



**Trinity
College
Dublin**

The University of Dublin

Decoding the Past: Ancient Genomics and the Histories of the Classic Maya Civilisation

Madeleine Murray

Supervisor: Dr Shigeki Nakagome

A thesis submitted for the degree of
Doctorate of Philosophy

School of Medicine,
Trinity College Dublin

January 2025

Table of contents

Decoding the Past: Ancient Genomics and the Histories of the Classic Maya	
Civilisation.....	1
Table of contents.....	2
Declaration.....	5
Abbreviations.....	6
Research outputs.....	7
Peer-reviewed research articles.....	7
Poster presentation.....	7
Oral presentation.....	7
Courses undertaken (10 ECTS required).....	7
Abstract.....	8
Overview.....	9
Chapter 1.....	11
I. The history of DNA and ancient DNA.....	11
II. Stories from ancient DNA.....	14
III. Lessons from the ancient past: the settling of the Americas.....	17
IV. A non-exhaustive presentation of American civilisations.....	22
V. Genomic perspectives on American civilisations.....	29
Chapter 2.....	31
Nomenclature.....	32
Archaeological contexts of seven Classic Copán individual sequenced in this study.	32
CpM0203 and CpM0506.....	35
CpM11.....	37
CpM12.....	37
CpM13.....	40
CpM34 and CpM4243.....	42
Ethics statement.....	43
Chapter 3.....	46
Abstract.....	46
Contributions:.....	46
Introduction.....	48
Origins of royal families of the Classic Copán.....	48
Archaeogenetics in the Maya region.....	50
Aims of Chapter 3.....	51
Results.....	52
I. Ancient genomes from the Classic Copán.....	53
I. 1. Genomic data generation of Classic Copán individuals.....	53
I. 2. Authentication, kinships, and molecular sexing.....	55
I. 3. Isotopic profiles.....	57
II. Uniparental markers of the Classic Copán.....	60
II. 1. Identifying paternal markers from the Classic Copán.....	60

II. 2. Interpreting paternal haplogroups from the Classic Copán.....	61
II. 3. Identifying maternal haplogroups from the Classic Copán.....	62
II. 4. Comparing Classic Copán maternal haplogroups to present-day Mesoamericans.....	63
II. 5. Contextualising the Classic Copán with ancient Maya.....	64
II. 6. Population structures of ancient Maya through mtDNA.....	66
III. Masking genomic segments derived from non-American ancestry in present-day Americans.....	69
III. 1. Non-American proportions in present-day Maya genomes from HOA.....	69
III. 2. Non-American proportions in present-day Maya genomes from SGDP.....	73
III. 3. Non-American segments in present-day Maya from HOA and SGDP.....	74
III. 4. Non-American proportions in present-day Americans.....	77
IV. Population structure within the Americas.....	78
IV. 1. Genetic structure in ancient and present-day Americans.....	78
IV. 2. Genomic profiles of ancient and present-day Americans.....	83
IV. 3. Population migration in the Americas.....	87
V. Genetic affinities in Maya.....	90
V. 1. Genetic affinities to Classic Copán.....	90
V. 2. Genetic continuity in the Maya region.....	93
VI. Genetic ancestries in ancient Maya.....	95
VI. 1. Admixture modelling of ancient and present-day Maya populations.....	95
VI. 2. Potential variation within the Classic Copán.....	98
Discussion.....	100
Materials and Methods.....	103
I. Individual processing, haplogroup determination and contamination validation.....	103
I. 1. Sampling and archeological contexts.....	103
I. 2. DNA extraction.....	103
I. 3. Sequencing and damage assessment.....	103
I. 4. Sequence data processing.....	104
I. 5. Relatedness with READ.....	104
I. 6. Molecular sex determination.....	105
I. 7. Contamination of individuals based on Y-chromosome.....	105
I. 8. Contamination of individuals based on mitochondrial DNA.....	105
I. 9. Y-chromosome haplogroup determination.....	105
I. 10. Mitochondrial haplogroup determination.....	106
II. Dataset preparation: pileup and masking.....	107
II. 1. Reads trimming.....	107
II. 2. Pileup of ancient with present-day genomes.....	108
II. 3. Masking with MOSAIC.....	109
III. Genome-wide analysis.....	111
III. 1. Principal Components Analysis (PCA).....	111
III. 2. Multidimensional scaling (MDS).....	111
III. 3. ADMIXTURE.....	111
III. 4. Estimated effective migration surface (EEMS).....	112
III. 5. F-statistics.....	112

III. 6. Admixture modelling by qpAdm.....	112
Chapter 4.....	114
Abstract.....	114
Contributions:.....	114
Introduction.....	115
Results.....	118
Discussion.....	122
Methods.....	125
1. Processing and downsampling of ancient genomes.....	125
2. Step 1 of double-imputation pipeline: GLIMPSE.....	125
3. Validation of imputation.....	125
Chapter 5.....	127
Abstract.....	127
Introduction.....	128
Results.....	131
I. Societal insights into ancient American communities.....	131
I. 1. Different kinship results in the Classic Maya of Chichén Itzá.....	131
I. 2. A possible patriarchal society for the hunter-gatherer Sumidouro of Brazil.....	134
I. 3. A family pedigree for early individuals from San Nicolas Island.....	136
I. 4. An intricate connection between maritime hunter-gatherers of San Nicolas, San Clemente, and Santa Catalina islands.....	138
I. 5. Cultural biased pedigree in Patagonia.....	141
I. 6. A surprising link between Anzick and an individual from Chile.....	144
II. Demographic impact of ancient American population.....	145
II. 1. Small IBD segments in ancient Americans.....	145
II. 2. Runs-Of-Homozygosity analysis.....	146
II. 3. Effective population size in the ancient Americas.....	149
Discussion.....	152
Methods.....	156
1. Data preparation and genotype imputation.....	156
2. Inferring IBD segments.....	157
3. Kinship analysis.....	157
4. Pedigree validation with simulation.....	158
5. Runs-Of-Homozygosity analysis.....	158
6. Population effective size estimation.....	158
Chapter 6.....	159
Acknowledgements.....	163
Supplementary figures.....	167
Supplementary tables.....	177
References.....	189

Declaration

I declare that this paper 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 paper 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. I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

I, as the author, declare no conflict of interest.

In this document, the word "Americas" denotes the entire American continent, from Alaska to Patagonia, rather than just a country like the USA, which is often implied in everyday conversation. Additionally, the terms "Indigenous American" or simply "American" refer to the Indigenous peoples who inhabited the American continents prior to the arrival of the Europeans in the XVth century, or their descendants, past and present. Those terms are preferred over the racial attributed term "Amerindian" or the politicised USA term "Native American". Scientists are encouraged to mention communities on how they wish to be cited (Tsosie *et al.*, 2020). In this regard, the population classification of "Eskimo" from public databases was revised to "Inuit", the people's own designation rather than the colonial term imposed by non-Indigenous people (Kaplan, n.d.). Similarly, the language group "Eskimo-Aleut" was rectified to "Eskaleut" (Bergsland, n.d.). Lastly, the classification "Mayan" was amended to "Maya" as the former denotes the language, while the latter refers to the people (Davies, 2016).

