail).

The Early Christian period in Ireland occurred during a turbulent time in European history, when mass migrations of peoples were occurring on the continent following the collapse of the Roman Empire, including Germanic, Slavic and Hunnic groups (Halsall 2007). In Britain, the withdrawal of Roman rule was quickly superseded by the Anglo-Saxon invasions, a migratory event which has been confirmed through genomic surveys of British Iron Age, Roman and Anglo-Saxon burials (Martiniano et al. 2016; Schiffels et al. 2016). There is also indications of an Irish invasion of western Scotland during the same period, which introduced Gaelic languages via the maritime Dál Riata over-kingdom and gradually subsumed Pictish cultures further north and east (Dillon & Chadwick 1967). This dynamic demographic backdrop, as well as the social upheavals that likely accompanied the introduction of a new belief system to Ireland, must be considered in the interpretation of Iron Age genomic dataset present here, and individuals sampled may not fully reflect the population structure that had existed on the island prior to these events.

The demographic effects of historical and proto-historical migrations on the modern populations of Ireland and Britain have recently been explored using novel haplotypic methods (Leslie et al. 2015, Byrne et al. submitted), revealing the impact of both continental movements into the islands (Anglo-Saxons, Vikings, Normans) and movements between the islands, which have left particularly strong signatures in the east and north of Ireland, as well as in Scotland. Furthermore, Byrne et al. (submitted) have demonstrated that haplotypic clustering methods loosely separate the modern Irish population on the basis of the historical provinces of Ireland (Ulster in the north, Leinster in the east, Munster in the south and Connacht in the west), whose boundaries, typically incorporating geographical barriers such as the Shannon river, have existed in some form since the very beginning of Irish written history. These records also indicate a society which was feudal and dynastic in nature, with great emphasis put male line descent. Indeed, patronyms are still the most common surname type in Ireland, and emerged early in Medieval records. Warfare was common between kingdoms, of which cattle-raiding constituted a key part, given their value as a key economic and sustenance source (O'Connell et al. 2016). It is perhaps remarkable that the modern Irish population has some of the highest incidences of lactase persistence in the world (Gerbault et al. 2011), alongside other variants with possible selective advantages, including mutations involved in iron-retention (Distante et al. 2004) and skin pigmentation (Rajeevan et al. 2012).

Methods

Sample Processing

A full description of molecular methodology and bioinformatic data processing is provided in Appendix II and summarised in the Methods section of Chapter Two.

Variant Calling and Dataset Preparation

A set of 2,733,477 autosomal transversions from the 1000 Genomes Phase 3 release (European MAF > 1%) were used for f_3 -statistic analysis, as described in Chapter Three. Randomised pseudo-haploid genotypes for these sites were called in ancient samples using the Pileup tool from GATK (McKenna et al. 2010) (See Appendix II for further detail). The Mbuti population from the SGDP dataset (Mallick et al. 2016) again provided an outgroup population. All dataset manipulation and merging was carried out in PLINK v1.90 (Chang et al. 2015).

For ChromoPainter analysis imputed ancient genotypes were merged with a dataset of modern British (n=2,020) and Irish (n=991) individuals sourced from Leslie et al. 2015; van Rheenen et al. 2016 and Byrne et al. submitted. Methods for imputation using Beagle software (Browning & Browning 2007) are detailed in Appendix II. The modern dataset was filtered for SNPs and individuals removed during QC in the source papers, leaving 3,008 individuals for analysis. The ancient dataset consisted of Irish individuals from the Early Modern (n=2), Iron Age (n=22), Bronze Age (n=19) and Chalcolithic (n=5) periods; as well as British individuals from the Anglo-Saxon (n=7), Roman (n=4) and Iron Age (n=3) periods. Also included were four Late Neolithic/Bronze Age individuals from Germany and Scandinavia, a Nordic Iron Age individual and a Czech sample from the Middle Ages (See Table 4.1). Before merging with the modern dataset ancient genotypes were filtered for a genotype probability (GP) above 99% and a MAF above 5% across the 1000 Genomes dataset. Notably, transitions were not excluded, despite their lower imputation accuracy (See Appendix II), in order to preserve a usable number of SNPs for analysis. After ancient and modern datasets were merged, SNPs not called across all individuals were excluded, leaving 104,271 autosomal sites for input into ChromoPainter.

Imputed genotypes were also used for ROH and F_{ST} analysis in PLINK with the same filters as described above (GP > 99%; MAF > 5%). Additionally, transition SNPs were also excluded.

Outgroup f_3 -Statistics

Outgroup f_3 -statistics were calculated using the AdmixTools package (Patterson et al. 2012). The Mbuti population of Central Africa were used as an outgroup, given their undetectable levels of Eurasian admixture (Gurdasani et al. 2015). Tests of the form $f_3(\text{Mbuti}; X, Y)$ were used to measure shared genetic

drift between several populations (EHG, Yamnaya, Irish_NE, Spain_CA and TRB) and a set of Irish and continental individuals, spanning the Late Neolithic to Bronze Age periods (region dependent). For Irish individuals, these values are presented in Fig. 4.4 for EHG, Yamnaya and Irish_NE populations. For continental and Irish individuals, test scores for f_3 -statistic configurations involving Neolithic/Chalcolithic populations (Irish_NE, Spain_CA and TRB) were regressed against one another and outliers identified based on studentised residuals (degrees freedom=92, CI=95%), with results presented in Fig. 4.5.

Only individuals with over 20,000 variant calls were considered in these analyses. For information on ancient individuals included in these tests and the population keys used see Tables 2.1, 3.1 and 4.1. The population of Yamnaya individuals (not shown in Table 4.1) was sourced from Allentoft et al. 2015 and Haak et al. 2015.

ChromoPainter Analysis

ChromoPainter analysis (Lawson et al. 2012) was performed on a dataset of 3,008 modern Irish and British individuals, alongside 68 imputed ancient samples (see above for dataset preparation). The dataset was first phased with SHAPEIT V2 (Delaneau et al. 2011) using the 1000 Genomes (Phase 3) as a reference panel (1000 Genomes Project Consortium et al. 2015). Phasing was done by chromosome using default settings and the GRCh37 build genetic map to estimate linkage disequilibrium. ChromoPainter was run under default settings to paint each ancient individual using all other individuals (-a 0 0), as detailed in Chapter Two. The output coancestry matrix was then used for PCA (Fig. 4.6), following instructions available at <http://www.paintmychromosomes.com/>.

Runs of Homozygosity Analysis

A total of 848,715 transversion sites called across 48 Irish Chalcolithic, Bronze and Iron samples were used for ROH (see above for selection criteria). Distributions of runs of homozygosity (ROH) were calculated using PLINK, with the same parameters described in Chapter Three. To explore recent inbreeding and population constrictions the average length of ROH tracts for each individual was plotted against the length of the longest tract observed (Fig. 4.7).

Phenotypic analysis

Given the substantial genetic continuity seen in Ireland since the Bronze Age (Figs. 1.4 and 4.6), a preliminary scan for possible selection events between the Irish Bronze and Iron Age periods was carried out. Weir & Cockerham's F_{st} (Weir & Cockerham 1984) was estimated using the --fst command in PLINK v1.90 (Chang et al. 2015) for each variant in the dataset described above (848,715 sites), with

Irish Chalcolithic/Bronze Age and Iron Age populations predefined via the --within option. The results are plotted in Fig. 4.8.

Through the use of imputed genotypes, approximate frequencies of variants found at relatively high incidence in the modern Irish population were also calculated in Mesolithic, Neolithic, Chalcolithic/Bronze Age, and Iron Age cohorts (Fig. 4.9).

Table 4.1 List of samples from the Late Neolithic to Early Modern periods analysed in the current chapter.

Sample identifiers and archaeological contexts are given. For samples from which both whole genome sequence data (WGS) and data from the 1240k SNP capture panel released in Mathieson et al. 2015 (Cap.) is available, the data type used in this study is noted in the sample ID. The inclusion or exclusion of each sample in the different analyses presented here is noted in the final two columns. Y haplogroup placements that may be confounded by post-mortem damage are bracketed. ‘Number of Variants’ refers to the total number of secure calls for a set of 2,733,477 autosomal transversions from the 1000 Genomes Phase 3 release. Radiocarbon calibrations are in years BC, unless AD is specifically stated. The European country of origin for samples is denoted by its first three letters. Yamnaya individuals used for f_3 -statistics are not shown here, but were sourced from Haak et al. 2015 and Allentoft et al. 2015.

Reference Code: ¹Allentoft et al. 2015; ²Haak et al. 2015; ³Mathieson et al. 2015; ⁴Gamba et al. 2014; ⁵Current Study; ⁶Martiniano et al. 2017; ⁷Günther et al. 2015; ⁸Schiffels et al. 2016; ⁹Martiniano et al. 2016.

Reference	Current Study ID	Other IDs	Country	Latitude	Longitude	Period	Date (cal BC) *not direct	Mitochondrial Haplogroup	Y Chromosome Haplogroup	Data Type	Number of Variants	F ₃ -Statistics	ChronoPainter
1	BB_Kne1	RISE567	Cze	50.12	14.26	CA	-	U5b2c	-	WGS	54890	Yes	No
1	BB_Kne2	RISE566	Cze	50.12	14.26	CA	2279-2033	H1	R1b1a1a2a1a	WGS	209802	Yes	No
1	Unetice_Mor1	RISE586	Cze	48.8	17.02	EBA	2300-1800*	K1b1a	-	WGS	263914	Yes	No
1	Unetice_Vel1	RISE577	Cze	50.16	14.31	EBA	2300-1800*	T2b	-	WGS	1204525	Yes	Yes
1	Nordic_BA3	RISE47	Den	56.97	9.55	EBA	1499-1324	I	R1b1a1a2	WGS	216301	Yes	No
1	Nordic_BA5	RISE21	Den	55.55	12.2	EBA	1426-1281	N1a1a1a2	..	WGS	54436	Yes	No
1	Nordic_LBA1	RISE276	Den	55.91	11.57	LBA	794-547	T2b	R1b1a1a2	WGS	176616	Yes	No
1	Nordic_LN2	RISE71	Den	56.68	10.03	LN	2196-2023	H3b	-	WGS	425470	Yes	No
1	Nordic_LN4	RISE42	Den	55.66	12.16	LN	2191-1972	H3v+16093	..	WGS	36304	Yes	No
1	Nordic_MN1	RISE61	Den	55.7	11.86	MN	2851-2492	J1e4	R1a1a1	WGS	569379	Yes	No
2,3	Alberstedt_LN	I0118	Ger	51.45	11.63	LN	2471-2246	HV	-	Cap.	893374	Yes	Yes
1	BB_Aus1	RISE559	Ger	48.33	10.9	CA	2461-2207	H46	-	WGS	217971	Yes	No
1	BB_Aus2	RISE560	Ger	48.33	10.9	CA	2500-2000*	U5a1a1	R1b1a1a2	WGS	94850	Yes	No
3	BB_BZH15	I1549	Ger	51.82	10.91	CA	2500-2050*	W1e1	-	Cap.	289420	Yes	No
3	BB_BZH2	I1546	Ger	51.82	10.91	CA	2500-2050*	U5a1b1	-	Cap.	38900	Yes	No
1	BB_Ost1	RISE563	Ger	48.69	13.02	CA	2572-2512	K1c1	R1b1a1a2a1a2	WGS	571106	Yes	No
1	BB_Ost2	RISE564	Ger	48.69	13.02	CA	2500-2000*	H-T16311C	R1b1a1a2a1	WGS	165383	Yes	No
3	BB_QLB26	I0805	Ger	51.79	11.15	CA	2467-2142	H1	R1b1a2	Cap.	109139	Yes	No
2,3	BB_QLB28	I0806	Ger	51.79	11.14	CA	2431-2150	H1	R1b1a2a1a2	Cap.	73055	Yes	No
2,3	BB_QUEXIII4	I0113	Ger	51.79	11.14	CA	2346-2033	J1e5	-	Cap.	239236	Yes	No
2,3	BB_QUEXIII6	I0112	Ger	51.79	11.14	CA	2457-2142	H13a1a2	-	Cap.	548786	Yes	No
2,3	BB_Rot3	I0060	Ger	51.45	11.54	CA	2428-2149	K1a2c	-	Cap.	33751	Yes	No
2,3	BB_Rot4	I0111	Ger	51.45	11.54	CA	2475-2204	H3ao	-	Cap.	205947	Yes	No
2,3	BB_Rot6	I0108	Ger	51.45	11.54	CA	2575-2299	H5a3	-	Cap.	344993	Yes	No
2,3	BZH_LN2	I0059	Ger	51.82	10.91	LN	2337-2138	H1	-	Cap.	244637	Yes	No
2,3	BZH_LN3	I0171	Ger	51.82	10.91	LN	2287-2041	U5a1a2a	-	Cap.	35402	Yes	No
2,3	Halberstadt_LBA	I0099	Ger	51.89	11.04	LBA	1193-979	H23	R1a1a1b1a2	Cap.	549243	Yes	No

2,3	Karsdorf_LN	I0550	Ger	51.27	11.66	LN	2570-2471	T1a1	-	Cap.	87999	Yes	No
2,3	Unetice_ESP2	I0114	Ger	51.42	11.68	EBA	2138-1952	I3a	I2a2	Cap.	326682	Yes	No
2,3	Unetice_ESP29	I0117	Ger	51.42	11.68	EBA	2272-2039	I3a	-	Cap.	119586	Yes	No
2,3	Unetice_ESP3	I0115	Ger	51.42	11.68	EBA	1954-1760	U5a1i	-	Cap.	325233	Yes	No
2,3	Unetice_ESP4	I0116	Ger	51.42	11.68	EBA	2134-1939	W3a1	I2c	Cap.	435954	Yes	No
2,3	Unetice_EUL41	I0803	Ger	51.17	11.85	EBA	2132-1942	H4a1a1a	-	Cap.	84300	Yes	No
2,3	Unetice_EUL57	I0804	Ger	51.17	11.85	EBA	2137-1965	H3	I2	Cap.	23169	Yes	No
2,3	Unetice_HAL16	I0047	Ger	51.89	11.04	EBA	2111-1891	V9	-	Cap.	363743	Yes	No
2,3	Unetice_QUEVIII6	I0164	Ger	51.79	11.14	EBA	2023-1894	U5b2a1b	-	Cap.	344210	Yes	No
3,4	BR1_Mako_WGS	BR1	Hun	47.17	20.83	EBA	2194-2082	K1c1	-	WGS /Cap.	1172365	Yes	No
3,4	BR2_Kyjatice_WGS	BR2	Hun	47.82	19.95	LBA	1266-1181	K1a1a	J2a1	WGS /Cap.	2600306	Yes	No
1	Hungary_MBA1	RISE349	Hun	46.36	20.99	MBA	2034-1784	T2b3	-	WGS	177727	Yes	No
1	Maros_Szo1	RISE371	Hun	46.22	20.2	MBA	2136-1941	U5a2b	-	WGS	120263	Yes	No
1	Maros_Szo2	RISE373	Hun	46.22	20.2	MBA	1886-1696	K1a2a	-	WGS	544092	Yes	No
1	Maros_Szo3	RISE374	Hun	46.22	20.2	MBA	1866-1619	T2b	G2a2a1a2a	WGS	332210	Yes	No
1	Vatya_Erd1	RISE480	Hun	47.34	18.9	MBA	1700-1500*	U5a2a	-	WGS	241940	Yes	No
1	Vatya_Erd2	RISE483	Hun	47.34	18.9	MBA	2000-1500*	H2a1	-	WGS	230038	Yes	No
1	Vatya_Erd3	RISE484	Hun	47.34	18.9	MBA	2000-1500*	T1a1	-	WGS	262092	Yes	No
1	Vatya_Erd4	RISE479	Hun	47.34	18.9	MBA	2000-1500*	T2b	I2a2a1a2a	WGS	1703163	Yes	No
1	Vatya_Sza1	RISE247	Hun	47.33	18.96	MBA	1746-1611	H11a	I2a2a1	WGS	460028	Yes	No
1	Vatya_Sza2	RISE254	Hun	47.33	18.96	MBA	2128-1909	J1c9	I	WGS	151734	Yes	No
5	Annaghkeen84		Ire	53.45	-9.19	EBA	1878-1682	HV0+195	-	WGS	1880563	Yes	Yes
5	Blackhill32		Ire	53.26	-6.57	EBA	2458-1983	H40	-	WGS	1428766	Yes	Yes
5	Bolinready79		Ire	52.6	-6.34	EBA	2193-1777	T2b	R1b1a1a2a1a2	WGS	1575501	Yes	Yes
5	Cherrywood1		Ire	53.25	-6.13	MLBA	1400-1000*	T2c1+146	-	WGS	1748085	Yes	Yes
5	Coolnatullagh1005		Ire	53.08	-9.03	EBA	1880-1610	U2e1b2	-	WGS	1792988	Yes	Yes
5	Grange10		Ire	53.73	-8.24	EBA	2201-2045	H6a1b4	-	WGS	1681045	Yes	Yes
5	Inchagreenoge134		Ire	52.59	-9.09	LBA	1270-1040	T2f	R1b1a1a2a1a2c1 a1k	WGS	1650318	Yes	Yes
5	Keenoge14		Ire	53.64	-6.42	EBA	2023-1775	H2a1e1b	-	WGS	1716031	Yes	Yes
5	Keenoge3		Ire	53.64	-6.42	EBA	2188-1887	K1a26	-	WGS	1243561	Yes	Yes
5	Keenoge7		Ire	53.64	-6.42	EBA	2200-1800*	U5b2b2	-	WGS	678017	Yes	No
5	Killuragh1		Ire	52.6	-8.33	CA	2344-2036	H	I2a2a1a1a2	WGS	1815839	Yes	Yes
5	Labbacallee212		Ire	52.17	-8.33	CA	2402-2138	H41a	-	WGS	1820605	Yes	Yes
5	Labbacallee213		Ire	52.17	-8.33	CAEBA	2500-1800*	W4	-	WGS	1842988	Yes	Yes
5	Moyvecla381		Ire	53.26	-8.84	EBA	2286-2039	H1b	R1b1a1a2a1a2c1 a5c3	WGS	1654347	Yes	Yes
5	Ploopluck6		Ire	53.22	-6.69	EBA	2037-1786	H1a1	R1b1a1a2a1a2c	WGS	2097767	Yes	Yes
5	Pollnagollum86		Ire	54.26	-7.81	CA	2300-2131	U4a3a	-	WGS	1670799	Yes	Yes
5	Pollnagollum90		Ire	54.26	-7.81	CAEBA	2500-1800*	T2b2b	-	WGS	1226637	Yes	No
5	Pollnagollum911		Ire	54.26	-7.81	CA	2349-2135	K1b1a1	R1b1a1a2a1a2c1 a3	WGS	1783737	Yes	Yes
5	Poulnabrone01		Ire	53.05	-9.14	EBA	2028-1758	W5b	-	WGS	1658273	Yes	Yes

A Genomic Compendium of an Island

5	Rathlin1		Ire	55.29	-6.19	EBA	2026-1885	U5a1b1c	R1b1a1a2a1a2c1a5	WGS	2614761	Yes	Yes
5	Rathlin2		Ire	55.29	-6.19	EBA	2024-1741	U5b2a2a	R1b1a1a2a1a2c1a(5c)	WGS	1987496	Yes	Yes
5	Rathlin3		Ire	55.29	-6.19	EBA	1736-1534	J2b1a	R1b1a1a2a1a2c1	WGS	2052686	Yes	Yes
5	RoughanHill125		Ire	52.99	-9.11	EBA	1943-1663	U5a2+16294	-	WGS	1617920	Yes	Yes
5	RoughanHill468		Ire	52.99	-9.11	EBA	1943-1663	U5a2+16294	-	WGS	1694918	Yes	No
5	Sliguff27		Ire	52.67	-6.97	EBA	2134-1701	U5b2b2	R1b1a1a2a1a2c1a	WGS	1741810	Yes	Yes
5	Stonepark3409		Ire	53.78	-9.15	EBA	2122-1898	H	-	WGS	1441124	Yes	Yes
5	Treanmacmurtagh116		Ire	54.06	-8.42	EBA	2015-1758	H1c	R1b1a1a2a1a2c1a5	WGS	1852747	Yes	Yes
1	Unetice_Cho1	RISE139	Pol	58.18	13.4	EBA	2135-1923	U2e1f1	..	WGS	51423	Yes	No
1	Unetice_Pol1	RISE145	Pol	50.91	17.17	EBA	2188-1958	H6a1b	-	WGS	57495	Yes	No
1	Unetice_Prz1	RISE150	Pol	16.96	50.92	EBA	1885-1693	U5a1b1	-	WGS	1294784	Yes	Yes
1	Unetice_Szc1	RISE154	Pol	16.94	50.95	EBA	1925-1765	K1a4a1	-	WGS	284940	Yes	No
1	Unetice_Woj1	RISE109	Pol	17.07	50.98	EBA	1954-1772	U4	-	WGS	410082	Yes	No
6	MonteGato104	MG104	Por	38.04	-7.61	BA	1740-1430	U5b3	R1b1a2	WGS	1536599	Yes	No
6	TV32032		Por	37.94	-7.64	BA	1750-1510	X2b+226	R1b1a2	WGS	1450082	Yes	No
6	TV3831		Por	37.94	-7.64	BA	1750-1510	H1+152	R1b1a2a1a2	WGS	1274805	Yes	No
6	ValeOuro10207	VO10207	Por	38.01	-8.15	BA	2000-1500*	U5b1+16189+@16192	-	WGS	530354	Yes	No
7	ATP9	ATP9	Spa	42.21	-3.31	MBA	1700-1568	U5b1b	-	WGS	887265	Yes	No
1	Nordic_BA1	RISE175	Swe	13.6	55.4	EBA	1395-1132	T1a1	I	WGS	213340	Yes	No
1	Nordic_BA2	RISE210	Swe	56	14.1	EBA	1432-1292	T2a1a	I	WGS	130983	Yes	No
1	Nordic_BA4	RISE207	Swe	56	14.1	EBA	1493-1302	J1c8a1	..	WGS	37411	Yes	No
1	Nordic_LN	RISE98	Swe	55.38	13.45	LN	2275-2032	K1b1a1	R1b1a1a2a1a1	WGS	2145140	Yes	Yes
1	Nordic_LN1	RISE179	Swe	13.6	55.4	LN	2010-1776	K1a3	I	WGS	106178	Yes	No
1	Nordic_LN3	RISE97	Swe	55.56	13.06	LN	2025-1885	K2a5	-	WGS	1005690	Yes	No
1	Slav_Bra1	RISE569	Cze	50.19	14.16	MA	660-770 AD	H1af	-	WGS	1211341	No	Yes
9	6DT21		Eng	53.95	-1.097	RM	71-350 AD*	J1c3c2	R1b1a2a1a2c2	WGS	1774539	No	Yes
9	6DT22		Eng	53.95	-1.097	RM	71-350 AD*	H2+195	R1b1a2a1a2b	WGS	1733206	No	Yes
9	6DT23		Eng	53.95	-1.097	RM	71-350 AD*	H6a1b2	R1b1a2a1a	WGS	1243029	No	Yes
9	6DT3		Eng	53.95	-1.097	RM	71-350 AD*	J1b1a1	R1b1a2a1a1	WGS	2086658	No	Yes
8	HI1	HI1	Eng	52.08	0.18	IA	160 BC – 26 AD	K1a1b1b	R1b1a2a1a2c	WGS	1547982	No	Yes
8	HI2	HI2	Eng	52.08	0.18	IA	170 BC – 80 AD	H1ag1	R1b1a2a1a2c1	WGS	2489544	No	Yes
8	HS1	HS1	Eng	52.08	0.18	AS	666–770 AD	H2a2b1	-	WGS	2117136	No	Yes
8	HS2	HS2	Eng	52.08	0.18	AS	631–776 AD	K1a4a1a2b	-	WGS	2169678	No	Yes
8	HS3	HS3	Eng	52.08	0.18	AS	690–881 AD	H2a2a1	-	WGS	1197122	No	No
8	L	L	Eng	52.1	0.28	IA	360–50	H1c	-	WGS	1235078	No	Yes
9	NO3423	NO3423	Eng	54.59	-1.32	AS	550-650 AD	H1a	I1	WGS	1670620	No	Yes
8	O1	O1	Eng	52.26	0.06	AS	420–570 AD	U5a2a1	-	WGS	2075917	No	Yes
8	O2	O2	Eng	52.26	0.06	AS	385–535 AD	H1g1	-	WGS	1819089	No	Yes
8	O3	O3	Eng	52.26	0.06	AS	395–540 AD	T2a1a	-	WGS	2448993	No	Yes

8	O4	O4	Eng	52.26	0.06	AS	400-545 AD	H1at1	-	WGS	2441931	No	Yes
5	Ballybunnon54		Ire	52.51	-9.68	IA	583-674 AD	U5b2c2b	R1b1a1a2a1a2c1 a2b2a	WGS	1774954	No	Yes
5	Ballycagh79		Ire	52.48	-9.68	IA	470-663 AD	K1a4a1c	-	WGS	1574445	No	Yes
5	Ballyglass44		Ire	53.71	-8.71	IA	80-610 AD*	V7a	R1b1a1a2a1a2c1 a1a1a(1)	WGS	1943553	No	Yes
5	Ballyglass45		Ire	53.71	-8.71	IA	80-610 AD*	K1a4a1	-	WGS	1815553	No	Yes
5	Ballyglass46		Ire	53.71	-8.71	IA	80-610 AD*	H1c3	-	WGS	1609888	No	Yes
5	Ballymacaward2		Ire	54.51	-8.24	IA	608-658 AD	J1c1a	-	WGS	1602111	Yes	Yes
5	Ballymacaward3		Ire	54.51	-8.24	IA	421-536 AD	H56	-	WGS	1695792	Yes	Yes
5	Claristown14		Ire	53.65	-6.28	IA	60-420 AD	U5a1h	R1b1a1a2a1a2c1 a3a2(a1d)	WGS	1775387	Yes	Yes
5	Courtmacsherry37		Ire	51.63	-8.69	IA	540-662 AD	T2b3+151	R1b1a1a2a1a2c1 a2b2a	WGS	1489593	No	Yes
5	Derrynamanagh01		Ire	53.31	-8.55	IA	200-700 AD*	T1a1+152	-	WGS	1679800	No	Yes
5	Derrynamanagh02		Ire	53.31	-8.55	IA	324-534 AD	J1c3g	-	WGS	1578077	No	No
5	Derrynamanagh03		Ire	53.31	-8.55	IA	200-700 AD*	J1c1a	-	WGS	2025994	No	Yes
5	Derrynamanagh04		Ire	53.31	-8.55	IA	200-700 AD*	K1a1b1g	-	WGS	1665238	No	Yes
5	Derrynamanagh05		Ire	53.31	-8.55	IA	324-534 AD	J1c3g	-	WGS	1655823	No	Yes
5	Derrynamanagh06		Ire	53.31	-8.55	IA	200-700 AD*	H1b	-	WGS	1952573	No	Yes
5	Derrynamanagh08		Ire	53.31	-8.55	IA	200-700 AD*	H5c	R1b1a1a2a1a2c1 (a1a1a1a1)	WGS	1752528	No	Yes
5	Derrynamanagh09		Ire	53.31	-8.55	IA	200-700 AD*	J1c3g	R1b1a1a2a1a2c1 a1a1a1a1	WGS	1586541	No	No
5	HillOfWard14		Ire	53.63	-6.89	IA	410-546 AD	K1c1	R1b1a1a2a1a	WGS	1831016	No	Yes
5	Jerpoint15		Ire	52.52	-7.17	IA	257-416 AD	H3k	-	WGS	1954039	No	Yes
5	Knockadoon275		Ire	52.52	-8.53	MD	1448-1635 AD	V15a	R1b1a1a2a1a2c	WGS	1868170	No	Yes
5	Knowth4		Ire	53.7	-6.49	IA	86-252 AD	K1a+195	-	WGS	1662644	No	Yes
5	Knowth10		Ire	53.7	-6.49	IA	175-50	H1aw	R1b1a1a2a1a2c1 a5b1a1a	WGS	1838639	No	Yes
5	Lagore14		Ire	53.51	-6.53	MD	1492-1665 AD	W1c	R1b1a1a2a1a2c1 a1a1a1a1a1a	WGS	2283053	No	Yes
5	Lehinch27		Ire	53.35	-7.61	IA	399-541 AD	K1a+195+14 6	-	WGS	1679749	No	Yes
5	Lehinch28		Ire	53.35	-7.61	IA	422-596 AD	H66a1	R1b1a1a2a1a2c1 (a*)	WGS	1882317	No	Yes
5	Rossnarec1		Ire	53.69	-6.5	IA	257-533 AD	J1c3b1	-	WGS	1946658	No	No
5	Rossnarec2		Ire	53.69	-6.5	IA	257-533 AD	J1c3b1	R1b1a1a2a1a2c1 a5c1	WGS	1987078	No	Yes
1	Nordic_IA1	RISE174	Swe	55.54	13.1	IA	427-611 AD	W1	..	WGS	1214760	No	Yes

Results

Sequencing Statistics

Detailed sequencing statistics and data authenticity results for the 54 samples analysed in the current chapter can be found in Appendix II and Electronic Data Table S2. All individuals were sequenced to between 0.9-2.2X, with the exception of three samples from Keenoge and Pollnagollum (Keenoge3, Keenoge7 and Pollnagollum90), which were sequenced to between 0.3-0.75X (See Fig. 1.1), and the high coverage Rathlin1 genome (10.5X).

The dataset encompasses a wide geographical and chronological range, which spans the Chalcolithic to Early Modern periods of the island. This includes four samples with direct Chalcolithic (CA) dates, Killuragh1, Pollnagollum86, Pollnagollum911 (Cave sites), and Labbacallee212 (Wedge Tomb); and two undated samples from Chalcolithic/Early Bronze Age contexts, Pollnagollum90 and Labbacallee213. A further 17 samples were directly dated to the Early Bronze Age (EBA). Two individuals were indirectly also placed in this period given their context (Keenoge7) or kinship to a dated EBA sample (RoughanHill125). The majority of these are from crouched inhumation burials in cists or flat graves, with the exception of three individuals from megalithic contexts in the Burren (Poul nabrone01, RoughanHill125 and RoughanHill468). Two Late Bronze Age (LBA) individuals are also included here, one dated directly to the period (Inchagreenoge134), and one indirectly dated using animal bone from the same context (Cherrywood1). Notably two potential LBA samples from crannóg sites (Derrymaquirk20 and Ballinderry6) are excluded from this chapter due to low sequencing coverage (0.03-0.14X) and lack of securely dated contexts. The Irish Iron Age/Early Christian (IA) population presented here consists of 25 samples dating from roughly 175 BC to 770 AD, eight of which, from Derrynamanagh and Rossnaree, are indirectly dated through kinship with directly dated individuals or through site context. Finally, two early modern (MD) samples are also included, dating to approximately 1,450-1,660 AD and overlapping with the period of plantations in Ireland. Further archaeological contexts for all samples are detailed in Appendix I.

Uniparental Analysis

No autosomal data is yet available for the Bell Beaker and Bronze Age periods of Britain and northwestern Europe and so, as in Chapter Four, the Irish dataset is compared here to the preliminary Y chromosome and mtDNA haplogroups reported for the regions in (Olalde et al. 2017), alongside published data from Gamba et al. 2014; Allentoft et al. 2015; Günther et al. 2015; Haak et al. 2015; Mathieson et al. 2015; Martiniano et al. 2016; Schiffels et al. 2016; Jones et al. 2017 and Martiniano et al. 2017. The results are displayed in Figs. 4.2 and 4.3.

Y Chromosome

It is immediately evident that Ireland witnessed the same drastic turnover in male lineages reported between the Neolithic (Fig. 3.2) and Bronze Age (Fig. 4.2) periods of Europe (Allentoft et al. 2015; Haak et al. 2015; Olalde et al. 2017). R1b haplogroups, specifically R1b-M269, are seen to dominate in the majority of western populations from the Beaker horizon onwards, showing a west-east gradient which reflects that observed in Europe today (Myres et al. 2011). Interestingly, eastern populations, such as Corded Ware and the related Andronovo and Sintashta steppe cultures, show a higher incidence of R1a, again reflecting the haplogroup's modern distribution. However, the earlier Yamnaya and Poltavka cultures of the Caspian-Pontic Steppe show discontinuity with these later groups, possessing a majority of R1b-M269 haplogroups, although these appear to belong to a separate subclade (CTS1078) than that observed in western Europe.

The only exception to the predominance of R1b in western Beaker populations is seen in Iberia, where R1b-M269 is not observed until the Bronze Age period (Martiniano et al. 2016), although two earlier incidences of R1b1 are seen in Bell Beaker samples (Olalde et al. 2017). R1b haplogroups are also at relatively lower frequency in the Bell Beaker population from Hungary, emphasising the heterogeneity of Beaker culture samples also seen in autosomal analyses (Haak et al. 2015; Olalde et al. 2017). Notably, R1b haplogroups are not seen in later Hungarian Bronze Age individuals, in a reversal of what is observed in Iberia. R1b chromosomes are also not observed in the the later Únětice Culture of Central Europe, which shows a complete resurgence in Mesolithic I2 lineages, relative to the preceding Corded Ware and Beaker horizons, albeit with a small sample size (n=3). I2 lineages reached substantial frequencies in the prior Neolithic period of the region, and are the sole observed haplogroup in Globular Amphora culture (See Chapter Four), a predecessor of Corded Ware. Notably, inflated WHG ancestry has been previously reported in the Únětice population relative to a Bell Beaker group from the same region (Haak et al. 2015). I2 haplogroups are also prevalent in the Swedish Late Neolithic and Bronze Age periods, suggesting some Neolithic male line continuity occurred in these regions, despite the influx of steppe-related ancestry that has been demonstrated (Allentoft et al. 2015).

The west and northwest of the continent, including Ireland, exhibit a more complete replacement of Neolithic I2 lineages, suggesting a large male-mediated bias in the mass introduction of steppe-related ancestry to Atlantic Europe. Interestingly, the only Irish I2 lineage observed from the Beaker horizon onwards belongs to a Chalcolithic sample from the southwest of the island (Killuragh1). This individual belongs to the clade I2-M284, which is somewhat restricted to Britain and Ireland today and common in Irish Late Neolithic samples (See Chapter Four for further discussion). The two I2 lineages observed in the British Beaker and Bronze Age periods also belong to M284, supporting some small level of male line continuity across the transition on the islands. Importantly, the only other male Chalcolithic sample from Ireland, Pollnagollum911 (2,349-2,135 cal BC), belongs to R1b-M269, pushing back the earliest incidence of this haplogroup on the island from the Early Bronze Age (Cassidy et al. 2016) to the preceding Beaker period.

The majority of R1b-M269 lineages seen in Bell Beaker and Bronze Age populations from France, the Netherlands, Britain and Ireland belong to P312, which shows a modern distribution focused west of the Rhine Basin (Myres et al. 2011). The relatively late (post-Beaker) appearance of this marker in Iberia, in comparison to Ireland and Britain, suggests the arrival of R1b haplogroups to the islands was not catalysed by direct population movement from Iberia, but instead via a more northern Beaker group. Notably, while absent in contemporaneous individuals from Iberia, R1b-P312 lineages are observed in Beaker samples from southern France and northern Italy (dating to *circa* 2,500-1,900 BC), indicating the haplogroup was not just restricted to the very north of Europe, but also areas associated with Mediterranean Beaker networks.

A subclade of R1b-P312, L21, is only observed in ancient British and Irish samples. Remarkably, L21 exhibits a similar restriction in modern day Europe, showing a very steep frequency peak in Ireland (Fig. 4.9), where it accounts for over a half of all Y chromosomes. This distribution is not uniform, with increased incidence of R1b-M269 seen in the south and west of the island (>90%) (Hill et al. 2000). L21 appears at a similarly high frequencies in the island's Chalcolithic/Bronze Age and Iron Age populations (Fig. 4.9), with only one sample from each period (Bolinready79 and HillOfWard14) not possessing the marker. The predominance of L21 in Irish populations seen throughout time, or more specifically a sublineage of the haplogroup defined by the mutation DF13, strongly suggests male line continuity on the island over the past four millennia. Britain does not show such an obvious continuous trend, with non-P312 subclades of R1b-M269, such as M405 and PF6570, previously observed on the continent only, appearing in the Iron Age population of the island. Notably, the single male Anglo-Saxon individual from Britain available for analysis belongs to a non-R1b haplogroup, I1, widespread in Nordic countries today (Martiniano et al. 2016). Moreover, an earlier decrease in L21 haplogroups is suggested between Britain's Beaker and Bronze Age periods, although it is unclear whether this is a result of low sample coverage (Olalde et al. 2017).

Within Ireland a large diversity in R1b-DF13 subclades is seen in both the Bronze and Iron Ages, with an expectedly higher number of downstream mutations observed for Iron Age samples (See Electronic Data Table S5). Notably, all northwestern Irish Iron Age individuals sampled (Ballyglass44, Derrynamanagh08 and Derrynamanagh09) were seen to belong to the R1b-M222 subclade or a lineage leading directly to it. This haplogroup peaks in northwestern Ireland today and has been previously associated with the early Medieval Uí Néill dynasty of the region (Moore et al. 2006). Intriguingly, the two southern Iron Age individuals sampled, Courtmacsherry37 and Ballybunnion54, also both share a subclade, R1b-CTS3087. Given the known emphasis placed on patrilineal descent in Gaelic Ireland, denser surveys of Iron Age Y chromosomal variation on the island may contribute greatly to the understanding of territorial boundaries and patronymic surname distributions that were recorded during the early historical period.

Most noticeably, in northwestern and central Europe, Ireland included, a small increase in Mesolithic U lineages is seen between the Neolithic and Bronze Age, including previously unobserved U5a1, U2e and U4 haplogroups, common in Latvian, Ukrainian and Russian HG individuals, as well as later steppe cultures (Figs. 2.1 and 4.3). In Iberia, an increase in U lineages is also seen across the transition, although these do not show any particular eastern affiliations. Remarkably, no U lineages are observed in Late Neolithic and Bronze Age samples from Scandinavia, nor have they been yet observed in individuals (n=6) from the preceding Neolithic TRB culture of Sweden (Skoglund et al. 2014; Mittnik et al. 2017), despite their prevalence in Mesolithic and Neolithic hunter-gatherer groups of the region.

Increased frequencies of mtDNA haplogroup I are also seen in several post-Neolithic populations from northern and central Europe, including Bronze Age individuals from Britain, Ireland, Denmark and Germany. Interestingly, high incidences of the lineage are seen in Hungarian Bell Beaker samples, although in other Bell Beaker populations the haplogroup is absent or extremely rare. The heterogeneous nature of Bell Beaker culture is also seen in the distribution of haplogroup X, present in Atlantic Beaker populations, but absent in groups further east, and haplogroup H, seen at substantial frequencies in all Beaker groups with the exception of the British cohort. In this respect, the British Beaker population also displays remarkable discontinuity with the preceding Neolithic and succeeding Bronze Age of the island.

Mitochondrial continuity is more apparent across the Neolithic to Chalcolithic/Early Bronze Age transition in Ireland, which interestingly shows a reversed trend to that seen in Beaker Britain, namely an inflation in H haplogroups and a decrease in haplogroup K (Fig. 3.3 and 4.3). In this way, the Irish Chalcolithic/Early Bronze Age shows more similarity to the later British Bronze Age, rather than Beaker period. The Irish Iron Age exhibits a similar mitochondrial profile to the island's Bronze Age population, although a definite decrease in U haplogroups is seen, alongside an inflation of J and K lineages. A similar trend is apparent in Britain, with a more notable increase in H haplogroups also observed.

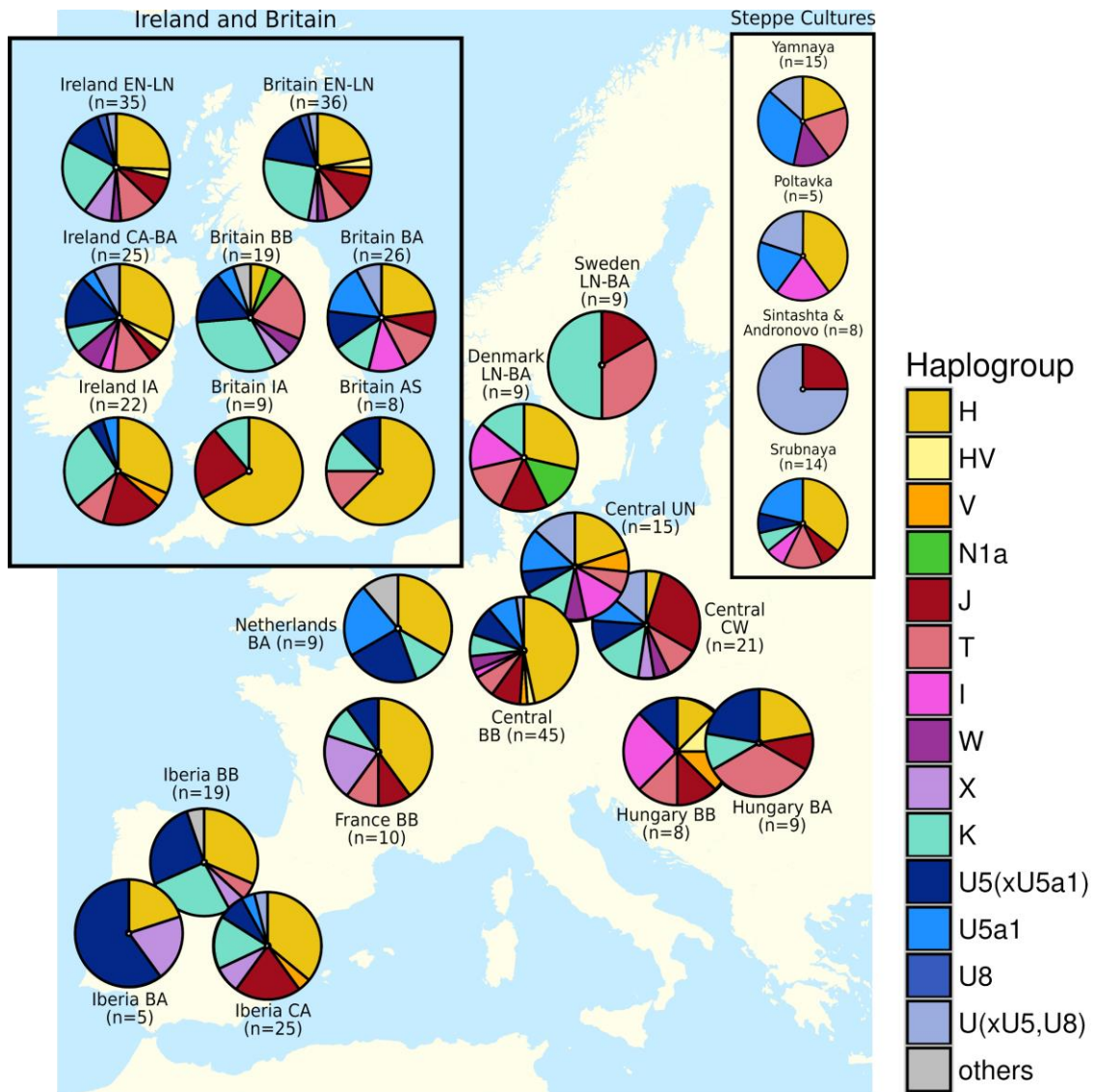


Figure 4.3. mtDNA haplogroup frequencies in Europe from the Corded Ware and Bell Beaker horizons onwards. Haplogroups frequencies for Neolithic and Iron Age/Early Christian populations from Britain and Ireland are also shown for comparative purposes, alongside four steppe populations. Haplogroup colours are consistent with those shown in Fig. 3.3. Population Key: Corded Ware (LN) - CW, Bell Beaker (CA) - BB, Únětice (EBA) - UN.

Dissecting the Neolithic to Bronze Age transition in Ireland and Europe using outgroup f_3 statistics and ADMIXTURE components

This section addresses three fundamental questions regarding the demographic upheaval that occurred during the transition to metallurgy in Ireland, with key conclusions headered below.

1. What was the absolute level of steppe-related introgression that entered Ireland during the Chalcolithic and Bronze Age periods, relative to other European regions?
2. Does the above measure show any variability within Ireland itself, given the bimodal nature by which aspects of the Beaker phenomenon entered the island?
3. What was the genetic contribution of local Neolithic populations to the succeeding metallurgical societies of Europe, and again do these signals show any variability within Ireland?

1. Approximately half of Irish Bronze Age ancestry is derived from the steppe.

Autosomal analyses in Chapter Two visualise the large influx of steppe-related ancestry that occurred into Europe and Ireland during the Corded Ware and Bell Beaker horizons (Allentoft et al. 2015; Haak et al. 2015; Cassidy et al. 2016), with the majority of Irish Early Bronze Age samples showing ADMIXTURE profiles (Figs. 1.2C and 4.4D) and PCA placements (Fig. 1.3D) similar to contemporaneous individuals from north central Europe and Scandinavia. If Yamnaya is taken as a direct source of introgression into these populations, percentage estimates of steppe ancestry can be approximated based on ADMIXTURE components. Specifically, the Caucasus-related component (teal), not present in Europe before this period, is present in Yamnaya and related steppe populations at a proportion of 75-80% of their total ancestry. The majority of remaining Yamnaya ancestry is composed of the European HG component (red), which is also present in European Neolithic populations and thus non-informative. However, the percentage at which the Caucasus-related component features in Late Neolithic and Bronze Age European individuals can be scaled up to include the remainder of Yamnaya ancestry hidden within the European HG component, giving an approximate estimate of steppe introgression in each sample. This was done using the equation:

$$\text{Caucasus Component (\%)} \div 0.775 = \text{Steppe Ancestry (\%)}$$

This method gives a range of 78-85% steppe-derived ancestry in the Corded Ware population, close to the value of 75%, reported in Haak et al. 2015. The majority of Irish Chalcolithic and Bronze Age samples fall between the range 48% and 64%, with highest values (>58%), observed for a number of northern (Rathlin1, Rathlin3, Pollnagollum90, Grange10) and eastern samples (Keenoge3, Blackhill32). This pattern can be visualised in Fig. 4.4D and will be discussed further in the following section. Similar ranges were observed for Únětice (53-65%) and Bell Beaker (42-56%) groups, again matching previously published estimates (Haak et al. 2015). This suggests Ireland, despite its peripheral position, did not receive a diluted level of steppe ancestry relative to populations within the Corded Ware contact zone, again emphasising the extensive nature of the migrations that occurred into the island at this horizon. In

contrast, the lowest levels of steppe ancestry were calculated for another peripheral Atlantic population, the Portuguese Bronze Age (Martiniano et al. 2016), where it featured at a proportion of only 14-16% (Fig. 1.2C).