Finally, I want to underline the distinction between genetic ancestry and identity, the latter of which is frequently rooted in social connections rather than biological relationships. Genetic findings should not be a threat to cultural identity and individuals retain the right to define themselves.

Abbreviations

AMOVA: Analysis of Molecular Variance

ANGSD: Analysis of Next Generation Sequencing Data

ANS: Ancient North Siberians

ANC-A: Ancestral A

ANC-B: Ancestral B

BCE: Before Common Era

BP: Before Present

CE: Common Era

EEMS: Estimating Effective Migration Surfaces

ENA: European Nucleotide Archive

HOA: Human Origins Array

IBD: Identical-By-Descent

ISOGG: International Society Of Genetic Genealogy

LGM: Last Glacial Maximum

LD: Linkage Disequilibrium

MAF: Minor Allele Frequency

MDS: Multidimensional scaling

MOSAIC: Mosaic Organises Segments of Ancestry In Chromosomes

mtDNA: mitochondrial DNA

NNA: Native North Americans

PCA: Principal Component Analysis

READ: Relationship Estimation from Ancient DNA

ROH: Runs-Of-Homozygosity

SGDP: Simons Genome Diversity Project

SNA: South Native Americans

UDG: uracil-DNA glycosylase

Research outputs

Peer-reviewed research articles

- **Genomic imputation of ancient Asian populations contrasts local adaptation in pre- and post-agricultural Japan**, by Niall P. Cooke, **Madeleine Murray**, Lara M. Cassidy, Valeria Mattiangeli, Kenji Okazaki, Kenji Kasai, Takashi Gakuhari, Daniel G. Bradley, Shigeki Nakagome. Published in *iScience* on the 21st of June 2024 (<https://doi.org/10.1016/j.isci.2024.110050>).
- **Ancient genomes reveal demographic trajectories during the Classic Maya period**, by **Madeleine Murray**, Seiichi Nakamura, Melvin Fuentes, Rolando Canizales, Omar Talavera, Salvador Varela, Zoila Madrid, Carlos Carbajal, Masahiro Ogawa, Minoru Yoneda, Lara M. Cassidy, Takashi Gakuhari, Shigeki Nakagome. Currently under revision in the *Cell Press* journals.

Poster presentation

- **More than a glimpse into the human past: genome-wide imputation of ancient genomes**, by Madeleine Murray, Niall P. Cooke, Lara M. Cassidy, Daniel G. Bradley and Shigeki Nakagome. Poster presentation at the conference “*Reconstructing the human past using ancient and modern genomics*” in September 2022 at the European Molecular Biology Laboratory in Heidelberg, Germany

Oral presentation

- “Ma thèse en 180 secondes”: national competition in Ireland in which participants present their thesis in 3 minutes in French to a non-scientific audience. I won 3rd prize from the jury, and also the public’s prize!



Courses undertaken (10 ECTS required)

- **Genomics Data Science** (30 ECTS at University of Galway)
- **Genome analysis workshop** (5 ECTS at TCD)
- **Research integrity and impact in open scholarship era** (5 ECTS at TCD)
- **Teaching and supporting learning for graduate teaching assistants** (5 ECTS at TCD)
- **Summer course in analysis of high throughput data for population genetics** (5 ECTS at the University of Copenhagen)

Abstract

The Maya civilisation, renowned for its advanced urban centres, sophisticated architectural achievements, and rich cultural heritage, stands as one of the most well-known and influential ancient societies. Understanding this history not only deepens our knowledge of past civilisations but also provides valuable insights into societal resilience—offering lessons that can help modern human societies adapt, survive, and thrive in the face of contemporary global challenges. However, the genetic history of the ancient Maya remains relatively understudied.

This thesis presents seven newly sequenced genomes from the city-state of Copán during the Classic Maya period, including a dynastic ruler and his sacrificial companion. Population structure analyses reveal genetic continuity in the Maya region from the Late Archaic period to the present, alongside an influx of highland Mexican ancestry appearing as early as the Early-to-Middle Classic period.

To enhance the resolution of ancient genomic data, this thesis also evaluates GLIMPSE, an imputation tool for low-coverage ancient genomes. Imputation was assessed across two high-coverage ancient individuals. GLIMPSE consistently achieved high accuracy, particularly at homogenous sites, while heterogenous sites exhibited improved performance above $0.1\times$.

Leveraging these methodological advancements, this thesis reconstructs demographic patterns across ancient American populations using imputed genomes. Analyses of identical-by-descent (IBD) segments and runs-of-homozygosity (ROH) confirm small founding population size in most hunter-gatherers, but not populations from the Early Holocene. In the Maya people, demographic reconstruction reveals substantial growth coinciding with the incorporation of fully domesticated maize.

These demographic insights also contribute to the broader understanding of civilisation decline and societal transformations: population size estimates in the Maya region directly depict a decline coinciding with the Maya collapse.

Overview

This thesis explores the histories of the Classic Maya civilisation through the lens of ancient genomics. It analyses newly sequenced genomes from Copán in Honduras, a major city-state at the southeasternmost extent of the Classic Maya civilisation, as well as recently published genomes from the Terminal Classic Maya at Chichén Itzá and the Colonial Maya from Campeche, both in Mexico. These analyses are contextualised within a broader temporal and geographic framework of over 700 ancient and present-day genomes from across the Americas. The thesis is structured into five distinct chapters, each contributing to a comprehensive understanding of the Classic Maya civilisation.

Chapter 1 explores the histories of American civilisations and highlights the potential of ancient genomics to uncover social structures, kinships, and population demography. This chapter sets the stage for the interdisciplinary approach adopted in the thesis.

Chapter 2 presents the archaeological contexts of individuals in Copán, Honduras. It provides a detailed investigation of the historical and cultural significance of the region, forming a crucial background for interpreting the genomic data.

Chapter 3 investigates the genetic affinity and ancestry of the Classic Copán population. This includes a detailed analysis of a potential member of the ruling dynasty and an associated sacrificial burial, using pseud-haploid ancient genomes from the Americas. The results reveal that while local ancestry has persisted in the Maya region over the last 4,000-5,000 years, gene flow from highland Mexican populations significantly shaped the genetic and social foundations of the Classic Maya civilisation. This chapter links genomic findings to broader historical narratives.

Chapter 4 highlights the effectiveness of genomic imputation, an increasingly valuable tool for addressing the inevitable limitations of low-coverage ancient genomes by computationally enhancing genomic resolution, with a specific focus on its application to the datasets from the Americas.

Chapter 5 makes use of the imputation pipelines validated in Chapter 4 for all ancient American genomes included in Chapter 3. After implementing a stringent cutoff for missingness in imputed data, 324 phased genomes were retained for identifying runs-of-homozygosity (ROH) and identical-by-descent (IBD) segments. These analyses uncover patterns of inbreeding, kinship networks, and population structures, providing deeper insights into the social organisation and demographic history of ancient populations.

Chapter 6, as the concluding remarks, discusses the power of ancient genomics in uncovering the rise and fall of the Classic Maya civilisation.