2. A southwestern refugium of Neolithic ancestry existed in Early Bronze Age Ireland.

A key exception to the above trend is seen in several southwestern Irish individuals, for whom substantially reduced steppe ancestry was estimated (24% in Labbacallee212 and 36-39% in RoughanHill468, Killuragh1 and Poul nabrone01), a phenomenon that is apparent in both ADMIXTURE (Figs. 1.2C and 4.4D) and PCA (Fig. 1.3D) plots. To confirm this signal and explore it in more depth, outgroup f_3 -statistics were used to compare the affinities of individual Irish Chalcolithic and Early Bronze Age samples to both the preceding Irish Neolithic population, as well as EHG and Yamnaya populations from Russia (Fig. 4.4A-C). For additional comparative purposes, individual ADMIXTURE proportions (K=12) from the analysis presented in Fig. 1.2C are plotted alongside the resulting f_3 -statistic values (Fig. 4.4D).

An immediate geographical distinction between the southwest and the remainder of the island is apparent for all f_3 -statistic tests. This divide is most visible when Yamnaya and EHG populations are considered (Fig. 4.4A-B), with the lowest levels of shared drift observed for individuals from the Burren (Poul nabrone, Roughan Hill and Coolnatullagh), Killuragh Cave and Labbacallee Wedge Tomb (Labbacallee213 excluded). A subtle reversal of this trend is seen when the Irish Neolithic population is tested (Fig. 4.4C), with highest affinities observed for Poul nabrone01 (Early Bronze Age) and Labbacallee212 (Chalcolithic).

Inflation of European Early Farmer (EEF) ancestry in southwestern individuals can also be observed in ADMIXTURE analysis (Fig. 4.4D), represented by the orange component. This EEF component accounts for a little less than a half of the ancestry in samples from Roughan Hill, Poul nabrone and Killuragh Cave. When scaled up to account for the European HG ancestry present in the Irish Neolithic population (~20%), a range of 56-58% Neolithic ancestry is indicated for these individuals, closely complementing the estimates of steppe ancestry calculated above. Labbacallee212 exhibits an even larger proportion of Neolithic ancestry (72% vs. 24% steppe-derived).

It is important to consider that these apparent geographical trends could also be driven by cultural (e.g. burial type) or chronological factors. Thus, it is worthy of note that when only directly dated Chalcolithic samples are considered a north-south division is still seen in all panels, with Labbacallee212 and Killuragh1 (2,402-2,213 and 2,344-2,036 cal BC), showing substantially less eastern introgression and higher levels of Neolithic ancestry relative to Pollnagollum86 and Pollnagollum911 (2,300-2,131 and 2349-2135 cal BC). This distinction is also clearly visible in PCA, where Killuragh1 and particularly Labbacallee212 place substantially closer to European Neolithic populations (Fig. 1.3D). Interestingly, two undated samples from Pollnagollum and Labbacallee both show a drop in Neolithic ancestry relative

to other samples from the same burial. As both sites have also yielded later Early Bronze Age radiocarbon determinations, there is a possibility these discrepancies may be driven by chronological discontinuity, which will require further dating to explore.

In the Burren, the three samples that show a definite decrease in steppe-related ancestry (two sisters from Roughan Hill and Poul nabrone01) all date towards the end of the Early Bronze Age period (1,943-1,663 and 2,028-1,758 cal BC). This strongly suggests a prolonged persistence of European Neolithic ancestry in this region of the island for many centuries after the transition to metallurgy. Importantly, these three Burren samples were all retrieved from reused or newly constructed megalithic burials, differentiating them from Early Bronze Age samples from other areas, which were taken from individualised inhumation burials (cist or flat grave). This distinction in funerary practice may be directly linked to their observable differences in ancestry. The affinities of the contemporaneous Coolnatullagh1005 individual (1,880-1,610 cal BC), retrieved from a singular cist burial in the Burren, supports such a possibility. This sample shows a more similar ADMIXTURE profile (Fig. 4.4D) and PCA placement (Fig. 1.3D) to other cist burials from further north and east, although somewhat reduced affinity to EHG and Yamnaya populations is still seen (Fig. 4.4A-B). If the ancestral distinction between these burial sites on the Burren is the result of cultural separation, the overall geographical trends observed here are likely to be maintained on denser sampling of the island, given the general restriction of megalithic wedge tombs to western Ireland, with highest concentrations seen in the southwest (Ó Nualláin 1989).

Only two later Bronze Age samples are included in this thesis, including a directly dated individual from the southwest coast (Inchagreenoge134). Notably, this sample (not shown in Fig. 4.4) does not exhibit the same excess of Neolithic ancestry seen in the earlier individuals of the region (36% Neolithic vs. 57% steppe-derived based on ADMIXTURE components).

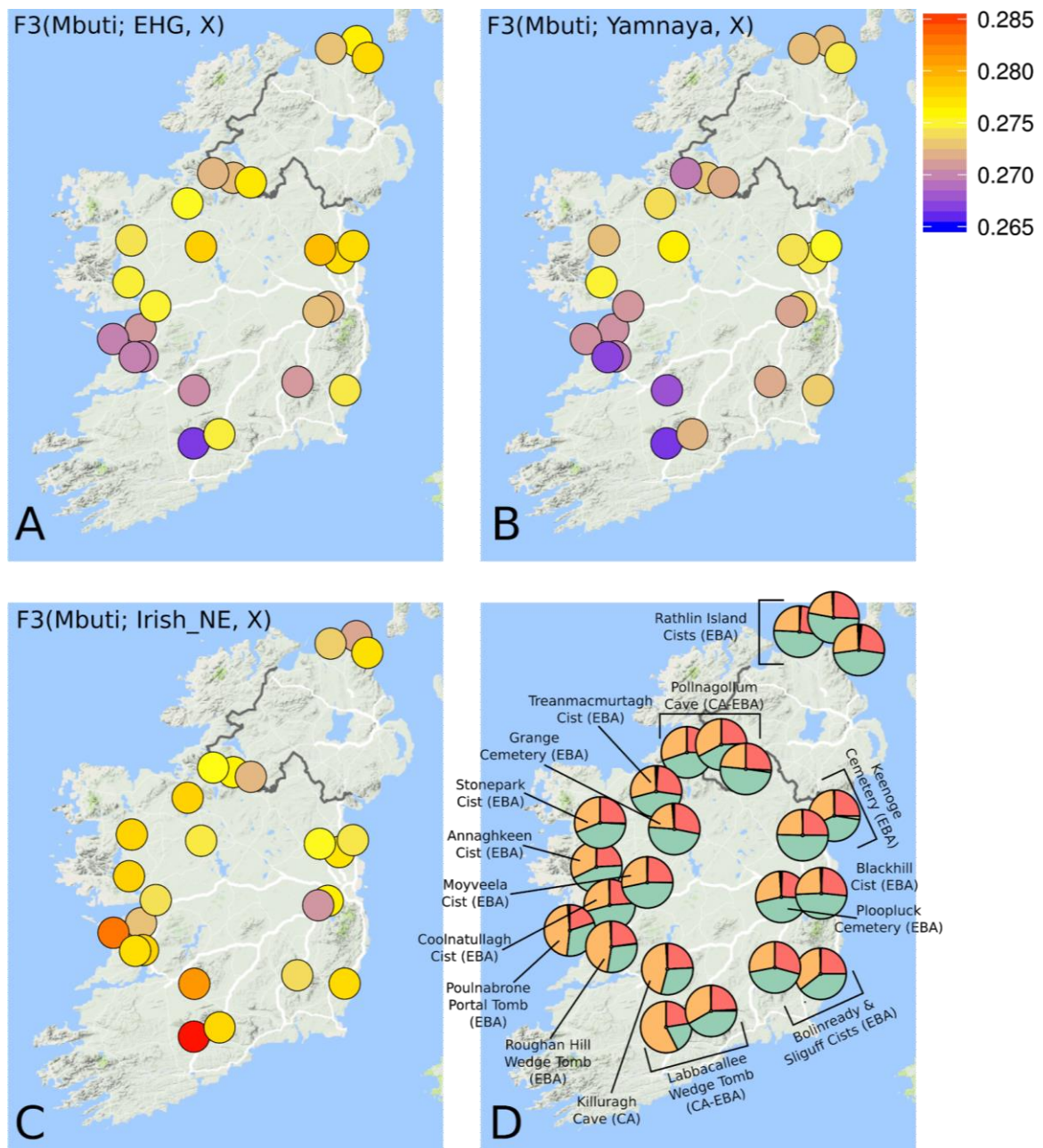


Figure 4.4. Steppe introgression into Irish Chalcolithic and Early Bronze Age individuals. Panels A-C show the level of shared genetic drift seen between Irish samples and EHG, Yamnaya and Irish Neolithic populations, measured using outgroup f_3 -statistics. Red indicates higher shared drift, with the same scale used for all panels. Panel D plots individual ADMIXTURE components for the same set of Irish samples, taken from the analysis shown in Fig. 1.2 ($K=12$). Red, orange and teal components are found at highest levels in WHG, European Neolithic and steppe populations respectively. Notably, a lower coverage individual from Keenoge and a related individual from Roughan Hill were not included in ADMIXTURE analysis.

3. Signals of local Neolithic continuity are present in the Irish, Iberian and Scandinavian peripheries of Europe.

The above sections suggest a significant residuum of Neolithic ancestry persists in later European and Irish Bronze Age populations, some more so than others, although the ultimate source of this ancestry in different regions (non-local vs. local) is not clear. Thus, the true extent of demographic replacement across the transition is difficult to ascertain. To explore whether preceding local farming populations (Irish_NE, TRB and Iberia_LNCA) contributed to the succeeding cultures of the same regions, outgroup f_3 -statistics of the form $f_3(\text{Mbuti}; X, \text{Neolithic/Chalcolithic Population})$ were used. The resulting values of shared genetic drift for a selection of Late Neolithic (North Central Europe and Scandinavia), Chalcolithic (Ireland and North Central Europe) and Bronze Age (Hungary, Iberia, Ireland, North Central Europe and Scandinavia) individuals were regressed against one another for the three populations, allowing the visualisation of geographical trends in affinities (Fig. 4.5). Importantly, it was possible to include a larger number of samples here than used for PCA and ADMIXTURE analysis, due to the increased size of the SNP set used (~2.7 million). Late Neolithic cultures with minimal levels of European Neolithic ancestry (Corded Ware and Battle-Axe) were excluded from analysis.

In all three regression plots a clear distinction can be seen between the Hungarian and Iberian populations relative to the remainder of the dataset. These samples share higher levels of drift with all Neolithic populations considered, reflecting their lower overall levels of steppe ancestry (Fig. 1.2C). Inflated sharing with Neolithic groups is also seen in a number of southwestern Irish individuals and is no doubt due to the same phenomenon. Surprisingly, three Swedish Bronze Age samples, not included in ADMIXTURE or PCA due to low genomic coverage, show some of the highest levels of shared drift with Neolithic populations. Importantly, the same pattern is not observed for samples of comparable coverage, suggesting this inflation is not simply an artefact of low genotype counts (>20,000 required for inclusion). Late Neolithic Swedish samples show more variability, while Danish samples fall at both extremes in each plot. This highlights the dynamicity of the Scandinavian region during the introduction of Corded Ware Culture in the Late Neolithic and its subsequent initiation into the European Bronze Age. The heightened affinity to Neolithic populations observed for some samples, as well as the higher than expected allelic sharing seen with TRB populations for a number of individuals, including the outliers Nordic_LN4 and Nordic_BA1, may indicate persistence and possible resurgence of local ancestry in Scandinavia across the Late Neolithic and Bronze Age, particularly in more remote Sweden, a pattern also reflected in Y chromosomal diversity.

The Irish population exhibits on average lower levels of allele sharing with Neolithic populations relative to Iberian, Scandinavian and Hungarian groups, clustering instead in the center of the plot alongside the majority of north central Europeans. They are most comparable to the Únětice Culture, with earlier Bell Beaker and German Late Neolithic individuals showing higher variability in affinities, including several individuals who show substantially deflated allele sharing with European Neolithic populations. When f_3 -statistic values for the tests $f_3(\text{Mbuti}; X, \text{Iberia_LNCA})$ and $f_3(\text{Mbuti}; X, \text{TRB})$ are regressed against

one another, Únětice and Irish populations, with the exception of southwestern samples, are indistinguishable. Strikingly however, the two groups are clearly pulled apart when the Irish_NE population are considered, with nearly all Irish Chalcolithic and Bronze Age samples falling above the regression line in comparisons to both TRB and Iberia_LNCA populations. The majority of continental samples place on or below the regression line for both plots, supporting some flow through of ancestry related to the Irish Neolithic into the island's later Chalcolithic and Bronze Age populations. Notably, samples from the southwest possess typically larger residuals than those from the rest of the island, suggesting potentially greater local input in this region, although without the inclusion of neighbouring British and French Neolithic populations, the ultimate source of this ancestry cannot be confirmed.

Iberian Bronze Age individuals also show some increased affinity to the Irish_NE population when values are regressed against those obtained for TRB. This trend is likely being driven by aversion to TRB, with the Irish Neolithic population acting as a better proxy for Iberian Neolithic ancestry, and the pattern reverses when Irish_NE and Iberia_LNCA are compared. In comparisons between TRB and Iberia_LNCA, Iberian Bronze Age samples show even greater divergence from the regression line and several clear outliers are identified. These patterns together suggest the large excess of Neolithic ancestry in the Iberian Bronze Age has some local contribution.

Finally, it is worth noting that individuals from north central Europe do not only vary widely in their absolute levels of allele sharing with Neolithic populations, but also in the size of their residuals, with particular variability seen in relation to the preceding TRB population. This stands in contrast to the Irish and Iberian Bronze Age clusters, individuals within which share similar trends in affinities, particularly as regards the earlier Iberia_LNCA and Irish_NE populations of their respective regions. This may simply be the result of Central Europe's position as a cultural and genetic melting pot during the Beaker and Bronze Age periods, with more continuous demographic turnover occurring here relative to the more peripheral regions.

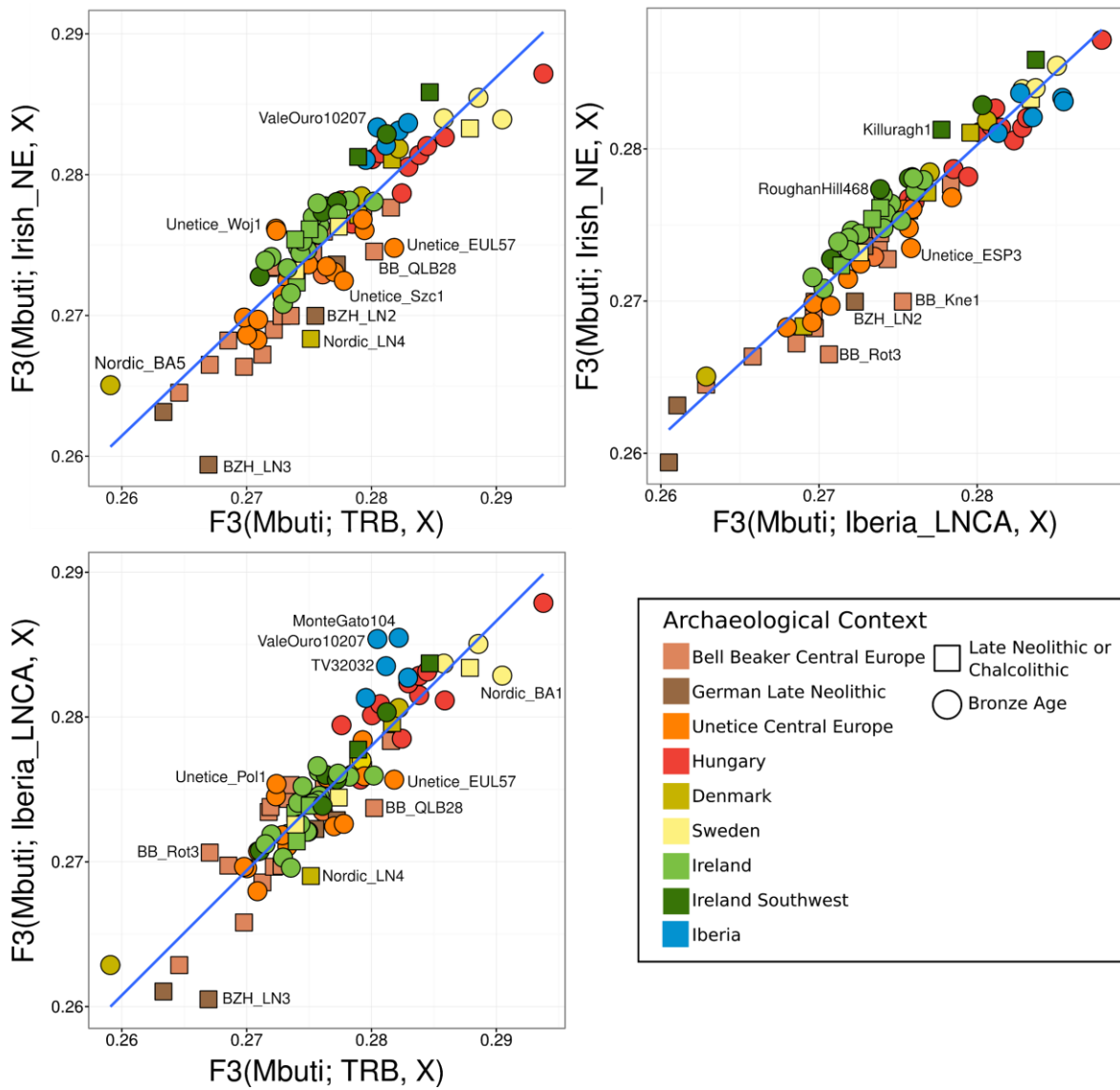


Figure 4.5. Patterns of allelic-sharing between Late Neolithic/Bronze Age individuals and preceding Neolithic populations. The f_3 -statistic values on the Y axes are regressed against those on the X. Outliers are identified and labelled based on studentised residuals (CI=95%).

Haplotypic exploration of ancient and modern populations from Britain and Ireland

Recent studies have utilised patterns of haplotypic sharing among modern British and Irish individuals to great effect, in order to explore finescale population structure within the islands (Leslie et al. 2015), Byrne et al. submitted). A similar analysis is performed on the same datasets here, with the novel addition of imputed genotypes from ancient Irish, British and northern European samples. The resulting components of variation, reveal not only spatial, but temporal trends in haplotypic affinity from the Irish Chalcolithic period onwards (Fig. 4.6). These are used here to visualise the ancestral similarities and differences found between and within Ireland's Early Bronze Age, Iron Age and modern populations.

Isolation at the British Peripheries

Several of the major principal components (PCs) identified explain little of the variation present in the ancient dataset. This includes the first and third PCs (Fig. 4.6A), which respectively segregate the populations of Orkney and Wales, while compressing the remainder of modern and ancient variation. The presence of strong haplotypic differentiation between these populations and the rest of Britain was reported in the original publication of the dataset (Leslie et al. 2015), and, in the case of Orkney, has been explained as the result of isolation and of Norse settlement. Notably, PC5 also solely describes Orcadian variation, and, alongside PC1, is not considered in further results and discussion.

The driving factor behind haplotypic divergence of Welsh populations is less clear. However, we note here that, alongside modern individuals from the border regions of Wales, the entire Irish Early Bronze Age population (several southwestern samples excluded) is also pulled away from the main cluster and in the direction of Welsh individuals on PC3 (Fig. 4.6A), suggesting they possess some haplotypic variation found in Wales that is absent in the remainder of the dataset. Surprisingly, the same increased affinity is not seen for the Iron Age Britons of Yorkshire and southeastern England, as may have been expected given both the persistence of Brittonic Language and culture in Wales after the Anglo-Saxon migrations, and the previously demonstrated affinity of the Yorkshire individuals to the modern Welsh population (Martiniano et al. 2016). This could indicate that the prolonged regional isolation of Wales, aided by its mountainous geography, stretches into the Bronze Age period, allowing the build up of the extensive haplotypic diversity seen in PC3.

The Establishment of Irish Haplotypic Diversity

The second component of variation (Fig. 4.6B-C) is unique in that it explains a large amount of variation present in both ancient and modern individuals. This corresponds to the primary split seen in fineSTRUCTURE analysis by Byrne et al. (submitted), which segregates Ireland and Britain into two distinct genetic islands, capturing what is defined as an Anglo-Celtic cline. Western Ireland and southeastern Britain form the two extremes of this component, with the Scottish population bridging the gap between the two clusters. Strikingly, ancient samples also separate out along this axis, with Irish

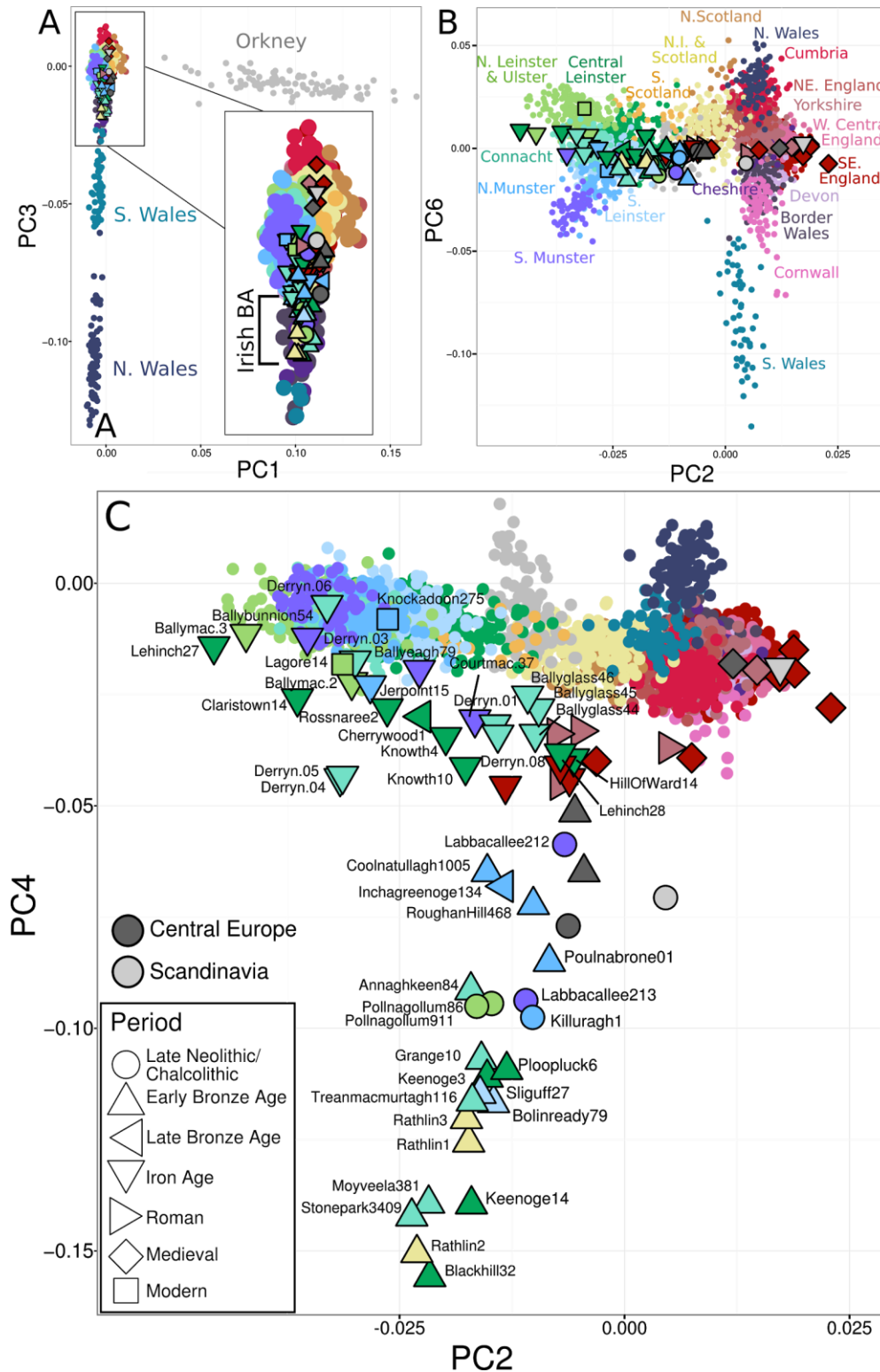


Figure 4.6. PCA of haplotypic similarity of ancient and modern individuals from Britain and Ireland. Modern individuals are coloured by geographical region (labelled in panel B), based on fineSTRUCTURE clustering assignments from Byrne et al. (submitted) Ancient individuals are outlined in black and coloured following the same geographical key. Six continental ancient samples are also included in grey. **A)** plots PC1 and PC3, which segregate the modern populations of Orkney and Wales from the remained of the dataset. A magnified image of ancient individuals is also shown. **B)** plots PC2 and PC6, which provide the most accurate geographical representation of the two islands. **C)** plots PC2 and PC4, the latter of which serves to distinguish Chalcolithic and Bronze Age samples from the remainder of the dataset.

individuals from both the Early Bronze Age and Iron Age periods falling further towards modern Irish variation than their British and continental counterparts (Fig. 4.6B-C).

Irish Iron Age samples extend the entire range of Irish variation on PC2, suggesting substantial continuity with the modern population. Irish Early Bronze Age samples show a more constricted distribution closer to the center of the plot, but still exhibit a systemic shift towards Irish Iron Age and modern populations, particularly those from individualised burials. The most parsimonious explanation for such observations is direct continuity between the Chalcolithic/Early Bronze Age and modern period in Ireland, with much of the haplotypic variation explained by PC2 forming in the intervening millennia, in a similar manner as suggested for Wales in PC3. While migration may be partially responsible for this structure, it is worth noting that the Irish Iron Age and modern population typically extends away, rather than towards, any potential external sources of variation in the dataset, including a contemporary Iron Age population from Britain, the most likely source of migration into Ireland between the Bronze Age and Early Christian periods. However, several exceptional Irish Iron Age samples exist, returned to in later sections.

The homogenisation of British population structure through admixture

In contrast to the gentle gradient of ancient Irish variation, British and continental individuals show a more punctuated distribution along PC2 (Fig. 4.6B-C), forming two clear clusters at both ends of modern British variation. Anglo-Saxons fall with southeastern English variation in this and all other PCs considered, alongside a Nordic Iron Age sample, reflecting the large genetic contribution of Germanic migrations to this part of the island (Leslie et al. 2015; Schiffels et al. 2016). Iron Age Britons comprise another tight grouping at the opposite end of British variation, emphasising the admixed nature of the modern population (Leslie et al. 2015; Martiniano et al. 2016; Schiffels et al. 2016). Early snapshots of continental introgression events may be represented by two samples that fall midway between the two groups, one from an Anglo-Saxon context (O3), which was reported as admixed in the original study (Schiffels et al. 2016), and the second from a Roman British population (6DT23), another member of which was demonstrated to be of likely Middle Eastern origin (Martiniano et al. 2016). Notably, no Irish Iron Age samples are seen to fall into this region of the PC space.

The compression of Iron Age British haplotypic variation close to the zero coordinate, relative to that of Ireland, suggests that PC2 may not effectively explain the majority of diversity present within this group, possibly due to their lack of representation within the larger admixed modern British cohort. In this respect, PC2 is perhaps best considered as explaining the distribution of Irish-related haplotypic variation in both modern and ancient individuals, which acts as somewhat of an imperfect proxy for Celtic ancestry in the neighbouring island of Britain, counterbalancing the Anglo-Saxon input. We caution that such a phenomenon may cause similar placement of individuals for unrelated demographic reasons. For example, the placement of Northern Irish and Scottish individuals between the two islands is proposed to be the result of numerous migrations in both directions, including the Gaelicisation of

Scotland *circa* 600 AD and the later Ulster plantations (Byrne et al. submitted). It is notable that no PC segregates Scotland from the rest of the dataset, suggesting the modern population has been mainly borne from admixture, rather than isolation, the reverse of what is proposed for Wales. Indeed, the more muted and systematic shift towards Irish variation of Welsh populations, whose diversity is better captured in PC3 and PC6, may represent more ancient shared Celtic ancestry between the groups. The tight clustering of three German Late Neolithic and Bronze Age individuals at the edge of ancient Irish variation, alongside the Iron Age British population, could also be due to a similar effect of older shared ancestry. Such an interpretation may find some temporal grounding in the differential placement of a Nordic Late Neolithic individual further towards the Germanic extreme of the plot.

Bimodal nature of Irish Chalcolithic and Bronze Age haplotypic variation

PC4 appears to capture much of the ancient haplotypic variation missing in PC2 (Fig. 4.6C), and exhibits a clear temporal trend. Modern variation is compressed around the zero mark, with Iron and Bronze Age individuals forming approximate layers along the component's axis. Most strikingly, ancient northern European populations from the Bronze Age onwards, hitherto near indistinguishable based on independent marker analysis despite wide geographical and chronological ranges (Figs. 1.2C-E and 1.3D-F), can now be clearly differentiated. This further emphasises the powerful potential of imputation and haplotypic methods for low coverage ancient data (Martiniano et al. 2017).

The majority of PC space is occupied by previously compressed Irish Chalcolithic and Early Bronze Age diversity, which exhibits clear geographical and chronological differentiation. All northern and eastern Early Bronze Age individuals segregate from both their southwestern contemporaries (Poulnabrone01, RoughanHill468 and Coolnatullagh1005), as well as earlier Chalcolithic samples from both the north and south of the island. Importantly, unlike previous divergences reported for independent marker analysis (Figs. 1.2C, 1.3D and 4.4), this pattern does not correlate with steppe ancestry or burial type. Indeed, both the southwestern Coolnatullagh1005 (cist burial) and Labbacallee213 (wedge tomb) individuals display substantial steppe-related input, as do the Chalcolithic individuals from Pollnagollum cave. Notably, the western cist burial of Annaghkeen also falls away from the majority of northern and eastern Bronze Age variation, despite a similar ancestral profile (Fig. 4.4D).

The clear haplotypic divergence between the Early Bronze Age population of Ireland and the preceding Chalcolithic, may reflect the archaeological discontinuity that occurred on the island during the transition and supports some migratory influx occurring alongside the uptake of more traditional elements of the Beaker package. The even larger divergence between these Early Bronze Age individuals and contemporaries from the southwest, suggests the geographical divide hinted at in the Chalcolithic on the basis of steppe ancestry (Fig. 4.4) was compounded, rather than reduced during the Bronze Age transition. Moreover, such an ancestral partition may have extended as far as the Late Bronze Age period, as evidenced by the close PCA placement of the southwestern Inchagreenoge134 (1,270-1,040 cal BC) nearby the neighbouring sites of Roughan Hill and Coolnatullagh (~1,950-1,600 BC). Unfortunately, the

scarcity of LBA remains from the remainder of the island makes such a signal impossible to confirm. Notably, the only other LBA individual considered (Cherrywood1), places within Irish Iron Age variation on both PCs 2 and 4 (Fig. 4.6C), necessitating a more cautious interpretation of the indirectly dated burial (1,400-1,000 cal BC), which is not considered further in results or discussion.

One potential confounding factor that must be considered in the above interpretations is the reproductive success of each ancient individual and their offspring, and how this may have impacted their haplotypic contributions to the modern day populations of the islands. While ultimately a stochastic process, variables such as social status, gender, health and longevity may all affect the likelihood of a single individual's haplotypes reaching high frequency in a descendent population. Although it is not possible gauge the extent to which these variables may be driving the patterns of haplotypic variation seen here, it is interesting to note that there is no clear divide between male and female Bronze Age burials on either PC2 or PC4.

Signals of Migration and Mobility in Iron Age Ireland

A millennium separates Inchagreenoge134 and the earliest Irish Iron Age burial (Knowth10), in which time considerable haplotypic variation is seen to have built up on the island. The Iron Age population also shows some correlation between PCs 2 and 4, showing a diagonal distribution between the British Iron Age population and modern individuals from the west of Ireland. Intriguingly, two contemporaneous individuals from the same burial site (Lehinch) form the extreme ends of this distribution, suggesting some level of mobility in the community to which they belonged. Lehinch28 and HillOfWard14 are in fact indistinguishable from British Iron Age variation, and may potentially represent recent arrivals from the neighbouring island.

The most likely candidate migrants from an archaeological perspective are two eastern individuals from Knowth (175-50 cal BC; 86-252 cal AD), whose burial rites are common in Britain and almost unknown in Ireland during the period (McGarry 2010). This interpretation is supported by their placement away from the main distributions of Irish Iron Age and modern variation. However, it must be noted that a number of other Irish Iron Age individuals place even further towards the British cluster, including three unrelated samples from Ballyglass Middle, the only Irish Iron Age site sampled that shows clear ancestral homogeneity among burials. The site itself is unusual in the relatively early date retrieved from unburnt bone (80-420 cal AD), at a time cremation was ubiquitous in Ireland. Two individuals from Derrynamanagh also show increased British affinities, but again the site shows wide differentiation on PCs 2 and 4, despite the close kinship among a number of samples (not included here). Previously undetected relatedness is also seen here between Derrynamanagh04 and Derrynamanagh05, later confirmed through IBD kinship analysis, which revealed the pair to be fourth degree relatives (Appendix II).

The overall lack of detectable geographical structure in the Irish Iron Age population, alongside the large minority of samples falling close to or within British Iron Age variation, must be considered in the context of the Late Iron Age period, a time when new Christian-type burial rites were being introduced to the island and substantial demographic upheaval was occurring in both Britain and Europe. In this respect, it must be emphasised that the selection of unburnt burials for aDNA analysis may be biasing sampling to more mobile communities with connections to overseas populations. Unfortunately, as for the Late Bronze Age sample set, the ubiquity of cremation burial in the preceding periods makes such potential confounders difficult to ascertain.

Haplotypic diversification after the Iron Age

PC6 shows the reverse trend to PC4, compressing ancient variation along the zero line, while allowing modern variation from all populations to fan out across the axis (Fig. 4.6B). This spread of modern haplotypic diversity shows something of a north-south trend, as identified in Byrne et al. (submitted). South Welsh and Cornish populations exhibit the largest amount of haplotypic variation and are followed on the axis by populations from southern Ireland, Devon and Border regions of Wales. Populations from the northern regions of Wales, England and Ireland, as well as Scottish groups, form the other extreme of PC6, with compression of the eastern populations of both islands apparent due to the homogenising effects of Anglo-related admixture (Byrne et al. submitted).

The clustering of ancient samples along the zero line suggests that, in Ireland at least, the majority of the geographical variation captured by PC6 postdates the Iron Age period. That said, a subtle shift towards northern groups is apparent in the Irish Iron Age, relative to both the preceding Bronze Age and British Iron Age. This is particularly apparent for individuals falling further towards Irish modern variation on PC2, and suggests some of the diversification captured by PC6 was already underway at this point in time. Two early modern Irish individuals from the Plantation period, postdating the Late Iron Age by roughly a millennia, are also plotted here and show extremely similar placement to modern individuals from the same regions, falling further 'north' and 'south' of the preceding axis of Iron Age samples. Further sampling of the British Iron Age and Medieval periods, specifically in Wales, will be required to interpret how such patterns are related to the clear divergence of northern and southern Celtic-speaking populations on the neighbouring island of Britain. Indeed, this novel preliminary analysis highlights the powerful temporal anchors ancient genomes can provide to spatial trends of regional genetic variation.

Further signals of change and continuity within the Irish population

Increased complexity of mating systems

Exploration of runs of homozygosity (ROH) in Irish Copper, Bronze and Iron Age individuals allowed for the detection of signals of inbreeding. Two measures were chosen, the length of the longest tract of homozygosity observed, and the average length of ROH tracts across the genome, which are plotted against one another in Fig. 4.7. A level of correlation is seen between the two values, with differential clustering of Copper, Bronze and Iron Age individuals apparent.

The Copper and Bronze Age populations possess longer average tract lengths, relative to the majority of Iron Age individuals, suggestive of larger, outbreeding mating networks in the later period. Furthermore, of the ten longest ROH tracts identified in separate samples, five belong to Chalcolithic individuals. This could indicate some population constriction occurred during the arrival of new metallurgical communities to the island. One southwestern individual in particular, Killuragh1, shows strong evidence of inflated parental relatedness. Notably, two other southern outliers are also seen in the Early Bronze Age period (Coolnatullagh1005 and Bolinready79). Importantly, such signals may also be the result of culturally-imposed consanguinity (e.g social caste systems, cousin marriage), rather than an absolute reduction in census size.

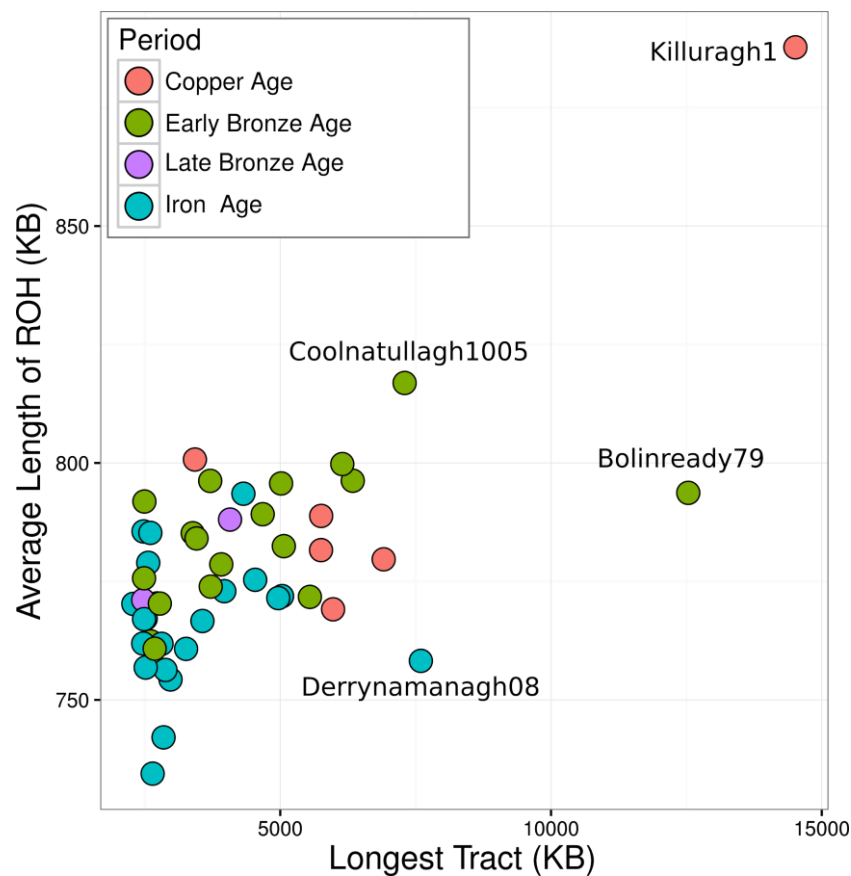


Figure 4.7. Signals of parental relatedness in Irish Copper, Bronze and Iron Age individuals. The length of the longest identified ROH tract is plotted against the average length of ROH for each individual genome.

A preliminary scan for signatures of selection between the Irish Bronze and Iron Ages

Given the demonstrable genetic continuity that has occurred in Ireland since the Early Bronze Age period, evidenced by haplotypic analysis (Figs. 1.4 and 4.6), we ask whether any variants show larger than expected differences in F_{ST} between the Chalcolithic/Bronze Age ($n=26$) and Iron Age ($n=22$) populations. Only three clear signals related to known genes are apparent, highlighted in Fig. 4.8.

The first of these is the *LCT* gene, which codes for the lactase protein required for the digestion of milk. In humans, lactase transcription is typically downregulated after weaning. However, two mutations in the upstream regulatory region of the gene have been linked to lactase persistence in certain populations. In Europeans, a single C to T change at rs4988235 is responsible for the phenomenon, allowing the consumption of milk into adulthood (Enattah et al. 2002). Frequencies are highest in northern European populations, with the phenotype reaching incidences of 89–96% in the populations of Scandinavia, Ireland and Britain (Ingram et al. 2009; Itan et al. 2010; Gerbault et al. 2011) and typical estimates of the selection coefficient associated with the locus in these populations range between 0.07–0.1 (Tishkoff et al. 2007; Itan et al. 2010). The mutation has been reported in a 12,000 year old Mesolithic European, although it was not found on the typical lactase persistence haplotype (Mathieson et al. 2017), as well as in Early Neolithic populations from Sweden and Spain (Malmstrom et al. 2010; Plantinga et al. 2012). Notably, it is absent in large samples of Neolithic Anatolians, Hungarians and LBK individuals (Burger et al. 2007; Gamba et al. 2014; Mathieson et al. 2015), as well as Yamnaya and other early steppe populations (Allentoft et al. 2015; Mathieson et al. 2015), suggesting that it existed as low frequency standing variation for many millennia before selection occurred.

To further explore the signal of selection around the *LCT* locus between the Bronze and Iron Age, frequencies for the derived allele at rs4988235 were calculated for the populations of both periods (Fig. 4.9). Strikingly, only three incidences of the mutation are found in the 52 Copper/Bronze Age chromosomes considered, one of which is found in the Cherrywood1 sample of questionable date. In contrast, frequencies close to 60% are seen in the Iron Age population. This suggests, that a sweep occurred within the island of Ireland itself after the establishment of the Bronze Age population. This resonates strongly with the massive emphasis placed on cattle herding in the earliest recorded Irish folklore and annals, a tradition that still has major bearing on the island's modern economy (O'Connell et al. 2016).

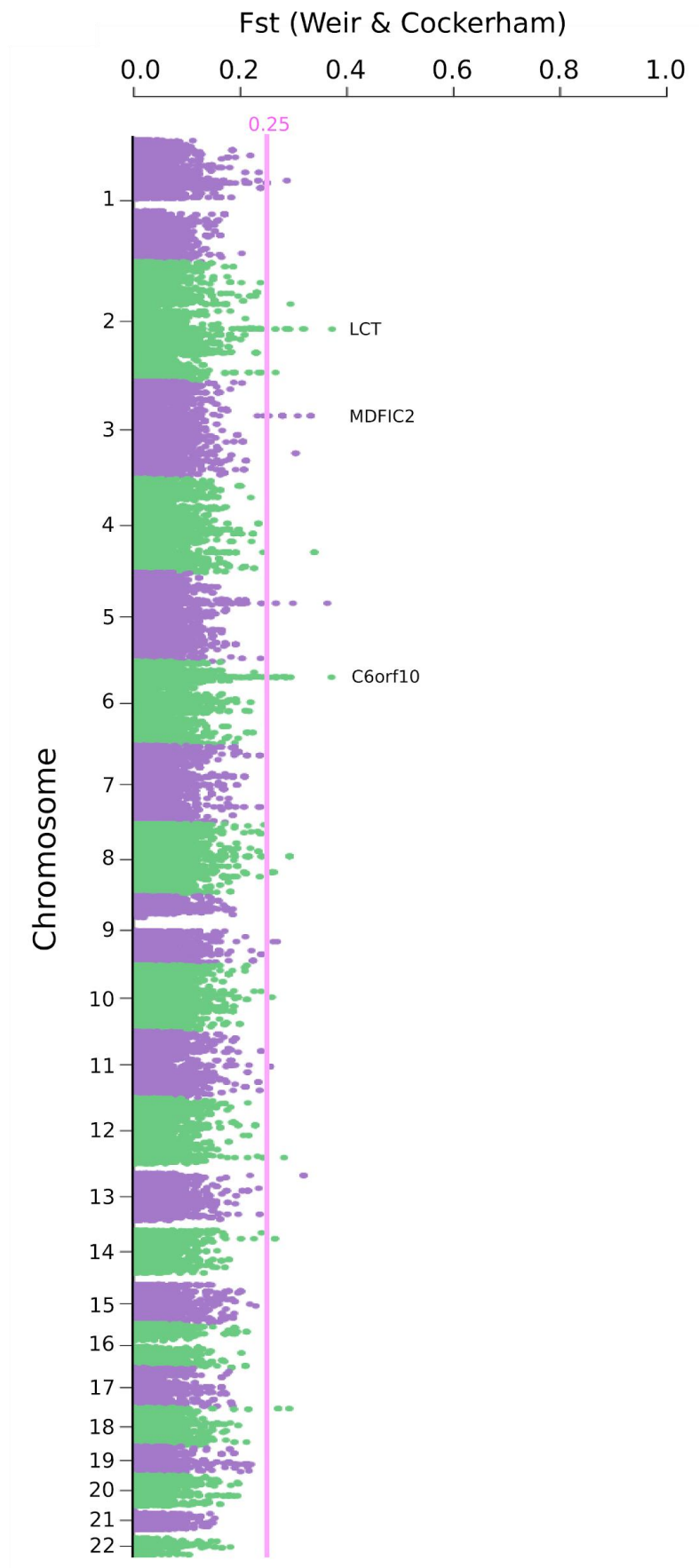


Figure 4.8. Distribution of F_{ST} values between Irish Chalcolithic/Bronze Age as a function of genomic position. A total of 848,715 biallelic SNP sites are considered.

The second signal is related to the C6orf10 gene, which codes for an uncharacterised protein. Notably, the locus falls within 70 Kb of HLA-DRA, a MHC class II antigen, which may indicate an immune function. The third signal is found within the MDFIC2 locus (MyoD family inhibitor domain containing 2), also a protein-coding region with no established function.

Other phenotypes of interest

Ireland is also unusual in displaying world maximum frequencies of a number of important genetic variants, including those involved in cystic fibrosis and haemochromatosis (Devaney et al, 2003; Distant et al. 2004). Interestingly, both Chalcolithic/Bronze And Iron Age populations displayed similar frequencies for the C282Y mutation (~13-18%), responsible for the most penetrative type of hemochromatosis, a recessive iron-retention disorder (Fig. 4.9). These estimates are very close to that of the modern Irish population (~15%), where the mutation enjoys its peak frequency today (Byrnes et al. 2001; Lucotte et al. 2003). Notably, the mutation has not yet been reported in any other ancient population, and was not detected here in any imputed individuals outside of Ireland until the British Iron Age and Anglo-Saxon periods, where it features at similar frequencies to the Irish Iron Age (~16-18%).