Chapter 1

Ancient DNA, the settling of the Americas, and ancient American civilisations

I. The history of DNA and ancient DNA

Although Charles Darwin proposed the theories of evolution and natural selection, and Gregor Mendel modelled the process of inheritance using mathematics in the 19th century (Brookfield, 1996), it was in 1953 that the molecular structure of DNA was solved through the efforts of Rosalind Franklin, Francis Crick and James Watson (Watson and Crick, 1953), with the Nobel prize in Physiology or Medicine being attributed to the latter two in 1962 (Maddox, 2003). Since, a series of technologies breakthroughs—development of Sanger sequencing for more rapid and more accurate decoding of nucleotides (Sanger et al., 1977), the ability to amplify targeted DNA fragments through Polymerase Chain Reaction (PCR) (Mullis and Faloona, 1987; Saiki et al., 1988), the first assembly of the human genome (International Human Genome Sequencing Consortium et al., 2001), and the reducing costs together with increased efficiency of Next-Generation Sequencing (NGS) (McCombie et al., 2019)—have led to the mainstream and large-scale production of genomes. Sequencing the first human genome in 2001, even as a haploid, took about 10 years harnessing the efforts of multiple labs worldwide (Shendure et al., 2017), while sequencing a full diploid individual can now be accomplished in just a few days with minimum input from a lab technician (McCombie et al., 2019). Following the assembly of the human reference genome, efforts have expanded to catalogue human genetic diversity across populations worldwide. Initiatives such as the HapMap Project (The International HapMap Consortium et al., 2003) and the 1000 Genome Project (The 1000 Genomes Project Consortium et al., 2015) laid the groundwork by providing dense maps of human genetic variation. More recent efforts, including the Simons Genome Diversity Project (SGDP) (Mallick et al., 2016) and the Human Origins Array (HOA) (Patterson et al., 2012) aim to catalogue common-to-rare genetic variation across populations globally. In parallel, the Human Genome Diversity Project (HGDP) (Cavalli-Sforza, 2005) contributed valuable data from underrepresented populations. Finally, as a strategy to

democratise the field, the capture approach offers a cost-effective alternative to whole-genome/shotgun sequencing (Pickrell and Reich, 2014).

Alongside the growth of genomic data, researchers have applied and refined statistical methods to decode population histories from patterns of genomic variation. For example, Principal Component Analysis (PCA) identifies population structures by extracting the most informative features that explain variation across individuals in high-dimensional genomic data (Patterson et al., 2006). ADMIXTURE is a widely used method for modelling different ancestral components at individual levels (Alexander et al., 2009). Population movement and isolation can be formally tested using f - and D -statistics (Patterson et al., 2012). However, many prehistoric ancestries are often underrepresented by genomic variation of present-day populations, underscoring the significance of directly sequencing ancient genomes extracted from archaeological human remains.

In the 1980s, it was discovered that DNA could survive *post-mortem* (Higuchi et al., 1984). However, ancient DNA is often degraded, resulting in short, fragmented sequences with depurination at terminal ends. Its preservation depends on various environmental factors, including soil minerals, pH (Burger et al., 1999; Kravanja et al., 2025), and fluctuations in humidity and temperature (Kistler et al., 2017). Given archaeological remains are buried underground, extracted DNA contains not only endogenous human DNA but also exogenous DNA from bacteria, viruses, fungi and plants, all living in the surrounding soil. Furthermore, anyone handling archaeological samples could introduce contemporary human DNA contamination (Kazarina et al., 2019; Mann et al., 2018). As a result, two decades ago, researchers still believed that sequencing a complete genome from extinct species or even an individual who died hundreds or thousands years ago was impossible (Der Sarkissian et al., 2015).

Still, several key discoveries have significantly advanced and popularised the field of ancient genomics through selecting denser bones, such as petrous or teeth, establishing dedicated facilities, avoiding substances like animal-based glues, incorporating UV-radiation for contamination removal, and using Uracil-DNA Glycosylase (UDG) treatment to correct DNA damage (Nägele et al., 2022).

The first successful sequencing of autosomal DNA was achieved in 2008 with woolly mammoths, producing a draft genome and insights into their genetic relationship

with elephants (Miller et al., 2008). This milestone was followed two years later by the sequencing of a 4,000-year-old paleo-Inuit genome from Greenland, marking the first whole-genome in humans (Rasmussen et al., 2010).

That same year saw the sequencing of genomes from two archaic hominins. The Neanderthal genome revealed evidence of admixture with non-African *Homo sapiens* (Green et al., 2010), while the genome of a previously unknown hominin, named Denisova, was identified (Reich et al., 2010). Based on just a finger bone and two teeth, without morphological data, the Denisovan genome uncovered an additional introgression with the ancestors of Asians and American *Homo sapiens* (Meyer et al., 2012).

Pioneering work in archaic hominin genomics, led by Svante Pääbo, culminated in his being awarded the 2022 Nobel prize in Physiology or Medicine. The prize recognised his “discoveries concerning the genomes of extinct hominins and human evolution” and “his achievement of what once seemed impossible” (<https://www.nobelprize.org/prizes/medicine/2022/press-release/>).

Today, the field of ancient genomics continues to break new ground. Ancient genomes are being sequenced more frequently as techniques and technologies evolve, pushing the boundaries of what is possible. The oldest sequenced specimens to date include 1-million-year-old mammoths (Van Der Valk et al., 2021), while 2-million-year-old environmental DNA extracted from sediments has provided new insights into the flora and fauna of ancient ecosystems (Kjær et al., 2022).

II. Stories from ancient DNA

Ancient DNA provides invaluable insights into our prehistoric past, illuminating periods predating written records and addressing gaps left by scarce archaeological and skeletal remains. Beyond uncovering archaic introgression into *Homo sapiens*, ancient DNA has resolved long-lasting debates on human history. For instance, it identified three distinct ancestral populations contributing to the genetic makeup of present-day European populations (Lazaridis et al., 2014), confirmed that the nomadic tribes associated with the Scythian culture were genetically diverse rather than homogenous (Unterländer et al., 2017), and dated the arrival of Indigenous peoples into the Americas to the Last Glacial Maximum (LGM) through a single wave of migration across Beringia, rather than multiple waves (Raghavan et al., 2015).

Recently, ancient genomics has expanded its focus from prehistoric to historic periods, complementing knowledge from historical records and sometimes challenging their veracity (Antonio et al., 2024). For example, genetic analysis of a Viking warrior burial revealed that an individual previously identified as male was biologically female, which questioned contemporary perceptions of gender roles and refuted the notion that female warriors in the Viking culture were purely mythical (Hedenstierna-Jonson et al., 2017). Similarly, ancient DNA revised the timeline of Anglo-Saxon migration to England, showing it began earlier than 8th-century historical accounts suggested and extended into the 11th century (Gretzinger et al., 2022).