That this seemingly harmful mutation has been maintained at high frequency in the Irish population for over 4,000 years is strongly indicative of some underlying positive effect, perhaps in the form of a heterozygote advantage, as has previously been suggested (Motulsky 1979). Hypothetical advantages include adaption to changes in dietary iron and immunity from infection, while its proximity to the HLA region has also led to the suggestion of genetic hitchhiking (Distant et al. 2004). The prevalence another HFE mutation (H63D) in the Irish Neolithic population (33%), still common in Europe today (10-20%) and responsible for a milder form of the disease, argues against a simple hitch-hiking effect, instead indicating there may have been some requirement on iron-retention itself. The H63D mutation is also prevalent in continental farming groups and is found in two WHG individuals (KO1 and Canes1). Given the divergent sustenance strategies between the two groups, the potential selective pressure may not be dietary-related.

We also explored changes in skin pigmentation between the Irish Chalcolithic/Bronze Age and Iron Age periods, with respect to both modern Ireland and the preceding Neolithic and Mesolithic periods. This revealed an overall trend in skin lightening across Irish prehistory. As reported in Chapter Three, both Mesolithic individuals were homozygous for the ancestral alleles at the two major known variants responsible for European skin depigmentation, rs1426654 (SLC24A5) and rs16891982 (SLC45A2) (Lamason et al. 2005; Soejima & Koda 2007). The later Neolithic cohort shows a sharp increase in the derived allele at rs1426654 (>90%), which is at near fixation in the majority of European populations today (Fig. 4.9) and also reaches appreciable frequencies in the many North African, Middle Eastern, South Asian and Central Asian populations. No incidences of the ancestral allele for rs1426654 are found in the Chalcolithic/Bronze Age and Iron Age cohorts.

The derived allele at rs16891982 shows a more varied in modern Europe, existing on a north-south cline, and is associated with lighter skin, hair and eye colour (Norton et al. 2007). It is known to have been under strong positive selection over the past 5,000 years, alongside other alleles involved in depigmentation (HERC2 and TYR), suggested to be a response to constraints on vitamin D synthesis in high latitude regions, which may have been compounded by dietary changes in the Neolithic to vitamin D-poor food (Wilde et al. 2014). Strikingly, while a change in the derived allele frequency at rs16891982 is seen between the Neolithic and Bronze Age periods (57% to 35%; Fig. 4.9), likely due to population turnover, an even sharper decrease is seen in the two thousand years between the Early Bronze Age and Late Iron Age (35% to 2%), an interval with little evidence of migratory influx. This is suggestive of some selection pressure for depigmentation acting upon the Irish population over the period.

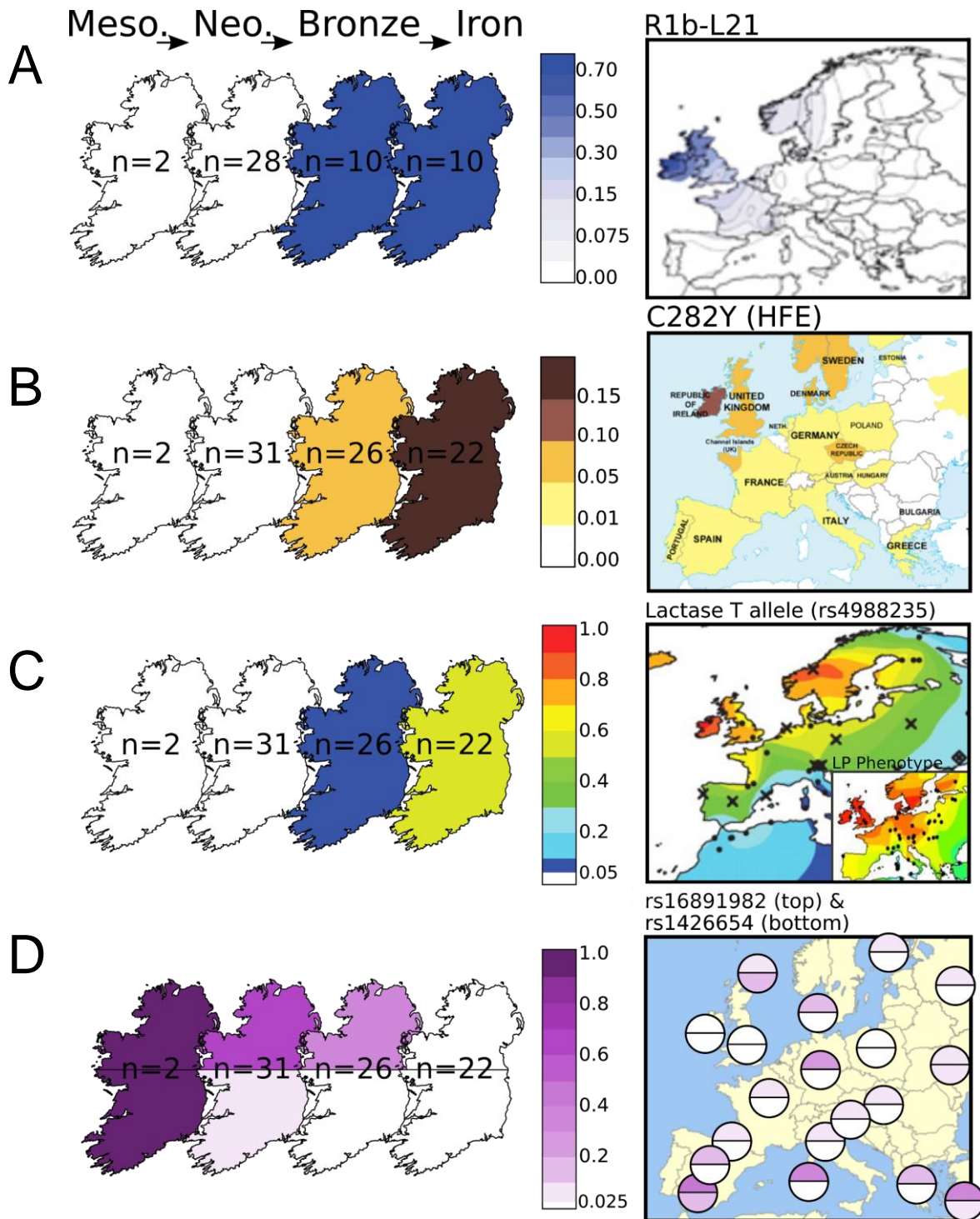


Figure 4.9. Establishment of variants at maximums in the modern Irish population. Frequencies for four key Irish prehistoric periods are shown, alongside the modern European distribution (same colour keys used). The number individuals used for frequency estimation is indicated for each ancient populations. **A)** Frequency of Y chromosome marker L21 (Map Reference: Myres et al. 2011). **B)** Frequency of the C282Y mutation associated with a severe form of haemochromatosis (Map Reference: Distanto et al. 2004). **C)** Frequency of the derived allele at rs4988235, responsible for lactase persistence in Europeans (Map Reference: Gerbault et al. 2011). **D)** Frequencies of the ancestral alleles at two SNPs responsible for European skin depigmentation, rs1426654 (SLC24A5) and rs16891982 (SLC45A2) (Map Reference: Rajeevan et al. 2012).

Conclusions

This conclusion attempts to address some of the key questions surrounding the Chalcolithic and Early Bronze Age transitions in Ireland and the later Iron Age period, in light of the novel results presented here. As with Chapter Four, the overarching aim is not to provide solid answers to these, but thoughtful discussion, which can go on to inform further archaeological and genomic research.

Did multiple migrations into Ireland occur during the Chalcolithic and Early Bronze Age?

As with the introduction of farming, Ireland sat between two overlapping zones of influence during the demographic influxes of the Chalcolithic and Early Bronze Age (Chapter Two), with aspects of both Atlantic and Northern European Beaker communities apparent in the island's material culture. These migrations find their ultimate source in pastoralist populations occupying the steppe regions above the Black Sea, whose distinctive Caucasus-related ancestry washed into Europe during the third millennium BC and contributed significantly to the genetics of northern populations (Allentoft et al. 2015; Haak et al. 2015), likely via the vector of Corded Ware culture. Early metallurgical technologies in Ireland derive from a more western source, arriving from Iberia via the maritime and river networks of Atlantic Europe. These networks also provided the avenue by which the Bell Beaker cultural phenomenon spread and diversified, hybridising with Corded Ware culture in more northern regions and giving rise to the classic individualised Beaker burial, which entered Britain *circa* 2,500 BC. During the same period, Beaker pottery, copper metallurgy and a new type of megalith appeared in Ireland, with the classic Beaker Burial traditions common in Britain only arriving on the island several centuries later, near the start of the Early Bronze Age, and mainly in the north and east.

A key question is whether this later introduction of individualised inhumation burial was the result of population migration (Mallory 2013), a scenario which finds substantial support in the haplotype-based analysis presented here (Fig. 4.6C). Early Bronze Age burials from the northern and eastern halves of the island clearly cluster with one another, segregating from both temporal contemporaries in the southwest and earlier Chalcolithic samples from the north and south, suggesting some higher degree of relatedness among these individuals. Importantly, this differentiation is not driven by levels of steppe introgression, as demonstrated by the divergence of southwestern and Chalcolithic individuals with inflated steppe ancestry (Labbacallee213, Coolnatullagh1005, Pollnagollum86 and Pollnagollum911) from the main Early Bronze Age cluster. It is then important to consider that the steppe ancestry present in both Chalcolithic Ireland and later southwestern individuals may derive from different sources than that present in the majority of sampled Early Bronze Age individual burials.

How big were the migratory influxes and where were their geographical origins?

The possibility of a secondary migration into Ireland near the start of the Early Bronze Age, several centuries after the initiation of the Chalcolithic, must be taken into account as a confounding variable when interpreting the uneven distribution of steppe-related ancestry on the island, which shows a clear

reduction in regions of the island furthest from Britain in both periods (Fig. 4.4). Specifically, three key factors warrant consideration. Firstly, how much steppe-related ancestry was present in the different migratory source populations? Secondly, what were the levels of migratory influx and to what extent were migrant ancestries diluted due to admixture with indigenous Irish groups? Finally, is the level of this dilution dependent on geographical or cultural factors, e.g. burial rite, which may be correlated themselves (Fig. 4.1)?

Unfortunately, no published genomic data exists from contemporary sites in northern France and Britain, the most likely source zones of Beaker culture and steppe ancestry in Ireland, making these questions difficult to address at present. However, large levels of steppe ancestry have been reported in the British Beaker population, with some reduction seen in the Bronze Age period, most likely the result of homogenisation with indigenous Neolithic groups (Olalde et al. 2017). This Beaker population also has high frequencies of R1b-L21 (Fig. 4.2.), making it not only a prime geographical and archaeological candidate for the new burial traditions of the Irish Early Bronze Age, but also a suitable genetic one. As we are currently unable to directly compare levels of steppe-related ancestry in the British and Irish populations, it is impossible to determine whether any dilution of this ancestry occurred upon entry to Ireland. Outgroup f_3 -statistics suggest some increased affinity between Irish Early Bronze Age individuals and the island's preceding Neolithic population, relative to continental groups (Fig. 4.5), however a comparative British Neolithic population will be required to confirm whether this is indeed the result of local Irish rather than British admixture. Thus, while the Irish Early Bronze Age population in the north and east derive roughly up to half of their ancestry from European Neolithic groups, the contribution of the Irish Neolithic population may be much lower.

The source(s) and scale of migration during the Irish Chalcolithic are somewhat more enigmatic. While Britain remains a candidate, the prevalence of megalithic burial and early appearance of copper mining in the southwest of the island suggests some migratory input may have occurred from Brittany. If the majority of Irish Chalcolithic steppe introgression is derived from Beaker groups in Britain then substantial dilution of this ancestry must have occurred upon entry to Ireland, given the large residuals of Neolithic ancestry seen in samples such as Labbacallee212 and Killuragh1 (56-72%). However, if the small steppe input seen in these individuals is derived from a yet unsampled Atlantic region, such as Brittany, the scale of migration is much harder to surmise, given the unknown ancestral profiles of these communities. That said, the inflated affinity of Killuragh1 to the Irish Neolithic population (Fig. 4.5), and his possession of a Y chromosome lineage common in Late Neolithic Ireland, may suggest some indigenous input.

The comparatively high levels of steppe ancestry in the northern Pollnagollum samples (Fig. 4.4), alongside the presence of R1b-L21, emphasise the expected heterogeneity of the Irish Chalcolithic population. Indeed, the entire period lasted at most several centuries, which is too short a timeframe for complete genetic homogenisation to occur. Importantly, migration from potentially different continental

and British regions may have been continuous, as Ireland became a hub of metallurgical production. It is then hard to determine whether the few individuals sampled here are derived predominantly from the larger Irish population at the time or more recent migrant communities, a factor which may be dependent on cultural variables such as burial rite. Moreover, the slightly inflated signals of inbreeding in Chalcolithic samples (Fig. 4.7), particularly Killuragh1, could be the result of demographic or societal stress in both indigenous and/or migrant communities.

Intriguingly, the low levels of steppe introgression seen in the two southern Chalcolithic samples is mirrored in three later individuals (two of whom are sisters) from megalithic contexts in the southwest (Fig. 4.4), who date to the end of the Early Bronze Age, resonating with the lower penetrance of cist and pit burials in these regions (Fig. 4.1). It is again unclear whether this inflation of Neolithic ancestry represents a survival of indigenous Irish communities or whether it is mostly derived from other Atlantic populations. However, as with preceding Chalcolithic, comparative outgroup f_3 -statistics do suggest some excess of ancestry related to the Irish Neolithic (Fig. 4.5). The differentiation in terms of steppe introgression between these samples and those from contemporaneous cist and pit burials across the island hints that the differing funerary traditions may be related to differing ancestral and cultural identities that persisted throughout the Early Bronze Age period, a prospect which can be further explored through the sampling of wedge tombs and reused megaliths in regions where individualised cist and pit burials are also common.

Interestingly, the individual from the Coolnatullagh cist burial, which is located within 15 km of the Roughan Hill and Poul nabrone megaliths, also shows inflated steppe-related ancestry relative to these samples (Fig. 4.4), despite his close haplotypic affinities to RoughanHill468 (Fig. 4.6). The lack of ancestral homogenisation in such a small region suggests either ongoing migration into the Burren during the period, or persistent social/cultural barriers to gene flow. The phenomenon of genetically differentiated groups living in close proximity or within the same society for many generations has already demonstrated in modern populations (Basu et al. 2015; Van Dorp et al. 2015) and its occurrence in ancient societies is not unfeasible. As no surviving endogenous DNA has yet been retrieved from cremated bone, such prospects are as daunting as they are intriguing and present a significant hurdle in the understanding of both the Irish Bronze and Iron Ages, particularly the periods of obvious cultural flux.

What changes occurred in the Irish population after its establishment in the Early Bronze Age?

Chapter Two established haplotypic continuity between Irish Early Bronze Age and modern population, supported strongly by male-line lineage frequencies. This continuity is also demonstrated in the more detailed haplotype-sharing analysis presented here (Fig. 4.6), although diversification and loss of haplotypic variation is apparent between both the Bronze and Iron Age, and the Iron Age and modern era. This emphasises the more recent population history that is captured by patterns of linkage disequilibrium in the genome and the importance of these methods in future surveys of ancient genetic

variation. Notably, the Irish Iron Age population shows substantially more haplotypic similarity with the modern Irish population, relative to the preceding Early Bronze Age. This increased similarity is also reflected in the frequencies of certain phenotypic variants involved in depigmentation and lactose tolerance (Fig. 4.9), which rose in frequency through time, allowing the UV-starved population to exploit much needed vitamins and nutrition from both the sun and milk. Other characteristic traits of the modern Irish population have remained stable across the millennia, such as the C282Y mutation, responsible for iron-retention disease, and the R1b-L21 lineage (specifically R1b-DF13).

Throughout prehistory mating networks in Ireland appear to have increased in complexity, leading to a decrease in inbreeding (Fig. 4.7), potentially the result of population growth. Some level of migration is also likely to have occurred into the island between the Early Bronze Age and Iron Age, as has been demonstrated between the Iron Age and modern era (Byrne et al. 2017). However, the lack of available unburned remains between these periods makes such possibilities hard to explore. Indeed, the cohort of Late Iron Age burials sampled here may represent a snapshot of a population in the throes of a new demographic flux, associated with the spread of Christianity during the rise and fall of Roman Britain. The sudden re-emergence of inhumation rituals after several millennia of near-universal crematory practices, again emphasises the potential of burial rite to bias aDNA sampling schemes. Iron age samples show no apparent geographic structure, relative to either each other or the modern Irish population, while a number of sites (e.g. Lehinch and Derrynamanagh) contain individuals with diverse haplotypic ancestries. Denser sampling of both the Iron Age and Early Medieval periods of Ireland and Britain may consolidate these patterns, and help relate them to the lineage histories and geographic distributions of various Y chromosome haplogroups and other genetic variants of interest.

References

- 1000 Genomes Project Consortium et al., 2015. A global reference for human genetic variation. *Nature*, 526(7571), pp.68–74.
- Allentoft, M.E. et al., 2015. Population genomics of Bronze Age Eurasia. *Nature*, 522(7555), pp.167–172.
- Anthony, D.W., 2007. *The Horse, the Wheel and Language. How Bronze-Age Riders from the Eurasian Steppes Shaped the Modern World*. New Jersey: Princeton University Press.
- Basu, A. et al., 2015. Genomic reconstruction of the history of extant populations of India reveals five distinct ancestral components and a complex structure. *Proceedings of the National Academy of Sciences of the United States of America*, 113(6), pp.1594–1599.
- Bayliss, A. & O’Sullivan, M., 2013. Interpreting chronologies for the Mound of the Hostages, Tara and its contemporary contexts in Neolithic and Bronze Age Ireland. In M. O’Sullivan, C. Scarre & M. Doyle (eds.), *Tara—from the past to the future: Towards a new research agenda*. Bray, Ireland: Wordwell, pp.26–104.
- Brindley, A.L., 2007. *The dating of food vessels and urns in Ireland*. Galway, Ireland: NUIG.
- Browning, S.R. & Browning, B.L., 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localised haplotype clustering. *American Journal of Human Genetics*, 81(5), pp.1084–1097.
- Burger, J. et al., 2007. Absence of the lactase-persistence-associated allele in early Neolithic Europeans. *Proceedings of the National Academy of Sciences of the United States of America*, 104, pp.3736–3741.
- Burgess, C., 1979. The background of early metalworking in Ireland and Britain. In M. Ryan (ed.), *The Origins of Metallurgy in Atlantic Europe = Proceedings of the Fifth Atlantic Colloquium, Dublin 1979*. Dublin: The Stationery Office.
- Byrne, R.P. et al., 2017 Celtic population structure and genomic footprints of migration. Manuscript submitted for publication.
- Byrnes V. et al., 2001. Genetic hemochromatosis, a Celtic disease: Is it now time for population screening? *Genet Test*, 5(2), pp.127–130.
- Carlin, N., 2011. Into the West: placing Beakers within their Irish contexts. In G. Kirkham & A.M. Jones (eds.), *Beyond the core: Reflections on regionality in prehistory*. Oxford: Oxbow, pp.87–100.
- Case, H., 1995. Irish Beakers in their European context. In J. Waddell & E. Shee Twohig (eds.), *Ireland in the Bronze Age: Proceedings of the Dublin Conferenc, April 1995*. Dublin: The Stationery Office, pp.14–29.
- Cassidy, L.M. et al., 2016. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proceedings of the National Academy of Sciences of the United States of America*, 113(2), pp.368–373.
- Chang, C.C. et al., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4, p.7.
- Chernykh, E.N., 1992. *Ancient Metallurgy in the USSR: The Early Metal Age*. Cambridge: CUP Archive.
- Chernykh, E.N., 2008. Formation of the Eurasian “steppe belt” of stockbreeding cultures: Viewed through the prism of archaeometallurgy and radiocarbon dating. *Archaeology, Ethnology and Anthropology of Eurasia*, 35(3), pp.36–53.
- Cunliffe, B.W., 2008. *Europe between the Oceans: 9000 BC-AD 1000*. New Haven-London: Yale University Press.
- Cunliffe, B.W., 2013. *Britain Begins*. Oxford: Oxford University Press.
- Cunliffe, B.W. 2015. *By Steppe, Desert, and Ocean: The Birth of Eurasia*. Oxford: Oxford University Press.
- Czebreszuk, J., 2014. *Similar but Different: Bell Beakers in Europe*. Netherlands: Sidestone Press.
- Delaneau, O., Marchini, J. & Zagury, J.-F., 2011. A linear complexity phasing method for thousands of genomes. *Nature Methods*, 9(2), pp.179–181.
- Devaney, J. et al., 2003. Cystic fibrosis mutation frequencies in an Irish population. *Clin Genet*, 63(2), pp.121–125.
- Dillon, M. & Chadwick, N.K., 1967. *The Celtic Realms*. London: Weidenfeld & Nicolson.
- Distante S., et al., 2004. The origin and spread of the HFE-C282Y haemochromatosis mutation. *Human Genetics*, 115(4), pp.269–279.
- Enattah, N.S. et al., 2002. Identification of a variant associated with adult-type hypolactasia. *Nature Genetics*, 30(2), pp.233–237.
- Fitzpatrick, A.P., 2013a. *The Amesbury Archer and the Boscombe Bowmen: Bell Beaker burials at Boscombe Down, Amesbury, Wiltshire*. England: Wessex Archaeology.
- Fitzpatrick, A.P., 2013b. The arrival of the Bell Beaker Set in Britain and Ireland. In J.T. Koch & B.W. Cunliffe (eds.), *Celtic from the West 2*. Oxford: Oxbow Books, pp.41–70.
- Flanagan, L., 1998. *Ancient Ireland: Life Before the Celts*. Dublin: Gill & Macmillan Ltd.
- Gamba, C. et al., 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nature Communications*, 5, p.5257.
- Gerbault, P. et al., 2011. Evolution of lactase persistence: an example of human niche construction. *Philosophical transactions of the Royal Society B: Biological Sciences*, 366(1566), pp.863–877.

- Guilaine, J.F., et al., 2001. Campaniformes et environnement culturel en France méditerranée. In F. Nicolis (ed.), *Bell beakers today: Pottery, People, Culture, Symbols in prehistoric Europe*, 1, pp.229–275.
- Günther, T. et al., 2015. Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proceedings of the National Academy of Sciences*, 112(38), pp.11917–11922.
- Gurdasani, D. et al., 2015. The African Genome Variation Project shapes medical genetics in Africa. *Nature*, 517(7534), pp.327–332.
- Haak, W. et al., 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522(7555), pp.207–211.
- Halsall, G., 2007. *Barbarian Migrations and the Roman West, 376–568*. Cambridge: Cambridge University Press.
- Harari, Y.N., 2014. *Sapiens: A brief history of humankind*. London: Harvill Secker.
- Harrison, R. & Heyd, V., 2007. The transformation of Europe in the third millennium BC: the example of “Le Petit-Chasseur I+ III”(Sion, Valais, Switzerland). *Præhistorische Zeitschrift*, 82(2), pp.129–214.
- Herity, M. & Eogan, G., 1977. *Ireland in Prehistory*. London: Routledge and Kegan Paul.
- Heyd, V., 2017. Kossinna’s smile. *Antiquity*, 91(356), pp.348–359.
- Hill, E.W., Jobling, M.A. & Bradley, D.G., 2000. Y-chromosome variation and Irish origins. *Nature*, 404(6776), pp.351–352.
- Ingram, C.J.E. et al., 2009. Multiple rare variants as a cause of a common phenotype: several different lactase persistence associated alleles in a single ethnic group. *Journal of molecular evolution*, 69(6), pp.579–588.
- Itan, Y. et al., 2010. A worldwide correlation of lactase persistence phenotype and genotypes. *BMC Evolutionary Biology*, 10, p.36.
- Jeunesse, C., 2015. The dogma of the Iberian origin of the Bell Beaker: attempting its deconstruction. *Journal of Neolithic Archaeology*, pp.158–166.
- Jones, C., Mcveigh, T.O. & Ó Maoldúin R., 2015. Monuments landscape and identity in Chalcolithic Ireland. *Landscape and Identity: Archaeology and Human Geography*, pp.3–25.
- Jones, E.R. et al., 2017. The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Current Biology*, 27(4), pp.576–582.
- Jones, E.R. et al., 2015. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications*, 6, p.8912.
- Kristiansen, K. et al., 2017. Re-theorising mobility and the formation of culture and language among the Corded Ware Culture in Europe. *Antiquity*, 91(356), pp.334–347.
- Kristiansen, K. & Larsson, T.B., 2005. *The Rise of Bronze Age Society: Travels, Transmissions and Transformations*. Cambridge: Cambridge University Press.
- Lamason, R.L., 2005. SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science*, 310(5755), pp.1782–6.
- Lawson, D.J. et al., 2012. Inference of population structure using dense haplotype data. *PLoS Genetics*, 8(1), p.e1002453.
- Leask, H.G. et al., 1935. The Labbacallee Megalith, Co. Cork. *Proceedings of the Royal Irish Academy. Section C. Archaeology, Celtic studies, history, linguistics and literature*, 43C, pp.77–101.
- Leslie, S. et al., 2015. The fine-scale genetic structure of the British population. *Nature*, 519(7543), pp.309–314.
- Lipson, M. et al., 2017. Parallel ancient genomic transects reveal complex population history of early European farmers. *bioRxiv*, p.114488.
- Lucotte, G. & Dieterlen, F., 2003. A European allele map of the C282Y mutation of haemochromatosis: Celtic versus Viking origin of the mutation? *Blood Cells Mol Dis*, 31(2), pp.262–267.
- Mallick, S. et al., 2016. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature*, 538(7624), pp.201–206.
- Mallory, J.P., 2013. *The origins of the Irish*. London: Thames & Hudson.
- Malmstrom, H. et al., 2010. High frequency of lactose intolerance in a prehistoric hunter gatherer population in northern Europe. *BMC Evol. Biol.*, 10, p.89.
- Martiniano, R. et al., 2016. Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nature Communications*, 7, p.10326.
- Martiniano, R. et al., 2017. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genetics*, 13(7), p.e1006852.
- Mathieson, I. et al., 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528(7583), pp.499–503.
- Mathieson, I. et al., 2017. The Genomic History Of Southeastern Europe. *bioRxiv*.
- McGarry, T., 2010. Late pagan and Early Christian burials in Ireland: some issues and potential explanations. In C. Corlett & M. Potterton (eds.), *Death and burial in Early Medieval Ireland in light of recent archaeological excavations*. Dublin: Wordwell,

- pp.173-185.
- McKenna, A. et al., 2010. The Genome Analysis Toolkit: a MapReduce framework for analysing next-generation DNA sequencing data. *Genome Research*, 20(9), pp.1297–1303.
- McLaughlin, T.R. et al., 2016. The Changing Face of Neolithic and Bronze Age Ireland: A Big Data Approach to the Settlement and Burial Records. *Journal of World Prehistory*, 29(2), pp.117–153.
- Mitnik, A., et al., 2017. The Genetic History of Northern Europe. *bioRxiv*.
- Motulsky A.G., 1979. Genetics of hemochromatosis. *N Engl J Med*, 301, p.1291.
- Moore, L.T. et al., 2006. A Y-chromosome signature of hegemony in Gaelic Ireland. *American Journal of Human Genetics*, 78(2), pp.334–338.
- Mount, C., 2011. The two burial traditions of the beaker period in Ireland 1.1. Dr. Charles Mount Blog,
- Murillo-Barroso, M. & Martín-Torres, M., 2012. Amber Sources and Trade in the Prehistory of the Iberian Peninsula. *European Journal of Archaeology*, 15(2), pp.187–216.
- Myres, N.M. et al., 2011. A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe. *European Journal of Human Genetics*, 19(1), pp.95–101.
- Needham, S., 1996. Chronology and periodisation in the British Bronze Age. *Acta Archaeologica*, 67, pp.121–140.
- Norton H.L., et al., 2007. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Mol Biol Evol*, 24(3), pp.710–722.
- O'Brien, E., 2003. Burial practices in Ireland: first to seventh centuries AD. In J. Downes & A. Ritchie (eds.), *Sea change: Orkney and northern Europe in the later Iron Age AD*. UK: Pinkfoot Press, pp.62-74.
- O'Brien, W., 2000. Megalithic tombs, metal resources and territory in prehistoric south-west Ireland. In A. Desmond, G. Johnson & M. McCarthy (eds.), *New Agendas in Irish Prehistory. Papers in Commemoration of Liz Anderson*. Ireland: Wordwell, pp.161–177.
- O'Brien, W., 2012. The Chalcolithic in Ireland: a chronological and cultural framework. In M.J. Allen, J. Gardiner & A. Sheridan (eds.), *Is There a British Chalcolithic? People, Place and Polity in the later Third Millennium*. Oxford: Oxbow, pp.211-225.
- O'Brien, W., 2014. *Prehistoric Copper Mining in Europe: 5500-500 BC*. Oxford: OUP
- O'Brien, W., 1999. *Sacred ground: megalithic tombs in coastal south-west Ireland*. Galway, Ireland: NUIG.
- O'Connell, M., Kelly, F. & McAdam, J.H (eds.), 2016. *Cattle in ancient and modern Ireland: farming practices, environment and economy*. Cambridge: Cambridge Scholars Publishing.
- Ó Maoldúin R., 2015. Preliminary Report on the excavation of Wedge tomb CL017-180002, Parknabinia, Co. Clare. Licence 15E0053.
- Ó Nualláin, S. 1989. *Survey of the Megalithic Tombs of Ireland, Volume 5*, Dublin: The Stationery Office.
- Olalde, I. et al., 2017. The Beaker phenomenon and the genomic transformation of northwest Europe. *bioRxiv*, p.28.
- Ottaway, B.S. & Roberts, B., 2008. The emergence of metalworking. In A. Jones (ed.), *Prehistoric Europe: theory and practice*. Chichester: Wiley-Blackwell, pp.193–225.
- Patterson, N. et al., 2012. Ancient admixture in human history. *Genetics*, 192(3), pp.1065–1093.
- Plantinga, T. S. et al., 2012. Low prevalence of lactase persistence in Neolithic South-West Europe. *European Journal of Human Genetics*, 20, pp.778-782.
- Price, T.D. et al., 2004. Strontium isotopes and prehistoric human migration: the Bell Beaker period in central Europe. *European Journal of Archaeology*, 7(1), pp.9–40.
- Rajeevan H. et al., 2012. ALFRED: an allele frequency resource for research and teaching. *Nucleic Acids Research*, 40(D1), pp.D1010-D1015.
- Ruiz-Taboada, A. & Montero-Ruiz, I., 1999. The oldest metallurgy in western Europe. *Antiquity*, 73(282), pp.897–903.
- Salanova, L., 2016. Behind the warriors: Bell Beakers and identities in Atlantic Europe (third millennium B.C.). In J.T. Koch & B.W. Cunliffe (eds.), *Celtic from the West 3*. Oxford: Oxbow Books
- Salanova, L., 2003. Les sépultures mégalithiques et le phénomène campaniforme. In V.S. Gonçalves (ed.), *Muitas gente, poucos antas? Origens, espaços e contextos do megalitismo*. Lisbon: Instituto Português de Arqueologia, pp.385–393.
- Salanova, L., 2004. The frontiers inside the western Bell Beaker block. In J. Czebreszuk (ed.), *Similar but Different. Bell Beakers in Europe*. Netherlands: Sidestone Press, pp.63-75
- Schiffels, S. et al., 2016. Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nature Communications*, 7, p.10408.
- Schulting, R. et al., 2008. Largantea and the Dating of Irish Wedge Tombs. *The Journal of Irish Archaeology*, 17, pp.1–17.
- Schulting, R.J. et al., 2012. New dates from the north and a proposed chronology for Irish court tombs. *Proceedings of the Royal*

Irish Academy, 112C, pp.1–60.

- Senna Martínez, J.C., 2013. Metals, Technique and Society. The Iberian Peninsula between the first Peasant Societies with Metallurgy and the “Urban Revolution”. In M.F. Guerra & I. Tissot (eds.), *A Ourivesaria Pré-Histórica do Ocidente Peninsular Atlântico*. Lisbon: *Project AuCORRE, National Museum of Archaeology*, pp.11–20.
- Sjögren, K.-G., Price, T.D. & Kristiansen, K., 2016. Diet and Mobility in the Corded Ware of Central Europe. *PloS One*, 11(5), p.e0155083.
- Skoglund, P. et al., 2014. Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science*, 344(6185), pp.747–750.
- Soejima, M. & Koda, Y., 2007. Population differences of two coding SNPs in pigmentation-related genes SLC24A5 and SLC45A2. *International Journal of Legal Medicine*, 121(1), pp.36-9.
- Tishkoff, S.A et al., 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics*, 39(1), pp.31-40.
- Underhill, P.A. et al., 2010. Separating the post-Glacial coancestry of European and Asian Y chromosomes within haplogroup R1a. *European Journal of Human Genetics*, 18(4), pp.479–484.
- Van Dorp, L. et al., 2015. Evidence for a Common Origin of Blacksmiths and Cultivators in the Ethiopian Ari within the Last 4500 Years: Lessons for Clustering-Based Inference.
- van Rheenen, W. et al., 2016. Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nature Genetics*, 48(9), pp.1043–1048.
- Vasilieva, I.N., 2011. The early neolithic pottery of the volga-ural region (based on the materials of the elshanka culture). *Archaeology, Ethnology and Anthropology of Eurasia*, 39(2), pp.70–81. *PLoS Genetics*, 10(8), e1005397.
- Waddell, J., 2010. *The Prehistoric Archaeology of Ireland*. Bray, Ireland: Wordwell.
- Waddell, J. & Ó Ríordáin, B., 1993. *The funerary bowls and vases of the Irish Bronze Age*. Galway, Ireland: Galway University Press.
- Weir, B.S. & Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution; International Journal of Organic Evolution*, 38(6), pp.1358–1370.
- Wencel, M.M., 2015. An Absolute Chronological Framework for the Central-Eastern European Eneolithic. *Oxford Journal of Archaeology*, 34(1), pp.33–43.
- Wilde, S. et al., 2014. Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y *Proceedings of the National Academy of Sciences of the United States of America*, 111(13), pp.4832–4837.

6. Final Discussion

The work presented in this thesis is hopefully but a primer for what is to come, as the fields of Irish archaeology and population genetics find ever more common ground. The broad demographic framework put forward in Chapter Two was dissected as far as the genomic dataset would permit in subsequent chapters. Interpretation was firmly anchored in observation, although several conclusion sections allowed for some speculation when current analyses offered only a tantalising glimpse of potentially wider geographical or cultural trends. I believe this is the privilege, and possibly misfortune, of any interdisciplinary researcher trying to marry science and the humanities. Archaeology has long straddled this divide and benefits greatly from the wide variety of cerebral approaches that are taken in its study. I am a firm believer that the best research asks more questions than it solves and in this spirit I want to devote this final discussion the future avenues of inquiry that can follow from the work presented here. To do this, I will highlight what I believe to be the main themes of Irish demographic history. Given that the constraints of geography and climate tend to encourage both historical and prehistorical repetition, many of these happen to be reoccurring ones.

The modern Irish genome can be disentangled into three separate streams of ancestry, each of which finds their origin in the same African population. These segregated upon exit from that great continent, with the first venturing early on to the hostile northern climes of Europe. Here it weathered the last glacial maximum, and formed part of the Mesolithic recolonisation efforts as the glaciers went into retreat. These hunter-gatherer communities abandoned big game and cave art in exchange for microliths and forested habitats, likely expanding from refugia near the Swiss piedmont and further east in the Balkans. One such population ended up on an impoverished western island, where they appear to have suffered from a severe demographic bottleneck and prolonged isolation. Genomic surveys of Mesolithic Britain and France will be required to establish the more exact origin(s) of this population, and whether continued contact occurred with other regions after initial colonisation of the island. A key question is whether the break in lithic technologies between the Early and Late Mesolithic also represents a break in genetic continuity, which will require further archaeological discovery to address. Furthermore, effective population size estimates of both the founder population and the Late Mesolithic population presented here should be possible upon high coverage sequencing of these samples. Divergence time estimates derived from high coverage British and continental Mesolithic genomes are also an exciting prospect. Finally, if more unburnt remains are identified, the population may present a somewhat unique case study on the effects of complete isolation and small effective population size on human genetic variation through time.

The second stream of Irish ancestry only entered Europe from Anatolia when it was equipped to tame its environs with new agricultural technologies and domestic species. Farming groups encompassed the continent in two migratory arms, enmeshing with the first strand of ancestry and meeting again in and

around the Paris Basin. Again, some of these populations, who possessed substantial contributions from the Mediterranean route of migration and northwestern hunter-gatherers, decided to make the voyage to the remote westerly isle. Here, they planted seeds of civilisation, from which grew momentous architecture and hierarchical social structure, although some small groups belonging to that first strand of ancestry may have been able to persist in more remote areas. Denser genomic surveys of the Irish Neolithic are possible and indeed imperative, given the underlying suggestions of population structure in both Y chromosomal and autosomal variation. High resolution haplotypic analyses on thousands of Neolithic individuals from both Ireland and the continent, somewhat unimaginable now, are likely to become a reality within a few years and will lead to unprecedented detail regarding the impact of geography and culture on ancestry. They will also provide better understanding as to the relationship between kinship and communal megalithic structures, and most importantly establish the rarity of outlying individuals in terms of phenomena such as sibling-inbreeding or recent Mesolithic introgression.

The third stream of Irish ancestry came through the Caucasus isthmus, merging with hunter-gatherer groups of the Pontic steppe. From this region a new and prolific form of mobile pastoralism emerged, which gave this ancestry access to western Europe through mass migrations across the northern European and Carpathian plains, carrying with it Indo-European languages and new ideologies. Atlantic maritime and river networks allowed this strand, which was already firmly intertwined in the previous two, to reach the western island. Here, the various ancestries appear to have fluctuated for more than a few centuries before reaching an equilibrium which they have largely maintained up until the present day. Full understanding of the clearly complex introduction of metallurgy and Beaker culture to Ireland will only come from extensive and unbiased sampling of the island across both the Chalcolithic and Early Bronze Age. However, the scarcity of unburnt remains from the Chalcolithic may prove a large obstacle to this ultimate goal, and the inability to sample cremated remains presents a potentially uncontrollable bias. Regardless, the release of datasets from France and Britain should provide some further clarity on the transition. A key question is whether Brittany can be identified as the direct source of Irish Chalcolithic ancestry, and whether this is the ancestry that we see persisting in the western megalithic burials of the island. Confirming the potential survival of local Neolithic ancestry is also a priority. Furthermore, surveys of Y chromosomal and autosomal variation in Britain and France will allow us to appreciate the full extent of direct continuity within the island of Ireland from this time onwards. Denser sampling of the Irish Early Bronze Age population may also allow for the identification of rarer disease causing variants that are still present in modern Ireland today.

As these three streams successively swept into Europe, one clear recurring is the persistence of older ancestries in the peripheral regions. For example, in Iberia and Scandinavia, Mesolithic and Neolithic ancestries and material cultures are seen to survive and contribute substantially to the later populations. Remarkably, Ireland appears to represent a microcosm of this phenomenon, with the remote southwest seemingly harbouring Irish Mesolithic ancestry long into the Neolithic period, as well as inflated levels of Neolithic ancestry together with megalithic traditions many centuries after the arrival of metallurgy,

individualised burial and steppe-related populations. This apparent isolation of the western regions in prehistoric times follows closely with what we observe across written Irish history and indeed in the modern day country, where the west coast represents one of the last bastions of Celtic languages. Haplotypic analyses of the modern population also reveal similar trends of western diversity through isolation and eastern homogenisation through migration, mainly from Britain. Importantly, the archaeological identification of isolated relic populations in Ireland throughout the prehistoric periods may be hindered by small population sizes and lack of surviving material culture, but could be uncovered through large-scale genomic surveys. Indeed, more visible archaeological populations may yield individuals with a recent ancestor belonging to the ghost population, as appears to be the case for Parknabinnia675.

Ireland does not only represent a peripheral final destination for the great cultural and demographic upheavals of European prehistory, but also a major meeting point, where Atlantic and Northern cultural spheres vie for influence. This contest typically extends to the neighbouring island of Britain, where the Northern European Plain usually exerts a firmer pull, particularly on the eastern coast. The result is a continuous west-east dichotomy, whether in the distribution of Maglemosian hunter-gatherers, megalith builders, classic Bell Beaker burials, R1b-L21 Y chromosomes or Celtic and Germanic languages. Throughout this tug-of-war, Ireland's connectivity with other regions waxes and wanes, although the overall trend is one of increasing interaction with both Britain and the continent. The remote outpost of an isolated hunter-gatherer population had risen to become a main center of megalithic culture and art by the Late Neolithic, with clear maritime connections to other Atlantic communities and an elite who were powerful and wealthy enough to be accorded potentially divine status. Later the island becomes a major exporter of raw and worked metal during the Chalcolithic and Bronze Age, although this status declines significantly as the Iron Age commences, while in Early Christian times Ireland again becomes a hub of art and learning. As cultural connectivity increases through the eras, so should migration both to and from the island. Dense genomic surveys of ancient Irish populations may not only help identify relic groups, but also recent migrant individuals or communities, who may have left little to no long term impact on the island's population. The Irish Iron Age/Early Christian and Medieval periods are potentially good candidates for such investigations, given the wider demographic upheavals that were occurring in Europe during the same interval.

The key priority then for ancient DNA research in Ireland is denser sampling. However, as petrous temporal bones are a finite resource this should not be done in haste. Furthermore, as the cost of sequencing decreases and exponentially more ancient genomes are published, the possibility of drowning in too much data becomes a danger. A bottleneck in researchers available to analyse these massive datasets thoroughly may lead to key results being missed in the deluge of genomes. A few years ago an entire thesis could be written on a single ancient genome, while now the trend is towards larger numbers of genomes and shrinking supplementals. Moreover, high impact studies require large numbers of novel genomes, a pressure which does not encourage re-examination of previously published data unless higher

coverage sequencing is undertaken. To counter these issues, well-thought-out sampling strategies and hypothesis-driven research may be the best way forward. In particular, localised sampling allows more incisive and thoughtful archaeological inquiry. Importantly, the expansion of modern and ancient reference datasets is increasing the amount of information that can be gleaned from even the most homogenous of ancient populations, as demonstrated by Chapters Four and Five of the current thesis. Combining these growing datasets with the imputation and haplotypic methods utilised here will provide unprecedented detail of fine-scale population structure in ancient populations, even those existing across relatively small geographical and temporal dimensions. When such grand vistas become a reality, the only true limitation to our understanding will be our knowing the right questions to ask.

Appendix I :
Archaeological Contexts and Sampling Information

Please note that a summary of the archaeological information for all samples can also be found in Electronic Data Table S1.

Table of Contents

Natural Sites	212
Caves	213
Pollnagollum Cave, Co. Fermanagh	213
Sramore Cave, Co. Leitrim	213
Killuragh Cave, Co. Limerick	214
Wetland Contexts	217
Stoneyisland, Co. Galway	217
Inchagreenoge, Co. Limerick	217
Lagore, Co. Meath	218
Derrymaquirk, Co. Sligo	218
Ballinderry, Co. Westmeath	218
Megalithic Tombs	220
Portal Tombs	220
Poul nabrone Portal Tomb, Co. Clare	220
Court Tombs	223
Cohaw Court Tomb, Co. Cavan	223
Parknabinnia Court Tomb, Co. Clare	223
Audleystown Court Tomb, Co. Down	224
Ballyalton Court Tomb, Co. Down	225
Mourne Park court tomb, Co. Down	225
Passage Tombs	228
Newgrange Passage Tomb Complex, Co. Meath	228
Carrowkeel Passage Tomb Complex, Co. Sligo	229
Passage Tomb Affiliated	230
Ballynahatty, Co. Down	230
Millin Bay, Co. Down	230
Linkardstown Cists	234
Baunogenasraid, Co. Carlow	234
Jerpoint West, Co. Kilkenny	236
Annagh, Co. Limerick	236
Ardcrony, Co Tipperary	236
Ashleypark, Co Tipperary	237
Norrismount, Co. Wexford	238
Wedge Tombs	240
Roughan Hill, Co. Clare	240
Labbacallee, Co. Cork	240
Chalcolithic and Early Bronze Age Inhumations	243
Cist Burials	243
Glebe, Rathlin Island, Co. Antrim	243
Sliguff, Co. Carlow	244
Vermount, Co. Carlow	244

Coolnatullagh, Co. Clare	244
Lisooly, Co. Donegal	245
Lisnamulligan, Co. Donegal	245
Glassamucky, Co. Dublin	246
Topped Mountain, Co. Fermanagh	246
Annaghkeen, Co. Galway	247
Moyveela, Co. Galway	247
Blackhill, Co. Kildare	247
Timolin, Co. Kildare	247
Waterunder, Mell, Co. Louth	248
Stonepark, Co. Mayo	248
Fourknocks II, Co. Meath	249
Culleens, Co. Sligo	249
Treanmacmurtagh, Co. Sligo	250
Stranagalwilly, Co. Tyrone	250
Bolinready, Co. Wexford	250
Cemeteries	253
Ploopluck (Osberstown), Co. Kildare	253
Keenoge, Co. Meath	253
Grange, Co. Roscommon	254
Late Iron Age and Early Medieval Burials	257
Courtmacsherry, Co. Cork	257
Ballymacaward, Co. Donegal	257
Derrynamanagh, Co. Galway	258
Ballybunnion, Co. Kerry	258
Ballyeagh, Co. Kerry	259
The Curragh, Co. Laois	259
Ballyglass Middle, Co. Mayo	259
Collierstown, Co. Meath	260
Claristown, Co. Meath	260
Johnstown, Co. Meath	261
Hill of Ward, Co. Meath	261
Knowth, Co. Meath	261
Rosnaree, Co. Meath	262
Lehinch, Co. Offaly	262
Tinnapark Demesne, Co. Wicklow	263
Other Sites	267
Gallow's Green, Co. Antrim	267
Cherrywood I, Loughlinstown, Co. Dublin	267
Cotterellsrath, Co. Kilkenny	268
Site I, Knockadoon, Lough Gur, Co. Limerick	268
References	270

Natural Sites

Caves

Pollnagollum Cave, Co. Fermanagh

Site Description

The Pollnagollum Cave, also known as Pollthanacarra, is a swallowhole in which a number of human and animal remains were discovered in 1972, leading to its excavation by the Ulster Museum (Doughty 1995). The assemblage of human bone was initially believed to consist of three individuals, but later re-examination identified remains from at least four people, including two adult males and one adult female and an adult of indeterminable sex (Dowd 2008). Three dates have been obtained from the human bone (2430–2140 cal BC, 2280–2035 cal BC and 2130–1890 cal BC: UBA-8154, UBA-8153 and UBA-8152) (Dowd 2015), suggesting their deposition during the Chalcolithic and Early Bronze Age periods. These may have been casualties who fell into the hole, though the lack of later dating incidents suggests these individuals, along with the animal remains recovered, were purposefully placed there (Dowd 2008).