Ancient genomics plays a pivotal role in resolving legal and ethical disputes. The case of Kennewick Man, discovered in 1996 in the United States and dated to approximately 9,000 years ago, illustrates this potential. Indigenous tribes, referring to the remains as “the Ancient One,” sought their repatriation under the Native American Graves Protection and Repatriation Act (NAGRA). However, earlier morphological analyses ruled out Indigenous American ancestry, leading to a judicial decision against repatriation. In 2015, ancient DNA analysis established Kennewick Man’s genetic affinity to Indigenous people in the Americas, resolving the dispute in favour of repatriation (Owsley and Jantz, 2001; Rasmussen et al., 2015).

Despite these advancements, ancient genomics has predominantly focused on histories in West Eurasia, leaving much of the world underexplored (De La Fuente

Castro and Figueiro, 2024). Over a 1,000 ancient DNA studies have been published (Andermann et al., 2020), most from West Eurasia, in particular Europe, with limited representation from Africa, Asia, Oceania and the Americas (Fig. 1.1) (Dalal et al., 2023; Liu et al., 2021). Expanding ancient DNA research beyond Europe is essential to counter the colonial biases that persist in archaeology and to enrich our understanding of global human history (“Diversity of thought and data enrich archaeology,” 2024). A key technological development in this effort is the “1240k capture” method, which enables genome-wide data recovery from ancient samples with low endogenous DNA by targeting specific regions of the human genome (Pickrell and Reich, 2014). This targeted approach offers a cost-effective alternative to whole-genome sequencing and was originally promoted as a way to democratise the field. However, for commercial reasons, access to the 1240k panel was not initially made broadly available, which unintentionally concentrated its use among groups collaborating with its developers and early adopters, thereby reinforcing disparities between well-funded laboratories in the Global North and under-resourced institutions in the Global South¹ (Ávila-Arcos et al., 2022).

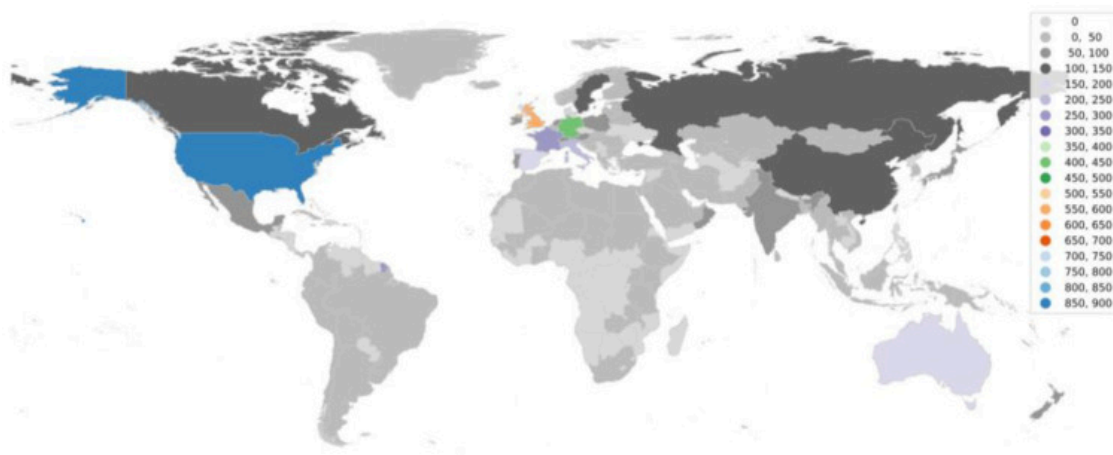


Figure 1.1. World map showing aDNA human-based research publications around the globe, as of 2022. Figure sourced from a previous study (Dalal et al., 2023).

The development of methods to extract and sequence ancient DNA contributed to a dramatic increase in the number of sequenced ancient genomes. Some researchers now compare this surge to a “factory-like” approach (Dalal et al., 2023). With access

¹ Global South: a term commonly used to refer to lower-income countries, many of which have been historically oppressed by colonialism

concentrated among a handful of Global North institutions, and the appeal of aDNA in the popular media coupled with the trend to publish such work in “high impact” journals, the field has been dubbed a “bone rush”—a race to collect and sequence ancient human remains. This has raised significant ethical concerns, particularly regarding the extractive nature of research practices. In many cases, ancient samples were collected from countries in the Global South without meaningful collaboration, recognition, or involvement of local scholars and communities. Such practices have been criticised as “parachute” or “helicopter” research, where the Global North researchers extract data and publish findings with little engagement with the people and regions from which those data originate (Ávila-Arcos et al., 2022; Dalal et al., 2023).

In addition to research work being largely conducted in the Global North, the majority of sequenced ancient genomes are of European origin. This geographical skew results in the underrepresentation of populations from other parts of the world, despite those regions harboring the majority of human genetic diversity (Liu et al., 2021). Those ethical concerns are especially relevant in the Americas, where the societal collapse of the 16th century—caused by colonisation, warfare, and disease—erased much of the region’s genetic diversity (Willerslev and Meltzer, 2021). By focusing on ancient populations from the Americas, this thesis seeks to reconstruct pre-colonial population histories and contribute to a more equitable and inclusive understanding of the region’s past.

III. Lessons from the ancient past: the settling of the Americas

The Americas were the last to be inhabited, with human arrival between 26,500 and 16,000 years ago (Silva et al., 2022). The first peoples arrived in the Americas from Asia on foot. During the Ice Age, cooler temperatures exposed land masses, creating a land bridge known as Beringia, which connected the two continents of Eurasia and the Americas in the north. The Last Glacial Maximum (LGM), occurring between 23,000 and 19,000 years ago, caused sea levels to significantly drop by approximately 50 metres compared to today, forming a 1,800 km-wide land bridge. This land bridge was accessible as early as 30,000 years ago, and remained traversable until about 12,000 years ago. While Beringia was largely ice-free, the harsher climatic conditions during the LGM may have limited movements. Additionally, much of northern North America was covered by the Cordilleran and Laurentide ice sheets, which blocked further exploration, confining early settlers to present-day Alaska (Llamas et al., 2016; Willerslev and Meltzer, 2021).

The first people to reach the Americas are considered descendants of a population formed through admixture between Ancient North Siberian (ANS) and East Asian ancestry around 23,000 years ago (Moreno-Mayar et al., 2018a, 2018b; Raghavan et al., 2014b). This group diverged from the ancestors of Asian populations approximately 25,000 years ago, a period marked by harsh climatic conditions (Llamas et al., 2016; Moreno-Mayar et al., 2018a). This early immigrant population included Ancient Beringians, exemplified by an 11,500-year-old individual USR1, as well as the ancestors of present-day Indigenous American populations. Somewhere south of the Yukon territory, the latter group split into two main branches, Ancestral-A (ANC-A) and Ancestral-B (ANC-B), around 15,500 years ago (Moreno-Mayar et al., 2018b, 2018a; Posth et al., 2018). A summary of these dispersal routes and lineages can be found in [Figure 1.2](#).