Sampling

Three samples were taken from the Pollnagollum assemblage (Fig. I.1), none of which had a direct date attached. This included a petrous bone (PG86), a molar taken from a mandible (PG90), and a loose molar (PG911). Two of these samples were subsequently dated during this project, returning overlapping estimations of 2300-2131 cal BC for PG86 (UBA-35073) and 2349-2135 cal BC for PG911 (UBA-35074).

Further Notes

As aDNA analysis identified two females and one male individual, it is likely that the fourth adult of indeterminable sex noted in the osteology report is female.

Sramore Cave, Co. Leitrim

Site Description

The precise location of Sramore Cave, somewhere along the the northern side of Sramore Mountain, near the Sligo-Leitrim Border, is currently unknown (Dowd 2015). Investigated by cavers in 1995, three human bones, a femur, humerus and mandible, were retrieved from the end of the main passage. Given the lack of an excavation at the site, little can be said regarding the archaeological context of the remains.

The human bones represent a single adult individual, aged over 20 years, thought likely to be male (Dowd 2015). Two radiocarbon dates have indicated this individual lived and died during the final stages of the Mesolithic period. The femur returned a date of 4040-3970 cal BC (5202 ± 36 BP; UB-6407) (Dowd 2008) and the mandible 4050-3980 cal BC (5227 ± 36 BP; UB-157772) (Kador 2010). These dates fall extremely close to the proposed establishment of agriculture on the island, raising questions as to whether the individual was associated with early farming or hunter-gatherer communities. Isotopic analysis suggests a terrestrial protein diet, showing a similar signature to Early and Late Mesolithic remains from Killuragh Cave (Dowd 2015).

Sampling

A canine (SRA62) (Fig. I.1), retrieved from the dated mandible, was used for ancient DNA analysis.

Killuragh Cave, Co. Limerick

Site Description

Killuragh cave is one of several small cave systems that face out onto the floodplain of the River Mulkear, a tributary of the River Shannon (Woodman et al. forthcoming). Part of an elevated limestone reef, it consists of two entrance passages that join together in a main chamber. Two main excavations occurred in 1993 (Director: J. O'Shaughnessy; License: 93E0175) and 1996 (Director: P. Woodman; License: 93E0175 ext.), from which prehistoric deposits dating from the Early Mesolithic through to the Bronze Age were recovered.

Mesolithic

Fourteen dates on human bones were retrieved from the site, three of which were Early Mesolithic in date (7060-6830 cal BC; 7030-6780 cal BC; 6900-6640 cal BC) and three belonging to the final stages of the Late Mesolithic (4660-4500 cal BC; 4580-4480 cal BC; 4350-4260 cal BC) (Woodman & McCarthy 2003; Delaney & Woodman 2004; Meiklejohn & Woodman 2012). In addition, a series of microliths, seemingly 'Early Mesolithic' in type, as well as a potential portion of a flint blade, tentatively linked to the Later Mesolithic, were recovered, although these cannot be definitively associated with the human remains.

Given the lack of settlement debris, it has been suggested that the Mesolithic association with Killuragh Cave was ritualistic in nature. Furthermore, there is no evidence to suggest complete bodies were interred within the system, arguing against its usage as a burial site. Instead, it has been suggested that hunter-gatherer groups placed isolated bones and artefacts at the mouth of the cave, which subsequently entered the system through non-human processes. Its persistent use across the Mesolithic period, coupled with

its inconspicuous nature, is indicative of an important location that remained in the collective consciousness through generations (Dowd 2015; Woodman et al, forthcoming)

Chalcolithic/Early Bronze Age

A second group of human remains were dated to the final Neolithic up to the beginning of the Bronze Age. These included an infant tibia (2460-2310 cal BC), an adult tibia (2028-1890 cal BC), an infant scapula (2290-2140 cal BC) and adult mandible (2280-2140 cal BC) (Meiklejohn & Woodman 2012; Dowd 2015). Early Bronze Age pottery was recovered from the cave, though this likely post-dates the human remains, as well as two hollow scrapers, possibly corresponding to the Late Neolithic phase of use.

Sampling

Two Late Mesolithic samples were taken for aDNA analysis. These were a molar (KGH2) extracted from a dated mandible (OxA-6749; 4350-4260 cal BC), and a second molar (KGH6), for which a direct date (OxA-6752; 4660-4500 cal BC) had been obtained from another tooth of the same context (Fig. I.1). From the Chalcolithic/Early Bronze Age remains a molar (KGH1) was extracted from the dated adult mandible for analysis (OxA-6748; 2280-2140 cal BC) (Fig. I.1).

Further Notes

Upon sequencing both KGH1 and KGH2 up to ~0.65X coverage, kinship analysis revealed both samples to belong to the same individual. The mandibular segments from which each tooth was taken were reexamined and it was noted that they could represent separate sections of the same singular mandible. The inflated genetic affinities of Killuragh1 (Study ID of combined KGH1 and KGH2 samples) to Irish Early Bronze Age individuals suggests that the OxA-6748 date is correct. OxA-6749 is likely either contaminated, or its association with the KGH2 mandible is the result of labelling error.



Figure SI.1. Samples taken from cave sites for aDNA analysis.

Wetland Contexts

Stoneyisland, Co. Galway

Site Description

In 1929, turf-cutters at a peat bog close to Lough Derg discovered a near-complete skeleton below approximately 3 meters of uncut peat. No official excavation of the site was ever carried out. Two primary descriptions of the remains were published in the 1930s (Shea 1931; Martin 1935). Seven radiocarbon dates in total have been determined from the skeleton (Hedges 1993; Brindley and Lanting 1995; Kador 2010), the most reliable of which place the individual between 4230-3800 BC, straddling the Mesolithic to Neolithic boundary, with a Neolithic assignment being argued by some (Brindley and Lanting 1995; Woodman 2000). However, without any available archaeological context the question has remained unresolved.

Sampling

A portion of the right petrous temporal bone (SI118) was cut from skull and used for ancient DNA analysis (Fig. I.2).

Further Notes

Unfortunately, this important sample did not yield sufficient endogenous DNA for any further analyses. Though 2 Y chromosome reads were retrieved out of a total of 16 sex chromosome reads, the identification of this sample as male remains tentative at best. Notably, all work in ancient lab facilities was carried out by female researchers, limiting potential for Y chromosome contamination. As some endogenous DNA survival is apparent, this sample may benefit from further rounds of sequencing in an attempt to determine sex and possible mtDNA haplogroup membership, which may shed some light on the Mesolithic or Neolithic ancestry of the individual.

Inchagreenoge, Co. Limerick

Site Description

The site is a natural spring, approximately half a meter deep, surrounded by limestone deposits. Within a shallow bed of peat, just inside the edge of the spring a human skull and mandible was found (C22). The skull was crushed and incomplete, with fibrous peat growth covering the surfaces (Taylor 2002). A radiocarbon determination from the bone returned a date of 2940 ± 30 BP (GrN-281898), or 1270-1040 cal BC.

Sampling

A disarticulated left petrous temporal bone (IG134) was sampled from the skull (Fig I.2).

Lagore, Co. Meath

Site Description

Lagore is the site of a famous crannog, a type of artificial island structure (Hencken 1936; Hencken 1950). The majority of crannogs are located in the northern half of the island, and many sites have evidence of multiple occupations throughout historic and prehistoric times. Lagore is noted as ‘Inis Locha Gabur’ in annal entries for the years 850 and 934 AD as a seat of local kings (Price 1950), and is larger than many other crannogs. Late Bronze Age activity has also been recorded at the site (Coles 1990). Human skulls have been associated with many wetland and crannog sites (Fredengren 2002), with some retrieved from Bronze Age stratigraphic layers. One such skull was excavated from the Lagore site by Hencken (E14:-).

Sampling

A petrous bone (LG14) from a human skull taken from the Lagore site was sampled for aDNA analysis (Fig I.2). Radiocarbon dating of this bone, carried out during the current project, yielded an early modern date of 1492-1665 cal AD (285 ± 33 ; UBA-35077).

Derrymaquirk, Co. Sligo

Site Description

This is a crannog site (BOYL 026) on Lough Gara, located in the Derrymaquirk townland, which has yielded radiocarbon determinations from the Late Bronze Age (Fredengren 2002). Two skulls were found nearby the crannog (E20: 731, 732).

Sampling

A petrous (DQ20) was taken from one skull (E20:731) for aDNA analysis (Fig I.2).

Ballinderry, Co. Westmeath

Site Description

Ballinderry 1 is the site of another crannog structure, excavated by Hugh Hencken in 1932 (Hencken 1936). In 1933, he excavated a second crannog, Ballinderry 2, (Hencken 1942). Ballinderry 2 showed

evidence of both Late Bronze Age and early medieval activity, and three skulls were retrieved from the site.

Sampling

The petrous of one skull (E6: 803) was taken for aDNA analysis (BLD6; Fig I.2).

Figure I.2. Samples taken from wetland sites for aDNA analysis.



Megalithic Tombs

Portal Tombs

Poulnabrone Portal Tomb, Co. Clare

Site Description

The Poulnabrone portal tomb, located in the upland limestone region of Co. Clare known as the Burren, was excavated in the mid-1980s as part of a conservation project, the results of which are detailed in Lynch 2014. These focused on the chamber of the tomb, the portico feature, and the south-west quadrant of the surrounding cairn. The unburnt, disarticulated remains of at least 35 individuals were recovered from the chamber, representing all age groups and both sexes.

There are approximately 184 known portal tombs in Ireland with their distribution mainly concentrated in the northern third of the country and in the southeastern counties of Leinster (Kytmanow 2009). Poulnabrone is one of a small group found in the mid-west. It is easily the best dated portal tomb in Ireland, with dating programmes carried out in both the late 1980s (Hedges et al. 1990; Lynch & Donnabháin 1994) and in 2012 (Schulting 2014), with a high degree of consistency between the two. The human remains ranged in date from approximately 3800 cal BC to 3200 cal BC. The earlier of these burial dates arguably place the construction of the tomb at the very beginning of the Irish Neolithic and make Poulnabrone one of the earliest known megaliths in Ireland. After its initial construction, burial activity appeared to continue, possibly intermittently, over the following 300-600 years. Radiocarbon data from other portal tombs in Ireland and Britain are extremely limited, with only single determinations from seven tombs available, many of which relate to post-Neolithic reuse.

Notably, the modelled start dates proposed by Lynch (Lynch & Beckett 2014) for Poulnabrone (circa 3885-3710 cal BC) show no overlap with the suggested period of initial use of court tombs (3700-3550 cal BC) (Schulting et al. 2012). They also predate the well defined 'house-horizon' (Cooney 2007; McSparron 2008), the start of which has been modelled as 3720-3660 cal BC. This supports the claim of portal tombs as one of the earliest forms of Neolithic megalithic burial in Ireland (Whittle 2004).

Sampling

Approximately 39 petrous bones were recovered from Poulnabrone, none of which have been directly dated. The majority of radiocarbon determinations for the 2012 dating programme were carried out on mandibles. For this reason three teeth from dated mandibles were sampled (PN107, PN112 and PN113),

from which Early Neolithic dates had been obtained (3928-3666 cal BC, 3696-3535 cal BC and 3942-3702 cal BC; OxA-26052, OxA-25949 and OxA-25950). Alongside this, seven petrous bones (PN01-07) were also sampled (contexts listed in Electronic Data Table S1), two of which were chosen to be dated as part of the current project. One of these, PN01, was chosen given its Early Bronze Age affinities, and returned a date of 2028-1758 cal BC (UBA-35065). The other, PN07, which was selected for its high endogenous content, dated to 3629-3371 cal BC (UBA-35065). Sample pictures are shown in Fig. I.3.

Further Notes

The sample PN01 (Gryke F28A) was identified as an outlier in terms of genomic affinities. The individual showed increased affinity to Bronze and Iron Age samples, relative to other Irish Neolithic samples, and was subsequently dated to the Early Bronze Age period.



Figure I.3. Samples taken from the Poul nabrone portal tomb for aDNA analysis.

Court Tombs

Cohaw Court Tomb, Co. Cavan

Site Description

Cohaw is a dual-court tomb, excavated by Howard Kilbride-Jones in 1949 (Kilbride-Jones & Keenan 1951). The highest density of court tombs in Ireland is seen in the north of the island, and Cohaw fits with this distribution. The two ‘court’ areas were connected by a long, five-chambered burial gallery, with the middle chamber sealed off. The burial assemblage included fragments of teeth and skull, as well as the cremated remains of a child, all found in Chamber 5, the best preserved of the tomb. The report on skeletal remains, by Prof. E. Keenan, states that the skull belonged to a young individual (probably male) aged between ten and fifteen years. A fine, carinated, round-bottom Neolithic bowl with a perforation under the rim was also recovered.

Sampling

A single petrous bone (CH448) existed from the site, attached to a large portion of the cranium, which was sampled for aDNA analysis (Fig. I.4). No contextual date was available for the bone and so it was submitted for AMS dating as part of this project, returning an Early to Middle Neolithic date of 3652–3514 cal BC (UBA-35070).

Parknabinnia Court Tomb, Co. Clare

Site Description

The Parknabinnia megalith (Cl. 153) is one of several atypical court tombs found in north Munster (Jones forthcoming). These sites all possess several characteristic features of court tombs, including a sequence of chambered galleries, two in the case of Parknabinnia, with entrances facing the east or northeast. However, a number of shared peculiarities in their morphology exist, including narrow straight-sided (rather than open) forecourts and relatively small sizes. These anomalies, as well as their relative isolation from the main distribution of court tombs in the northern third of the island, warrants their consideration as a distinct morphological group (Jones forthcoming).

Situated on Roughan Hill, the Parknabinnia tomb forms part of an extensive Neolithic landscape in the Burren, which also includes the Poul nabrone portal dolmen, discussed above (Jones & Walsh 1996; Jones 2003). Radiocarbon dating indicates that its initial usage occurred *circa* 3,700–3,570 BC, a time when many other court tombs across the country were also being constructed (Schulting et al. 2012). However, unlike the majority of court tomb sites, Parknabinnia continued to be used into the Late Neolithic (GU-10575: 2905–2620 BC). The burial rite consisted of successive primary inhumations, which were

rearranged and disturbed during later interments (Beckett 2011). The rear chamber (Chamber 2) appears to have been blocked off some time between *circa* 3,100-2,800 BC, though later deposition continued on in the front chamber.

Disarticulated human remains, the majority inhumed, were retrieved from both chambers of the tomb, as well as numerous artefacts, including Carinated Bowl pottery and leaf-shaped arrowheads. A minimum of twenty individuals were identified, comprising fifteen adults and five sub-adults (Beckett 2011).

Sampling

A total of 12 right petrous bones were sampled from both chambers of the Parknabinnia court tomb, as well as its entrance (Fig I.5). Contexts are detailed in Electronic Data Table S1. Three of these samples, PB186, PB672 and PB675, were dated as part of this project and yielded dates of 3518-3355 cal BC, 3639-3514 cal BC and 3100-2910 cal BC respectively (UBA-35072, UBA-35067 and UBA-35064).

Further Notes

Two individuals, Parknabinnia357 and Parknabinnia675 were found to be third degree relatives (Appendix Two; Fig. II.9). Furthermore, the individual Parknabinnia675 (3100-2910 cal BC) shows a signal of recent, likely local, hunter-gatherer introgression.

Audleystown Court Tomb, Co. Down

Site Description

The Audleystown court tomb was initially discovered in 1946 and excavated in 1952 by Pat Collins of the Archaeological Survey (Collins et al. 1954). It can be classified as a dual-court tomb, or ‘double horned cairn’. It is located in the northeast of the island and thus falls within the main distribution of Irish court tombs.

A large assemblage of human remains was recovered from the tomb, but the majority have since been lost during fire or flood at Queen’s University Belfast. However, some bone is still accounted for, and this has yielded five Chalcolithic/Early Bronze Age dates, (Schulting et al. 2012). Food Vessel pottery was also found within the tomb. This is in keeping with a more widespread trend of Chalcolithic/Bronze Age reuse of court tombs after a period of inactivity.

Sampling

Two undated petrous bones (AT1 and AT7) were located and sampled from the surviving assemblage (Fig I.4).

Ballyalton Court Tomb, Co. Down

Site Description

The Ballyalton tomb is a single court grave, located close to the Audleystown site. A large assemblage of faunal remains and a small amount of human bone were retrieved from the site from the lower deposits of the inner and outer chambers (Evans & Davies 1934). A particular Neolithic ceramic type, Ballyalton Bowls, is named for this site (Case 1961). A series of radiocarbon determinations indicated both Middle Neolithic and Chalcolithic/Bronze Age use of the site, with the earliest date falling within the main range for construction of court tombs in Ireland (UB-7191: 3650–3520 cal. BC) (Schulting et al. 2012).

Sampling

Two dated teeth, BAL329 and BAL3211 (B76 and B42; 3650-3520 cal BC and 3650-3385 cal BC; UB-7191 and UB-7192) (photos not available) and a single undated petrous bone, BAL14 (Fig I.4), were sampled from the assemblage .

Mourne Park court tomb, Co. Down

Site Description

The Mourne Park court tomb is located in the same region as the Audleystown and Ballyalton sites. Upon its excavation in the 1930s, an assemblage of unburnt human bone was found among and under the stones that blocked the inner portal of Chamber II (Davies 1938). No successful radiocarbon dates have been retrieved for the site (Schulting et al. 2012).

Sampling

A single undated petrous bone (MP18) was sampled from the assemblage (Fig I.4).

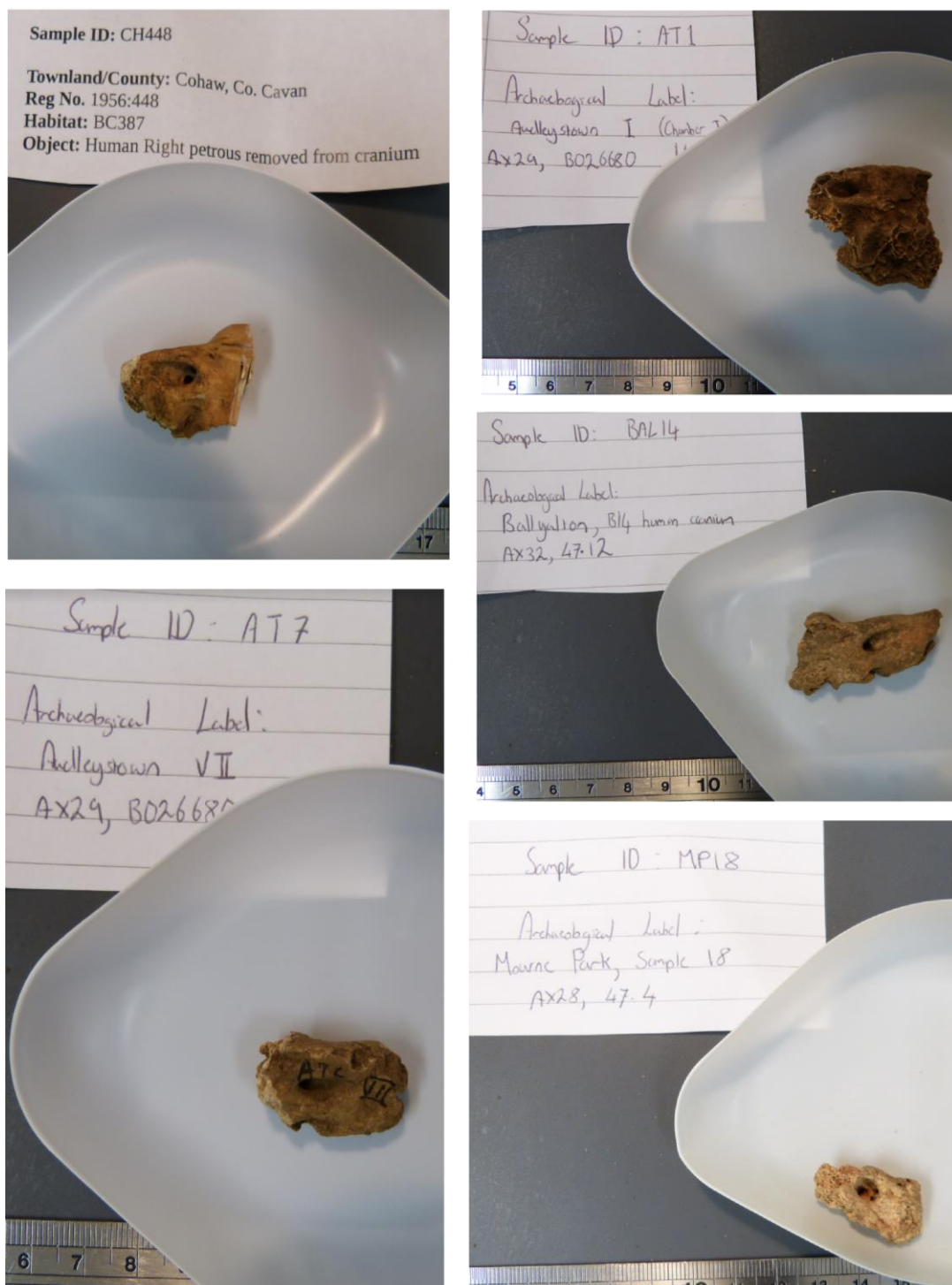


Figure I.4. Samples taken from northern court tombs for aDNA analysis.



Figure I.5. Samples taken from the Parknabinnia court tomb for aDNA analysis.

Passage Tombs

Newgrange Passage Tomb Complex, Co. Meath

Site Description

Newgrange is one of three relatively large passage tombs, out of a complex of over 40, and various other assorted prehistoric monuments, that together form the UNESCO world heritage site of Brú na Boinne. The site stands on an east to west running ridge, one of a series of low glacial hollocks which descend towards the Boyne River's floodplain. Three smaller passage-graves, sites K and L to the west, and site Z to the east, accompany Newgrange on that ridge. Other remains, including tumuli, standing stones and enclosures occur in the immediate vicinity and appear concentrated between the main monument and the Boyne (O'Kelly 1982).

The main Newgrange tomb consists of a passage and chamber, built from large slabs without mortar. A large cairn of loose stones covers the tomb. Bone fragments representing a small number of people, as well as grave goods typical of Irish passage-graves, were found on the floor of the tomb during excavation. Outside of the tomb, in the cairn slip and beyond, there were extensive finds associated with later settlement of the Late Neolithic/Beaker periods, as well as Roman objects. Eleven radiocarbon dates obtained for the initial report (O'Kelly 1982) suggest the main mound dates to approximately 3331-2916 cal BC. Both unburnt and burnt material were collected from inside the main tomb. This included at least one pair of unburnt petrous temporals.

By the time of site Z's construction the main mound at Newgrange had begun to decay, so site Z is later, though by how much is uncertain. Site Z was almost completely destroyed, with nearly all structural stones removed and smashed. During its destruction unburnt human burials must have been encountered either in its upper part, or in the upper part of the deposits in the chamber and passage. Skulls or skull fragments would have been recognisable amongst these bones. Three pits were dug outside the kerb through the then existing turf and the skulls, mandibles and other recognisable human bones were thrown into them with masses of stone fragments. The pits were then closed over with soil.

A full description of human remains from site Z was made by Professor Erskine and summarised by O'Kelly (O'Kelly et al. 1978). The material from the main chamber and passage of site Z is in very fragmentary condition, however there is mention of a temporal bone belonging to a female, as well as 2 petrous temporal bones, alongside other skull fragments. Skeletal material was also recovered from a "skull pit" SW of KI. This consists of large fragments of 3 adult male skulls and one infant skull. Three temporal bones were identified, 2 of left and 1 of right orientation.

Sampling

A single petrous bone (Fig. I.6) from the pair found in the main chamber of Newgrange was sampled (NG10). These were located in a box labelled 'cremated bone', however the report on the human skeletal remains lists these two petrous bones as unburnt and includes a picture, allowing them to be positively identified. Three petrous temporals were located from Site Z, however it was unclear whether these belong to the 'skull pits' assemblage or that from the main chamber. Two of these are left in orientation and one is right, matching the description of the petrous temporals recovered from the skull pit. The 2 left petrous temporals (Fig. I.6) were chosen for aDNA analysis (NGZ1 and NGZ2). NG10 and NGZ1 were submitted for radiocarbon dating as part of this project, yielding dates of 3339–3028 cal BC (OxA-36079; 4473±29 BP) and 3332–2921 cal BC (OxA-36080; 4421±30 BP) respectively.

Carrowkeel Passage Tomb Complex, Co. Sligo

Site Description

Carrowkeel is typically identified as one of the major passage tomb cemeteries on the island, the others being the Brú na Boinne, Carrowmore, and Loughcrew complexes. It consists of a core cluster of 14 monuments built atop limestone summits, situated within a wider megalithic landscape in and around the Bricklieve Mountains.

The main archaeological survey of the passage graves took place in 1911, carried out by R.A.S Macalister, and involved the investigation of 14 cairns, of which eight were excavated (Macalister et al. 1912). Large quantities of human remains, both cremated and unburnt, were retrieved from the tombs at this time, and put in storage. These were re-discovered in 2001 and are the subject of a number of recent and forthcoming studies (Hensey et al. 2014; Geber et al. 2016; Geber et al. 2017; Kador et al. forthcoming). A minimum of 29 individuals are represented, including 19 adults (nine male and four female) and 10 sub-adults, several of whom show evidence of dismemberment (Geber et al. 2017).

Radiocarbon dating indicates the complex was in use from the Middle Neolithic through to the very end of the period (*circa* 3600-2500 BC), with continuous activity likely. From the 40 dates generated, three overlapping phases of use were identified (Kador et al. forthcoming). The first is defined by four cremated antler samples, which indicate early activity prior to 3300 cal BC, although the tomb from which these remains were originally obtained is unknown. The second phase corresponds to the suggested main period of Irish passage tomb construction and use (O'Sullivan 2005; Cooney et al. 2011; Schulting forthcoming), *circa* 3400-3000 BC, and is represented by a number of human and red deer samples. The third phase begins with the deposition of a large amount of skull fragments, *circa* 2900 BC, after which time deposition becomes seemingly less frequent but continues until the end of the Neolithic period, demonstrated by a radiocarbon determination of 2636-2469 cal BC from an unburnt skull.

Sampling

A total of eight petrous bones were sampled from the Carrowkeel assemblage (Fig I.7). Of these six were retrieved from Cairn K, one from Cairn H and one from Cairn M. Two of these (CAK16 and CAK375) were cremated, and were subsequently seen to contain no surviving human DNA. Of the other four samples, four yielded substantial endogenous contents and were subsequently dated for inclusion in a study by Kador et al. (forthcoming). Two of these (CAK532 and CAK533) overlapped between the second and third phases of usage (3015-2891 cal BC and 3086-2903 cal BC; OxA-35326 and OxA-35327), while the other two fell securely within the third phase (CAK530 and CAK68; 2883-2635 cal BC and 2834-2469 cal BC; OxA-35325 and UBA-30808).

Passage Tomb Affiliated

Ballynahatty, Co. Down

Site Description

The Ballynahatty megalith was discovered in 1855, in a field to the north of the well-known henge monument, the Giant's Ring, and comprised a subterranean circular stone burial chamber. The burial structure was sub-circular and its interior had been divided into a series of compartments – the skull of a female (A.64), with an age-at-death of 18-35 years, was recovered from Compartment D and was associated with a collection of disarticulated bone and fragments of cremated bone (MacAdam & Getty 1855). Radiocarbon dating of the cranium produced a result of 3343-3020 cal BC (UB-7059; 4465 ± 38 BP) at 95% probability (Schulting et al. 2012), which places the individual into the later Middle Neolithic, or possibly Late Neolithic, periods. Her burial context, based on 19th century drawings and reports, is unusual in that its morphology does not readily fit within any of the recognised megalithic tomb traditions, although is most similar to passage grave structures. Human bone from compartments E/F has also been recovered and returned radiocarbon dates similar in range to the Millin Bay megalith (Schulting et al. 2012), within the second half of the fourth millennium, in keeping with the structure's developed passage tomb affinities.

Sampling

Initially, two dated teeth (no photos available) were sampled for the site, BA342 and BA346 (2882-2629 cal BC and 3501-3116 cal BC; UB-6723 and UB-7194). Later, the petrous temporal from the female skull (A.64), BA64, was sampled (Fig I.6).

Millin Bay, Co. Down

Site Description

The Millin Bay megalith consists of an elongated central cist set within a cairn, with affinities to the passage tomb tradition. Notably however, the tomb also neighbours a number of court tomb sites. A large assemblage of disarticulated human remains, numbering at least 15 individuals was recovered from the site (Collins et al. 1955). This material was until recently lost, but was rediscovered during a sampling trip for this project. Prior to the rediscovery of the main assemblage, several dates had been retrieved for a small amount of bone from the tomb (Schulting et al. 2012), falling within the second half of the fourth millennium BC, fitting with the megalith's developed passage tomb affinities (artwork and Carrowkeel ware pottery). However, one date was significantly earlier than expected, 3781±3651 cal. BC (OxA- 16598). This may indicate the deposition of ancestral relics, which predate the tombs construction, or alternatively, may be the result of chemical treatment with Bedacryl, a petroleum-based substance (Schulting et al. 2012).

Interestingly, cut marks relating to dismemberment were seen on some bones, assumed to be part of a post-mortem funerary rite, as seen at Carrowkeel (Geber et al. 2017).

Sampling

Prior to the rediscovery of the assemblage, two dated teeth, MB101 and MB9889 (3500-3130 cal BC and 3500-3127 cal BC; OxA-1610 and OxA-16107), were sampled from the surviving material (no photos available). After rediscovery, two undated petrous bones (MB4 and MB6) were sampled (Fig I.6). MB6 yielded a high amount of endogenous DNA and was subsequently dated as part of this project. The resulting radiocarbon date was within the expected range of use for the tomb (4548±51; 3376-3090 cal BC; UBA-35071).

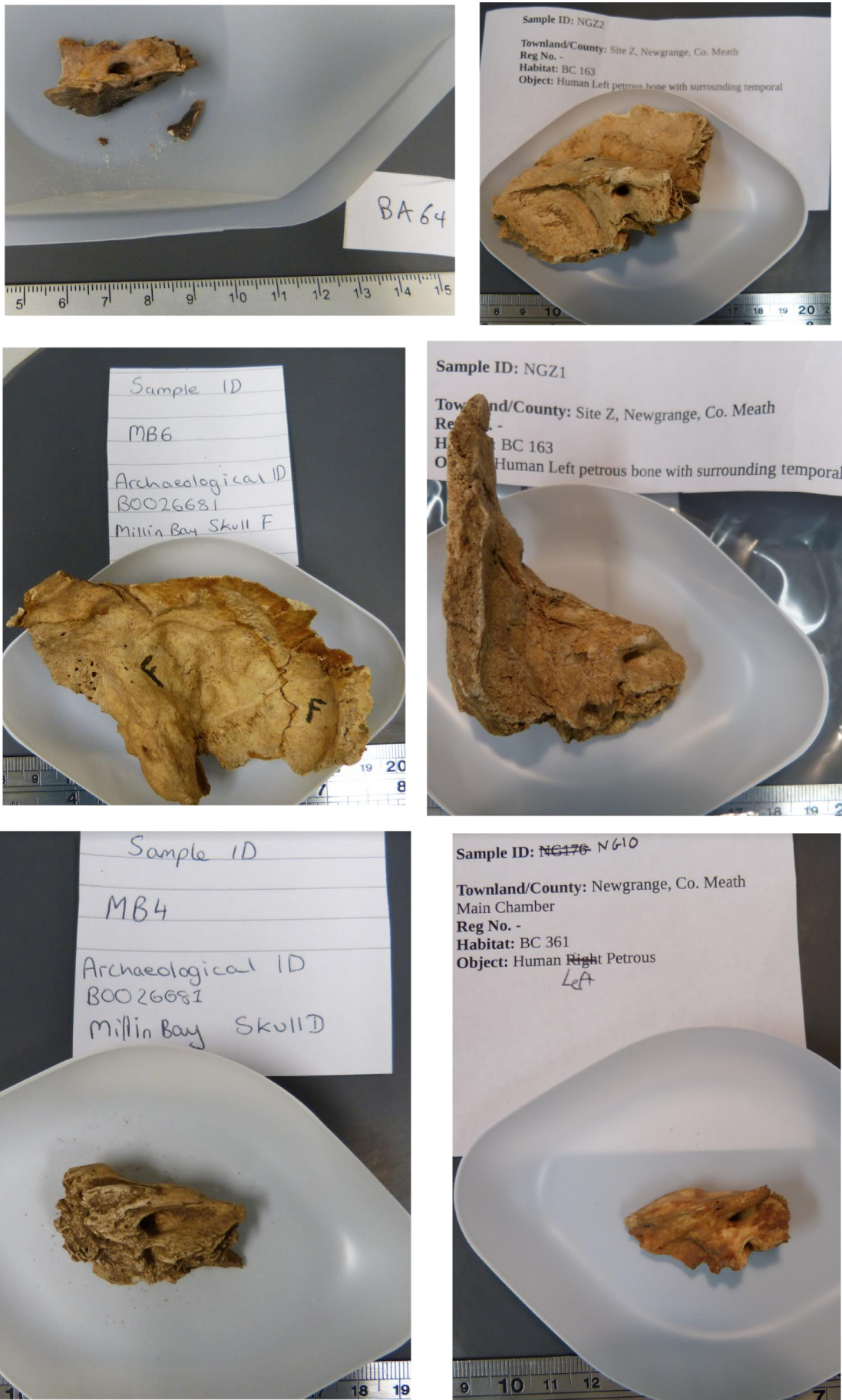


Figure I.6. Samples taken from eastern passage and passage-like tombs for aDNA analysis.

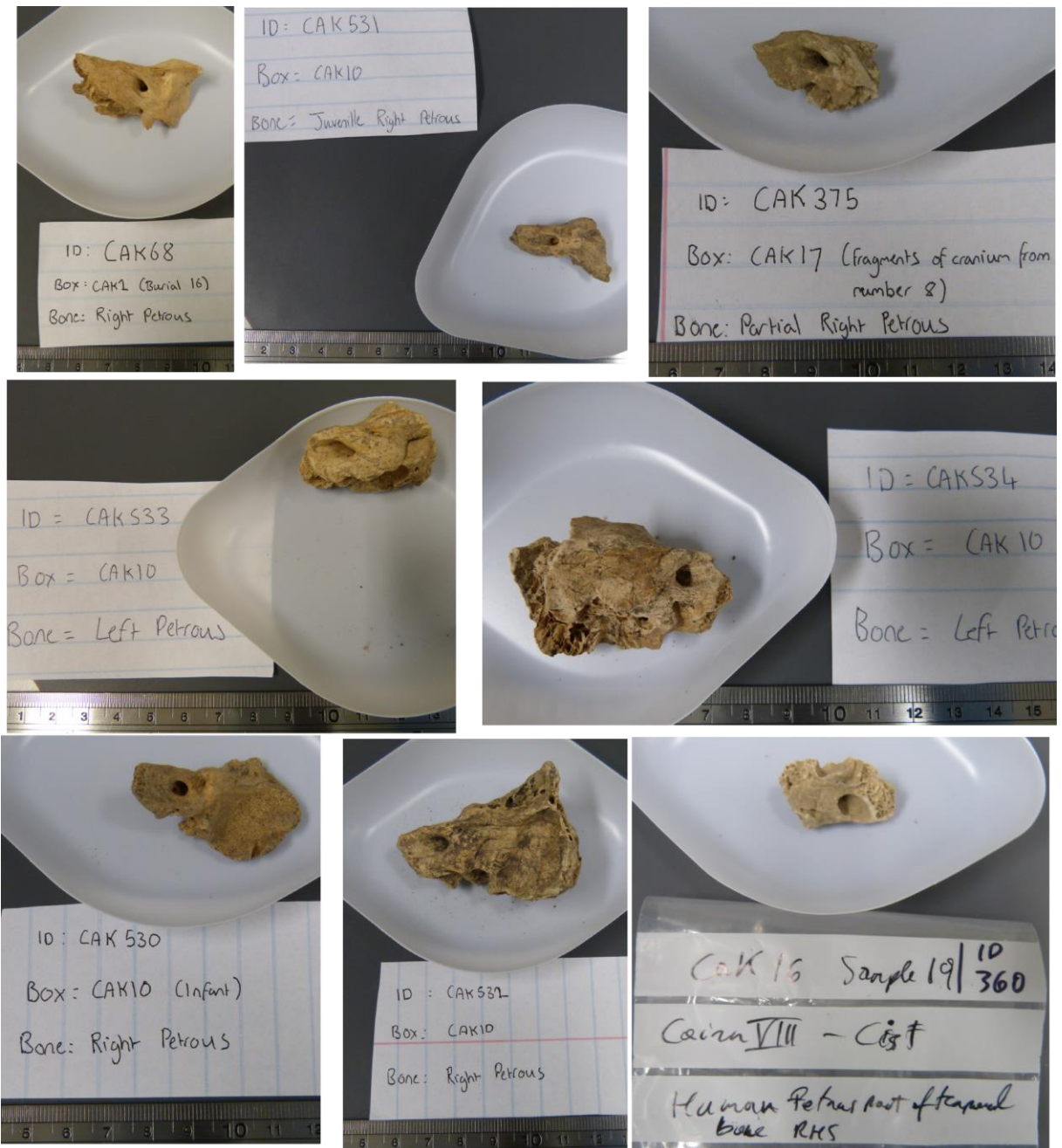


Figure I.7. Samples taken from the Carrowkeel passage tomb for aDNA analysis.

Linkardstown Cists

Linkardstown cists are a small group of tombs, mostly restricted to the south east of the country. They are distinguished by the placement of a central polygonal cist within a small mound and a specific set of grave goods (Brindley & Lanting 1989; Waddell 2010). Only eleven or so have been identified to date, and all tend to follow a similar burial rite - the interment of a single, or several, individuals, usually adult males accompanied by juveniles, with a distinct type of pottery associated.

The chronology of Linkardstown burials was debated for some time, with many placing them at the end of the Neolithic, when megalithic tombs traditions began to decline. However, the dating of human remains from these tombs upturned this theory, placing these burials in the earlier part of the Middle Neolithic (3600-3300 cal BC) (Brindley & Lanting 1989). Their origins remain an enigma, though they may represent some form of social elite, or possibly an invasive culture.

Baunogenasraid, Co. Carlow

Site Description

The site was excavated between 1972 and 1973 by Barry Raftery (Raftery 1974). The primary phase consisted of a single-burial, sub-megalithic structure with a low covering mound. The bones were the unburnt, disarticulated remains of a single individual and had been deposited neatly in a pile in the north-eastern corner of the chamber. The long bones were extended parallel to each other in an east-west direction and the pelvis was placed partly across the ends of some of the long bones, being flanked by several of the others. A short distance beyond the pelvis towards the south-west lay the remains of the skull; this was broken and incomplete and part of the jawbone lay beside it. The bones were examined by Professor Erskine. He reported that the bones belonged to an adult male of large stature. A decorated, round bottom pot was also discovered. The individual was later dated to 3631-3384 cal BC (GrN-11362) (Brindley & Lanting 1989).

The monument described above is Phase I of the Baunogenasraid site. It was abandoned for some time before subsequent reuse, with secondary burials being inserted in the upper levels of the mound.

Sampling

A petrous bone (BG72) was sampled from the central cist burial for aDNA analysis (Fig. I.8).

Jerpoint West, Co. Kilkenny

Site Description

The Jerpoint West megalith consisted of a polygonal cist built on ground surface, under a circular mound (Ryan 1973). It was discovered during the bulldozing of a low mound near to a quarry. Grave goods included a polished bone pin, a portion of a leaf-shaped arrowhead and Neolithic ceramics. An inhumation of a young adult male (E93:15), *circa* 152 cm in height, was found within the cist, as well as the cremated remains of one individual (E93:14). The skeletal remains of a child (E93:16) of approximately six years were also discovered within the bulldozed spoils. The original site context of these bones is not known.

The young adult male inhumation was subsequently dated and returned two non-overlapping radiocarbon determinations, one of which, GrN-11897 (4305±40 BP), was discarded. The other, OxA-2680 (4770±80 BP; 3694-3369 cal BC), is taken to be correct (Brindley & Lanting 1989).

Sampling

Two petrous bones from different individuals were identified from the site and taken for aDNA analysis. However, a labelling mix up led the petrous bone of the inhumed young adult male (E93:15) to be associated with the label given to the cremated remains, E93:14, and so the sample was given a laboratory ID of JP14. Moreover, the remains of the juvenile remains from the spoils, E93:16, were associated with the E93:15 label and so the sample was given the ID 'JP15'.

This confusion was only discovered at a late stage, during population genetic analysis and molecular sex determination. For this reason the lab IDs have not been changed to reflect the individuals they represent.

A sample from the JP15 (E93:16) petrous bone was sent for radiocarbon dating and returned a determination of 1688±29 (257-416 AD; UBA-35069), placing the burial within the Late Iron Age period.

Further Notes

A labelling mix up occurred detailed in the above section.

Annagh, Co. Limerick

Site Description

Excavations of the cave at Annagh were carried out in 1992 (Cahill & Sikora 2011). The site consisted of an oval chamber with the entrance blocked with a pillar like slab. Two pottery vessels of Linkardstown/Drimnagh type were discovered along with three human burials and portions of at least two others. All of the remains were identified as male.

Burial 1 (92E47:1) was of a crouched adult male, aged 55+, radiocarbon dated to 4670 ± 70 BP (GrA-1703) calibrated to 3639-3138 BC. Burial 2 (92E47:2) was of similar age, sex and position to Burial 1, and radiocarbon dated to 4780 ± 60 BP (GrA-1704), calibrated to 3660-3375 BC. Burial 3 (92E47:3) was a male of at least 30 and returned a radiocarbon date of 4810 ± 60 BP (GrA-1707), calibrated to 3706-3379 BC. Burial 4 (92E47:4) consisted of the remains of at least 2 individuals, most likely males, returning dates of 4640 ± 60 (GrA-1708) and 4840 ± 60 (GrA-1709), calibrated to 3633-3123 and 3767-3384 BC respectively. It was noted that the cranial remains in this disarticulated scatter consisted of a complete mandible.

There is a strong possibility of kinship among these samples; two, individuals, Burials 1 and 3, had congenitally absent third molars; burials 1 and 4 shared very similar expressions of complete mylohyoid bridging, a relatively uncommon trait; burials 1, 2 and 3 had similar acetabular calcaneal facets; and the two individuals from Burial 4 and Burial 3 all had double anterior calcaneal facets. These shared traits could reflect responses to similar environmental stimuli, but they could also reflect a higher than usual degree of consanguinity among these individuals.

Sampling

Currently only one individual from Annagh has been sampled (ANN1), Burial 1 (92E47:1). This was taken from the petrous portion of the individual's skull (Fig I.8), which was embedded in limestone.

Ardcrony, Co Tipperary

Site Description

The site is a Linkardstown burial, consisting of a massive central cist at the centre of a denuded cairn. Two disarticulated and unburnt skeletons, identified by Prof. C.A. Erskine as men of 17 and 45 years, lay on tile paved floor, one on either side of a shouldered, round-bottomed, highly decorated shallow bowl (Wallace 1977).

The disarticulated remains of the youth (NMI, reg. No. E180:2) lay along the western side of the cist, while the bones of the older man (E180:3), rested on a thin layer of silt in the eastern side of the cist. A femur from the skeleton of the youth was dated (Brindley et al. 1983), returning a determination of 4675 ± 35 BP (GrN - 9708), calibrated (2σ) to 3625-3366 cal BC.

Sampling

Petrous temporal bones were located for the male youth only (Original reg. no. E180:2; new reg. no. E183:2), one of which was sampled for aDNA extraction, ARD2 (Fig. I.8).

Ashleypark, Co Tipperary

Site Description

The site was excavated in 1980 (Manning et al. 1985) and is of Linkardstown type. A trapezoidal megalithic structure or tomb was covered by the cairn, containing the remains of an adult male and a child, accompanied by plain and decorated Neolithic pottery. Much animal bone was found among the stones which filled the rest of the structure, and the bones of another child were recovered from beneath these stones.

The human remains within the chamber were those of an elderly adult male (Burial 1) and a child of 4-5 years (Burial 2), both in a disarticulated and jumbled state. On the east side of the outer part of the tomb Burial 3 was discovered. This consisted of the bones of a child of approximately eight months. Some of the bones had slipped down between the stones, and the skull, in four main pieces, was scattered over an area 40cm by 30cm. It is therefore uncertain as to what state the bones were in when originally deposited.

The skull of Burial 1, a robust adult male aged about 60 years, consisted of 41 fragments, most of which could be assembled to form the temporal bones, a large part of the cranial vault and the right zygomatic bone with an adjoining portion of the maxilla. Burial Number 2, a child of 4-5 years, gives no mention of skull bones, only mandible and teeth are identified. Burial number 3 is the remains of an infant aged 8 months. Many of the skull sutures were unfused, and the remains include the right temporal bone and the left petrous temporal bone. The adult inhumation (Burial 1) returned a date of 4765 ± 40 BP (GrN-11036) calibrated (2σ) to 3635-3522 cal BC (Brindley & Lanting 1989).

Sampling

Petrous bones were identified for Burial 1 and Burial 3 (ASH1 and ASH3) and were sampled for aDNA analysis (Fig I.8).

Norrismount, Co. Wexford

Site Description

The site consisted of a mound, covering a central cairn, in turn covering a cist (Lucas 1950). When the capstone was removed the cist was found to be filled to depth of about 15 cm with a layer of fine yellow clay which had filtered in between the stones. Professor Keenan examined the bones from the cist and reported that -

“The find contains fragments of the following human bones: Skull - small fragments of the jaws (including teeth) and the side wall.”

Part of the broken skull protruding from the clay at the Northeast end of the cist was the only portion of the skeleton which was visible. Close to the skull were found a few fragments of pottery and a further two near the other end of the cist.