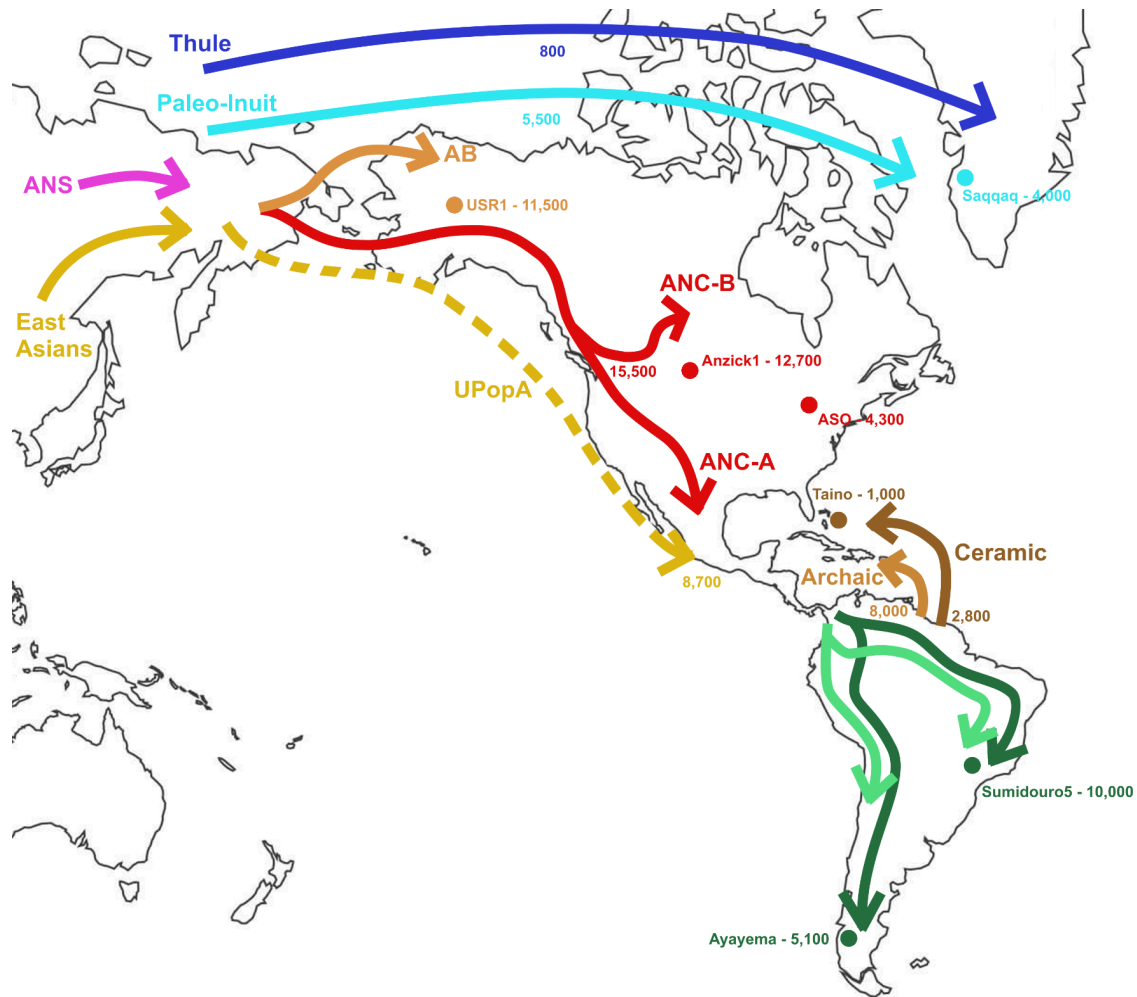


Figure 1.2. Population demography into the Americas, with dates and geographic locations of high-coverage ancient genomes. Arrows trace dispersal routes, with dashed arrows for the ghost population UPopA. High coverage genomes are marked by a dot in the colour of its dispersal route given its ancestry. Dates are in years before present and indicate dispersal route timing, or branch splitting if located at a branch intersection. The geographical outlines are slightly smoothed. The figure is modified from a previous study (Willerslev and Meltzer, 2021).

Groups ancestral to the ANC-A and ANC-B lineages likely dispersed southward from present-day Alaska via two potential routes: the ice-free corridor between the melting Cordilleran and Laurentide glaciers, and west of the Cordilleran sheet along the Pacific coast (Moreno-Mayar et al., 2018b). Geological evidence indicates that the corridor was ice-free from around 14,000 years ago, but ancient DNA from fossil bison and lake sediment indicate that plants and animals required to support hunter-gatherer subsistence were not available until about 13,000 years ago (Heintzman et al., 2016; Pedersen et al., 2016). Yet, this corridor is unlikely to have been the initial route for

human migration, given that human activity from Monte Verde in south Chile has been dated to between 18,500 and 14,500 years ago (Dillehay et al., 2015). Instead, early human dispersal likely occurred along the Pacific coast, which was ice-free and resource-rich between 16,000 and 15,000 years ago, allowing sufficient time to reach and settle in South America (Raghavan et al., 2015; Rasmussen et al., 2014; Willerslev and Meltzer, 2021).

The initial population, isolated in Beringia during the LGM, was relatively small, only reaching a few tens of thousands at most (Llamas et al., 2016). Mitogenomic data reveal a sudden burst in diversification and a significant increase in population size starting around 16,000 years ago. Between 16,000 and 13,000 years ago, the population expanded approximately 60-fold, coinciding with warming temperatures that marked the end of the LGM. This rapid growth aligns with an initial expansion along the Pacific coast route rather than through the ice-free corridor (Llamas et al., 2016).

After the LGM ended and the Cordilleran and Laurentide ice sheets melted, the American continent became more accessible (Llamas et al., 2016). Populations carrying the ANC-B lineage remained in northern North America, primarily in Canada and the Arctic, while those with the ANC-A lineage migrated southward, dispersing into southern North America and reaching Central and South America (Moreno-Mayar et al., 2018a; Willerslev and Meltzer, 2021). Some North American populations, such as a 4,300-year-old individual from the ancient Southwestern Ontario (ASO) population, show genetic evidence of admixture between ANC-A and ANC-B lineages (Moreno-Mayar et al., 2018b; Reich et al., 2012; Scheib et al., 2018).

By 12,000 years ago, rising sea levels submerged Beringia, making it impossible to travel from Siberia to America on foot. Crossing the Bering and Chukchi seas required advanced seafaring technologies and skills to navigate the storm-prone and ice-bound waters (Willerslev and Meltzer, 2021). In this context, the Arctic and Canada experienced two additional waves of migrations from Siberia. The first wave was the arrival of the Paleo-Inuit approximately 5,500 years ago (in cyan on [Fig. 1.2](#)), represented by “Saqqaq”, a 4,000-year-old individual from Greenland (Raghavan et al., 2014a). The second was the Thule expansion around 800 years ago (in blue on [Fig. 1.2](#)), who eventually replaced Paleo-Inuit and are generally considered to be the ancestors of present-day Inuit (Willerslev and Meltzer, 2021).

Additionally, there is evidence suggesting gene flow from an unknown source of ancestry, referred to as “Unsampled Population A” (UPopA) (in dashed yellow on [Fig. 1.2](#)). UPopA remains undiscovered, is neither Ancient Beringian nor ANC-A or ANC-B, and likely represents a lineage diverged from the one that originated from the admixture of ANS and East Asian ancestry between 30,000 and 22,000 years ago. Gene flow from UPopA into Mesoamerican² populations likely occurred around 8,700 years ago. This timing suggests that UPopA-related admixture took place after the initial settlement of South America, as the oldest ancient genomes from South America date to approximately 10,000 years ago (Moreno-Mayar et al., 2018b).

Populations dispersed into South America through two primary routes: the eastern route along the Atlantic coast and the western route using the Andean mountains as a corridor. For instance, the 10,000-year-old Sumidouro individuals from Brazil and the 5,100-year-old Ayayema individual from Patagonia cluster within the same clade, suggesting that they descended from the same initial migration into South America, which can trace back to the ANC-A lineage specifically related to Anzick-1 (Moreno-Mayar et al., 2018b).

However, genetic evidence from more recent South American individuals supports admixture between populations like Sumidouro and groups from Mesoamerica, suggesting that a second wave of ancestry, originating from Mesoamerica, entered South America and intermixed with the descendants of the initial migration (Moreno-Mayar et al., 2018b). The second migration also belonged to the ANC-A branch, but was unrelated to Anzick-1. It followed the same dispersal routes—along the Atlantic coast in the east and through the Andean mountain range in the west—and involved widespread admixture with the initial populations. Given that evidence of genetic continuity with present-day South American populations begins to emerge in ancient genomes dated between 9,000 and 5,800 years ago, this admixture process had largely occurred by then. Interestingly, this genetic influx did not reach Patagonia. The

² Mesoamerica, or Middle America, is a region encompassing south Mexico, Belize, Guatemala, El Salvador and western Honduras. While its boundaries can be defined using cultural/linguistic or geographic criteria, in this thesis, I define Mesoamerica based on climatic features. By avoiding present-day political borders and uncertain cultural/linguistic affiliations from the past, I believe this approach to be more appropriate for ancient populations. Positioned between the arid desert of Aridoamerica and the humid tropics of the Isthmo-Colombian Area, Mesoamerica alternates dry and wet seasons. These boundaries, however, were not fixed—for instance, Aridoamerica expanded southward during the drought of 1,000-1,200 CE (Villa-Islas et al., 2023)—and such shifts are taken into account when contextualising ancient populations throughout this thesis.

5,100-year-old Ayayema individual represents an exception, as their genetic profile reflects continuity with the earlier Sumidouro-like populations (Posth et al., 2018; Silva et al., 2022; Willerslev and Meltzer, 2021).

Humans reached the Caribbean islands by crossing over to the Lesser Antilles, located closest to the South American mainland, and then progressing through islands, eventually reaching Cuba and the Bahamas. There were at least two waves of migrations into the Caribbean: the first around 8,000 years ago leading to the Archaic phase, and the second around 2,800 years ago, marking the beginning of the Ceramic Age in the region. Although both groups coexisted on the islands, there is no evidence of admixture between them (Nägele et al., 2020; Schroeder et al., 2018; Willerslev and Meltzer, 2021).

IV. A non-exhaustive presentation of American civilisations

As peoples moved into the Americas, they domesticated endemic plants, establishing farming settlements. Over time, their languages diversified into distinctive families, and they developed complex civilisations characterised by large urban centres. Among the most renowned civilisations in the Americas are the Inca, Aztecs, and Maya, as summarised in [Table 1.1](#) (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Aztecs*, 2019; *Children of the sun: the Incas*, 2019; *Children of the sun: the Maya*, 2019).

The Inca civilisation thrived in the Andean mountains of South America. Initially settled in the Cuzco valley, the Inca expanded into a vast empire between 1,425 and 1,532 CE, extending their influence over present-day Ecuador, Bolivia, Peru and parts of Columbia, Chile and Argentina. At its height, the empire ruled over 200 tribes (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Incas*, 2019; Cartwright, 2014). The emperor, regarded as the son of the Sun god, was a divine ruler (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Incas*, 2019). He and his descendants were considered sacred and lived in manner befitting their divine status—eating from gold plates, wearing gold-fringed clothing, and residing in a golden palace of the Sun temple (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; Cartwright, 2014). From the capital Cuzco, the emperor governed the empire with the aid of four governors, who divided the land into three parts: one for priests, another for the emperor and nobility, and the third for commoners (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; Cartwright, 2014). Commoners, primarily farmers, cultivated potatoes, maize, beans, peanuts, manioc, chilli peppers, tomatoes, squash, cotton and quinoa, transforming steep hillsides into cultivable terraces for agriculture (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Incas*, 2019). In addition to granting two thirds of their harvest to priests and nobles, commoners paid tribute³. This required households to contribute labour, including military service, road construction, mining, or administrative tasks (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Incas*, 2019; Cartwright, 2014). In return, the empire provided food

³ Tribute: tax collected by a centralised government

during famines, care for the sick and elderly, and housing for their people (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; Cartwright, 2014). The Incas meticulously tracked tributes, taxes, and crop yields using *quipus*, an ingenious system of coloured strings and knots maintained by experts. Couriers, running along the empire's elaborate 24,000-mile road systems, carried *quipus* and oral messages to keep the emperor informed. Remarkably, a message could traverse the 2,300 miles separating Quito, at the northern edge of the empire, to Santiago, at the southern edge of the empire, in just 23 days (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Incas*, 2019; *Engineering in the Andes Mountains: History and Design of Inca Suspension Bridges*, 2005; Cartwright, 2014). The roads themselves showcased advanced engineering, with tunnels, cliffside steps, suspension bridges, stone walls to keep sand at bay, and waystations for travelers. Thousands of workers maintained this infrastructure (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Incas*, 2019; *Engineering in the Andes Mountains: History and Design of Inca Suspension Bridges*, 2005; Cartwright, 2014). Inca stonemason demonstrated extraordinary skills, as seen in temples and governmental buildings, where massive stones were fitted seamlessly, like pieces from a jigsaw puzzle, without the need of cement or mortar. The Sun Temple in Cuzco is a striking example, though how the Incas transported such enormous stones across mountainous terrain without wheels remains a mystery (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Incas*, 2019).

The Aztec empire was located in Mesoamerica, with Tenochtitlán as its capital. Founded as a small settlement around 1,250 CE, Tenochtitlán was built on the swampy shores of lake Texcoco. By scooping mud from the lakebed, the Aztec created a habitable island (*Amplify Core Knowledge Language Arts. Grade 5*, 2022). On this lake, they also developed innovative chinampas, or floating gardens, to cultivate crops, such as maize, beans, pumpkins, squash, tomatoes, avocados, amaranth, and chillies. The high agricultural productivity of these floating gardens led to rapid population growth, transforming the small settlement into a thriving metropolis within 200 years (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Aztecs*, 2019; Van Dop Dejesus, 2015). As Lake Texcoco contained saltwater, the Aztecs built three wide causeways that connected the city to the mainland and supplied fresh water to