Sampling

A single disarticulated petrous (NM61) was sampled from the site (Fig. I.8).



Figure I.8. Samples taken from Linkardstown cists for aDNA analysis. (Note: The archaeological tags photographed with JP14 and JP15 are incorrect. See the main text for further detail.)

Wedge Tombs

Roughan Hill, Co. Clare

Site Description

During the summer of 2015 a wedge tomb (CL017-180002) on Roughan Hill, in Parknabinnia townland, Co. Clare, was excavated by staff and students of The Irish Fieldschool of Prehistoric Archaeology, part of a larger project on the area. (Ó Maoldúin 2015). A number of other wedge tomb sites have been identified within this region.

The tomb was dilapidated and missing several structural monoliths. The chamber itself was small in size, approximately 2x1 metres in length and width and no more than 1.2 metres in height. It was also surrounded by a circular stone cairn, approximately 7 metres in diameter. It appears to have been the first structure built on the site. Cremated and unburnt human remains were found dispersed inside the chamber, with no notable individual deposits or grave goods identified. Post excavation work is ongoing. Preliminary dates suggest an initial phase of Chalcolithic use, associated with the cremated deposits, followed by Early Bronze Age activity, marked by dates from unburnt remains (Ros Ó Maoldúin, pers.comm.).

Sampling

Four petrous bones (RH98, RH125, RH468 and RH473) were retrieved during excavation and sampled for aDNA analysis (Fig. I.9). One of these, RH468, was subsequently sent for dating and returned an Early Bronze Age determination (3483 ± 56 ; 1943-1663 cal BC; UBA-35062).

Further Notes

Two samples, RH98 and RH468, were determined to be sisters (Appendix Two; Fig. II.9).

Labbacallee, Co. Cork

Site Description

This wedge tomb, excavated by Leask and Price (Leask et al. 1935), consists of a chamber and sealed end chamber, within a wedge-shaped cairn. Pottery of beaker type was also reported, although its characterisation is doubtful (Brindley pers. comm.). The bones represent at least five individuals. Skeletons B and C were found scattered over the floor of the outer larger chamber. Skeleton C is that of a child. There was evidence of disturbance of these burials after decomposition, with a skull (Skull A) found sitting on its base within the middle of the deposit. This skull did not belong to skeletons B or C, both of whom have skulls already accounted for. Skeleton E consists of a few fragments of bone found

in a cist outside the main monument. Two burials were found inside the smaller sealed chamber, one represented by a few fragments of cremated bone and the second being a female inhumation with the skull and neck bones missing (Skeleton D). It is thought possible that Skull A belongs to this individual.

Skeleton D appears to be the primary interment with a calibrated date range of 2456-2424 and 2402-2138 cal BC (GrN-11359). The ranges for skeletons B and C are 2458-2102 and 2092-2038 cal BC (OxA-2759), and 2202-1874, 1842-1814 and 1804-1776 cal BC (OxA-2760) respectively (Brindley & Lanting 1991). It is unknown whether these burials occurred at intervals or within a relatively short time period.

Sampling

Three petrous bones from Labbacallee were located, two being a pair belonging to the female Skull (Skull A; E5:212). The third petrous had no identification other than site name and registration number (E5:213), making it impossible to link to any of the published descriptions of skeletons. A single petrous from Skull A, LB212, and the unidentifiable petrous, LB213, were both sampled for aDNA analysis (Fig. I.9). Both samples were submitted for AMS dating as part of this project. A sample of E5:213 was submitted for radiocarbon dating as part of this project, however it failed quality control.

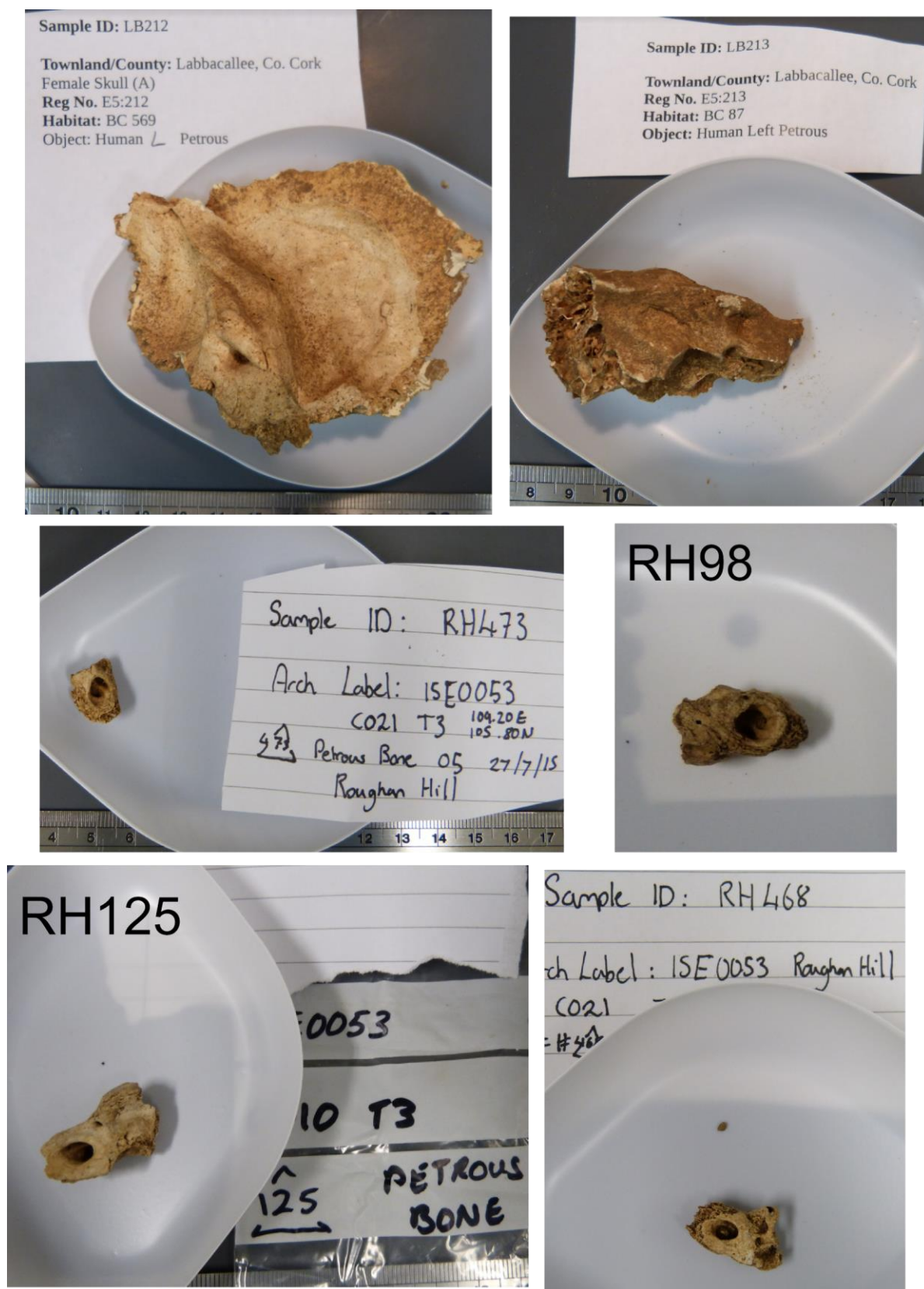


Figure I.9. Samples taken from wedge tombs for aDNA analysis.

Chalcolithic and Early Bronze Age Inhumations

Cist Burials

Glebe, Rathlin Island, Co. Antrim

Site Description

The three individuals from Rathlin Island were discovered during excavations in Church Bay in 2006 (Sloan et al. 2008). The style of burial, the associated pottery and radiocarbon dates derived from the three individuals indicated that they were Early Bronze Age in date. C127 was a 40-60 year old male who had been buried in the interior of a stone cist (C124). He was lying on his left side in a crouched position along a north-south orientation, with the head to the north. He was a robust individual who would have had a height of approximately 5'11". His bones revealed signs of an active lifestyle and osteoarthritis, Schmorl's nodes and torsion of a number of lumbar spinous processes were apparent in his vertebrae, while os acromiale was visible in his left scapula. The tuber of his left calcaneus displayed healed lytic lesions which may have been caused by a soft tissue injury of his heel at some stage during his life (Murphy & McGranaghan 2008). The remains of a complete tripartite bowl Food Vessel were recovered adjacent to his lower legs in the south-east corner of the cist (Carver 2008).

C122 (Sk 1) and C105 (Sk 2) were both young adult males whose disarticulated remains were recovered in association with a disturbed cist (C112) adjacent to the complete cist (C124). Disarticulated bone from an adult female was also recovered from the deposit. Signs of infection or inflammation were visible in the right parietal and mastoid region of the cranium of C122 (Sk 1) and in the right scapula of C105 (Sk 2) (Murphy & McGranaghan 2008)). The practice of using a single cist for the interment of multiple individuals was also identified during excavations on Rathlin during the 1980s when Kenny Wiggins discovered a cist that contained the remains of at least five individuals (Grave 2). Interestingly, the radiocarbon dates of C122 (Sk 1) and C105 (Sk 2) revealed that it would have been impossible for them to have been buried at the same time. As such, it is possible that the cist had been re-opened at some stage to facilitate further burials or that one of the individuals was an 'ancestor' whose remains had been curated for a period of time prior to their burial in the cist (Sloan et al. 2008).

Sampling

Petrous bones were sampled from the C127 (RM127), C122 (RSK1) and C105 (RSK2) individuals (Fig I.10). Early sampling of a phalanx (RM127-2) and two molars (RSK1-2 and RSK2-2) from the same individuals had also been undertaken for the project, but did not yield substantial endogenous contents. No pictures for these earlier samples are available.

Sliguff, Co. Carlow

Site Description

The site consists of a short rectangular cist containing the crouched inhumation of a young adult male (1974:27), accompanied by a tripartite bowl (Waddell 1990). After an initial anomalous date, the redating of the human remains returned a date of 3570 ± 70 BP (OxA-2671), which calibrates to 2134-1701 cal BC. Brindley (Brindley 2007) places the bowl in the second phase of development of the bowl tradition, stage 2, which is dated to 2080-1980 cal BC.

Sampling

A right disarticulated petrous temporal from the individual was sampled (SG27; Fig. I.10).

Vermount, Co. Carlow

Site Description

The site consisted of a single short cist containing two inhumation burials, a flint blade and a ribbed bowl. The bones (P1955:23) most likely belong to two adult females of small build, and appear to be contemporaneous. A sample of bone returned a radiocarbon date of 3575 ± 40 BP (GrA-24158), which calibrates to 2032-1775 cal BC. This is in concordance with the assignment of the ribbed bowl to stage 3 in the development of the bowl tradition (Brindley 2007), which covers the period between 1980 cal BC to 1930/20 cal BC (Cahill & Sikora 2011).

Sampling

Two right of petrous bones (VM231 and VM232) were identified and sampled (Fig. I.10).

Coolnatullagh, Co. Clare

Site Description

The Coolnatullagh site (97E0204:5) consists of a cairn within which a central cist was constructed. The structure was damaged during farming operations in 1997 and subsequently excavated by James Eogan (Eogan 2002). The presence of beaker pottery and radiocarbon dating (2460-2140 cal BC) indicated Chalcolithic usage, with later Early Bronze Age activity also apparent. The fill from the central cist contained unburnt bone from an adult and child and cremated bone from an adult. A radiocarbon determination from the unburnt subadult (F1005), returned a date of 1880-1610 BC (OxA-10572).

Sampling

The petrous bone of F1005 was taken for aDNA analysis (CN1005; Fig I.10).

Lisooly, Co. Donegal**Site Description**

The site consists of up to four short Early Bronze Age cists, one of which was empty, save for a femur, and another of which contained a cremation. The remaining two cists, Graves 1 and 4, both contained inhumations. Grave 1 contained the remains of a young adult (04E0505:2), probably male, while Grave 4 contained a juvenile individual (04E0505:6), possibly male (Cahill & Sikora 2011). Bipartite bowls accompanied both individuals. A sample of bone from Grave 1 returned a radiocarbon date of 3755 ± 40 (GrA-29067), which calibrates to 2290-2030 cal BC.

Sampling

Although there is mention of petrous temporal bones for both graves 1 and 4 in the site report, skull fragments of only one individual could be identified in the NMI archives. These remains had no registration number attached. However, based on description we believe these bones to belong to the individual from Grave 1. The skull consists mainly of the right side of the calvarium in one intact piece. This includes the right temporal bone and petrous, which are both slightly decayed. The petrous portion (Fig. I.10) was removed from the skull within the NMI archives using a Dremel saw (LC04).

Lisnamulligan, Co. Donegal**Site Description**

The site is composed of two Early Bronze Age cist burials, found in the same field, one of which (Grave 2), contained the remains of an adult individual (1989:46) with an accompanying upright bipartite bowl. A sample of bone from grave 2 returned a radiocarbon date of 3640 ± 70 BP (OxA-2667), which calibrates to 2205-1776 cal BC. Brindley (Brindley 2007) places the bowl from grave 2 in stage 1 of the development of the bowl tradition, which covers the period between 2160-2080 cal BC (Cahill & Sikora 2011).

Sampling

The burial contained a single disarticulated petrous with a small portion of attached temporal (Fig. I.10), which was taken for aDNA analysis (LM46).

Glassamucky, Co. Dublin

Site Description

The site consists of a short rectangular cist containing a crouched inhumation (1978:343) and tripartite bowl (Kelly 1998). This bowl was classified as a bipartite vase by Ó Ríordáin and Waddell, but was reclassified by Brindley (Brindley 2007). A sample of the unburnt bone yielded a radiocarbon date of 3765 ± 40 BP (GrN-11899), which calibrates to 2296-2036 cal BC. Brindley places the bowl in the final phase of development of the bowl tradition, stage 3, which is dated to 1980-1920 cal BC.

Sampling

A single petrous bone (Fig. I.11), with a significant portion of surrounding temporal bone, was identified as belonging to the burial, and taken for sampling (GM343).

Topped Mountain, Co. Fermanagh

Site Description

The site consists of a cairn on the summit of a hill, within which a sub-rectangular cist had been constructed (Waddell 1990). The inhumed, possibly crouched, remains of an adult male (1898:11) were found within the cist, alongside a cremation. A bronze dagger with a gold band and a tripartite vase were both found associated with the burial. The grave was subject to extensive radiocarbon dating, however only one date is considered reliable, that from carbonate belonging to the cremation, which returned a date of 3570 ± 40 BP (GrA-14761), or 2029-1774 cal BC. Brindley (Brindley 2007) places the vase in the second phase of development of the vase tradition, stage 2, which is dated from 1920-1830 cal BC.

Sampling

One petrous was identified from the site (Fig. I.11), which was taken for aDNA analysis (TOP11).

Annaghkeen, Co. Galway

Site Description

The site consists of a short cist containing an inhumation, discovered in May 1984 during ploughing activity (Cahill and Sikora 2011). The inhumation (1984:34) was that of a juvenile individual (14-15 yrs), accompanied by a bipartite vase, a flint blade, some pig bone and mussel shells. Collagen from the skeleton yielded two dates. The first (4210 ± 60 BP; GrN-13580) was dismissed as too early (Brindley 2007). The second determination gave a date of 1937-1537 BC (3440 ± 70 BP; OxA-2665). A number of other Bronze Age burial monuments have been discovered in the same townland, most notably a

tumulus, excavated in 1924 (Ó Ríordáin and Waddell 1993). The mound contained a rectangular cist, within which the skeletons of an adult male and a child were discovered alongside a ribbed bowl.

Sampling

A petrous bone from the inhumation (1984:34) was sampled for aDNA analysis (AK84; Fig. I11).

Moyveela, Co. Galway

Site Description

A short rectangular cist was excavated in 1928 and found to contain the inhumed remains of a youth (1928:381), possibly female (Waddell 1990). The burial was accompanied by a bipartite vase. Dating of the collagen from the unburnt bone returned a date of 3520 ± 40 BP (GrN-12273), which calibrates to 2286-2039 cal BC. Brindley (Brindley, 2007) places the vase in first phase of the development of the vase tradition, stage 1, which covers the period between 2020/1990-1920 cal BC.

Sampling

A single petrous bone (Fig. I.11) was sampled (MV381).

Blackhill, Co. Kildare

Site Description

The site consists of a short rectangular cist, within which were discovered the inhumed remains of a child (L1932:1A) with a necked bipartite bowl (Waddell 1990). A sample of the unburnt bone yielded a radiocarbon date of 3770 ± 70 BP (OxA-2668), which calibrates to 2458-1983 cal BC. Brindley (Brindley 2007) places the bowl in the second phase of development of the bowl tradition, stage 2, which is dated to 2080-1980 cal BC.

Sampling

A pair of petrous bones were identified as belonging to the burial, one of which was attached to a large portion of cranium. The second was attached to a section of surrounding temporal bone (Fig. I.11), which we took for sampling (BH32).

Timolin, Co. Kildare

Site Description

In 1981 a rescue excavation of a short cist discovered during ploughing activity was undertaken by Raghnaill Ó Floinn (Cahill and Sikora 2011). A second cist was discovered in 1982 and excavated by Nessa O'Connor. The first cist (Grave 1) was found to contain the inhumed disarticulated remains of an adult male of 25-30 years (1981:342), alongside a complete tripartite bowl and fragments of a bipartite bowl. Collagen from the cist yielded a radiocarbon date of 3700 ± 30 BP (2200-1980 cal BC; GrN-11364) (Brindley 2007). Brindley places the accompanying bowls in stage 2 of the development of the bowl tradition, suggesting the pottery can be placed within a period of 100 years centering on the date 2030 BC.

Sampling

A single petrous bone (TM81) was sampled from the male inhumation (1981:342), see Fig. I.11.

Waterunder, Mell, Co. Louth

Site Description

Up until recently, Beaker period burials in Ireland had only been discovered within the context of megalithic monuments, primarily wedge and court tombs. However, a small number of pit burials (notoriously difficult to identify) associated with the Beaker period have in recent times been excavated, only one of which, Mell, contains unburnt remains. The site consists of two successive phases of occupation including c. 500 sherds of Beaker pottery and an inhumation burial. The crouched inhumation burial, that of an adult female, was discovered on high ground, approximately 600m north west of the occupation deposits (McQuade 2005). It was in a partially stone-lined pit grave and in a poor state of preservation. A sample of tooth returned a date of 2488–2206 cal BC (Wk-17463; 3894 ± 50 BP).

Sampling

The crown of a tooth (ML1), the root of which had been removed for radiocarbon dating, was taken for ancient DNA analysis.

Stonepark, Co. Mayo

Site Description

The site consisted of a short rectangular cist, disturbed in the late 19th century, and excavated in 1933. The remains of an adult male (1933:3409), in a possibly crouched position, were found within, alongside

a bipartite vase and a 'curiously shaped fossil coral' (Waddell 1990). A radiocarbon date on the unburnt bone gave a result of 3625 ± 30 BP (GrN-12275), which calibrates to 2122-1898 cal BC. Brindley (Brindley 2007) places the vase in the earliest phase of development of the vase tradition, stage 1, which is dated to 2020/1990-1920 cal BC.

Sampling

A single petrous was located with a significant portion of surrounding cranium. This was removed and taken for aDNA analysis (SP3409; Fig. I.11).

Fourknocks II, Co. Meath

Site Description

Fourknocks II is a complex site, located 50m to the east of the Fourknocks I passage grave, consisting of a cairn, several ditches and a short, roofed megalithic passage. A mound was built over these earlier structures. While primarily a Neolithic site, Fourknocks II was re-used as a cemetery in the Chalcolithic and Early Bronze Age, with possibly as many eight burials discovered, including four cists (Hartnett and O'Sullivan 1971). Of primary interest to our project is the well preserved burial in Cist 2, that of a female aged approximately 25 years (E18:142). Reuse of megalithic sites in the Chalcolithic and Early Bronze Age is a well established phenomenon. Recent dating programmes of Fourknocks II suggest the possibility of Chalcolithic activity, identifying this individual as a primary candidate for aDNA analysis.

Sampling

A single disarticulated petrous bone, with some surrounding temporal bone, is associated with E18:142 and was taken for aDNA analysis (FK142; Fig I.11).

Culleens, Co. Sligo

Site Description

The site consists of a short Early Bronze Age cist containing an inhumation (2009:23) and a Tripartite vase. The remains belong to a female individual. Two radiocarbon determinations from the bone gave a mean date of 3685 ± 50 BP, or 2210-1920 cal BC. Brindley (Brindley 2007) places the vase in the earliest phase of development of the vase tradition, stage 1, which is dated to 2020/1990-1920 cal BC (Cahill & Sikora 2011).

Sampling

A single petrous with some surrounding temporal bone was located and sampled (CL23; Fig I.11).

Treanmacmurtagh, Co. Sligo

Site Description

The site consists of a cairn, which contained a short cist, located on a mountain ridge close to the Carrowkeel complex. The remains of a crouched inhumation (1966:116) were found within, belonging to a child of 10-12 years old. The burial was accompanied by a small anomalous bowl (Brindley 2007), placed in front of the skull. This vessel was classified as a vase by Waddell (Waddell 1990). The cist also contained the cremated remains of an adult. Two sherds of a second, probably bowl-shaped vessel were also found. A radiocarbon determination from the unburnt bone yielded a date of 3665 ± 35 BP (GrN-11356), which calibrates to 2015-1758 BC.

Sampling

A pair of petrous bones were identified for this burial, the right of which had a grey surface colour, possibly the result of exposure. The left petrous was sampled (TR116; Fig. I.11)

Stranagalwilly, Co. Tyrone

Site Description

The Stranagalwilly site lies close to the foot of the Sperrin Mountains, and consists of three burial cists, which were excavated between 1961-1962, and a fourth outlier, excavated later in 1969 (Waterman et al. 1993). The cists contained bowl pottery and were ascribed to the Early Bronze Age. However, two anomalous dates, 406-367 cal BC and 2467-2298 cal BC, have been reported from the INSTAR database (2317 ± 23 and 3896 ± 28 ; UBA-13544 and UBA-13543) and should be treated with skepticism.

Sampling

A petrous bone (ST63; Fig. I.11) was sampled from the juvenile individual with the anomalous Iron Age date (406-367 cal BC).

Bolinready, Co. Wexford

Site Description

The site consists of a short rectangular cist, within which were discovered the inhumed remains of a young adult male (1965:79) (Prendergast 1968). The upper portion of a child's fibula was also discovered. A sample of the unburnt bone yielded a radiocarbon date of 3620 ± 60 BP (GrN-9321), which calibrates to 2193-1777 cal BC. The remains were associated with a tripartite bowl, which Brindley (Brindley 2007)

places in the second phase of development of the bowl tradition, stage 2, which is dated to 2080-1980 cal BC.

Sampling

A pair of petrous bones were identified as belonging to the burial, from which the righthand (Fig. I.11) was taken for aDNA analysis (BR79).

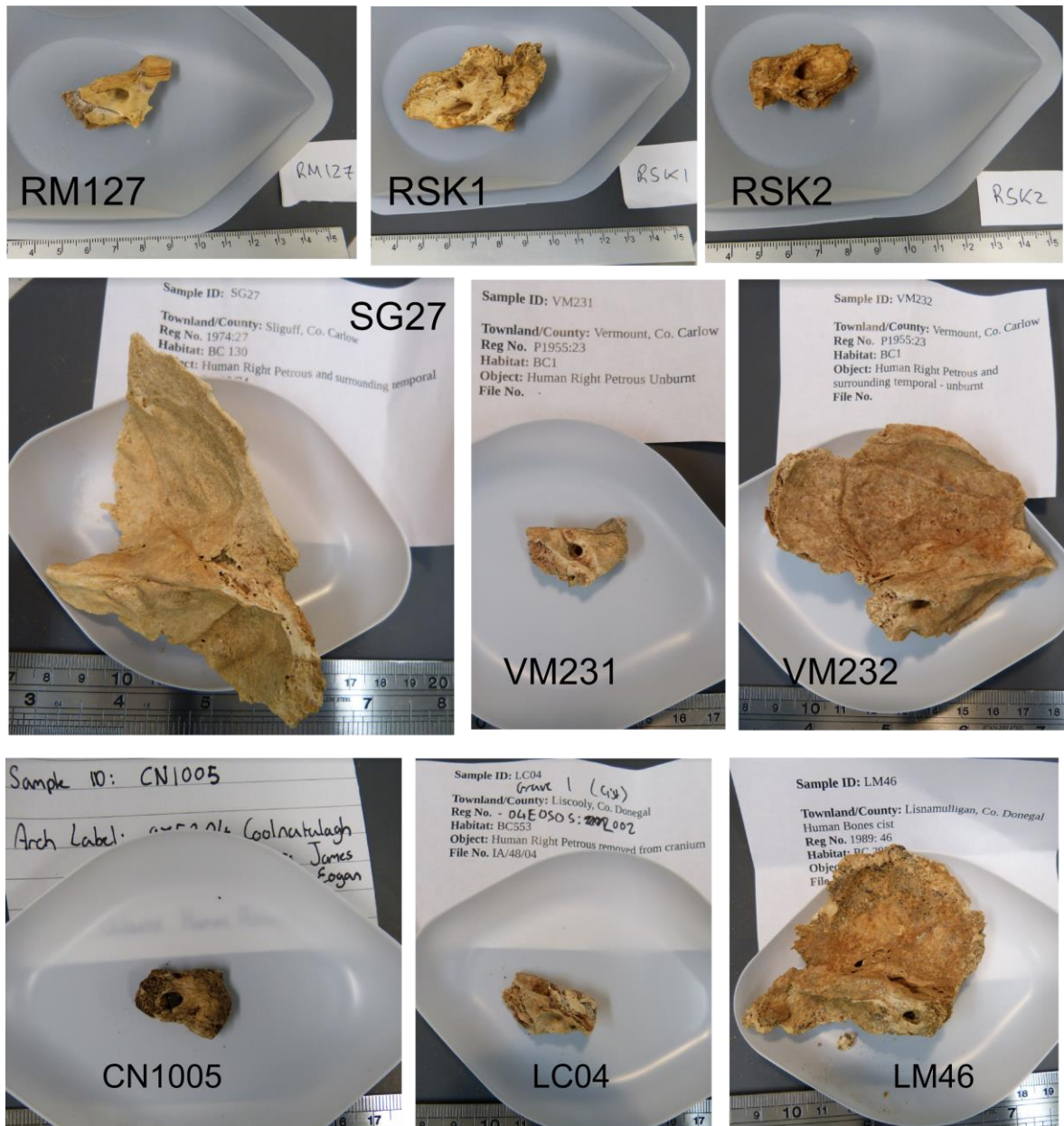


Figure I.10. Samples taken from cist inhumation burials for aDNA analysis.



Figure I.11 Samples taken from cist inhumation burials for aDNA analysis.

Cemeteries

Ploopluck (Osberstown), Co. Kildare

Site Description

The site consists of a flat cemetery, uncovered during quarrying between the years 1933 and 1935. At least six graves and possibly as many as ten were recorded. These contained nine inhumations and one cremation. Five clearly defined graves are noted in Mount et al. 1998. Four skeletons were also discovered by quarry workers in the sandpit. These were disturbed and subsequently reburied in two groups, with three skeletons in one pit and most of the fourth skeleton in the other. The remains were re-excavated and sent to the Museum in 1935, where they were registered under 1935:3-5 and 1935:6. The skeletal remains registered under 1935:3-5, thought to belong to three individuals, were demonstrated by L. Buckley to belong in fact to a single young adult male. It is unclear whether three skeletons did in fact exist, with only one being recovered by the excavators. A sample of bone from this individual returned a radiocarbon date of 3586 ± 39 BP (2111-1826 cal. BC 2 σ , UB-3973). The remains registered under 1935:6, represent an incomplete adult male, unburnt bone from whom returned a determination of 3578 ± 39 BP (2037-1786 cal BC 2 σ , UB-3974). Two sherds of a shouldered bowl were also found, but it is not known which individual they accompanied.

Sampling

A single petrous bone was identified for both the individual registered under 1935:3-5 and the individual under 1935:6, both of which were sampled (PP35 and PP6; Fig. I.12).

Keenoge, Co. Meath

Site Description

The Keenoge site is an Early Bronze Age flat cemetery consisting of six cist burials and eight to ten pit burials, which in total contained the remains of 26 individuals. The site was excavated between 1926-1936 and is described in Mount & Buckley 1997. The finds included nine complete bowls, one to two vase urns, an encrusted urn, a jet necklace, one bronze knife and a number of flint objects. We will focus here on the five graves from which petrous bones were located for the project.

Grave 3 is a pit burial containing the inhumed remains of an adult female (1929:14-15) and an adult cremation (1929:16b). The burial was associated with two bowls, one constricted and one simple. A bronze tanged knife, a thumb scraper, and several stone and flint artefacts were also recovered. Brindley (2007) reports that a sample of unburnt bone from this grave returned a date of 3685 ± 45 BP (GrN-12272), which calibrates to 2200-1947 cal BC. However, the registration ID is reported for this bone as

1929:19. This ID in fact corresponds to the inhumation in Grave 5. Brindley also reports a date for the inhumation from Grave 5, but finds it to be anomalous. For this reason we do not take this date to be reliable. Carbonate from the cremation was also dated, giving a result of 3640 ± 50 BP (GrA-17385), calibrating to 2188-1887 cal BC. Both bowls were placed into the second phase of development of the bowl tradition, stage 2, which is dated to 2080-1980 cal BC (Brindley 2007).

Grave 4 is a pit burial, which contained the inhumation of an adult male (1929:16) in a crouched position. The grave also contained a small amount of cremated bone (1929:17).

Grave 7 is a small roughly square cist, which contained the inhumed remains of an infant (1929:24), not more than six months old and some bones from an adult.

Grave 12 is a pit burial, containing the unburnt remains of one adult male, one adult female and a child, all recorded under the registration number 1929:72. No grave goods were found associated with any of these three burials.

Grave 14 is a partially lined pit burial recorded as containing a single crouched inhumation (1929:72b). However, examination of the remains revealed that two inhumations, that of an adult male and a juvenile, were present, as well as the cremated remains of two further individuals. The crouched burial was associated with an anomalous bowl, placed by Brindley (Brindley 2007) in the final phase of development of the bowl tradition, stage 3, which is dated to 1980-1930/20 cal BC. A sample of the unburnt bone returned a radiocarbon date of 3565 ± 35 (GrA-2132), which calibrates to 2023-1775 cal BC.

Sampling

A single petrous temporal bone was located for Grave 3, while a pair of petrous bones were located for each of Grave 4, Grave 7, Grave 12 and Grave 14 (KO3, KO7, KO12 and KO14). All petrous bones were disarticulated, with several having a portion of surrounding temporal bone attached. We sampled the right petrous bone from Grave 4, Grave 7 and Grave 14, and the left petrous bone of Grave 12, as well as the single petrous bone from Grave 3. See Fig. I.12 for sample photos.

Grange, Co. Roscommon

Site Description

The site consisted of a tumulus that when excavated revealed a cemetery of eight cist graves and six secondary pit burials (Ó Ríordáin et al., 1997). Six of the cists contained bowl-type pottery and human remains. Cremations were the principal type of bone deposit found, however two cists contained unburnt material. Grave 2, a rectangular cist, contained the unburnt remains of two individuals (E65:41), those of an adult male and child, with two bipartite bowls. Grave 3, also a rectangular cist, contained the

crouched inhumation of a child (E65:10) with a barely constricted bowl. Grave 2 returned two radiocarbon dates of 3600 ± 40 BP (GrA-13331) and 3620 ± 80 BP (OxA-2761), while Grave 3 gave radiocarbon determinations of 3740 ± 40 BP (GrA-13332) and 3770 ± 70 BP (OxA-2664). These calibrate roughly to 2200-1757 cal BC for Grave 2, and 2284-2030 cal BC for Grave 3.

Sampling

A pair of petrous bones were located for Grave 3, while a singular petrous bone attached to a significant portion of surrounding temporal was found for Grave 2. A petrous bone from each grave was sampled (GR41 and GR10; Fig. I.12).

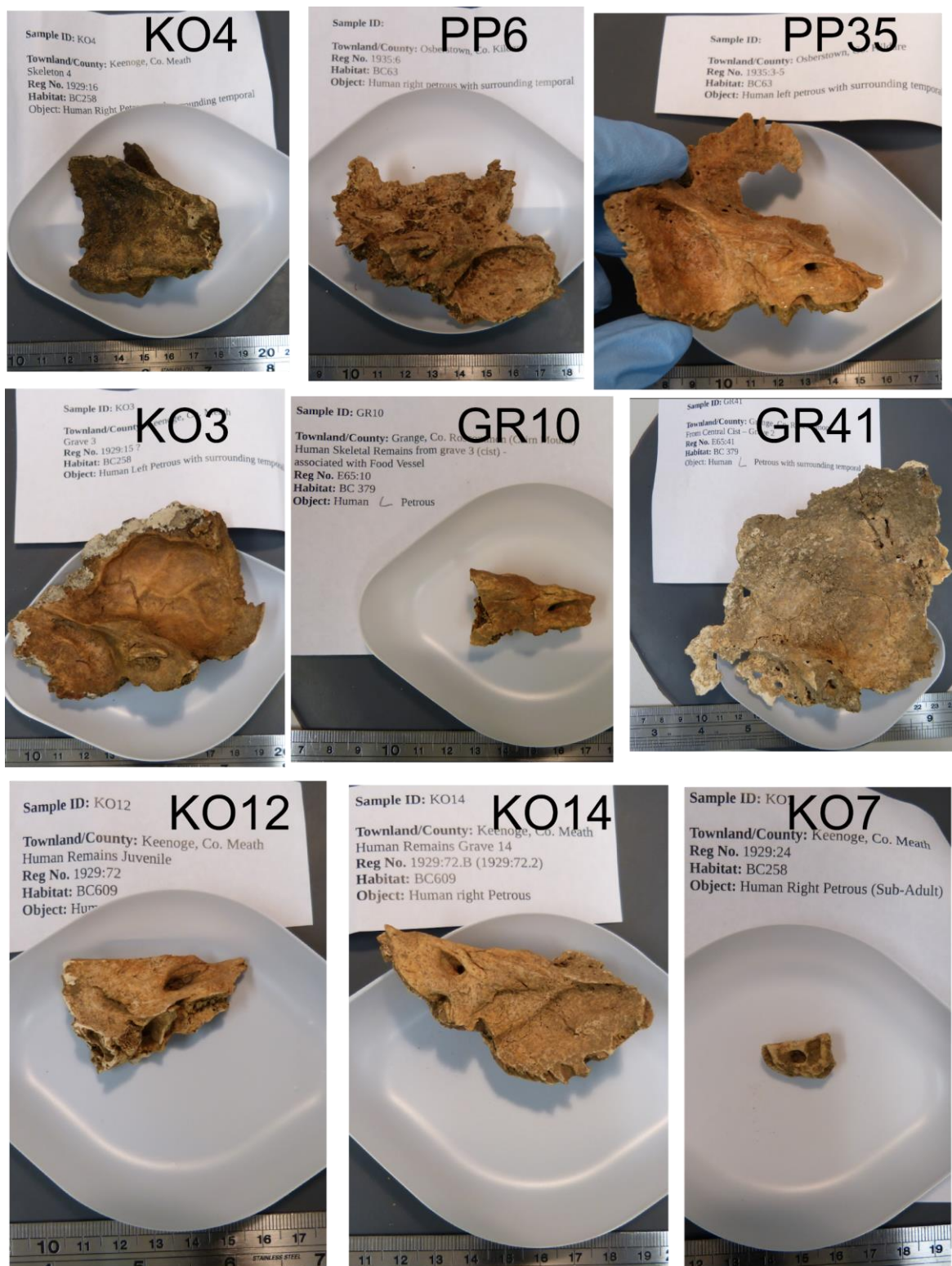


Figure I.12. Samples taken from Bronze Age cemeteries for aDNA analysis.

Late Iron Age and Early Medieval Burials

Courtmacsherry, Co. Cork

Site Description

The site consists of a lintel grave, within which the extended inhumation of an adult male (1966:38) was found. There were no accompanying artefacts. A sample of bone yielded a radiocarbon determination of 1445 ± 45 BP (GrA-24500), or 540-662 AD (Cahill & Sikora 2011).

Sampling

A single petrous bone with some attached temporal was identified as belonging to the site. However, it is stored under the registration number 1966:37, rather than 1966:38. As there are no other human remains associated with this site we presumed this to be the adult male individual and took the petrous bone for aDNA analysis (CM37; Fig I.13).

Ballymacaward, Co. Donegal

Site Description

This site consists of a Bronze Age cairn that was reused on at least four separate occasions for burial in the Iron Age and Early Christian period. This is an excellent example of an ancestral boundary *ferita*. The first two occasions of reuse involved the insertion of cremations in the Iron Age. The next stage of reuse occurred in the fifth to sixth century AD, involving the construction of slab-lined cist graves, followed by a final phase of reuse in the seventh century AD, represented by the insertion of unprotected dug graves (O'Brien 1999). All burials appear to belong to females, with one possible exception. Petrous temporal bones were identified for Burials 2 and 3.

From Robert M Chapple's *Catalogue of Radiocarbon Determinations & Dendrochronology Dates* (sites.google.com/site/chapplearchaeology/home) the following information could be gleaned -

“Final Phase – women’s graveyard 9 extended inhumations, laid west-east, in unprotected dug graves. Date from one (B2, F4) - 20 years. Not seen publication, just typescript copy supplied by author”

“Human bone from extended inhumation burial laid west-east (head west) in slab-lined long-cist grave (B3, F6). Elderly female (60+ years) with post-menopausal osteoporosis of the spine. Not seen publication, just typescript copy supplied by author”

We presume these dates to refer to Burials 2 and 3. The dates are recorded as 1406 ± 20 (UB-4170) and 1592 ± 20 (UB-4171) respectively, which calibrate to 608-659 AD and 421-536 AD.

Sampling

Three petrous temporals were identified, a pair belonging to Burial 2 (Feature 5), with registration number 97E0154:4, and one belonging to Burial 3 (Feature 6), with registration number 97E0154: (written on the bone it looks like 97E0154:5). We sampled one petrous bone from each burial (BW2 and BW3; Fig. I.13).

Derrynamanagh, Co. Galway

Site Description

The remains of several late Iron Age burials were discovered during farm works at a large rath (GA085-045), in Derrynamanagh, Co. Galway and excavated by Professor Etienne Rynne in 1969. The site is approximately 12.5km west of Athenry and only 4km from the rath of Feerwore, the proposed original site of the Turoe Stone. It is situated within a steep bend of the Rford river and surrounded by bog. The rath consists of a bank and external earthwork, which form a subcircular enclosure. The burials included two single inhumations in pits and one multiple burial in a stone lined cist. An unpublished report of the skeletal material, written by Professor Stephen Shea of the NUIG Anatomy Dept. (1970), identified the remains within each of the single burials to be of young children, approximately 2-3 years in age, while the multiple burial contained the remains of at least 13 adults.

Sampling

Nine petrous bones, DM01-DM09, were located and sampled from the site (Fig. I.15)

Further Notes

Extensive kinship was found among individuals at the site (Appendix Two; Fig. II.9). A pair of sisters were identified, DM02 and DM05, to whom DM09 is a second degree male relative. Given all three individuals share the same mitochondrial haplogroup, it is likely DM09 is related to the sisters via the female line. The possible scenarios include a maternal uncle, a nephew (another sister's child), a half-sibling who shares the same mother, but not a grandfather or grandchild. All three individuals share a fourth degree relative, DM04, while DM09 shares a separate fifth degree relative, DM08.

Ballybunnion, Co. Kerry

Site Description

The site consists of a long cist containing an inhumation burial. The excavation of this cist revealed two further inhumation burials, one within a lintel grave, and the other unprotected. Grave 2 contained the inhumed remains of an adult male (1987:54), which dated to 1400±35 BP, or 583-674 AD (Cahill & Sikora 2011).

Sampling

A petrous bone belonging to the adult male (1987:54) was sampled (BB54; Fig I.13).

Ballyeagh, Co. Kerry**Site Description**

The site consists of a lintel grave containing an inhumation burial, close to the Ballybunnion site. The remains belonged to an adult female (1979:81), which were dated to 1455 ± 45 BP, or 470-663 AD (Cahill & Sikora 2011).

Sampling

A skull with one disarticulated petrous was located under the site name, however the registration ID was missing from the sample label. As only one set of remains has been recovered for the site we assumed it belongs to the female individual (1979:81) and took the petrous for aDNA analysis (Fig. I.13).

The Curragh, Co. Laois**Site Description**

Five burials were excavated in a sandhill near Killeshin, Co. Laois, representing an early medieval cemetery of substantial size (Cahill & Sikora 2011). The remains of at least 14 skeletons were recovered. Grave 2 contained the remains of an adult female, who suffered from spina bifida. A sample of bone from this burial returned a radiocarbon date of 1370 ± 40 , calibrating to AD 598-767.

Sampling

A single tooth (CR01) was sampled from the adult female inhumed in Grave 2 (Fig. I.13).

Ballyglass Middle , Co Mayo**Site Description**

The site most likely represents an Iron Age ringbarrow re-used for burial approximately 400 to 600 AD. The skeletal remains were disturbed during quarrying and original burial positions cannot be determined. The remains of an adult male and female, as well as a juvenile of 3-5 years were identified. Radiocarbon determinations for the adult male and female returned dates of 1765 ± 70 BP (80-420 AD; OxA-3871) and 1590 ± 60 (330-610 AD; OxA-3872) respectively.

Sampling

A pair of petrous bones and two right petrous bones were identified as belonging to the site. All were labelled under “Kiltullagh Hill, td. Ballyglass Middle” and are reported in Cribbin et al. 1994. A skull was also found under the label “Kiltullagh” , which we believe is associated with a separate report (McCormick et al. 1995), and is not of interest to this project. We sampled the three right petrous bones (BW44, BW45 and BW46; Fig. I.13) identified from (Cribbin et al. 1994).

Collierstown, Co. Meath

Site Description

The late Iron Age/early medieval cemetery site at Collierstown (E3068) was uncovered during construction of the M3 motorway and excavated by Rob O’Hara (O’Hara 2009). The site was associated with a ring-ditch, and likely not ecclesiastical in nature. Radiocarbon dates suggest the burial ground went out of use in the seventh or eighth century. This is in keeping with a wider trend of burial within ancestral grounds, small enclosures and mounds (thought to have been known as *fertá*) throughout the fifth to seventh centuries. Only after this time did the majority of the population begin to bury their dead in enclosures associated with churches.

Sampling

Three samples were used for aDNA analysis (CLT622, CLT627 and CLT640). No photos are available.

Claristown, Co. Meath

Site Description

The site consists of an Early Iron Age ring ditch, which was possibly dismantled and replaced with a cairn ring (Russell et al. 2002). Fifteen inhumation burials were found associated with the site, the majority of which were extended in a supine position, with the head looking to the west. Burial 14 was that of an adult male individual, placed centrally within the ring ditch, in a lintel grave. A radiocarbon determination from the remains returned a date of 1790 ± 80 BP (Beta-185252), or 60-420 AD. Given the large margin of error for this radiocarbon result, as well as the burial context, it is suggested Burial 14 dates to the end, rather than middle, of this range (McGarry 2007).

Sampling

A pair of petrous bones were located for the individual from Grave 14, one of which was sampled (CT14; Fig. I.14).

Johnstown, Co. Meath

Site Description

The Johnstown site is an early medieval enclosure with two separate accompanying burial grounds (Clarke 2002). The first, which contained the majority of burials, numbering 424 in total, was within the phase 3 enclosure. This was likely used as a burial ground until pre-famine times. The second cemetery site was within the 'Scaruppa' area, traditionally recognised as the location of the cillín burial-ground. Sixty-one infant and two adult burials were found here, likely of later date than the first burial ground (100-200 years ago). The site was also had evidence of settlement, such as refuse pits and smelting pits. The excavation was completed in 2002.

Sampling

Three samples were used for aDNA analysis (JNT400, JNT616 and JNT645). No photos are available.

Hill of Ward, Co. Meath

Site Description

The Hill of Ward is located within the megalithic Brú na Boinne complex, which has seen persistent human activity since the Neolithic period. The site is composed of four concentric earthworks, that enclose an inner area approximately 150 metres in diameter. It was excavated in 2014 by Stephen Davis, who identified three phases of use. The first phase corresponds to its construction in the Late Bronze Age (1200-800 BC) and the last to the Late Iron Age/Early Christian period. A child burial dating to the later phase of the site (1580 ± 30 ; 410-546 cal AD; Beta-388456) was found next to a ritual deposit of cow bones. There is no evidence of other burial on the site, save some scattered Bronze Age remains, and the infant was buried with care. Given the ritualistic nature of the site, this suggests the child may have belonged to someone of significance in the society.

Sampling

A petrous bone (HW14) was taken from the child burial for aDNA analysis (Fig. I.14).

Knowth, Co. Meath

Site Description

Knowth is a passage grave site in the Brú na Boinne complex, which saw further burial activity in the Iron Age period (Eogan 2012). Fourteen crouched inhumation burials were discovered which could be assigned Late Iron Age dates based on accompanying grave goods (beads, gaming pieces, dice), burial

position, or radiocarbon determinations (100 BC - 300 AD), with a further six unclassified, but likely Iron Age in date. Two of these inhumations, E70:189 (Burial 4) and E70:147 (Burial 10), had particularly early radiocarbon dates, 86-253 AD (1830±30 BP; GrN-15369) and 175-50 BC (2095±20 BP; GrN-15372) respectively, falling within a period where the near universal burial practice in Ireland was cremation. Burial 4 consisted of the remains of an approximately twenty year old individual, buried close to the kerb of Passage Tomb 2. Burial 10 was located to the south of Passage Tomb 1. The remains were poorly preserved and no grave goods were present.

Sampling

Petrous bones from Burials 10 (KNW10) and 4 (KNW4) were taken for aDNA analysis.

Further Notes

KNW10 and KNW4 were determined to be male and female respectively (Fig. I.14).

Rossnaree, Co. Meath

Site Description

The site consists of a sub-rectangular grave pit that may have been covered by a small cairn (Cahill and Sikora 2011). There was no evidence of stone lining within the grave itself. It was investigated by Dr Joseph Raftery in 1942 and found to contain the remains of three adult females and an infant (1942:1917.2), accompanied by a decorated silver earring. The remains of the infant, not more than six months old, were found over the pelvic area of the most complete adult skeleton (Skeleton 1), belonging to a female aged 26-45 years. Collagen from Skeleton 1 yielded a radiocarbon determination 1660±40 BP, or 257-553 AD (GrA-24354).

Sampling

A petrous bone was taken from Skeleton 1 (RNR1) and the infant remains (RNR2) for aDNA analysis (Fig. I.14).

Further Notes

RNR1 and RNR2 were found to be mother and child on the basis of kinship analysis (Appendix Two; Fig. II.9).

Lehinch, Co. Offaly

Site Description

The site consists of both an Early Bronze Age burial and early medieval graves, with a ring-ditch in close proximity, suggesting continuity of burial tradition from the prehistoric period (Cahill and Sikora 2011).

A minimum of six adults (three males, one female and two of indeterminate sex) and one juvenile were identified among the inhumations, suggesting the site was a generalised cemetery, possibly for a family or clan. Grave 3, an oval pit lined with stone, contained the inhumation of an adult of indeterminate sex (1979:27). Grave 4, another stone-lined oval pit, contained the remains of an older adult male (1979:28).

Sampling

Petrous bones from both Grave 3 (LH27) and Grave 4 (LH28) were taken for aDNA analysis (Fig. I.14).

Further Notes

Molecular sex determination identified the remains in Grave 3 to belong to a female.

Tinnapark Demesne, Co. Wicklow**Site Description**

The site consists of at least eight graves, the majority of which were lintel lined, dating from as early as the mid-fifth century AD (Cahill & Sikora 2011). Grave 4 contained an extended inhumation of an adult individual of unknown sex (1963:97). The skeleton lay in a supine position with its head to the west. A sample of the bone yielded a radiocarbon determination of 1405 ± 45 BP, or 558-684 AD (GrA-24787).

Sampling

A single left petrous bone was identified as belonging to the individual in Grave 4 and sampled (TD97; Fig. I.14).

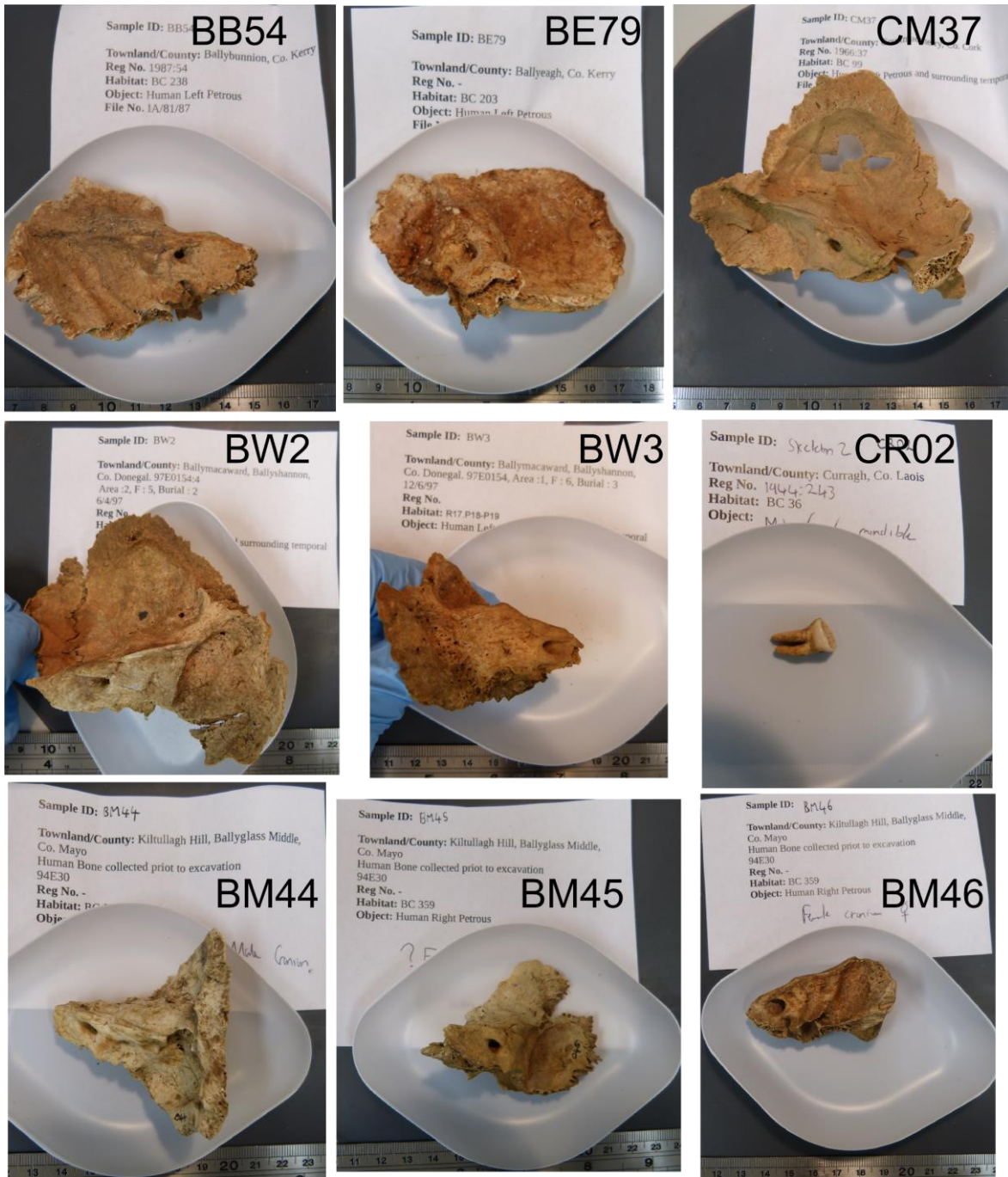


Figure I.13. Samples taken from Iron Age inhumations for aDNA analysis.

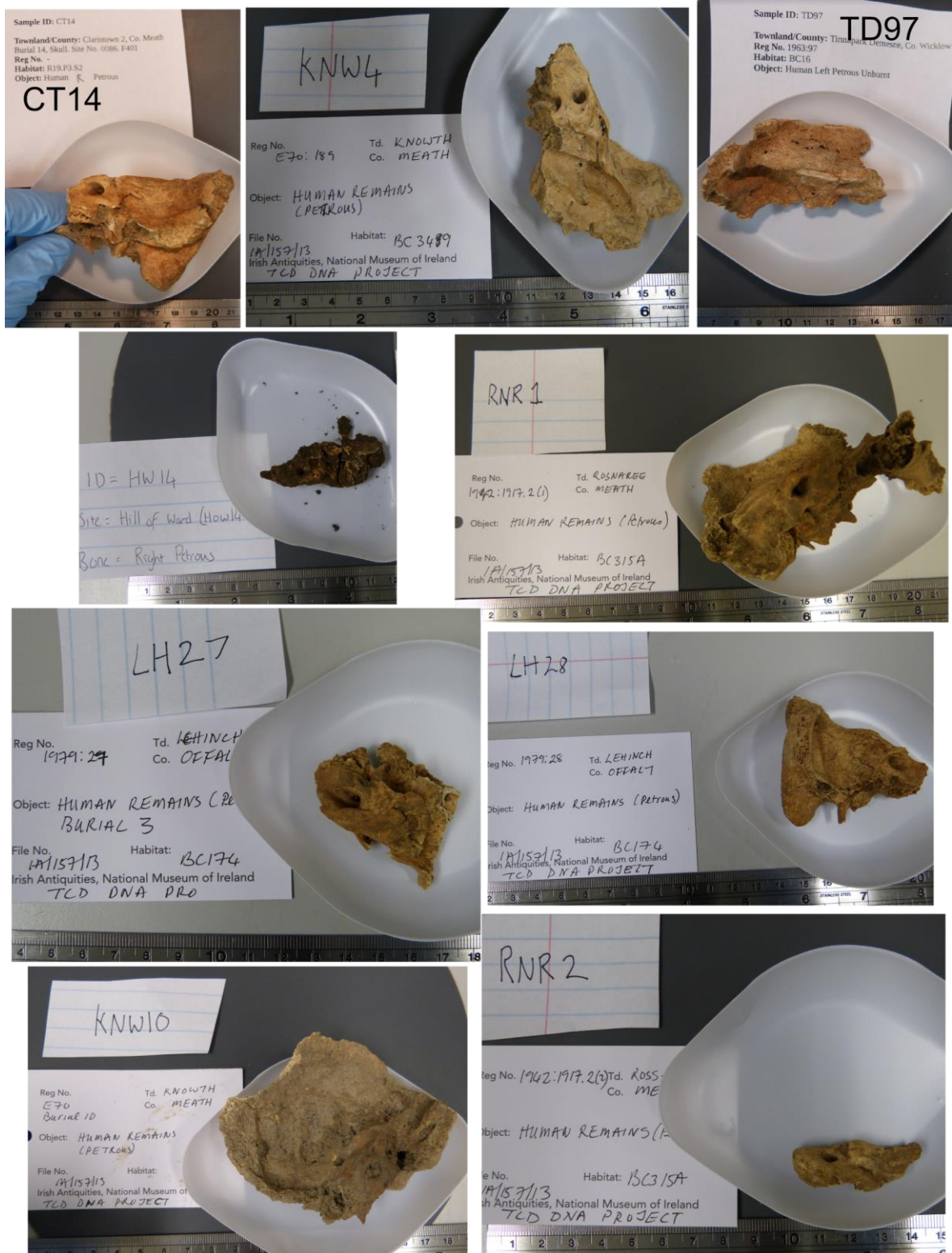


Figure I.14. Samples taken from Iron Age inhumations for aDNA analysis.



Figure I.15. Samples taken from the Derrynamanagh Iron Age site for aDNA analysis.

Other Sites

Gallow's Green, Co. Antrim

Site Description

Gallow's Green is a hangman's site in Carrickfergus, Co. Antrim. The remains of a male aged 35-50 years, with a radiocarbon determination of 91+/-21 BP (1734-1806 cal AD; UB-18958), was retrieved from a nearby beach. Osteo-biographical analysis is ongoing (Eileen Murphy pers. comm.).

Sampling

The left fourth metacarpal from the individual was taken for aDNA analysis (GG01). No picture of the sample is available.

Cherrywood I, Loughlinstown, Co. Dublin

Site Description

The site consists of a ring ditch containing a mound within the central area. An extended juvenile inhumation was discovered beneath the mound, with no indication of a grave cut into the subsoil (Ó Néill 2000). A pit was found approximately 1.3m south of the feet of the burial, within which five burnt cattle teeth were found. One of these teeth returned a radiocarbon determination of 2970±70 BP, which roughly calibrates to 1400-1000 BC, thought to represent activity after 1300 cal BC. The excavator believed the cattle tooth to be broadly contemporary with the inhumation. Given the ubiquity of cremation during the Late Bronze Age this site may represent one of the only unburnt burials identified from the period.

Sampling

A pair of petrous bones were identified as belonging to the burial. The right petrous bone was substantially degraded, with only a small portion of the bone remaining and so the left was sampled (CW1).

Cotterellsrath, Co. Kilkenny

Site Description

Cotterellsrath is the site of an enclosure structure, one of several sites that mark an association between hillforts and linear earthworks in Kilkenny (Gibbons 1990). Several unpublished excavations have occurred in the townland, which have yielded some cremated bone from pits.

Sampling

A cremated petrous bone fragment (KIL1) was taken for aDNA analysis.

Site I, Knockadoon, Lough Gur, Co. Limerick

Site Description

The area around Lough Gur lake is rich in archaeology. This includes a wedge tomb, from which human remains yielded dates spanning between 2480-1900 cal BC, and a series of occupation sites on Knockadoon peninsula. All these occupation sites produced ceramics known to date from the Neolithic through to the Late Bronze Age.

Site I, excavated by Ó Ríordáin (Ó Ríordáin 1954), is one such occupation site, consisting of the remains of a small hut. Class I, Class II, beaker and bowl tradition pottery were found in and around the hut. An extended burial, laying on its right side, was found associated with the site. The remains belong to a young person of about sixteen years and indeterminable sex. There were no grave-goods, but in the loose soil immediately under the skeleton several fragments of Class II pottery were found. Although the burial has not been dated directly, it is unlikely that it is contemporary with the later periods (medieval or early modern) of which there is evidence of activity in the vicinity. The occurrence of Class II pottery near the burial and in the post-hole at its foot suggests an early date. The dating of Class II ware is discussed in Cleary et al. (2003).

“ In terms of dating Class II ware, the probability is that flat-based coarse ware was first used in the Late Neolithic and was found in association with Class I and Beaker period ceramics. The pottery was manufactured into the Late Bronze Age, by which time the fabric was coarse and the vessels largely undecorated. It is, however, difficult to separate the flat-based coarse pottery into convenient chronological units or to introduce a new terminology for the pottery, as the vessels were produced from the same raw materials and used the same manufacturing techniques. Contexts producing only undecorated flat-based, coarse-textured wares are probably Late Bronze Age in date. “

Sampling

A pair of articulated petrous bones were identified as belonging to the burial from Site I, one of which was sampled and yielded high endogenous DNA content. However, upon dating for this project, the individual was found to belong to the Norman or Early Modern period (1448-1635 cal AD; UBA-35068).

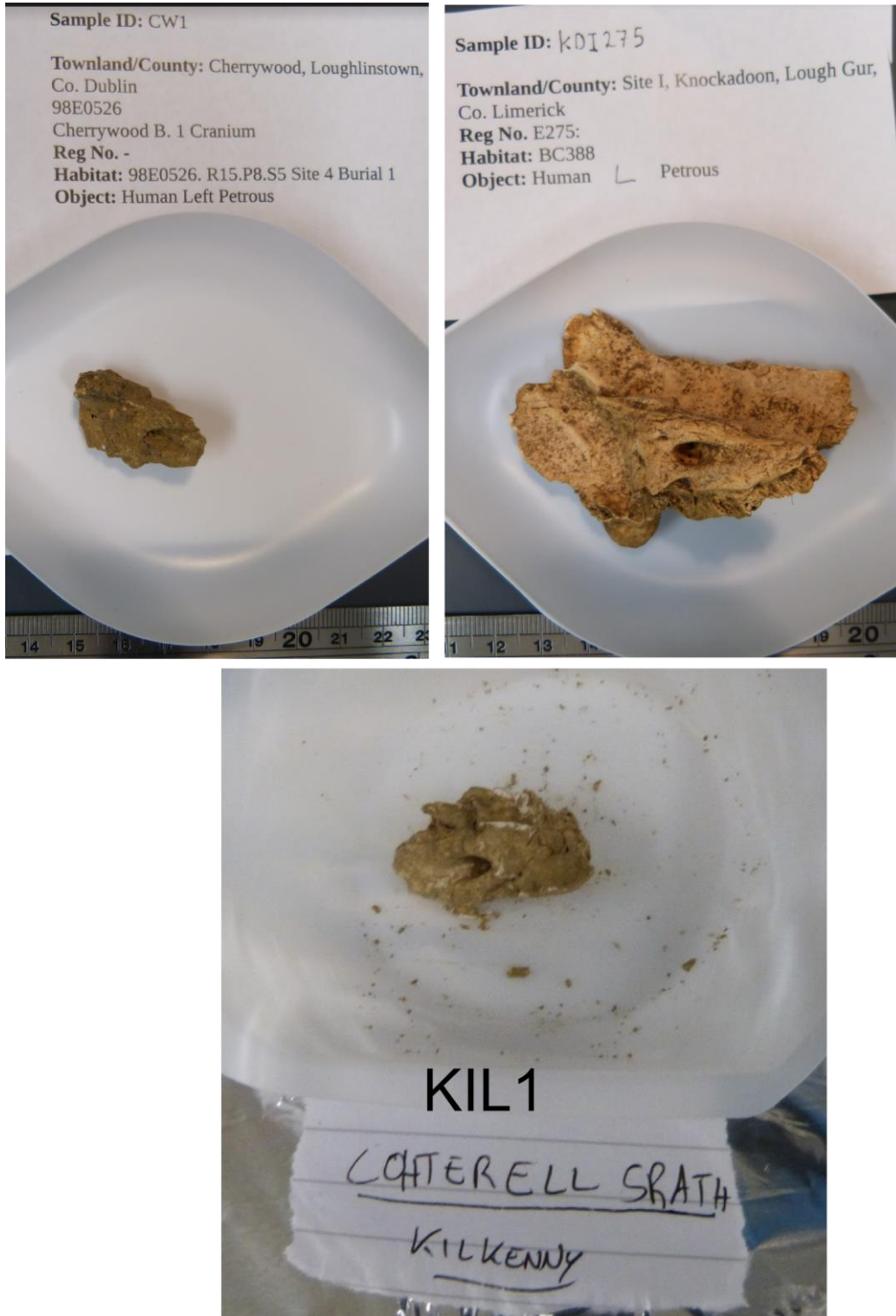


Figure I.16. Samples taken from miscellaneous sites for aDNA analysis.

References

- Beckett, J., 2011. Interactions with the Dead: A Taphonomic Analysis of Burial Practices in Three Megalithic Tombs in County Clare, Ireland. *European Journal of Archaeology*, 14(3), pp.394-418.
- Brindley, A.L., 2007. *The dating of food vessels and urns in Ireland*. Galway, Ireland: Galway University Press.
- Brindley, A.L. & Lanting, J.N., 1989. Radiocarbon Dates for Neolithic Single Burials. *The Journal of Irish Archaeology*, 5, pp.1-7.
- Brindley, A.L. & Lanting, J.N., 1991. Radiocarbon Dates from Wedge Tombs. *The Journal of Irish Archaeology*, 6, pp.19-26.
- Brindley, A.L. & Lanting, J.N., 1995. Irish bog bodies: the radiocarbon dates. In R.C. Turner & R.G. Scaife (eds.), *Bog bodies: new discoveries and new perspectives*. London: British Museum Press, pp.133-136.
- Brindley, A.L., Lanting, J.N. & Mook, W.G., 1983. Radiocarbon Dates from the Neolithic Burials at Ballintruer More, Co. Wicklow and Ardcroney, Co. Tipperary. *The Journal of Irish Archaeology*, 1, pp.1-9.
- Cahill, M. & Sikora, M., 2011. *Breaking Ground, Finding Graves: Reports on the Excavations of Burials by the National Museum of Ireland, 1927-2006*. Bray, Ireland: Wordwell.
- Carver, N., 2008. The Food Vessel, pp. 66-68 in Sloan, B., A Bronze Age Cist Burial at Glebe, Rathlin Island, County Antrim. *Ulster Journal of Archaeology*, 67, pp.60-83.
- Case, H., 1961. Irish Neolithic pottery: distribution and sequence. *Proceedings of the Prehistoric Society*, 27, pp.174-233.
- Clarke, L., 2002. An early medieval enclosure and burials, Johnstown, Co. Meath. *Archaeology Ireland*, 16(4), pp.13-15.
- Cleary et al., 2003. Enclosed Late Bronze Age Habitation Site and Boundary Wall at Lough Gur, Co. Limerick. *Proceedings of the Royal Irish Academy*, 103C(4), pp.97-189.
- Coles, B.J., 1990. Anthropomorphic wooden figures from Britain and Ireland. *Proceedings of the Prehistoric Society*, 56, pp.315-33.
- Collins, A.E.P, Morton, W. & Scott, J.H., 1954. The excavation of a double horned cairn at Audleystown, Co. Down. *Ulster Journal of Archaeology*, pp.7-56.
- Collins, A.E.P. et al., 1955. *Millin Bay: A Late Neolithic Cairn in Co. Down: With an Account of the Human Skeletal Material*. Belfast: HM Stationery Office.
- Cooney, G., 2007. In Retrospect: Neolithic activity at Knockadoon, Lough Gur, Co. Limerick, 50 years on. *Proceedings of the Royal Irish Academy*, 107C, pp.215-225.
- Cooney, G. et al., 2011. Ireland. In A. Whittle, F. Healy & A. Bayliss (eds.), *Gathering Time: Dating the Early Neolithic Enclosures of Southern Britain and Ireland*. Oxford: Oxbow Books, pp.562-669.
- Cribbin, G. et al., 1994. A Destroyed Late Iron-Age Burial from Kiltullagh, Ballyglass Middle td., Co. Mayo. *Emania*, 12, pp.61-65.
- Davies, O., 1938. Excavations at Mourne Park. *Proceedings of the Belfast Natural History and Philosophical Society*, 1, pp.18-26.
- Delaney, M. & Woodman, P. C., 2004. Searching the Irish Mesolithic for the humans behind the hatchets. In H. Roche et al. (eds.) *From megaliths to metal: essays in honour of George Eogan*. Oxford: Oxbow, pp.6-11.
- Doughty, P.S., 1995. The vertebrate fauna of Pollnagollum of the Boats, Co. Fermanagh. *Irish Speleology*, 15, pp.54-60.
- Dowd, M.A., 2008. The use of caves for funerary and ritual practices in Neolithic Ireland. *Antiquity*, 82(316), pp.305-317.
- Dowd, M., 2015. *The Archaeology of Caves in Ireland*. Oxford: Oxbow Books.
- Eogan, G., 2012. *Excavations at Knowth Volume 5: The archaeology of knowth in the first and second millennia AD*. Dublin: Royal Irish Academy.
- Eogan, J., 2002. Excavations at a cairn at Coolnatullagh townland, County Claremore. *North Munster Antiquarian Journal*, 42, pp.113-150.
- Evans, E.E. & Davies, O., 1934. 111. Excavation of a Horned Cairn at Ballyalton, Co. Down. *Proceedings of the Belfast Natural History and Philosophical Society*, pp.88-90.
- Fredengren, C., 2002. *Crannogs*. Bray, Ireland: Wordwell
- Geber, J., et al., 2016. Reassessing the Age, Sex and Metric Variation of Commingled Human Remains from a 1911 Excavation of a Neolithic Passage Tomb Complex in North-West Ireland. *International Journal of Osteoarchaeology*, 27(2), pp.131-142.
- Geber, J., et al., 2017. Facilitating Transitions: Postmortem Processing of the Dead at the Carrowkeel Passage Tomb Complex, Ireland (3500-3000 cal. B.C.). *Bioarchaeology International*, 1(1-2), pp.35-51.
- Gibbons, M., 1990. The archaeology of early settlement in county Kilkenny. In W. Nolan & K. Whelan (eds.), *Kilkenny: History and Society*. Dublin: Geography Publication, pp.1- 32.
- Hartnett, P. J. & O'Sullivan, W., 1971. The Excavation of Two Tumuli at Fourknocks (Sites II and III), Co. Meath. *Proceedings of the Royal Irish Academy*, 71C, pp. 35-89
- Hedges, R.E.M. et al., 1990. Radiocarbon dates from the Oxford AMS system: Archaeometry Datelist 10. *Archaeometry*, 32(1),

- pp.101–108.
- Hedges, R.E.M. et al., 1993. Radiocarbon dates from the Oxford AMS system: Archaeometry Datelist 17. *Archaeometry*
- Hencken, H. O'N., 1936. Ballinderry crannog no. 1. *Proceedings of the Royal Irish Academy*, 43C, pp.103–239.
- Hencken, H. O'N., 1942. Ballinderry crannog no. 2. *Proceedings of the Royal Irish Academy*, 47C, pp.1–76
- Hencken, H. O'N., 1950. Lagore crannog: an Irish royal residence of the seventh to tenth century A.D. *Proceedings of the Royal Irish Academy*, 53C, pp.1–248
- Hensey et al., 2014. A century of archaeology – historical excavation and modern research at the Carrowkeel passage tombs, County Sligo. *Proceedings of the Royal Irish Academy*, 114C, pp.57–86
- Jones, C., 2003. Neolithic beginnings on Roughan Hill and the Burren. In I. Armit et al. (eds.), *Neolithic Settlement in Ireland and Western Britain*. Oxford: Oxbow, pp.188-194.
- Jones, C., forthcoming. Regional traditions and distant events - Parknabinnia and other atypical court tombs in North Munster, western Ireland. In J. Müller & M. Hinz (eds.), *Megaliths, Societies, Landscapes - Early Monumentality and Social Differentiation in Neolithic Europe*.
- Jones C. & Walsh P. 1996. Recent Discoveries on Roughan Hill, County Clare. *Journal of the Royal Society of Antiquaries of Ireland*, 126, pp.86–107.
- Kador, T., 2010. Searching for Neolithic migrants at the introduction of agriculture to Ireland. Unpublished report submitted to the Heritage Council.
- Kador et al., forthcoming. Human population dynamics at the Carrowkeel passage tomb complex, Co. Sligo, Ireland.
- Kelly, E., 1998. Burials at Glassamucky, Glenasmole Valley, Co. Dublin. In C. Manning (ed.), *Dublin and beyond the Pale : studies in honour of Patrick Healy*. Bray, Ireland: Wordwell in association with Rathmichael Historical Society. pp.113-16.
- Kilbride-Jones, H.E. & Keenan, E., 1951. Double Horned Cairn at Cohaw, County Cavan. *Proceedings of the Royal Irish Academy*, 54C, pp.75–88.
- Kytmanow, T., 2009. *Portal tombs in the landscape: the chronology, morphology and landscape setting of the portal tombs of Ireland, Wales and Cornwall*. UK: Doctoral Thesis, Queen's University Belfast.
- Leask, H.G. et al., 1935. The Labbacallee Megalith, Co. Cork. *Proceedings of the Royal Irish Academy*, 43C, pp.77–101.
- Lucas, A.T., 1950. Neolithic Burial at Norrismount, Co. Wexford. *Journal of the Royal Society of Antiquaries of Ireland*. Royal Society of Antiquaries of Ireland, 80(2), pp.155–157.
- Lynch, A., 2014. *Poul nabrone: An early Neolithic portal tomb in Ireland*. Bray, Ireland: Wordwell.
- Lynch, A. & Donnabháin, B.O., 1994. Poul nabrone portal tomb. *The Other Clare (Ennis: Shannon Archaeological and Historical Society)*, 18, pp.5–7.
- MacAdam, R. & Getty, E., 1855. Discovery of an ancient sepulchral chamber. *Ulster Journal of Archaeology*, 3, pp.358–365.
- Macalister, R.A.S., Armstrong, E.C.R., & Praeger, R.L.I., 1911/1912. Report on the Exploration of Bronze-Age Carns on Carrowkeel Mountain, Co. Sligo. *Proceedings of the Royal Irish Academy*, 29C(1911/1912), pp.322-23.
- Manning, C. et al., 1985. A Neolithic Burial Mound at Ashleypark, Co. Tipperary. *Proceedings of the Royal Irish Academy*, 85C, pp.61–100.
- Martin, C.P., 1935. *Prehistoric Man in Ireland*. London: Macmillan.
- McCormick, F. et al., 1995. A pagan-Christian transitional burial at Kiltullagh. *Emania*, 13, pp.89–98.
- McGarry, T., 2007. Irish late prehistoric burial practices: continuity, developments and influences. *Trowel*, 11, pp.7–25.
- McGarry, T., 2010. Late pagan and Early Christian burials in Ireland: some issues and potential explanations. In C. Corlett & M. Potterton (eds.), *Death and burial in Early Medieval Ireland in light of recent archaeological excavations*. Dublin: Wordwell, pp.173-185.
- McQuade, M., 2005. Archaeological Excavation of a Multi-Period Prehistoric Settlement at Waterunder, Mell, Co. Louth. *Journal of the County Louth Archaeological and Historical Society*, 26(1), pp.31–66.
- McSparron, C., 2008. “Have you no homes to go to?” *Archaeology Ireland*, 22(3), pp.18–22.
- Meiklejohn, C. & Woodman, P.C., 2012. Radiocarbon dating of Mesolithic human remains in Ireland. *Mesolithic Miscellany*, 22(1), pp.22–41.
- Mount, C. & Buckley, L., 1997. Adolf Mahr's Excavations of an Early Bronze Age Cemetery at Keenoge, County Meath. *Proceedings of the Royal Irish Academy*, 97C(1), pp.1–68.
- Mount, C., Buckley, L. & Lynch, P., 1998. Five Early Bronze Age Cemeteries at Brownstown, Graney West, Oldtown and Ploopluck, County Kildare, and Strawhall, County Carlow. *Proceedings of the Royal Irish Academy*, 98C(2), pp.25–99.
- Murphy, E. & McGranaghan, C., 2008. Osteological report on human remains from Glebe, Rathlin Island, County Antrim, pp. 72-83 in Sloan, B., A Bronze Age cist burial at Glebe, Rathlin Island, County Antrim. *Ulster Journal of Archaeology*, 67, pp.60–83

- O'Brien, E., 1999. Excavation of a multi-period burial site at Ballymacaward, Ballyshannon, Co. Donegal. *Donegal Annual*, 51, pp.56–61.
- O'Hara, R., 2009. Collierstown 1: A late Iron Age-Early Medieval enclosed cemetery. In M.B. Deevy & D. Murphy (eds.), *Places along the way: first findings on the M3*. Bray, Ireland: Wordwell, pp.83–100.
- O'Kelly, M.J., 1982. *Newgrange: archaeology, art and legend*. London: Thames & Hudson.
- O'Kelly, M.J. et al., 1978. Three Passage-Graves at Newgrange, Co. Meath. *Proceedings of the Royal Irish Academy*, 78C, pp.249–352.
- Ó Maoldúin R., 2015. Preliminary Report on the excavation of Wedge tomb CL017-180002, Parknabinia, Co. Clare. Licence 15E0053.
- Ó Néill, J., 2000. Science and Technology Park, Cherrywood: multi-period. In I. Bennett (ed.), *Excavations 1998*. Bray, Ireland: Wordwell, pp.40-2.
- Ó Ríordáin, B., Brindley, A.L. & Ó Donnabháin, B., 1997. A Bronze Age Cemetery Mound at Grange, Co. Roscommon. *The Journal of Irish Archaeology*, 8, pp.43-72.
- Ó Ríordáin, B. & Waddell, J., 1993. *The Funerary Bowls and Vases of the Irish Bronze Age*. Galway, Ireland: Galway University Press.
- Ó Ríordáin, S.P., 1954. Lough Gur excavations: neolithic and bronze age houses on Knockadoon. *Proceedings of the Royal Irish Academy*, 56C, pp. 297-459.
- O'Sullivan, M., 2005. *The Mound of the Hostages, Tara*. Bray, Ireland: Wordwell.
- Prendergast, E., 1968. Cist Burial at Bolinready, Co. Wexford. *Journal of the Royal Society of Antiquaries of Ireland*, 98(2), pp.161–162.
- Price, L., 1950. The history of Lagore, from the annals and other sources, pp.18–35 in Hencken, H. O'N., Lagore crannog: an Irish royal residence of the seventh to tenth century A.D. *Proceedings of the Royal Irish Academy*, 53C, pp.1–248
- Raftery, B., 1974. A Prehistoric Burial Mound at Baunogenasraid, Co. Carlow. *Proceedings of the Royal Irish Academy*, 74C, pp.277–312.
- Raftery, B., 1994. *Pagan Celtic Ireland: The Enigma of the Irish Iron Age*. London: Thames & Hudson.
- Russell, I.R., Mossop, M. & Corcoran, E., 2002. Claristown 2: a cemetery and later cairn; excavations for the Drogheda Bypass–M1 Motorway. *Ríocht na Míde*, 13, pp.23–31.
- Ryan, M.F., 1973. The Excavation of a Neolithic Burial Mound at Jerpoint West, Co. Kilkenny. *Proceedings of the Royal Irish Academy*, 73C, pp.107–127.
- Schulting, R.J. et al., 2012. New dates from the north and a proposed chronology for Irish court tombs. *Proceedings of the Royal Irish Academy*, 112C, pp.1–60.
- Schulting, R.J., 2014. The dating on Poul nabrone. In Lynch, A., *Poul nabrone: An early Neolithic portal tomb in Ireland*. Bray, Ireland: Wordwell, pp.93-113.
- Shea, S., 1931. Report on the human skeleton found in Stoney Island Bog, Portumna. *Galway Archaeological and Historical Society*, 15, pp.73-79.
- Sloan, B. et al., 2008. A Bronze Age Cist Burial at Glebe, Rathlin Island, County Antrim. *Ulster Journal of Archaeology*, pp.60–83.
- Taylor, K., 2002. Inchagreenoge: *fulachta fiadhb*, ritual deposit of a human skull, wooden artifacts, post-medieval trackway. In I. Bennett (ed.), *Excavations 2002*. Bray, Ireland: Wordwell, pp.322–324.
- Waddell, J., 1990. *The Bronze Age Burials of Ireland*. Galway, Ireland: Galway University Press.
- Waddell, J., 2010. *The Prehistoric Archaeology of Ireland*. Bray, Ireland: Wordwell.
- Wallace, P.F., 1977. A prehistoric burial cairn at Ardcroney, Nenagh, Co. Tipperary. *North Munster Antiquarian Journal*, 19, pp.3-20.
- Waterman, D.M. et al., 1993. A Bronze Age Cist Cemetery at Stranagalwilly, Co. Tyrone. *Ulster Journal of Archaeology*, 56, pp.44–60.
- Whittle, A., 2004. Stones that float to the sky: portal dolmens and their landscapes of memory and myth. In V. Cummings & C. Fowler (eds.), *The Neolithic of the Irish Sea*, pp. 81–90.
- Woodman, P.C. (2000) Getting back to basics: transitions to farming in Ireland and Britain. In T.D. Price (ed.), *Europe's First Farmers: Lessons in the Transition to Agriculture*. Cambridge: Cambridge University Press, pp.219-259.
- Woodman, P., Dowd, M., O'Shaughnessy, J., forthcoming. Archaeological excavations at Killuragh Cave, Co. Limerick: a persistent place in the landscape from the Early Mesolithic to the Late Bronze Age.
- Woodman, P. & McCarthy, M., 2003. Contemplating same awful(ly interesting) vistas: importing cattle and red deer into prehistoric Ireland. In I. Armit et al. (eds.), *Neolithic settlement in Ireland & western Britain*. Oxford: Oxbow, pp.31-39.

Appendix II :

Molecular and Bioinformatic Methodology

Table of Contents

Molecular Methodology	274
1. Overview	275
2. Anti-Contamination Procedures	276
3. Sample Drilling and Bone Powder Isolation	276
4. DNA Extraction and Purification	279
5. Library Creation and Amplification	280
6. Low Coverage Sequencing (Illumina MiSeq)	280
7. USER Enzyme Treatment	280
8. Higher Coverage Sequencing (Illumina HiSeq2000 or HiSeq2500)	282
Bioinformatic Methodology	288
9. Overview	288
10. Raw Data Processing	290
11. Data Authenticity	290
12. Molecular Sex Determination	294
13. Mitochondrial Genome Analysis	294
14. Y Chromosome Analysis	295
15. Variant Calling and Modern Dataset Preparation	300
16. Exploration of accuracy and reference bias for pseudo-haploid and imputed calls	302
17. Kinship Analysis	304
References	312

Molecular Methodology

1. Overview

Over the course of this project a total of 144 skeletal elements from 140 individuals were sampled for ancient DNA analysis, sourced from 66 archaeological sites situated in 27 counties across Ireland (See Fig. 1.1). All samples were subject to the same initial screening pipeline, with minor variations added through time as methodologies were improved upon. This pipeline can be partitioned into four main phases -

1. Drilling and Bone Powder Isolation
2. DNA Extraction and Purification
3. Library Creation and Amplification
4. Low Coverage Sequencing (Illumina MiSeq)

This initial screening allowed the endogenous human DNA content of each sample to be estimated. Based on these values, selected sample extracts were treated for molecular damage, with the exception of the four genomes published in Cassidy et al. 2016, and subsequently taken forward for further rounds of library creation and sequencing. This can be summarised in three steps -

5. USER enzyme treatment
6. Second Library Creation and Amplification
7. Higher Coverage Sequencing (Illumina HiSeq2000 or HiSeq2500)

The details and results of each phase are described below.

2. Anti-Contamination Procedures

All sample processing was carried out in clean-room facilities dedicated to ancient DNA research at Trinity College Dublin. Full body suits, face masks, gloves and hairnets were worn throughout drilling, extraction and library creation processes. All tools and working surfaces were consecutively cleaned with bleach, DNA-ExitusPlus, Ethanol and long exposures to UV light. Upon arrival, all samples were also subject to UV irradiation for 20 minutes on each side.

A number of controls were introduced at different stages and brought forward to Illumina MiSeq sequencing to monitor possible contamination events. This included

1. Air Controls : Empty Eppendorf tubes left open for an hour in the fume hood between drilling sessions to test for bone powder cross contamination.
2. Water Controls : 1ml of ddH₂O added to the cylindrical metal container (Mixer Mill MM 400, Retsch) used to powderise bone, and shaken at high frequency. These water controls were created between drilling sessions to test for bone powder cross contamination.
3. Extraction controls : 1ml of ddH₂O carried throughout each round of extractions.
4. Library controls : 16.5µl of ddH₂O carried through each round of library creation.
5. PCR controls : 3µl of ddH₂O included in various rounds of library amplifications.

3. Sample Drilling and Bone Powder Isolation

Sampling was carried out with an overall aim of minimising damage to the archaeological material, while maximizing the information yield that could be gained from it. During the course of this project it was demonstrated in Gamba et al. 2014 that the petrous portion of the temporal bone, one of the densest elements of the skeleton, consistently yields higher levels of endogenous ancient DNA relative to other skeletal elements. Thus, where possible, this bone was selected for sampling. In some instances, if no petrous bone was available and the archaeological site was of contextual importance, teeth were also utilised. Overall, a total of 117 petrous temporals and 19 teeth were processed. In addition, 6 long bones, a metatarsal and a phalanx were also screened in the very early stages of the project.

Before any alteration, all samples were photographed extensively (See Appendix I for Images) and given a unique lab ID, typically combining an abbreviation of the townland it originated from and any attached numerical archaeological ID available. Sample drilling was then carried out in a fume hood, lined with bleached tinfoil. All bone fragments, powder and debris not brought forward for extraction were collected and returned to archaeological bagging. Drilling strategies varied depending on the skeletal element being sampled, discussed in turn below.

1. Petrous Temporal Drilling

The dense otic capsule region of the petrous temporal, that surrounds the cochlea, vestibule and semicircular canals, was targeted for aDNA extraction. A triangular shaped wedge section encompassing this area was cut from the surrounding bone using a Dremel diamond wheel saw and pliers (Fig. II.1A). Wherever possible, the remaining petrous bone was left fully intact and articulated. Once isolated the wedge was extensively cleaned using appropriate Dremel diamond bits and cutters. A Mixer Mill MM 400 (Retsch) was then used to gradually powderise the bone. Multiple rounds of powderisation allowed successive outer layers of the bone to be removed until only the densest region remained, usually that found above and behind the carotid canal. The quality of the powder, in terms of colour and texture, was seen to steadily increase with each powder aliquot retrieved, with the whitest and finest powder obtained from the very core of the bone (see Fig. II.1B for an example). Aliquots from these final stages of powderisation were used for extraction.

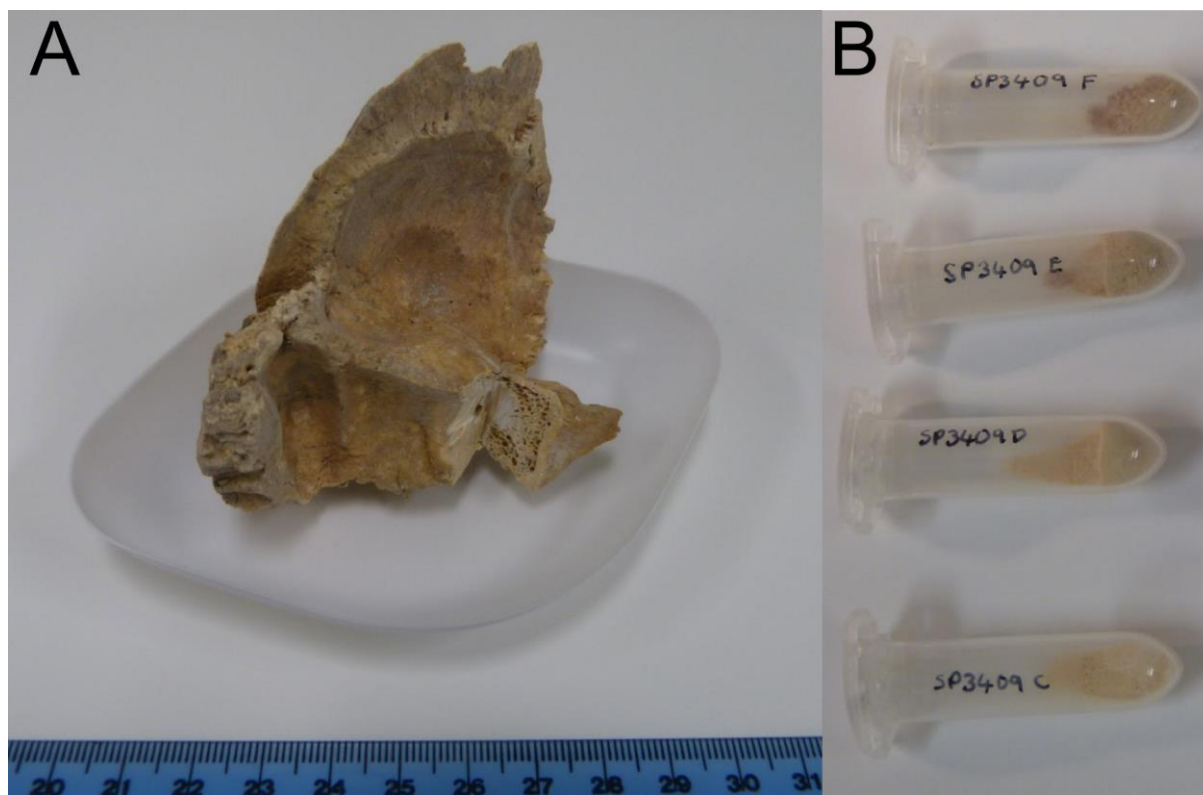
2. Tooth Drilling

It has been demonstrated that teeth roots far outperform crowns in terms of aDNA preservation (Adler et al. 2011), and so, when available, roots were sampled. This process involved cutting the root from the remainder of the tooth using a clamp and a Dremel diamond wheel saw. Once isolated the outer surface of the root was lightly cleaned using appropriate Dremel diamond bits and cutters. The inner pulp and dentine layers were then eroded away where possible, leaving the outer cementum layer isolated. This was then powderised in a Mixer Mill MM 400 (Retsch). The cementum has been previously shown to yield up to 14 time more endogenous DNA than the inner dentine (Damgaard et al. 2015). Several crowns were also sampled. Here, the enamel surface was cleaned and the inner pulp removed. The dentine and enamel layers were then powderised for extraction

3. Long Bone, Metatarsal and Phalanx Drilling

In the earliest stages of the project a number of limb, hand and finger bones were sampled. This process involved cutting a vertical section of bone from the center of the shaft using a Dremel diamond wheel saw. This was then heavily cleaned with a high speed Dremel cutter and more porous regions were removed. The remaining bone was subsequently powderised.

Figure II.1. Bone powder isolation from the petrous temporal. A) An example of the wedge shaped section removed from a petrous temporal for aDNA extraction. The marbled surface of the dense region sampled is clearly visible. B) Successive aliquots (F to C) of bone powder obtained from the otic capsule. Tube C contains powder from the very core of the dense section. The improvement in colour and texture is apparent.



4. DNA Extraction and Purification

A range of extraction methods have been developed for aDNA analysis, with those based on silica-column purification proving the most efficient, particularly in terms of short read recovery (Gamba et al. 2016). The silica-column method outlined in Yang et al. 1998 is followed here, using modifications introduced by MacHugh et al. 2000. The protocol is described briefly below.

Approximately 150mg of bone or tooth powder was suspended in 1ml of lysis buffer (20mM Tris HCL, pH 7.4; 47.5mM EDTA, pH 8; 0.7% Sarkosyl NL30; 0.65 U/ml Proteinase K, recombinant, PCR Grade), thoroughly vortexed and then incubated at 37°C for 24 hours, rotating at 750rpm. Tubes were subsequently centrifuged at 13,300rpm for 10 minutes and the supernatant, labelled *Extraction 1*, removed. A fresh 1ml of lysis buffer was added to the remaining pellets and the procedure repeated twice over, resulting in second and third extracts. Third extracts have been noted to yield the highest levels of endogenous DNA, though usually at the cost of overall DNA concentration (Marta Verdugo pers.comm.) Notably, the four samples from Cassidy et al. 2016, processed before this information was available, were subject only to two of these extraction rounds, with longer incubation periods (24 hours at 55°C, followed by 24 hours at 37°C).

Sample extracts, typically the third, but occasionally, for experimental purposes, the second or first also, were subsequently purified. This required the transfer of each extract to an Amicon Ultra-4 Centrifugal Filter Unit 30kDA, where it was diluted in 3ml of 10mM Tris-EDTA Buffer and centrifuged at 2,500rpm until 250µl of the solution remained (12-20mins). The resulting flow through was discarded and another 3ml of 10mM Tris-EDTA buffer was added to the filter. Samples were then spun again at 2,500rpm until a final volume of 100µl was obtained (15-25mins).

This volume was then transferred from the filter to a silica column (MinElute PCR Purification Kit, Qiagen, Hilden, Germany) and purified following the manufacturer's instructions. Briefly, PB buffer was added to the silica column at approximately 5 times the volume of the extract. The column was then spun at 13,000rpm for 1 minute and the flow through discarded. Following this, 750µl of PE buffer was added to the column, which was spun at 13,000rpm for a subsequent two minutes, or until there was no visible liquid left. Finally, Elution buffer (EB) at volume of 55µl was added to the column membrane and the sample DNA released from the filter by centrifugation at 13,000rpm for 1 minute into a fresh 1.5ml centrifuge tube. Specific details of extraction for each sample are available in Electronic Data Table S2.

5. Library Creation and Amplification

Libraries were prepared using 16.5-30µl of DNA extract following the method outlined in Meyer & Kircher 2010, with the following modifications as in Gamba et al. 2014. No DNA shearing by sonication was carried out due to the highly fragmented nature of ancient DNA molecules. All DNA purification steps were performed using the MinElute PCR Purification kit protocol as described above. Blunt end repair was done using NEBNext End Repair Module (New England BioLab Inc.), with all components scaled to 70% of the manufacturer's recommended volume, giving a final reaction volume of 70µl. This volume was then incubated for 15 min at 25°C followed by 5 min at 12°C. The final modification was the termination of Bst polymerase activity after the adapter fill-in reaction through the addition of a heat inactivation step (80°C for 20 min).