fountains, a sewage system, and public toilets. Waste was repurposed as compost for the floating gardens, while urine was used for dyeing fabrics and tanning leather (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Aztecs*, 2019; Garlinghouse, 2022; Van Dop Dejesus, 2015). A bustling marketplace, visited daily by 20,000 to 40,000 people, became a hub of economic activity. Farmers from surrounding areas and long-distance marchants from the entire empire sold goods, equipment, and slaves under shaded arcades. Supervisors at the market inspected goods, collected taxes, and swiftly adjudicated cases of theft and forgery. Transactions were often conducted using cocoa beans or, for higher-value items, grains of gold (*Children of the sun: the Aztecs*, 2019; Hodge et al., 1993; Van Dop Dejesus, 2015). At the heart of Tenochtitlán stood the Sacred District, dominated by the ruler's palace and a large pyramid with two temples. One temple was dedicated to the Sun god and the other to the Rain god (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Aztecs*, 2019). The Aztecs believed that the Sun god required nourishment to complete his daily journey across the sky. To prevent the world's end, they performed daily sacrifices, often of war captives, though tribute in the form of humans was occasionally demanded from conquered territories. Other tributes included cacao, rubber, seashells, cotton, feathers, and precious stones. Tribute payments allowed local rulers to remain in power under Aztec military dominance. The Aztec army was highly skilled, in part due to the compulsory education system that trained boys for combat from an early age. A sophisticated writing system of logograms and syllabic signs enabled efficient communication and governance. Combined with their military strength, this allowed the emperor to exert firm control over the empire. Between 1,430 CE, and their defeat by Spaniards in 1,521 CE, with assistance from local rebellions, the Aztecs expanded their empire to encompass over 80,000 square miles and more than 500 cities (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; Garlinghouse, 2022; Van Dop Dejesus, 2015).

The Maya prevailed in the southeastern region around the Yucatán Peninsula, encompassing southern Mexico, all of Guatemala and Belize, and western portions of Honduras and El Salvador. Unlike the Inca and the Aztecs, the Maya were not unified under a single emperor but were instead divided into numerous city-states, each governed by its own royal dynasty. These city-states engaged in both warfare and trade.

Traded goods included jaguar skins, feathers, jade, obsidian, and cultivated crops such as maize, chillies, tomatoes, pumpkins, cocoa, avocados, and beans (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; Ferguson, 2007; Martínez de Velasco Cortina and Vega Villalobos, 2015; Sharer, 2009). At the heart of each city-state was an elevated acropolis, which served as the public square and housed temples, pyramids, and palaces (*Children of the sun: the Maya*, 2019; Martínez de Velasco Cortina and Vega Villalobos, 2015). Maya architecture is notable for its intricate decoration, including stelae⁴ and hieroglyphic writing that adorned plazas and buildings. A prime example of this is the hieroglyph staircase in the city-state of Copán. Most Maya inscriptions begin with a Long Count date, a linear calendar that records time since the world was created by the gods. In daily life however, the Maya used two cyclical calendars: the “Haab”, a solar calendar of 365 days divided into 18 months of 20 days plus a special 5-day month called “Wayeb”; and the “Tzolk’in”, a divine calendar divided into 20 periods of 13 numbers. By simultaneously rotating the two calendars as wheels (Fig. 1.3), there are 18,980 unique combinations of glyphs, consisting of a cycle of 52 years known as the “Calendar Round”. As Maya rituals were strongly connected to astronomy, these calendars guided ceremonies, such as during the Wayeb (Davies, 2012; Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015). To track such events, the Maya applied their own counting system. Their numbers were made up of three characters: a dot, a line, and a zero (Fig. 1.3). To count from 0 to 5, one would need to write zero, one dot, two dots, three dots, four dots, and a horizontal line. And to count from 6 to 9, one to four dots would be drawn right on top of the line, 10 being two horizontal lines. As the Maya used a base twenty system, the number 15 would be written as three horizontal lines, but the number 20 would require another glyph with a dot for the top glyph and a zero for the bottom glyph. Continuing on higher numbers, their number 80 would consist of four dots as the glyph above, and a zero for the glyph below. Another example is a single dot for the glyph above (corresponding to twenty), and four dots directly placed above two horizontal lines, to read the number 34 altogether. Similarly, the number 234 is written as a dot on top of two horizontal lines for the glyph above, and four dots on top of two horizontal lines for the glyph below (Sharer, 2009).

⁴ Stelae: large blocks of monolithic stone carved by the Maya

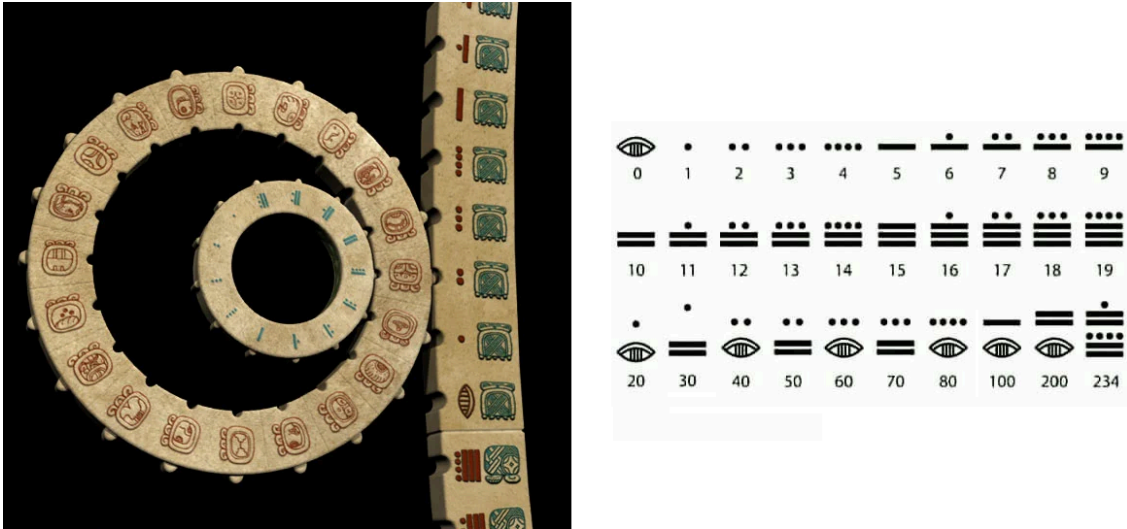


Figure 1.3. Maya calendars and numbers. The Tzolk'in divine calendar rotates inside the Haab solar calendar, itself rotating inside the Long Count calendar (left). The Haab calendar is decorated with 20 glyphs, while the Tzolk'in calendar is numbered from 1 to 13. Maya numbers from 0 to 19, as well as higher numbers, illustrates the counting system (right). Pictures are sourced from an archeological blog (Davies, 2012).