The library amplification reaction was set up for all samples and controls using Accuprime Pfx Supermix (Life Technology), primer IS4 (0.2µM), a specific indexing primer (0.2µM) and 3µl of sample library, with a total reaction volume of 25µl. Amplification took place under the following thermal cycling conditions: 5 min at 95°C; 12 cycles of 15 sec at 95°C, 30 sec at 60°C and 30 sec at 68°C; and a final extension step of 5 min at 68°C. The resulting PCR product was purified using the MinElute PCR Purification kit protocol described previously. Quantification and quality assessment of the amplified libraries was performed initially on an Agilent 2100 Bioanalyser, using a DNA-1000 chip and following manufacturer's protocol, which was subsequently replaced with the Agilent TapeStation system, using D1000 screentape. Repeat amplifications with adjusted numbers of cycles were carried out when necessary, dependent on DNA concentrations achieved.

Details of library creation and amplification for all samples are available in Electronic Data Table S2.

6. Low Coverage Sequencing (Illumina MiSeq)

Amplified libraries were screened on an Illumina MiSeq platform (TrinSeq, Trinity Genome Sequencing Laboratory, Trinity College Dublin, Ireland) using 65 or 70 bp single-end sequencing and following the manufacturer's instructions for multiplex sequencing. The sequence data retrieved was then used to estimate endogenous DNA contents for sample extracts. The final endogenous contents, estimated after mapping quality filtering and PCR duplicate removal, for the best extract of each sample are displayed in Table I.1, with a more detailed breakdown of endogenous contents given in Electronic Data Table S2. Fig. II.2 gives an overview of endogenous content, with relation to time and skeletal element used.

As previously noted, the petrous bone was seen to produce consistently higher yields compared to teeth and other bones tested. Three individuals from Rathlin Island (Rathlin1, Rathlin2 and Rathlin3) were sampled twice, from both the petrous temporal and another skeletal element (two teeth and one

phalanx). Petrous bone endogenous contents were seen to fall between 38.81-62.25%, in striking contrast to the tooth and phalanx samples, which performed exponentially worse (0.36-5.10%) (Table II.1). However, despite the observed superiority of the petrous bone, several tooth samples provided exceptionally high levels of human DNA, emphasising their potential as an alternative for aDNA analysis.

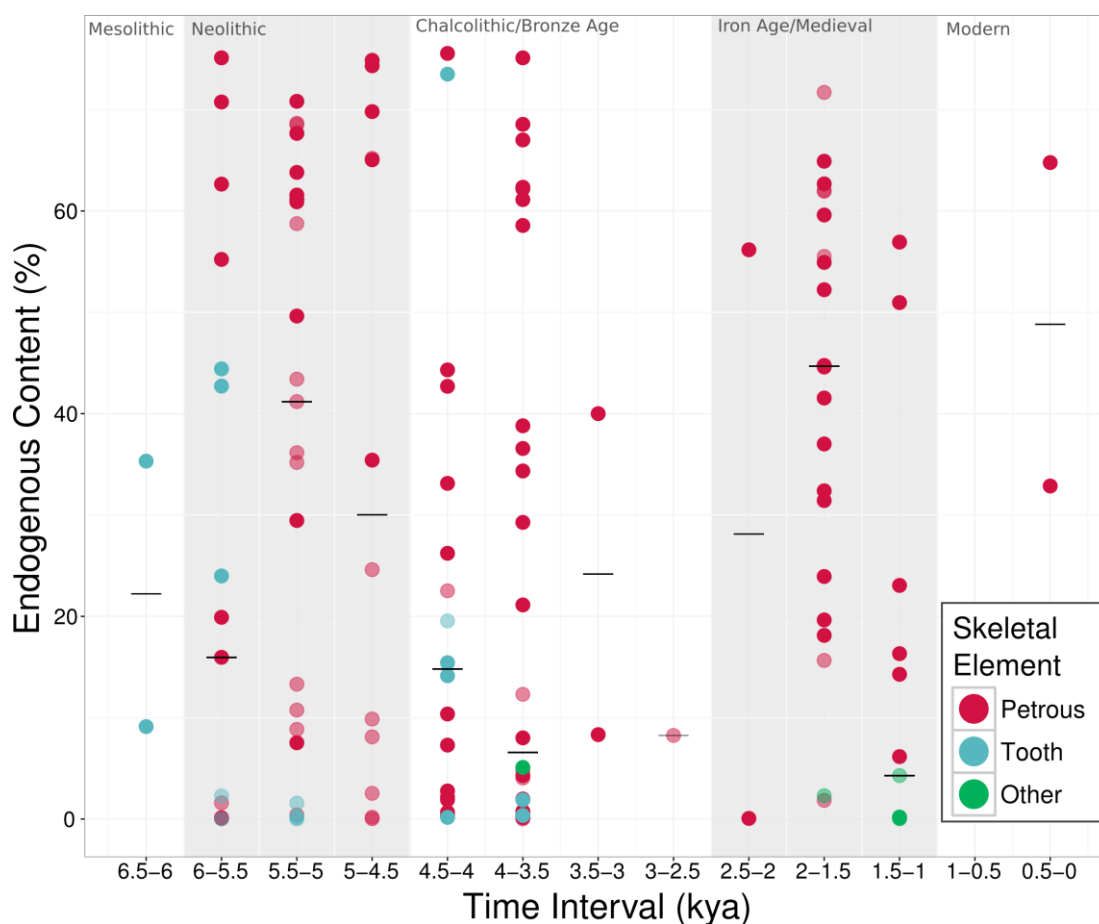


Figure II.2 Endogenous contents for 144 Irish bone and tooth samples. Endogenous content is estimated after imposition of a mapping quality filter (MQ >30) and PCR duplicate removal. Transparent data points do not have a direct radiocarbon determination and are dated by context. Black bars indicate median values for each time interval.

A significant number of Early Bronze Age samples produced endogenous DNA yields close to 0%, despite seemingly good skeletal preservation. Indeed, several were extremely dense and clean. Such low yields are not seen for Iron age or Neolithic samples, with the exception of several cremated samples tested, suggesting the lower success rates of Bronze Age petrous bones may be the result of unknown burial practices, such as funerary pyres, specific to that era. While the failed Bronze Age petrous bones show no evidence of cremation, indirect exposure to high temperatures may result in DNA destruction. Alternatively, differing soil conditions or unrecorded preservation treatments post-excavation may have also contributed to low endogenous DNA yields.

7. USER Enzyme Treatment

Extracts with sufficient endogenous content were then treated with USER Enzyme, a mixture of Uracil DNA glycosylase (UDG) and the DNA glycosylase-lyase Endonuclease VII, in an effort to repair post-mortem molecular damage present in ancient DNA. USER enzyme at a volume of 5 μ l was added to 16.25 μ l of each extract and incubated for 3 hours at 37°C. Following this step, library creation was carried out identically to that described above, with an initial volume of 21.25 μ l. In preparation for higher coverage sequencing, multiple PCR reactions, described above, were carried out on each “UDG” library using different barcoded indices. In an effort to increase complexity, the number of PCR cycles was kept to a minimum for each sample. Details of “UDG library” creation and amplification for all samples are available in Electronic Data Table S2.

8. Higher Coverage Sequencing (Illumina HiSeq2000 or HiSeq2500)

In total 94 samples were brought forward from the screening stage. The majority of these were sequenced to approximately 1-2X coverage. Notable exceptions are two individuals from Cassidy et al. 2016, sequenced to roughly 11X coverage, and 8 samples who, due to low DNA concentrations and endogenous contents, were restricted to 0.05-0.75X coverage.

Later identity-by-descent estimates (Fig. II.X) revealed two of the final samples (KGH1 and KGH2) to belong to the same individual, bringing the total number of genomes included in final analyses to 93. Amplified libraries were sent for higher coverage sequencing (100 bp single-end) at Macrogen Inc., Seoul, Korea, with the exception of the four genomes from Cassidy et al. 2016, which were sequenced at Beckman Coulter Genomics, MA, USA. Further sample sequencing information is detailed in Electronic Data Table S2.

Table II.1. Endogenous human DNA contents for 144 skeletal elements sampled. Endogenous content is estimated following the removal of reads below a mapping quality of 30 and PCR duplicates.

Study ID	Lab ID	Skeletal Element	Extraction	Final Endogenous Content (%)
Annaghkeen84	AK84	Right Petrous Temporal	Extract 2	29.28
Annagh1	ANN1	Left Petrous Temporal	Extract 3	60.91
Ardcrone2	ARD2	Right Petrous Temporal	Extract 3	29.47
Ashleypark1	ASH1	Right Petrous Temporal	Extract 3	15.96
Ashleypark3	ASH3	Right Petrous Temporal	Extract 3	55.23
-	AT1	Right Petrous Temporal	Extract 3	0.04
-	AT7	Left Petrous Temporal	Extract 3	0.11
-	BA342	Molar (Crown)	Extract 2	0.26
-	BA346	Molar (Crown)	Extract 2	0.13
Ballynahatty	BA64	Right Petrous Temporal	Extract 2	67.68
-	BAL14	Left Petrous Temporal	Extract 3	0.13
-	BAL3211	Molar (Crown)	Extract 2	2.27
-	BAL329	Molar (Crown)	Extract 2	0.01
Ballybunnion54	BB54	Left Petrous Temporal	Extract 3	56.93
Ballyeagh79	BE79	Left Petrous Temporal	Extract 3	16.33
Baunogenasraid72	BG72	Right Petrous Temporal	Extract 3	70.76
Blackhill32	BH32	Right Petrous Temporal	Extract 3	44.32
Ballinderry6	BLD6	Right Petrous Temporal	Extract 2	8.25
Ballyglass44	BM44	Right Petrous Temporal	Extract 3	18.13
Ballyglass45	BM45	Right Petrous Temporal	Extract 3	59.61
Ballyglass46	BM46	Right Petrous Temporal	Extract 3	32.39
Bolinready79	BR79	Right Petrous Temporal	Extract 3	8.02
Ballymacaward2	BW2	Right Petrous Temporal	Extract 3	23.06
Ballymacaward3	BW3	Left Petrous Temporal	Extract 3	44.74
-	CAK16	Right Petrous Temporal	Extract 3	0.03
-	CAK375	Right Petrous Temporal	Extract 3	0.18
Carrowkeel530	CAK530	Right Petrous Temporal	Extract 3	35.42
Carrowkeel531	CAK531	Right Petrous Temporal	Extract 3	24.61
Carrowkeel532	CAK532	Right Petrous Temporal	Extract 3	69.81
Carrowkeel533	CAK533	Right Petrous Temporal	Extract 3	65.06
Carrowkeel534	CAK534	Left Petrous Temporal	Extract 3	8.11
Carrowkeel68	CAK68	Right Petrous Temporal	Extract 3	74.87
Cohaw448	CH448	Right Petrous Temporal	Extract 3	19.91

-	CL23	Left Petrous Temporal	Extract 3	2.78
-	CLT622	Longbone	Extract 2	0.17
-	CLT627	Longbone	Extract 2	0.05
-	CLT640	Longbone	Extract 2	0.21
Courtmacsherry37	CM37	Left Petrous Temporal	Extract 3	14.29
Coolnatullagh1005	CN1005	Left Petrous Temporal	Extract 3	68.56
-	CR02	Molar (Root)	Extract 3	0.18
Claristown14	CT14	Right Petrous Temporal	Extract 3	41.55
Cherrywood1	CW1	Left Petrous Temporal	Extract 3	8.34
Derrynamanagh01	DM01	Right Petrous Temporal	Extract 3	71.70
Derrynamanagh02	DM02	Right Petrous Temporal	Extract 3	23.94
Derrynamanagh03	DM03	Right Petrous Temporal	Extract 3	55.52
Derrynamanagh04	DM04	Right Petrous Temporal	Extract 3	37.01
Derrynamanagh05	DM05	Right Petrous Temporal	Extract 3	31.43
Derrynamanagh06	DM06	Right Petrous Temporal	Extract 3	61.95
-	DM07	Right Petrous Temporal	Extract 3	1.85
Derrynamanagh08	DM08	Right Petrous Temporal	Extract 3	62.02
Derrynamanagh09	DM09	Right Petrous Temporal	Extract 3	19.65
Derrymaquirk20	DQ20	Fragment of Petrous Temporal	Extract 1	15.65
-	FK142	Right Petrous Temporal	Extract 3	2.18
-	GG01	Metacarpal	Extract 2	2.30
-	GM343	Left Petrous Temporal	Extract 3	0.33
Grange10	GR10	Left Petrous Temporal	Extract 3	26.23
-	GR41	Left Petrous Temporal	Extract 3	0.34
HillOfWard14	HW14	Right Petrous Temporal	Extract 3	44.61
Inchagreenoge134	IG134	Right Petrous Temporal	Extract 3	40.01
-	JNT400	Longbone	Extract 2	0.03
-	JNT616	Longbone	Extract 2	0.17
-	JNT645	Longbone	Extract 2	4.29
Jerpoint14	JP14	Right Petrous Temporal	Extract 3	75.11
Jerpoint15	JP15	Left Petrous Temporal	Extract 3	62.68
Knockadoon275	KDI275	Left Petrous Temporal	Extract 3	32.86
Killuragh1	KGH1	Molar (Root)	Extract 3	15.43
Killuragh1	KGH2	Molar (Root and Crown)	Extract 3	14.15
Killuragh6	KGH6	Molar (Root)	Extract 3	9.13
-	KIL1	Left Petrous Temporal	Extract 3	0.01
Knowth10	KNW10	Right Petrous Temporal	Extract 3	56.18

Knowth4	KNW4	Right Petrous Temporal	Extract 3	44.76
-	KO12	Left Petrous Temporal	Extract 3	4.04
Keenoge14	KO14	Right Petrous Temporal	Extract 3	67.01
Keenoge3	KO3	Left Petrous Temporal	Extract 3	10.37
-	KO4	Right Petrous Temporal	Extract 3	0.13
Keenoge7	KO7	Right Petrous Temporal	Extract 3	12.31
Labbacallee212	LB212	Left Petrous Temporal	Extract 3	7.30
Labbacallee213	LB213	Left Petrous Temporal	Extract 3	22.52
-	LC04	Right Petrous Temporal	Extract 3	0.68
Lagore14	LG14	Right Petrous Temporal	Extract 3	64.78
Lehinch27	LH27	Left Petrous Temporal	Extract 3	52.23
Lehinch28	LH28	Left Petrous Temporal	Extract 3	50.97
-	LM46	Left Petrous Temporal	Extract 3	0.23
-	MB101	Canine	Extract 2	1.57
-	MB4	Left Petrous Temporal	Extract 3	0.41
MillinBay6	MB6	Left Petrous Temporal	Extract 3	49.64
-	MB9889	Alveolar process adjacent Molar	Extract 2	0.01
-	ML1	Molar (Crown)	Extract 3	0.17
-	MP18	Right Petrous Temporal	Extract 3	0.13
Moyveela381	MV381	Right Petrous Temporal	Extract 3	42.70
Newgrange10	NG10	Left Petrous Temporal	Extract 3	65.20
NewgrangeZ1	NGZ1	Left Petrous Temporal	Extract 3	9.87
-	NGZ2	Left Petrous Temporal	Extract 3	2.54
-	NM61	Right Petrous Temporal	Extract 3	1.59
Parknabinnia1239	PB1239	Right Petrous Temporal	Extract 3	68.52
-	PB1327	Right Petrous Temporal	Extract 3	13.31
-	PB1794	Right Petrous Temporal	Extract 3	10.75
Parknabinnia186	PB186	Right Petrous Temporal	Extract 3	61.58
Parknabinnia2031	PB2031	Right Petrous Temporal	Extract 3	41.19
Parknabinnia357	PB357	Right Petrous Temporal	Extract 3	63.81
Parknabinnia443	PB443	Right Petrous Temporal	Extract 3	74.33
Parknabinnia581	PB581	Right Petrous Temporal	Extract 3	61.20
Parknabinnia672	PB672	Right Petrous Temporal	Extract 3	62.66
Parknabinnia675	PB675	Right Petrous Temporal	Extract 3	70.82
-	PB754	Right Petrous Temporal	Extract 3	8.85
Parknabinnia768	PB768	Right Petrous Temporal	Extract 3	36.15

Pollnagollum86	PG86	Left Petrous Temporal	Extract 3	75.56
Pollnagollum90	PG90	Molar (Root)	Extract 3	19.55
Pollnagollum911	PG911	Molar (Root)	Extract 3	73.51
Poulnabrone01	PN01	Right Petrous Temporal	Extract 3	75.12
Poulnabrone02	PN02	Right Petrous Temporal	Extract 3	43.41
Poulnabrone03	PN03	Right Petrous Temporal	Extract 3	58.75
Poulnabrone04	PN04	Right Petrous Temporal	Extract 3	67.64
Poulnabrone05	PN05	Right Petrous Temporal	Extract 3	35.19
Poulnabrone06	PN06	Right Petrous Temporal	Extract 3	68.66
Poulnabrone07	PN07	Right Petrous Temporal	Extract 3	7.53
Poulnabrone107	PN107	Molar (Root)	Extract 3	42.73
Poulnabrone112	PN112	Molar (Root)	Extract 3	44.42
Poulnabrone113	PN113	Molar (Root)	Extract 3	24.00
-	PP35	Left Petrous Temporal	Extract 3	4.29
Ploopluck6	PP6	Right Petrous Temporal	Extract 3	21.13
RoughanHill125	RH125	Right Petrous Temporal	Extract 3	34.35
RoughanHill468	RH468	Petrous Temporal	Extract 3	58.57
-	RH473	Petrous Temporal	Extract 3	4.76
-	RH98	Petrous Temporal	Extract 3	0.55
Rathlin1	RM127	Right Petrous Temporal	Extract 2	61.13
Rathlin1	RM127-2	Phalanx	Extract 2	5.10
Rossnaree1	RNR1	Left Petrous Temporal	Extract 3	54.93
Rossnaree2	RNR2	Right Petrous Temporal	Extract 3	64.91
Rathlin2	RSK1	Left Petrous Temporal	Extract 2	38.81
Rathlin2	RSK1-2	Molar	Extract 2	0.36
Rathlin3	RSK2	Left Petrous Temporal	Extract 2	62.35
Rathlin3	RSK2-2	Molar	Extract 2	1.90
Sliguff27	SG27	Right Petrous Temporal	Extract 3	62.19
-	SI118	Right Petrous Temporal	Extract 3	0.08
Stonepark3409	SP3409	Left Petrous Temporal	Extract 3	33.13
Sramore62	SRA62	Canine (Root)	Extract 3	35.32
-	ST63	Fragment of Petrous	Extract 3	0.07
-	TD97	Left Petrous Temporal	Extract 3	6.17
-	TM81	Right Petrous Temporal	Extract 3	1.91
-	TOP11	Left Petrous Temporal	Extract 3	1.98
Treanmacmurtagh116	TR116	Left Petrous Temporal	Extract 3	36.57

-	VM231	Right Petrous Temporal	Extract 3	0.76
-	VM232	Right Petrous Temporal	Extract 3	0.16

Bioinformatic Methodology

9. Overview

A total of 24,129,187,961 sequence reads were retrieved from Irish bone samples. These raw data were all subject to the same bioinformatic processing pipeline, detailed below, which can be divided into several phases. Published ancient sequence data were also processed, following similar procedures.

1. Adapter trimming and quality control
2. Read alignment
3. Mapped read filtering
4. Indel realignment
5. Base quality adjustment to account for post-mortem damage
6. Data authentication

Once processed and filtered, the reconstructed ancient genomes could then be used for population genetic analysis, which again can be divided into a number of stages.

1. Sex determination
2. Uniparental markers
 - a. Mitochondrial analysis
 - b. Y Chromosome analysis
3. Autosomal dataset preparation
 - a. Pseudo-haploid variant calling
 - b. Diploid variant calling
 - c. Imputation
 - d. Merging with modern datasets
4. ADMIXTURE (Chapter Three)
5. Principal component analysis (Chapter Three)
6. *D*- and *f*-statistics (Chapters Three, Four and Five)
7. ChromoPainter (Chapters Four and Five)
8. Runs of homozygosity (Chapters Three, Four and Five)
9. Phenotypic analysis (Chapters Three and Five)
10. Kinship analysis with lcMLkin (Chapter Four)

Detailed methods for stages 5 to 9 are found within the main text. The remaining analyses are discussed here.

10. Raw Data Processing

Irish Data

Data was returned in FASTQ format and subject to quality control using the FastQC suite (Andrews 2010). The majority of aDNA molecules are short enough for single-end reads to carry part of the P7 adapter at the 3' end. These residual adapter sequences were trimmed from reads using cutadapt v1.2.1 (Martin 2011), with parameters enabled to discard reads under 34bp in length after trimming (-m 34) and to allow a minimum overlap of 1bp between the read and adapter (-O 1). Quality trimming was also performed on the ends of reads with the same program, where necessary.

All remaining trimmed sequences were subsequently mapped to the human reference genome (hg19/GRCh37) with the mitochondrial genome replaced by the revised Cambridge reference sequence (Accession number NC_012920.1). BWA version 0.7.5 (Li & Durbin 2009) was used for alignment with the seed disabled (-l 16500) as recommended in Schubert et al. 2012, the edit distance set to 0.02 to allow for more substitutions (-n 0.02) and the amount of gaps allowed raised (-o 2).

Read groups were added using Picard Tools v1.101 (<http://broadinstitute.github.io/picard/>). Reads were sorted, filtered for a mapping quality of 30 or above and PCR duplicates removed using Samtools v0.1.19 (Li et al. 2009). For non-USER treated libraries, DNA damage patterns were assessed on a sample level using mapDamage 2.0 (Jónsson et al. 2013) and the base quality scores of likely deaminated positions in reads were rescaled based on these patterns, as well as their initial qualities and their position in the read. For the majority of samples this applied to Illumina MiSeq sequenced libraries only, with the exception of the four genomes from Cassidy et al. 2016, for which all libraries were merged and rescaled.

BAM files were merged to sample level using the MergeSamFiles Picard tool. The RealignerTargetCreator and IndelRealigner tools from GenomeAnalysisTK v2.4-7 (McKenna et al. 2010) were used to locally realign reads. Finally, as an additional safeguard to the high frequency of misincorporation sites found at read ends, each read had the base quality of two base pairs at both its 5' and 3' ends reduced to a PHRED score of 2.

Sequencing statistics for individual libraries and samples can be found in Electronic Data Table S2.

Published Data

Published ancient data processed alongside the Irish dataset for use in downstream analyses, taken from twenty-one previous studies - Keller et al. 2012, Gamba et al. 2014, Lazaridis et al. 2014, Olalde et al. 2014, Skoglund et al. 2014, Allentoft et al. 2015, Günther et al. 2015, Haak et al. 2015, Jones et al. 2015, Mathieson et al. 2015, Olalde et al. 2015, Broushaki et al. 2016, Fu et al. 2016, Hofmanova et al. 2016, Kılınç et al. 2016, Lazaridis et al. 2016, Martiniano et al. 2016, Schiffels et al. 2016, González-Forbes et al. 2017, Jones et al. 2017 and Martiniano et al. 2017.

With the exception of data from Gamba et al. 2014, Jones et al. 2015, Martiniano et al. 2016 and Martiniano et al. 2017, all downloaded data (FASTQ or BAM format) was realigned to the reference genome following the same parameters described above and filtered in an identical manner. Data from the above four listed studies was not realigned, but subject to any additional filters applied to the Irish data (e.g. read length minimum of 34bp) that were not applied in the original studies. Moreover, all published data not subject to UDG treatment had base qualities rescaled using mapDamage 2.0 software (Jónsson et al. 2013).

11. Data Authenticity

Sequencing analysis of molecular controls

DNA concentrations of controls were too low to detect and so could not inform equimolar sample pooling for sequencing. Instead, between 0.33 and 1µl of each control was added to the pool of samples it was processed alongside and sequenced on the Illumina MiSeq platform, which outputs approximately 20 million reads. Sequencing pools for different runs ranged from 15µl to 236µl in total volume, resulting in a wide variety of control contributions to the concentration. In spite of this variation, fairly uniform low numbers of reads were retrieved per control across runs (Median = 842; Mean = 1,221.62; Min = 128; Max = 6,318). Interestingly, metagenomic analysis (Fig. II.1) revealed the vast majority of these reads to be human rather than bacterial in origin, a testament to the sterile conditions of the lab space and equipment. This miniscule level of human contamination is likely to originate from the researcher or from other samples, but, as demonstrated in the following sections, has not compromised data integrity.

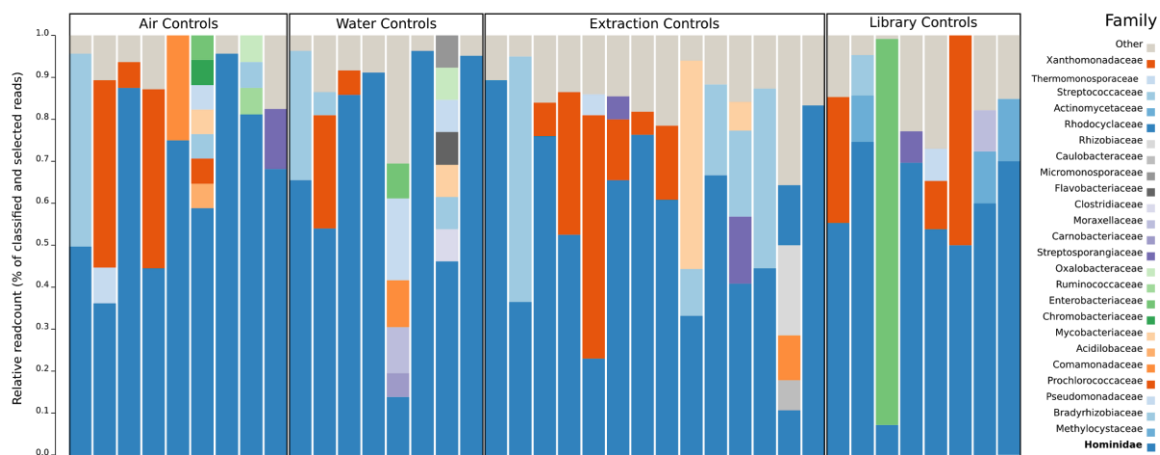


Figure II.1. Metagenomic analysis of controls carried through molecular stages of aDNA analysis. Human sequences (Blue) account for the majority of reads obtained from controls.

Patterns of molecular damage and sequence length distribution

Using mapDamage 2.0 (Jónsson et al. 2013) it was assessed whether non-USER treated data from each sample displayed typical patterns of aDNA post-mortem deamination short sequence length and C to T changes at the 5'-end of molecules (Brotherton et al. 2007; Briggs et al. 2007). Only reads with a minimum mapping quality of 30 were taken into account.

An increase in C to T and G to A changes towards the 5' and 3' ends of reads was observed in all samples (Fig. II.2). No relationship between misincorporation frequency and endogenous content or archaeological period could be ascertained.

Sequence length distributions showed a clear peak between 35-60 bp though this was more difficult to identify for lower coverage individuals with few read numbers (Fig II.3). This fits with previous observations that ancient samples display average sequence lengths of less than 100bp (Shapiro & Hofreiter 2014). No relationship between archaeological period and sequence length distribution was seen. Finally, as samples were sequenced using both 65-70 bp and 100 bp single-end sequencing, peaks at these points could be seen, representing all sequences greater than these lengths that were truncated in their respective sequencing runs.

Mitochondrial contamination

In order to detect the presence of possible modern mitochondrial contamination, the proportion of secondary bases not matching the sample consensus sequence was calculated at both haplogroup defining positions and private sample mutations (see section 5) identified using HAPLOFIND (Vianello et al. 2013), as previously described in Sánchez-Quinto et al. 2012; and Gamba et al. 2014. Bases with a quality below 30 were not considered. The contamination rate was defined as the percentage of the total number of secondary bases across all sample sites considered over the total base count across these positions (%C). Given the influence of sequencing errors, heteroplasmy and post-mortem molecular damage these estimates represent an upper limit of contamination. To address the issue of post-mortem damage the analysis was repeated removing sites where the defining SNP was a G or a C in the sample consensus sequence (%C-MD). No strong evidence of contamination was identified in any sample (Table II.2), with results discussed further in Chapter Two.

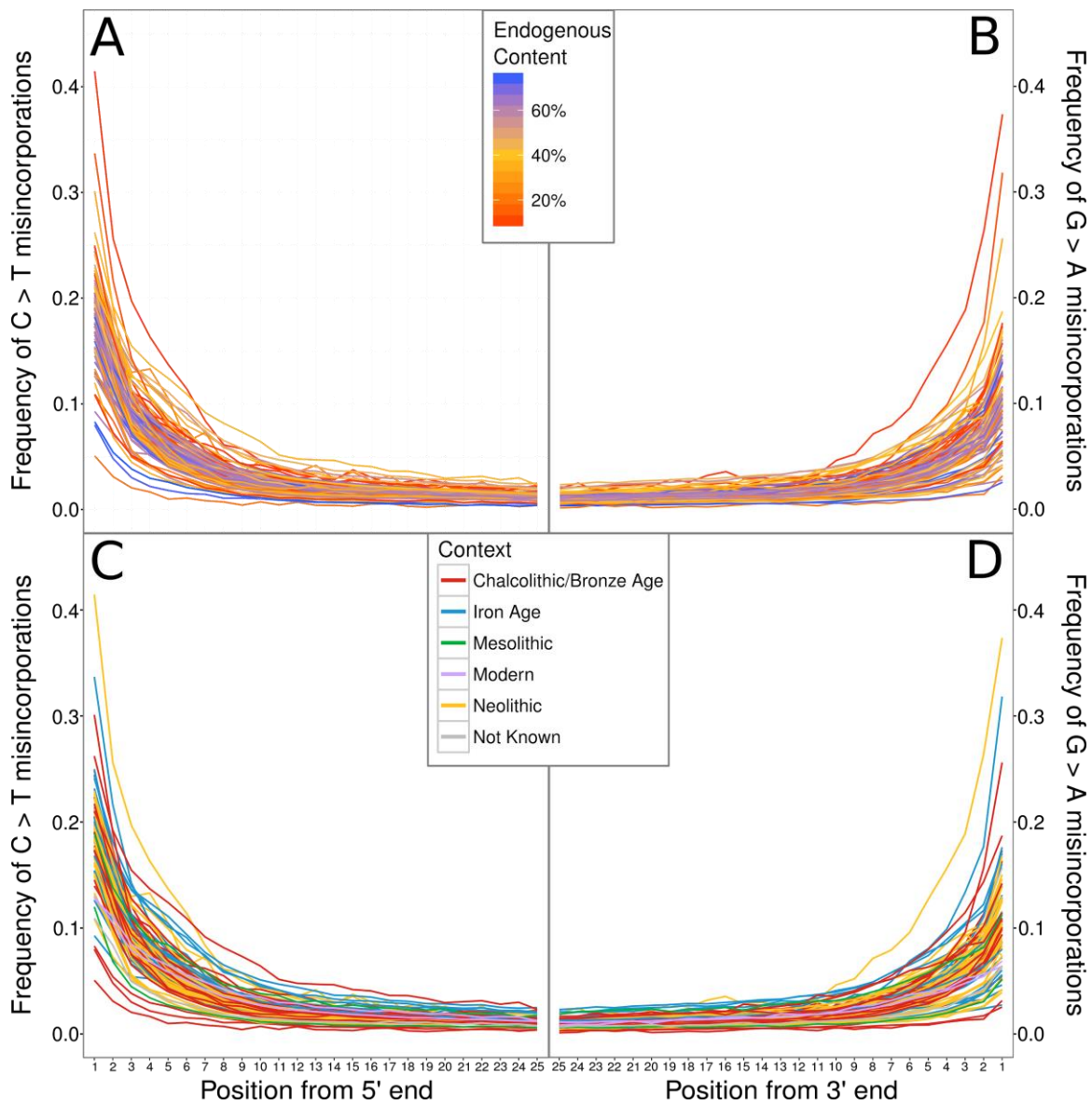


Figure II.2. Sequence misincorporation patterns for ancient Irish samples. (A) and (C) show the frequency of C to T transitions at the 5' ends of reads while (B) and (D) show the frequency of G to A transitions at the 3' ends of reads. In (A) and (B) samples are coloured by endogenous DNA content, and in (C) and (D) samples are coloured by archaeological period.

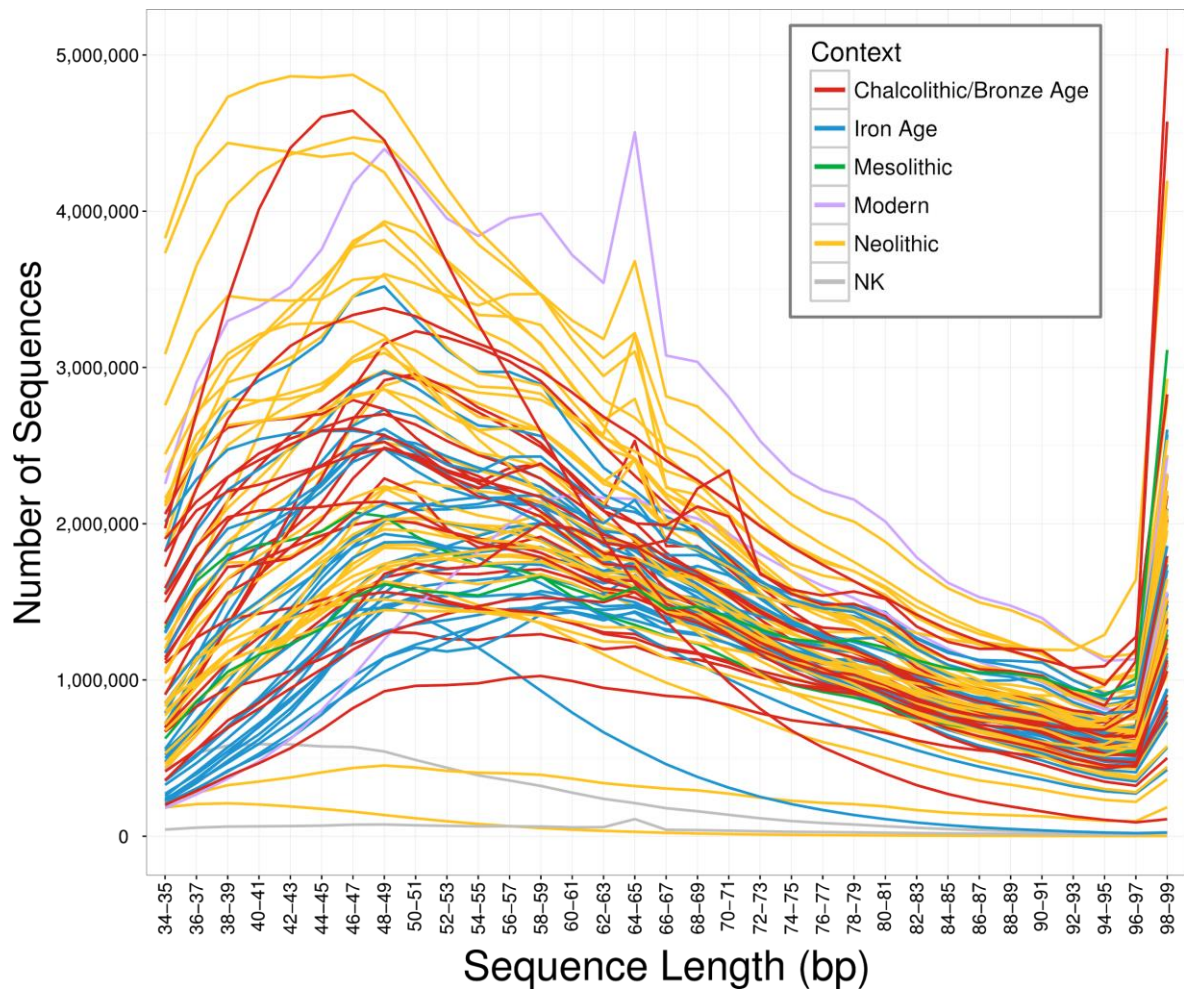


Figure II.3. Sequence length distributions for ancient Irish samples. The percentage of total reads found for any given read length is shown for each sample. Samples are coloured by archaeological period. Two high coverage individuals ($\sim 10X$), previously published in Cassidy et al. 2016, are not shown here for ease of visualisation, but display a similar pattern.

12. Molecular Sex Determination

Osteoarchaeological analysis can provide accurate sex determination of ancient individuals based on differences in the metric dimensions of female and male skeletal anatomy (Buikstra & Ubelaker 1994; Seidemann et al. 1998). Genetic sex determination can both confirm the findings of these studies and also take their place when an insufficient amount of skeletal material is available or the individual in question is a juvenile. Following the methods of Skoglund et al. 2013, we attempted to identify the genetic sex of all 140 individuals sequenced, including those with minimal endogenous DNA. This was done by calculating the ratio of reads aligning to the Y chromosome to reads aligning to both sex chromosomes (R_y). Only reads with a mapping quality over 30 were considered. Samples were assigned female if the upper bound of their confidence interval (CI) for R_y was lower than 0.016, while males required an R_y CI lower bound of above 0.075. Samples for which less than 100 reads aligned to the sex chromosomes were not assigned a sex, regardless of the R_y observed, while samples with between 100 and 300 aligned sex chromosome reads were given a tentative assignment. Results are displayed in Table II.2, with more detailed information on read numbers found in Electronic Data Table S3.

13. Mitochondrial Genome Analysis

To determine mitochondrial coverages and haplogroups, raw sample reads aligned to the human reference genome and revised Cambridge mitochondrial reference sequence (no mapping quality filter) were realigned to the mitochondrial reference alone. Data processing then proceeded in an identical fashion to that described in the above data processing sections. This circumvented the loss of mitochondrial reads that also aligned to similar sequences within the autosomal genome (NUMTs), as multiple mapping locations result in low mapping qualities. The mitochondrial coverage of each sample was then obtained using Qualimap v2.1.1 (Okonechnikov et al. 2015). Coverages are presented in Electronic Data Table S4. Notably, this realigned data was not used for estimates of mitochondrial contamination (see section 3) as misaligned NUMT sequences can inflate mismatches.

To determine sample haplogroups, consensus mitochondrial genome sequences were called using Samtools mpileup (with parameters -B and -Q 30) and vcfutils.pl (vcf2fq) (Li et al. 2009) as in Skoglund et al. 2014. HaploFind (Vianello et al. 2013) was used to identify defining mutations and assign sample haplogroups, which were then ascertained through the direct observation with IGV (Robinson et al. 2011). Only positions with a coverage of 6X or over were taken into account, with the exception of three individuals, BLD6, CAK534 and DQ20, the low coverage of which required relaxed filter of 2X. Haplogroup assignments are displayed in Table II.2, with more detailed mutation and marker information found in Electronic Data Table S4.

14. Y Chromosome Analysis

Of the 93 individuals analysed, 53 were found to be male. To determine the Y chromosome haplogroup for each of these individuals their allelic state at each SNP in the ISOGG database of Y chromosomal markers was assessed, omitting SNPs marked as “under investigation”. The Pileup tool in GATK v2.4 (McKenna et al. 2010) was used to report the base calls for all reads covering each site. Both a minimum base and mapping quality of 30 was required for consideration. Haplogroup assignments are displayed in Table II.2, with more detailed marker information shown in Electronic Data Tables S5-7.

Table II.2. Sex determination and haplogroup assignment for 140 ancient Irish samples. Mitochondrial contamination estimates are also shown - %C, represents the upper limit of contamination, while %C-MD is the estimated contamination rate when sites vulnerable to molecular damage are excluded. Y haplogroup placements that may be confounded by post-mortem damage are bracketed with an asterisk (*).

Sample	Molecular Sex	mtDNA Haplogroup	%C (Upper Limit)	%C-MD	Y Haplogroup
AK84	Female	HV0+195	0.04	0.05	NA
ANN1	Male	K1a+195	0.15	0.00	I2a2a1
ARD2	Male	J2b1a	0.08	0.13	I2a1b1a1
ASH1	Male	K2a9	0.99	0.80	I2a2a1a1a2
ASH3	Male	T2c1d1	0.08	0.09	I2a1a(2*)
AT1	Unknown	-	-	-	-
AT7	Unknown	-	-	-	-
BA342	Unknown	-	-	-	-
BA346	Unknown	-	-	-	-
BA64	Female	HV0+195	0.59	0.22	NA
BAL14	Unknown	-	-	-	-
BAL3211	Male?	-	-	-	-
BAL329	Unknown	-	-	-	-
BB54	Male	U5b2c2b	0.14	0.08	R1b1a1a2a1a2c1a2b2a
BE79	Female	K1a4a1c	0.12	0.08	NA
BG72	Male	K1a4a1	0.13	0.12	H2a
BH32	Female	H40	0.25	0.51	NA
BLD6	Female	V3c	0.85	0.00	NA
BM44	Male	V7a	0.12	0.06	R1b1a1a2a1a2c1a1a1a(1*)
BM45	Female	K1a4a1	0.25	0.26	NA
BM46	Female	H1c3	0.05	0.08	NA
BR79	Male	T2b	0.07	0.11	R1b1a1a2a1a2
BW2	Female	J1c1a	0.10	0.09	NA
BW3	Female	H56	0.30	0.28	NA
CAK16	Unknown	-	-	-	-
CAK375	Unknown	-	-	-	-
CAK530	Female	W5b	0.21	0.23	NA
CAK531	Male	H1	0.00	0.00	I2a2(a1a1a*)
CAK532	Male	J1c3	0.15	0.12	I2a2a1a1a
CAK533	Female	H	0.16	0.13	NA
CAK534	Male	X2b4	4.05	4.76	I
CAK68	Male	H	0.08	0.05	I2a2a1

CH448	Male	H1	0.07	0.08	I2a1b1a1
CL23	Female?	-	-	-	-
CLT622	Unknown	-	-	-	-
CLT627	Unknown	-	-	-	-
CLT640	Unknown	-	-	-	-
CM37	Male	T2b3+151	0.19	0.28	R1b1a1a2a1a2c1a2b2a
CN1005	Female	U2e1b2	0.15	0.16	NA
CR02	Female	-	-	-	-
CT14	Male	U5a1h	0.25	0.31	R1b1a1a2a1a2c1a3a2(a1d*)
CW1	Female	T2c1+146	0.15	0.24	NA
DM01	Female	T1a1+152	0.12	0.10	NA
DM02	Female	J1c3g	0.60	0.59	NA
DM03	Female	J1c1a	0.10	0.08	NA
DM04	Female	K1a1b1g	0.22	0.18	NA
DM05	Female	J1c3g	0.10	0.11	NA
DM06	Female	H1b	0.05	0.05	NA
DM07	Female?	-	-	-	-
DM08	Male	H5c	0.03	0.05	R1b1a1a2a1a2c1(a1a1a1a1*)
DM09	Male	J1c3g	0.13	0.18	R1b1a1a2a1a2c1a1a1a1a1
DQ20	Female	H1c	2.91	4.44	NA
FK142	Female?	-	-	-	-
GG01	Male?	-	-	-	-
GM343	Unknown	-	-	-	-
GR10	Female	H6a1b4	0.19	0.12	NA
GR41	Unknown	-	-	-	-
HW14	Male	K1c1	0.06	0.08	R1b1a1a2a1a
IG134	Male	T2f	0.21	0.15	R1b1a1a2a1a2c1a1k
JNT400	Unknown	-	-	-	-
JNT616	Unknown	-	-	-	-
JNT645	Male	-	-	-	-
JP14	Male	T2c1d1	0.14	0.19	H2a
JP15	Female	H3k	0.03	0.04	NA
KDI275	Male	V15a	0.17	0.16	R1b1a1a2a1a2c
KGH1/KGH2	Male	H	0.14	0.19	I2a2a1a1a2
KGH6	Male	U5b2a	0.09	0.05	I2a1b(-CTS1802)
KIL1	Unknown	-	-	-	-

KNW10	Male	H1aw	0.02	0.05	R1b1a1a2a1a2c1a5b1a1a
KNW4	Female	K1a+195	0.08	0.10	NA
KO12	Male?	-	-	-	-
KO14	Female	H2a1e1b	0.06	0.04	NA
KO3	Female	K1a26	0.36	0.22	NA
KO4	Unknown	-	-	-	-
KO7	Female	U5b2b2	0.51	0.34	NA
LB212	Female	H41a	0.05	0.05	NA
LB213	Female	W4	0.06	0.04	NA
LC04	Unknown	-	-	-	-
LG14	Male	W1c	0.14	0.13	R1b1a1a2a1a2c1a1a1a1a1a
LH27	Female	K1a+195+146	0.25	0.36	NA
LH28	Male	H66a1	0.09	0.00	R1b1a1a2a1a2c1(a*)
LM46	Unknown	-	-	-	-
MB101	Female	-	-	-	-
MB4	Unknown	-	-	-	-
MB6	Male	J1c3	0.23	0.13	I2a2a1a1a
MB9889	Unknown	-	-	-	-
ML1	Unknown	-	-	-	-
MP18	Unknown	-	-	-	-
MV381	Male	H1b	0.09	0.12	R1b1a1a2a1a2c1a5c3
NG10	Male	U5b1+16189+16192	0.22	0.30	I2a2a1(a1a*)
NGZ1	Female	X2b+226	0.06	0.00	NA
NGZ2	Unknown	-	-	-	-
NM61	Unknown	-	-	-	-
PB1239	Female	K1a1a	0.05	0.03	NA
PB1327	Male	-	-	-	-
PB1794	Female	-	-	-	-
PB186	Male	X2b+226	0.14	0.15	I2a1(a*)
PB2031	Male	K1a2b	0.31	0.36	I2a1(a2*)
PB357	Female	U8b1b	0.30	0.24	NA
PB443	Male	K1b1a1	0.15	0.12	I2a1b1a1
PB581	Male	T2b	0.09	0.05	I2a2a1a1a2
PB672	Male	T2c1d+152	0.12	0.10	I2a1(a*)
PB675	Male	H1	0.17	0.19	I2a1(a*)
PB754	Female	-	-	-	-

PB768	Male	H4a1a1	0.07	0.14	I2(a1a*)
PG86	Female	U4a3a	0.09	0.00	NA
PG90	Female	T2b2b	0.15	0.21	NA
PG911	Male	K1b1a1	0.26	0.18	R1b1a1a2a1a2c1a3
PN01	Female	W5b	0.11	0.08	NA
PN02	Male	U5b1c1	0.32	0.40	I2a2a1
PN03	Male	K1a1	0.00	0.00	I2a2a1
PN04	Male	H1+16189	0.12	0.13	I2a2a1
PN05	Male	K1a+195	0.26	0.18	I2a2a1a1a2
PN06	Male	H	0.02	0.00	I2a(xI2a2)
PN07	Male	U5b1c	0.07	0.04	I2a2a(1a1a*)
PN107	Male	U4a2f	0.15	0.15	I2a2a1(a1a*)
PN112	Male	U5b2b	0.23	0.04	I2a2a1a1
PN113	Female	H4a1a1a	0.06	0.09	NA
PP35	Male?	-	-	-	-
PP6	Male	I1a1	0.26	0.36	R1b1a1a2a1a2c
RH125	Female	U5a2+16294	0.06	0.04	NA
RH468	Female	U5a2+16294	0.07	0.03	NA
RH473	Female	-	-	-	-
RH98	Female?	-	-	-	-
RM127	Male	U5a1b1e	0.52	0.11	R1b1a1a2a1a2c1a5
RNR1	Female	J1c3b1	0.10	0.15	NA
RNR2	Male	J1c3b1	0.14	0.09	R1b1a1a2a1a2c1a5c1
RSK1	Male	U5b2a2a	0.78	0.18	R1b1a1a2a1a2c1a(5c*)
RSK2	Male	J2b1a	0.67	0.21	R1b1a1a2a1a2c1
SG27	Male	U5b2b2	0.18	0.22	R1b1a1a2a1a2c1a
SI118	Unknown	-	-	-	-
SP3409	Female	H	0.19	0.36	NA
SRA62	Male	U5a2d	0.14	0.10	I2a2(xI2a2a,xI2a2b)
ST63	Unknown	-	-	-	-
TD97	Male	-	-	-	-
TM81	Unknown	-	-	-	-
TOP11	Unknown	-	-	-	-
TR116	Male	H1e	0.27	0.41	R1b1a1a2a1a2c1a3
VM231	Unknown	-	-	-	-
VM232	Unknown	-	-	-	-

15. Variant Calling and Modern Dataset Preparation

The vast majority of published ancient genomes to date possess sequencing coverage too low for accurate diploid genotype calling. To circumvent this issue two methods were employed, pseudo-haploid genotype calling and genotype imputation. Diploid genotype calling was also carried out for two high coverage Irish samples (Ballynahatty and Rathlin1) and four other high coverage ancient genomes.

15.1 Pseudo-haploid genotype calling

The pileup tool in GATK v2.4 (McKenna et al. 2010) was used to report ancient sample base calls at 594,896 known SNP positions from the Human Origins Panel (Lazaridis et al. 2014) and a set of 2,733,477 autosomal transversions from the 1000 Genomes Phase 3 release (European MAF > 1%) (1000 Genomes Project Consortium et al. 2015)). Bases needed a minimum quality of 30 to be considered and sites with three or more different bases present were removed. A single base call was then picked at random from each site and duplicated to create a diploid homozygous genotype at that position for that individual. Creating genotype calls in this manner has the effect of increasing the appearance of drift artificially on the lineage specific to the individual, but this drift should not occur in any particular direction and thus should not substantially affect inferences about relationships between individuals, although see Section 8.2.4 for further discussion. This strategy was used for all ancient genomes regardless of coverage, to avoid any possible bias it may introduce compared to other genotype calling procedures.

Ancient pseudo-haploid calls for the Human Origins Panel were merged first with each other and then with the modern dataset using PLINK v1.90 (Chang et al. 2015). In cases where the allele of an ancient individual matched neither of the alleles given in the Human Origins dataset, we removed this position in the ancient sample, as they are most likely the result of postmortem damage or sequencing error. This dataset was used for PCA and ADMIXTURE analysis (Chapter Two).

Ancient pseudo-haploid calls for the 1000 Genomes set of transversions were merged with each other and with Mbuti population from the SGDP dataset (Mallick et al. 2016). Again, sites with allelic mismatches were removed. This dataset was used for f - and D -statistic analyses (Chapters Three, Four and Five).

15.2 Imputation

Imputation of 84 ancient Irish genomes (4 low coverage and 5 related individuals were excluded - DQ20, BLD6, CAK531, CAK534, RNR1, RH125, PB357, DM02 and DM09), 90 published ancient genomes (WGS) of approximately 1X coverage or higher, and 16 SNP captured samples, was carried out using Beagle 4.0 (Browning & Browning 2007), following a similar approach to Gamba et al. 2014; Jones et al. 2015; Martiniano et al. 2016 and Martiniano et al. 2017. Also included were five high coverage genomes, NE1, BR2 (Gamba et al. 2014), Loschbour, Stuttgart (Lazaridis et al. 2014) and Kotias (Jones et al. 2015),

downsampled to 2X coverage by selecting random reads with SAMtools v.0.1.19 (Li et al. 2009). Furthermore, chromosome 22 of NE1 was also downsampled to 1X, and used here to test imputation accuracy.

Genotypes for all biallelic autosomal SNPs in the 1000 Genomes phase 3 dataset (~77.8 million sites) were called in samples using the Genome Analysis Toolkit (GATK) UnifiedGenotyper tool (McKenna et al. 2010). The resulting VCF files were filtered to:

1. Add equal likelihoods for missing data (i.e. reset to zero)
2. Add equal likelihoods for genotypes which could be the result of damage (TT at CT SNPs, AA at GA SNPs)
3. Add equal likelihoods at heterozygous sites where one of the alleles may be due to damage.

Sample genotype calls were subsequently merged by chromosome and imputed in 15,000 marker windows using the 1000G phase 3 haplotypic reference panel and genetic map files provided by the BEAGLE website (<http://bochet.gcc.biostat.washington.edu/beagle/>), using the below command.

```
java -Xmx26667m -Djava.io.tmpdir=temporary_files_folder/ -jar beagle.r1399.jar gl=sample_genotypes_chr$.vcf.gz
ref=/reference_files/chr$.1kg.phase3.v5a.vcf.gz map=/reference_files/genetic_mapGRCh37/plink.chr$.GRCh37.map
out=test/chr$_imputed_calls nthreads=8 gprobs=true chrom=${positionX-positionY}
```

This resulted in 30,675,833 markers imputed in 190 individuals. All VCF file manipulation was carried out using VCFtools (Danecek et al. 2011)

15.3 Diploid genotyping

The six high coverage genomes were studied in relation to a modern SNP dataset based on samples from Li et al. 2008; Behar et al. 2010 and Henn et al. 2011, genotyped using the Illumina 660W array in Hellenthal et al. 2014. SNP positions from this dataset were mapped to build 37 (GRCh37) of the human genome and we excluded genotypes with a minor allele frequency below 0.5% and a genotyping rate below 99% across all samples. The final dataset consisted of a total of 474,491 analysed autosomal SNPs in 1530 individuals from 95 different populations, including Ireland.

Diploid genotype calls were generated at these positions for six high coverage samples (Stuttgart, Loschbour, BR2, NE1, Ballynahatty and Rathlin1), using the UnifiedGenotyper tool in GATK v2.4 (McKenna et al. 2010). Only bases with a quality above 30 were considered for calling (-mbq 30). Genotype calls were filtered for a depth of coverage of 10X or above and a genotype quality of 30 or above. Diploid genotypes for the ancient samples were subsequently merged individually with the modern dataset and used in ChromoPainter analysis (Chapter Two).

16. Exploration of accuracy and reference bias for pseudo-haploid and imputed calls

16.1 Imputation Accuracy

In order to assess imputation accuracy, imputed genotypes for chromosome 22 of the downsampled NE1 genome (1X), were compared to diploid genotypes of the same individual, called as in section 7.3, with a more conservative depth of coverage filter of 15X. Results are displayed in Fig. II.4 for a variety of minor allele frequency (MAF) and genotype probability (GP) thresholds.

As expected, homozygote sites imputed with better accuracy than heterozygote sites, with higher imputed genotype probabilities leading to higher imputation accuracy across all MAF filters. For low MAF filters (< 5%) transitions sites were seen to impute at approximately the same accuracy as transversion sites for homozygote genotypes. However, a large discrepancy was seen in the accuracy levels of transitions and transversions for heterozygote sites. This was seen across all MAF filters and GP thresholds, although it decreased somewhat with increasing MAF. Notably, a more subtle reversal of this trend was seen at high MAFs (>5%) for homozygote reference sites.

Based on these results, optimal filters of 5% MAF, 99% GP and exclusion of transition sites were chosen for the majority of ROH, F_{ST} and ChromoPainter analyses in Chapters Three, Four and Five. Exceptions are detailed in the main text.

Finally, to test whether differing ancestral backgrounds or UDG treatment can impact on imputation accuracy, imputed genotypes (chromosome 21) from four downsampled (2X) high coverage samples were compared. These included two European Neolithic individuals of similar ancestry, one treated with UDG (LBK) and one untreated (NE1), alongside an ancestrally divergent Mesolithic European, Loschbour, and a later Bronze Age Hungarian (BR2). No significant difference was identified between individual accuracies (Fig. II.5).

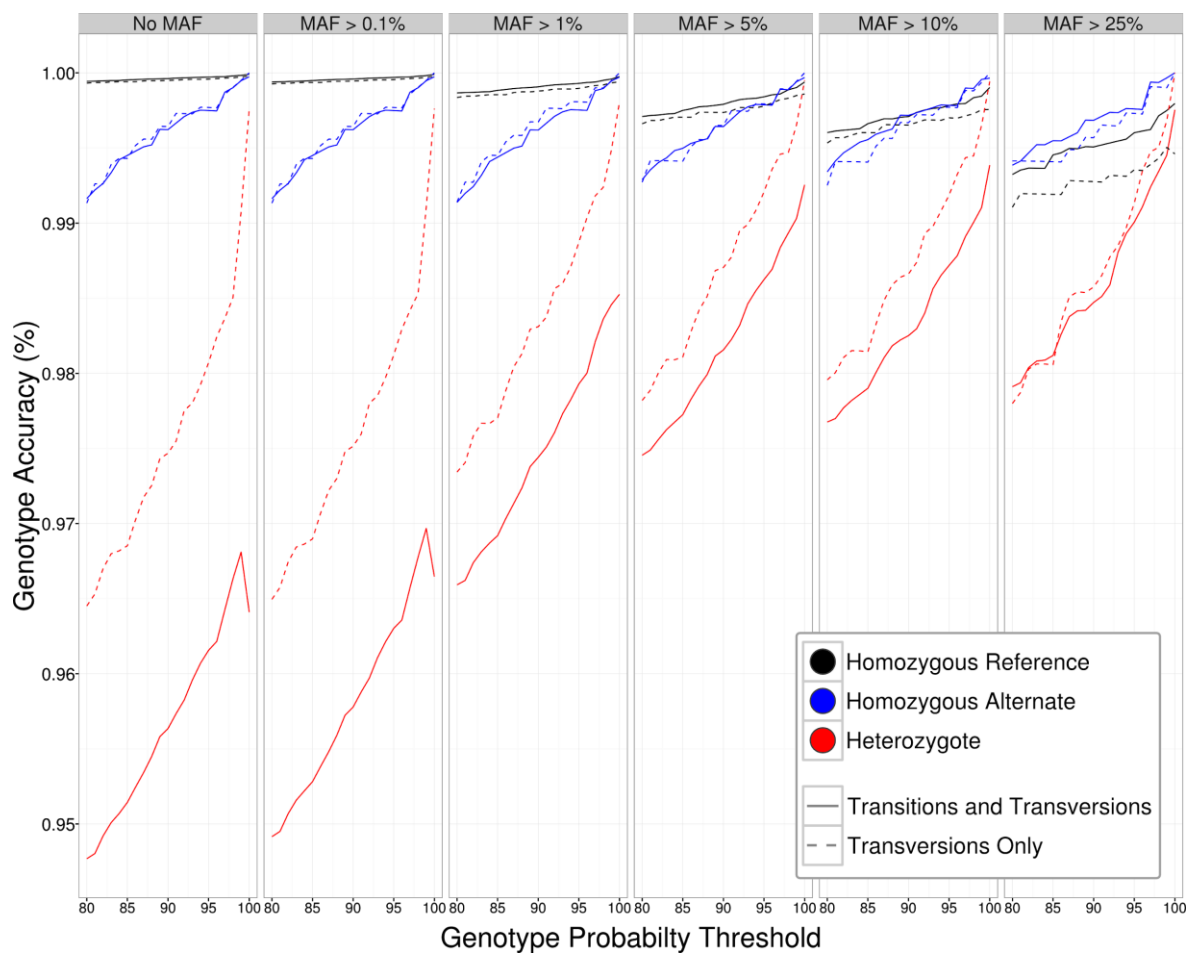


Figure II.4. Imputation accuracy for chromosome 22 of the high coverage NE1 genome downsampled to 1X. The levels of accuracy seen across all SNPs (solid line) is compared to that seen for transversions only (dashed line). Accuracies at different genotype probability (GP) thresholds and minor allele frequency (MAF) filters are shown for the three different genotype categories. MAF filters are based on overall frequency in the 1000 Genomes phase 3 dataset.

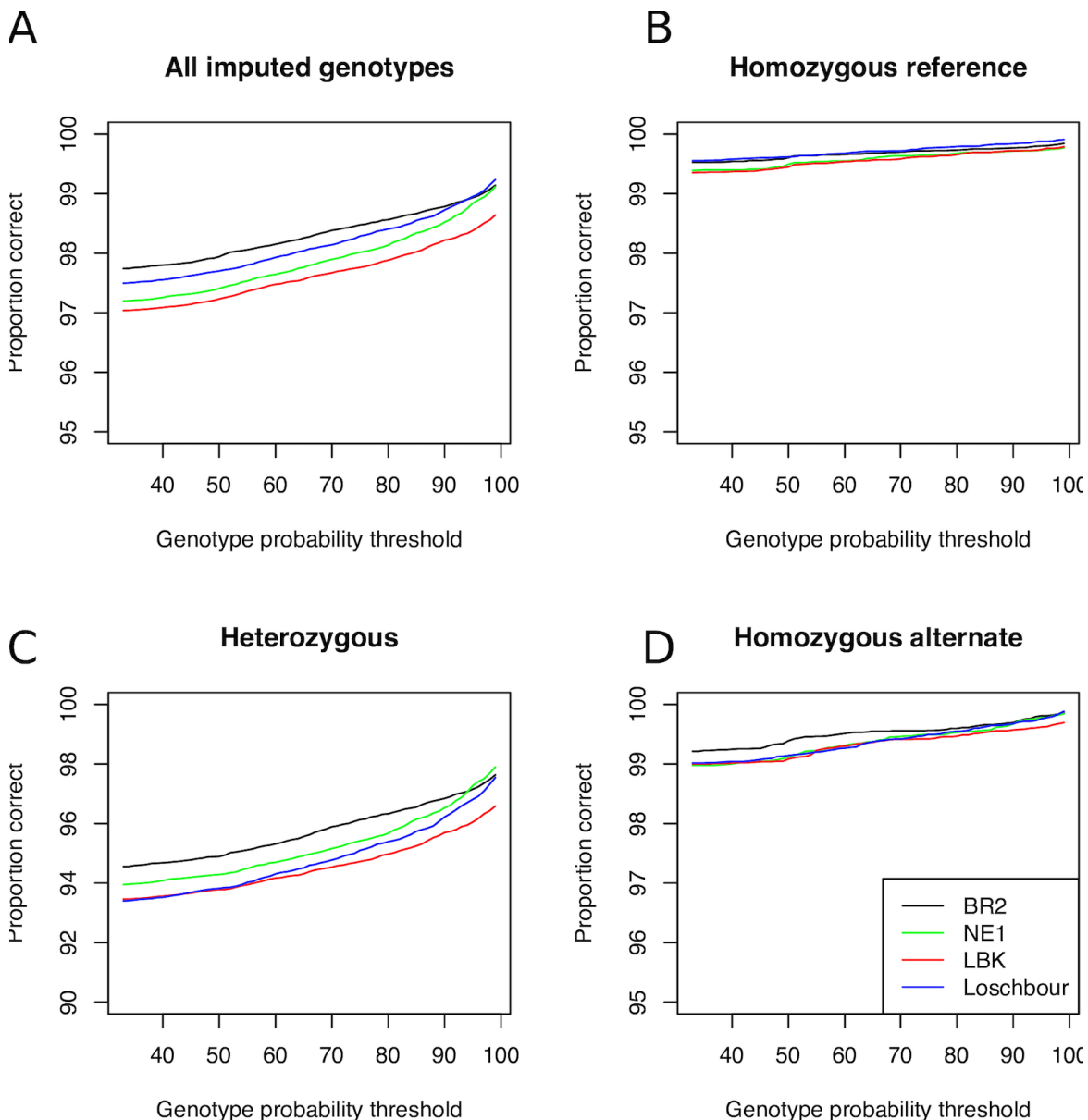


Figure II.5. Estimation of imputation accuracy on chromosome 21. Comparison of variant calls obtained for BR2, NE1, Loschbour and Stuttgart at full coverage with genotypes from the same 4 individuals downsampled to 2X and subsequently imputed. Accuracy in **(A)** all 3 types of genotypes; **(B)** homozygous reference; **(C)** heterozygous and **(D)** homozygous alternate.

16.2 Assessing bias in both imputed and pseudo-haploid calls towards 1000G reference panel populations using D-statistics

We set out to explore potential bias in imputed data resulting from the choice of 1000 Genomes reference panel populations. This was achieved using D -statistics of the form $-D(\text{Chimp}, 1\text{KG population}; \text{directly called ancient}, \text{imputed ancient})$. We also noted significant biases in pseudo-haploid data and discuss these below.

It is important to note that Chimp may be an imperfect outgroup for tests of reference/imputation bias, dependent on whether the majority of ancestral alleles present in chimp are also common human reference alleles. If increased imputation accuracy or alignment quality is present for reference alleles, inflated shared drift may occur between chimp and imputed/pseudo-haploid individuals, relative to diploid calls, which preserve rarer variation. For this reason, we emphasise relative patterns of bias across difference reference panel populations, rather than overall magnitudes. The majority of tests were carried out using five high coverage (>15X) ancient samples, all of whom had been down-sampled to 2X for imputation.

16.2.1 Pseudo-haploid bias to reference panel is stronger than imputation bias

First, we examined inflated shared drift with reference populations for imputed calls, filtered for a genotype probability of 0.99, relative to both pseudo-haploid and true diploid calls. Pseudo-haploid calls were generated at each of the ~30 million imputed sites for the full coverage (>15X) BAM file as described in section 7.1. Diploid genotypes were called for the same sites as described in section 7.3. To control for inaccuracy caused by low MAF we first only examined SNPs that had an overall MAF of 25% or above in the 1000G dataset (~2.7 M sites). To avoid the confounding effects of damage, transversion SNPs alone were also tested (~890 K sites).

The results of this test, $D(\text{Chimp}, 1000\text{G Reference population}; \text{Direct Call}, \text{Imputed Call})$, are displayed in Fig. II.6. Positive scores indicate a higher similarity between the imputed data and reference populations relative to directly called genotypes, while for negative scores the reverse is true. It is immediately apparent that the largest bias is that of pseudo-haploid calls towards the reference panel, when compared to imputed calls (Figure 1; $Z = -4.623$ to -18.80). Also, comparison between diploid and imputed calls, shows definite imputation bias for certain populations, although overall it is smaller and sometimes negative in magnitude ($Z = -4.098$ to 4.009). A relative increase in affinity in all imputed samples to European reference populations is observed, and also to South Asian populations for the Caucasus HG individual, Kotias. Similar results were obtained for transversion SNPs, with some reduction of imputation bias observed (Fig. II.6).

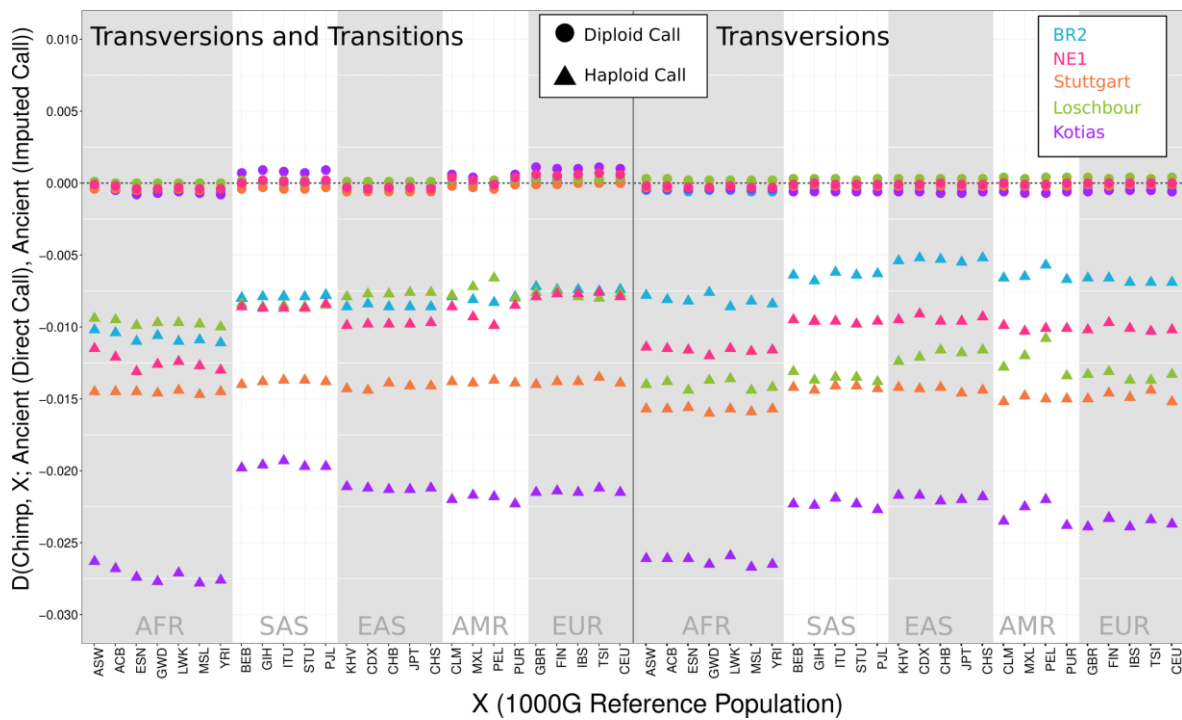


Figure II.6. Affinity of imputed calls to reference panel populations, relative to pseudo-haploid and diploid calls, for five high coverage ancient samples. Results are shown both for all sites and just transversions alone in two separate panels. A world minor allele frequency of 25% has been applied. 1000 Genomes population and super-population names are noted along the X axis.

16.2.2 Imputation bias increases for rarer variants

To further explore the inflated affinity of imputed calls to specific reference panel populations, compared to diploid calls, the test $D(\text{Chimp}, 1000\text{G Ref}; \text{Diploid call}, \text{Imputed call})$ was performed for four different MAF filters (World MAF of 25% and 5%; and European MAF of 25% and 5%), with the results presented in Fig. II.7. Again, these tests were also carried out on transversion SNPs alone.

The World MAF filter of 5% resulted in the largest magnitudes of reference bias ($Z = 2.20$ to 16.60), as well as the largest differences in bias effects, dependent on the ancestry of the reference panel population. The lowest bias effects were seen for a European MAF of 25%, for which the majority of values obtained were in fact negative. These results suggest that common variants, and more specifically variants that are common in the reference populations most closely related to the tested ancient samples, are less prone to bias, most likely due to increased imputation accuracy.

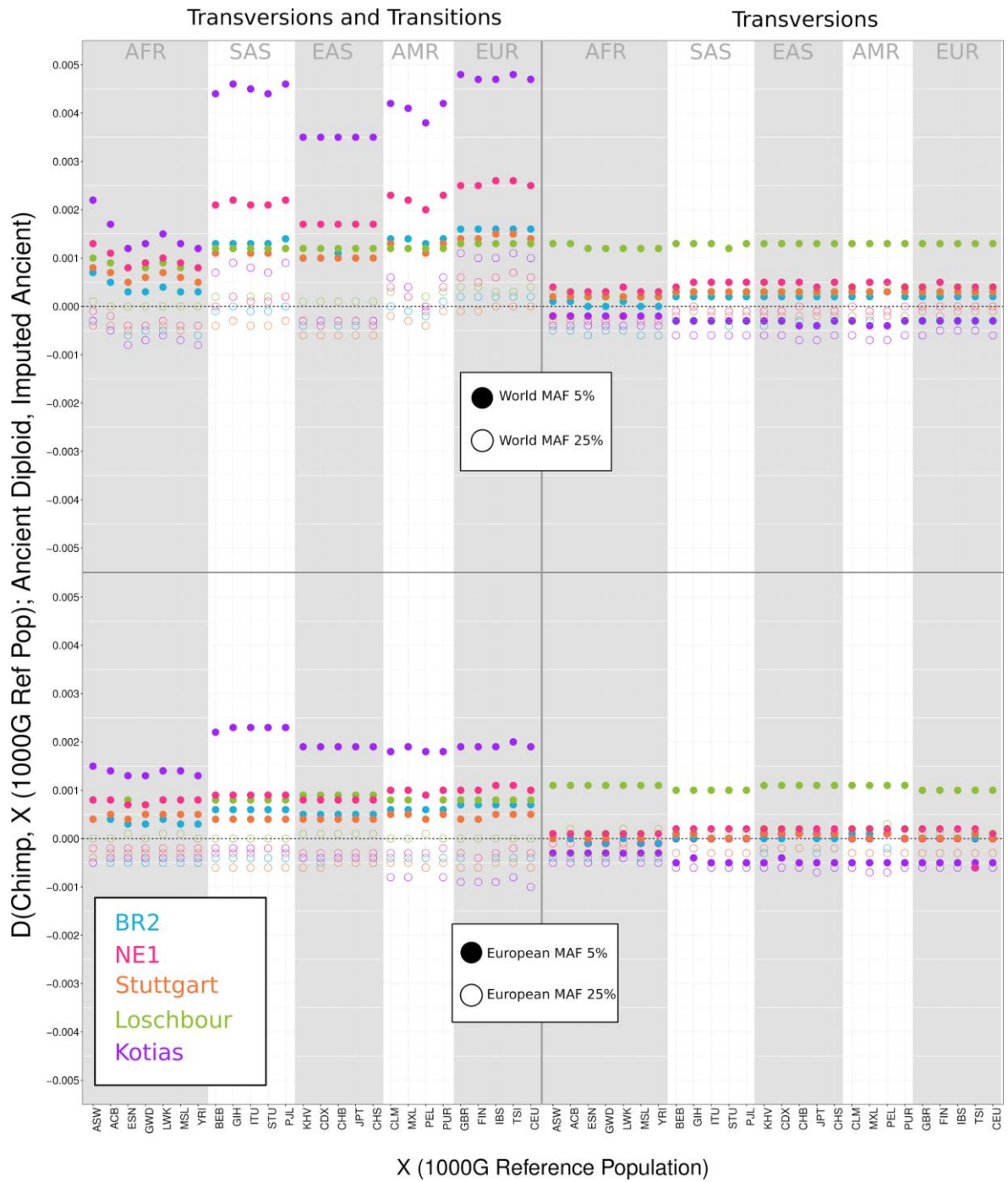


Figure II.7. Affinity of imputed calls from five high coverage ancient samples to reference panel populations, relative to diploid calls, for a series of MAF filters. Results are shown both for all sites and just transversions alone in two separate panels. Top panels display world MAF filters of 25% and 5%. Bottom panels display European MAF filters of 25% and 5%. 1000 Genomes population and super-population names are noted along the X axis.

16.2.3 Imputation bias inflates sample affinity to closely related populations from the reference panel.

The overall pattern of imputation bias tends to stay relatively similar across changing MAFs, despite fluctuation in magnitudes (Fig II.7). This pattern appears somewhat predictable, in that ancient samples will show the most bias to the populations most closely related to them - Europeans for the majority of ancient samples and also South Asians for Kotias. Consistently lower bias is obtained for African reference populations, a reversal for what is seen for pseudo-haploid calls (Fig II.6).

At lower MAF filters further bias for more specific populations is apparent in some samples. For the Neolithic individuals, Stuttgart and NE1, heightened affinity to modern day Iberians (IBS) and Tuscans (TSI) is seen. For the Caucasus HG, Kotias, affinity to Punjabi (PJI) and Gujarati (GIH) is inflated relative to other South Asian populations, while among Europeans, British (GBR) and Tuscans (TSI) show the strongest bias to this individual.

The exception to this rule is seen when a European MAF of 25% is applied, which appears to reduce the overall bias of ancient samples to Europeans, relative to other reference populations, but notably not to South Asians, in the case of Kotias. However, differing affinities for the various European panel populations is still apparent, dependent on ancient sample ancestry (e.g TSI and IBS for Neolithic individuals).

16.2.4 Exclusion of genotype information increases imputation bias

When only transversion sites are considered these biases are substantially reduced in magnitude and also show a more even distribution across the reference populations. This is the likely result of decreased imputation accuracy for transitions due to our exclusion of potential damage sites from the dataset prior to imputation. Interestingly, Loschbour does not show such a reduction. The reason for this is unclear, although we note European HGs show a lower affinity with modern populations, relative to other ancient samples tested, which may be impacting results. Kotias, on the other hand, shows an extreme reduction in bias when only transversions are considered, indicating that prior genotype information for this sample markedly increases imputation accuracy.

16.2.5 Pseudo-haploid bias

A common practice is to construct individual genotypes for analysis by sampling only one allele at each polymorphic site. These pseudo-haploid calls introduce bias which we explored with the test $D(\text{Chimp}, 1000\text{G Ref}; \text{Diploid Call}, \text{Pseudo-haploid Call})$ (Figure II.8) with two different MAF filters imposed (World MAF of 25% and 5%). Only transversion sites were considered to avoid the confounding effects of damage. Significant biases to the reference panel were observed ($Z = -0.768$ to 3.65 and 1.44 to 4.099 for World MAFs 25% and 5% respectively). Importantly, bias was distributed unevenly with regard to reference population tested. The majority of samples show strongest bias to African populations, followed by Europeans. Decreasing MAF from 25% to 5% increased bias for the majority of tests,

Loschbour being a notable exception. However, decreasing MAF also reduced the relative bias towards European reference populations, although bias to African populations remained the highest.

Given that the randomised nature of pseudo-haploidization should not in and of itself create bias to certain populations over others, the problem may stem from alignment bias to the reference genome used, an issue that has been previously noted (Günther, 2016). Reads possessing alleles alternate to the reference sequence may receive lower mapping qualities, and subsequently be filtered out, which in turn would bias haploidization towards reference alleles. This bias would be especially pronounced in ancient data, where high levels of post mortem damage already decrease alignment quality (Schubert et al. 2012). A preponderance of reference alleles, and loss of rarer variation, relative to diploid genotypes, would inflate affinity of pseudo-haploid samples to modern populations, particularly to those whose ancestry composes the human reference genome. Importantly, the application of MAF filters may affect biases in unpredictable ways, depending on the ancestries of the samples and populations involved, as observed in Fig. II.8

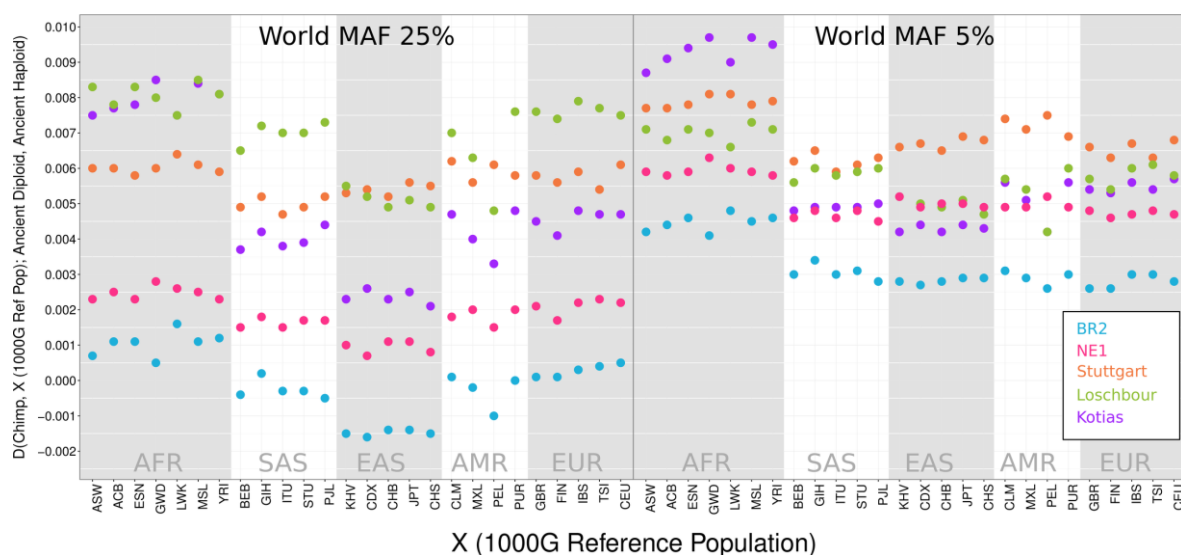


Figure II.8. Affinity of pseudo-haploid calls to reference panel populations, relative to diploid calls, for five high coverage ancient samples. Results are shown for world MAF filters of 25% and 5%. Only transversion SNPs are considered. 1000 Genomes population and super-population names are noted along the X axis.

17. Kinship Analysis

To explore patterns of kinship within the archaeological sites sampled here, IBD sharing was estimated using lcMLkin software, specifically designed for low coverage sequencing data (Lipatov et al. 2015). Genotype likelihoods for 1,978,478 positions (transversions with a European MAF above 5% in the 1000G panel) in 94 Irish samples were called using the SNPbam2vcf.py tool, recommended for lcMLkin

(https://github.com/COMBINE-lab/maximum-likelihood-relatedness-estimation/tree/master/src_python/SNPbam2vcf).

Sites were subsequently thinned to be 25 kb apart and a further MAF filter of 5% in the ancient Irish population, leaving 103,558 SNPs for analysis. lcMLkin was run with the parameters set to sum over all genotypes (-g all), given their raw likelihoods (-l raw).

The resulting coefficient of relatedness (\hat{P}_i score) was plotted against the estimated k_0 (probability of two diploid individuals sharing 0 alleles that are identical by descent) for all pairs of individuals, allowing relationships to be inferred. Two samples, KGH1 and KGH2, were seen to belong to the same individual and sequencing data from these were subsequently merged for population genetic analysis. Furthermore, three pairs of individuals were seen to have a coefficient of relatedness of approximately 0.5, i.e. what would be expected between siblings or parents and offspring. Examination of k_0 could further distinguish between these relationships. Two pairs, RH125/RH468 and DM02/DM05, were found to have a k_0 value close to 0.25, which would be expected for siblings, while RNR1/RNR2 fell closer to the 0 value expected for parent-offspring relationships. Moreover, all pairs shared the same mitochondrial haplogroups, indicative of shared female line descent.

A number of pairs showed lesser degrees of kinship. The pair of sisters, DM02 and DM05, were found to have a second degree male relative, DM09. Given all three individuals share the the same mitochondrial haplogroup (J1c3g), it is likely that DM09 is related to the sisters via the female line. The possible scenarios include a maternal uncle, a nephew (another sister's child) or a half-sibling who shares the same mother. DM09 cannot be a grandchild to both sisters, so this relationship can be excluded, while if female line descent is accepted, DM09 can be ruled out as a grandfather to the pair. All three individuals share a fourth degree relative, DM04, while DM09 shares a separate fifth degree relative, DM08. Two individuals from the Parknabinnia Court Tomb (PB357 and PB675) were also seen to be third degree relatives.

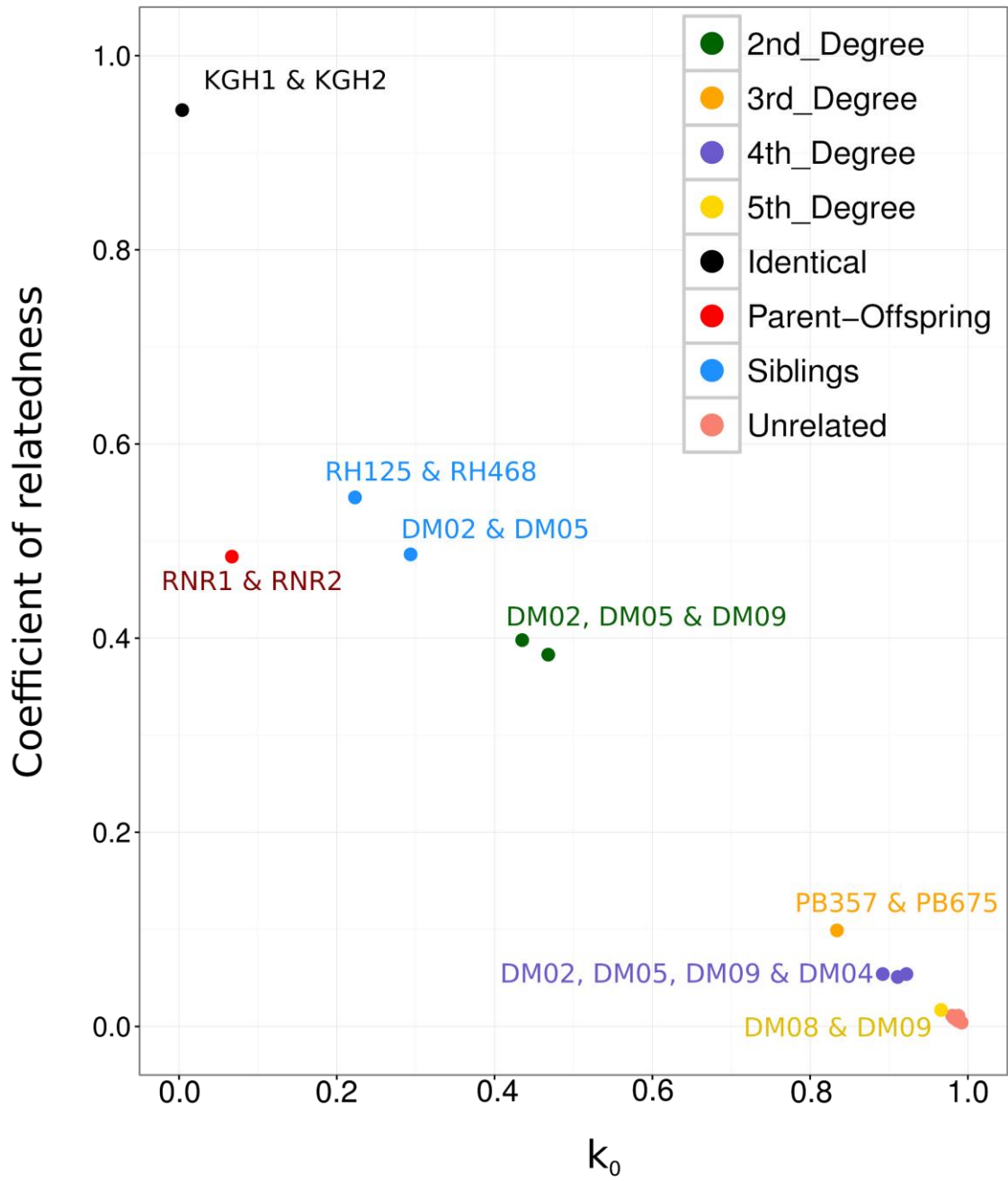


Figure II.9. Coefficient of relatedness, r , against k_0 for pairs of ancient Irish samples. Related samples are labelled.

References

- 1000 Genomes Project Consortium et al., 2015. A global reference for human genetic variation. *Nature*, 526(7571), pp.68–74.
- Adler, C.J. et al., 2011. Survival and recovery of DNA from ancient teeth and bones. *Journal of Archaeological Science*, 38(5), pp.956–964.
- Allentoft, M.E. et al., 2015. Population genomics of Bronze Age Eurasia. *Nature*, 522(7555), pp.167–172.
- Andrews, S., 2010. FastQC: a quality control tool for high throughput sequence data. Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Behar, D.M. et al., 2010. The genome-wide structure of the Jewish people. *Nature*, 466(7303), pp.238–242.
- Briggs, A.W. et al., 2007. Patterns of damage in genomic DNA sequences from a Neandertal. *Proceedings of the National Academy of Sciences of the United States of America*, 104(37), pp.14616–14621.
- Brotherton, P. et al., 2007. Novel high-resolution characterisation of ancient DNA reveals C> U-type base modification events as the sole cause of post mortem miscoding lesions. *Nucleic acids research*, 35(17), pp.5717–5728.
- Broushaki, F. et al., 2016. Early Neolithic genomes from the eastern Fertile Crescent. *Science*, 353(6298), pp.499–503.
- Browning, S.R. & Browning, B.L., 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localised haplotype clustering. *American journal of human genetics*, 81(5), pp.1084–1097.
- Buikstra, J. E. & Ubelaker, D., 1994. Standards for data collection from human skeletal remains. *Arkansas archeological survey research series*, 44. Fayetteville: Arkansas.
- Cassidy, L.M. et al., 2016. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proceedings of the National Academy of Sciences of the United States of America*, 113(2), pp.368–373.
- Chang, C.C. et al., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4, p.7.
- Damgaard, P.B. et al., 2015. Improving access to endogenous DNA in ancient bones and teeth. *Scientific reports*, 5, p.11184.
- Danecek, P. et al., 2011. The variant call format and VCFtools. *Bioinformatics*, 27(15), pp.2156–2158.
- Fu, Q. et al., 2016. The genetic history of Ice Age Europe. *Nature*, 534(7606), pp.200–205.
- Gamba, C. et al., 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nature communications*, 5, p.5257.
- Gamba, C. et al., 2016. Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. *Molecular ecology resources*, 16(2), pp.459–469.
- González-Fortes, G. et al., 2017. Paleogenomic Evidence for Multi-generational Mixing between Neolithic Farmers and Mesolithic Hunter-Gatherers in the Lower Danube Basin. *Current Biology*, 27(12), pp.1801–1810.e10.
- Günther, T. et al., 2015. Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proceedings of the National Academy of Sciences*, 112(38), pp.11917–11922.
- Günther, T., 2016. Do we always need to remap published aDNA data? Torsten Günther Blog: Population genomics | ancient DNA. Available at: <http://www.tgnthr.wordpress.com>.
- Haak, W. et al., 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, 522(7555), pp.207–211.
- Hellenthal, G. et al., 2014. A genetic atlas of human admixture history. *Science*, 343(6172), pp.747–751.
- Henn, B.M. et al., 2011. Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proceedings of the National Academy of Sciences of the United States of America*, 108(13), pp.5154–5162.
- Hofmanová, Z. et al., 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proceedings of the National Academy of Sciences of the United States of America*, 113(25), pp.6886–6891.
- Jones, E.R. et al., 2015. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature communications*, 6, p.8912.
- Jones, E.R. et al., 2017. The Neolithic Transition in the Baltic Was Not Driven by Admixture with Early European Farmers. *Current Biology*, 27(4), pp.576–582.
- Jónsson, H. et al., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, 29(13), pp.1682–1684.
- Keller, A. et al., 2012. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nature Communications*, 3, p.698.
- Jónsson, H. et al., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, 29(13), pp.1682–1684.
- Kılınc, G.M. et al., 2016. The Demographic Development of the First Farmers in Anatolia. *Current Biology*, 26(19), pp.2659–2666.
- Lazaridis, I. et al., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature*, 513(7518), pp.409–413.

- Lazaridis, I. et al., 2016. Genomic insights into the origin of farming in the ancient Near East. *Nature*, 536(7617), pp.419–424.
- Li, H. et al., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), pp.2078–2079.
- Li, H. & Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14), pp.1754–1760.
- Li, J.Z. et al., 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science*, 319(5866), pp.1100–1104.
- Lipatov, M. et al., 2015. Maximum Likelihood Estimation of Biological Relatedness from Low Coverage Sequencing Data. bioRxiv, p.023374.
- MacHugh, D.E. et al., 2000. The extraction and analysis of ancient DNA from bone and teeth: a survey of current methodologies. *Ancient biomolecules*, 3(2), pp.81–103.
- Mallick, S. et al., 2016. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature*, 538(7624), pp.201–206.
- Martiniano, R. et al., 2016. Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nature communications*, 7, p.10326.
- Martiniano, R. et al., 2017. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS genetics*, 13(7), p.e1006852.
- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), pp.10–12.
- Mathieson, I. et al., 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, 528(7583), pp.499–503.
- McKenna, A. et al., 2010. The Genome Analysis Toolkit: a MapReduce framework for analysing next-generation DNA sequencing data. *Genome research*, 20(9), pp.1297–1303.
- Meyer, M. & Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor protocols*, 2010(6), p.db.prot5448.
- Okonechnikov, K., Conesa, A. & García-Alcalde, F., 2015. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*. Available at: <http://dx.doi.org/10.1093/bioinformatics/btv566>.
- Olalde, I. et al., 2014. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature*, 507(7491), pp.225–228.
- Olalde, I. et al., 2015. A Common Genetic Origin for Early Farmers from Mediterranean Cardial and Central European LBK Cultures. *Molecular Biology and Evolution*, 32(12), pp.3132–42.
- Robinson, J.T. et al., 2011. Integrative genomics viewer. *Nature biotechnology*, 29(1), pp.24–26.
- Sánchez-Quinto F., et al., 2012. Genomic affinities of two 7,000-year-old Iberian hunter-gatherers. *Current Biology*, 22(16), pp.1494–1499.
- Schiffels, S. et al., 2016. Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nature Communications*, 7, p.10408.
- Schubert, M. et al., 2012. Improving ancient DNA read mapping against modern reference genomes. *BMC genomics*, 13, p.178.
- Seidemann, R.M., Stojanowski, C.M. & Doran, G.H., 1998. The use of the supero-inferior femoral neck diameter as a sex assessor. *American journal of physical anthropology*, 107(3), pp.305–313.
- Shapiro, B. & Hofreiter, M., 2014. A paleogenomic perspective on evolution and gene function: new insights from ancient DNA. *Science*, 343(6169), p.1236573.
- Skoglund, P. et al., 2013. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *Journal of archaeological science*, 40(12), pp.4477–4482.
- Skoglund, P. et al., 2014. Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science*, 344(6185), pp.747–750.
- Vianello, D. et al., 2013. HAPLOFIND: a new method for high-throughput mtDNA haplogroup assignment. *Human mutation*, 34(9), pp.1189–1194.
- Yang, D.Y. et al., 1998. Technical note: improved DNA extraction from ancient bones using silica-based spin columns. *American journal of physical anthropology*, 105(4), pp.539–543.