In addition to inscribing texts on monuments, the Maya produced numerous books, including codices that preserved their religious and historical traditions, along with calendars and astronomical observations. Tragically, many of these books were burned by Catholic priests during the Spanish conquest, who regarded them as heretical. However, the Maya managed to preserve some of their knowledge by adopting the Latin alphabet, allowing their traditions and rites to endure to the present day. The Maya worshipped their gods through grand ceremonies held at the acropolis, which often included feasts and blood sacrifices. Kings played a central role in these rituals, as they were believed to mediate between the gods and their people. Bloodletting, for instance, was a significant royal duty, as depicted on a stone carving from Dos Pilas (*Children of the sun: the Maya*, 2019; Martínez de Velasco Cortina and Vega Villalobos, 2015; Scherer, 2007). Around 900 CE, many Maya city-states experienced a rapid decline, a phenomenon known as the “Maya collapse.” This event marked the transition from the Classic (250-900 CE), considered the golden age of Maya civilisation, to the Postclassic Maya periods (900-1,521 CE) (Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015). Several hypotheses have been proposed to explain the collapse, with

the most plausible being a combination of factors. The timing of the collapse coincided with severe and prolonged droughts, which likely caused famines and social unrest, as kings failed to fulfill their role as appeasers of the gods. These pressures may have led to wars, revolts, and possibly epidemics linked to widespread malnutrition. Analyses of sediment cores indicate a series of droughts between 800 and 1,000 CE, accompanied by increased offerings and sacrifices in temples dedicated to the rain god, as well as heightened references to warfare (Andrews and Fash, 2005; *Children of the sun: the Maya*, 2019). During this time, cities in the southern parts of the Maya region, such as Copán, Tikal, and Palenque, declined, paving the way for the rise of northern cities like Chichén Itzá. Chichén Itzá emerged as a regional centre, expanding its influence over the eastern and western coasts of the Yucatán Peninsula and likely controlling coastal trade during the Postclassic period (Masson, 2012; Olga Hernández Espinoza and Márquez Morfín, 2015).

Table 1.1. Summary of characteristics of the Inca, Aztec, and Maya civilisations

	Period (CE)	Geography	Gouvernance	Sacrifices	Construction	Writing
Inca	1,425-1,532	Andean mountains	Centralised (Cuzco)	No	Building without mortar & elaborate roads with suspension bridges	No (but quipu knots to keep track of taxes)
Aztec	1,430-1,521	Mesoamerica	Centralised (Tenochtitlán)	Yes (daily)	City built on swamp with aqueducts and pyramids	Yes (with logograms and syllabic signs)
Maya	Classic: 250-900 Postclassic: 900-1,519	Mesoamerica	Several city-states	Yes (occasionally)	Pyramids and stelae with hieroglyphs	Yes (with complex hieroglyphs)

One reason that these three civilisations are so well-known is that Spanish conquistadors documented them upon their arrival in the Americas during the Colonial period. Equipped with superior weapons and bringing foreign diseases like smallpox, the Spaniards triggered the collapse of all three cultures (*Amplify Core Knowledge Language Arts. Grade 5*, 2022; *Children of the sun: the Aztecs*, 2019; *Children of the sun: the Incas*, 2019; *Children of the sun: the Maya*, 2019; Cartwright, 2014; Garlinghouse, 2022; Van Dop Dejesus, 2015). Warfare, combined with devastating epidemics, reduced Indigenous populations of Mesoamerica and South America by

90-95% from 1,492 CE onwards. This demographic collapse and subsequent loss of genetic diversity pose significant challenges for genetic research (Livi-Bacci, 2006; Silva et al., 2022). To address this gap, efforts must be made to reconstruct this lost heritage using ancient DNA from archeological remains predating the arrival of Spanish conquistadors.

V. Genomic perspectives on American civilisations

Central and South America are home to some of the most iconic cultures, including the Inca Empire, Aztec, and Maya. While archeological remains provide valuable insights, they alone cannot fully unravel the complexities of these civilisations. Ancient DNA complements this knowledge, offering a unique window into the genetic, cultural, and social dynamics of past societies.

Genomic research on the Inca Empire has provided key insights into the role of ritual practices in shaping their society. A recent study, for example, sequenced individuals associated with the Qhapaq hucha ceremony, a ritual of high-altitude child sacrifice in the Inca Empire (De La Fuente Castro et al., 2024). Among the three individuals analysed, one child exhibited genetic affinity to coastal Peru, a finding further supported by the textiles and footwear he was wearing. However, isotopic analysis of his tooth enamel revealed a terrestrial diet, devoid of marine foods in the last year and a half before his death, suggesting long-distance movement before his final burial at the Qhapaq hucha ceremony.

Genomic studies of individuals buried at Machu Picchu, the royal estate of Pachacuti, have demonstrated the remarkable diversity of the community serving the empire's rulers (Salazar et al., 2023). Analysis of 34 individuals revealed that servants were drawn from throughout the empire, with genetic ties to the Andean highlands, the Peruvian coast, and even Amazonian regions of Ecuador and Colombia. Many individuals exhibit mixed genetic ancestries, suggesting intermarriage between groups from distant regions. These findings complement historical accounts of the empire's ability to mobilise and unite diverse populations within the Inca state, showcasing how genomics can validate and expand on traditional archeological and historical interpretations.

Genomic evidence has also shed light on the impact of Aztec imperialism on conquered populations (Mata-Míguez et al., 2012). A study of ancient mitochondrial DNA from Xaltocan, the capital of a pre-Aztec city-state, revealed a significant genetic shift following the Aztec conquest. Maternal lineages from the pre-conquest period were replaced by new lineages in the post-conquest era, implicating large-scale population movement and reorganisation under Aztec rule. This genetic turnover is

consistent with historical accounts of Aztec strategies for managing their empire, which included the forced resettlement of conquered peoples and the establishment of trade networks that spanned vast distances.

For the Maya, genomic analyses have revealed intricate connections between mythology, ritual practices, and biological kinship (Barquera et al., 2024). A study of a mass burial near the Sacred Cenote at Chichén Itzá uncovered genome-wide data from 64 subadult individuals, all males, revealing numerous close kinships among them. This included two pairs of monozygotic twins, a discovery with profound cultural implications given the prominence of twins in Mayan mythology. The Popol Vuh, a sacred K'iche' text, recounts the adventures of the Hero Twins, who undergo cycles of sacrifice and resurrection to defeat the gods of the underworld. The presence of twins in a subterranean burial context, viewed by the Maya as entrances to the underworld, aligns with the ritual themes of duality, sacrifice, and rebirth central to Mayan mythology. Ancient genomics also revealed that one-quarter of the individuals in this burial were closely related, suggesting that kinships played a significant role in their selection for sacrifice. Combined with dietary evidence indicating shared consumption patterns, these findings suggest that sacrificial rituals may have been carefully orchestrated family or community events. This integration of genetic data with mythological and archaeological evidence provides a nuanced view of Mayan ritual practices and their cultural significance.

These studies demonstrate the power of ancient genomics to uncover patterns of mobility, diversity, and cultural integration during ancient civilisations in the Americas.

Chapter 2

Archaeological contexts of the Classic Maya State, Copán in Honduras

Located in present-day western Honduras, Copán was one of the most powerful Maya city-states. Through militarily, trade, and cultural influence, the city-state of Copán dominated the southeastern Maya region (Fig. 3.1) (Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015). Renowned for its well-preserved sculptures and stelae, the ruins of Copán stand as a testament to its former grandeur.

The earliest human occupation at Copán goes back to the Early Preclassic (2,000-1,000 BP), or possibly even the Late Archaic period (5,600-3,700 BP), characterised by small farming communities reliant on maize-based food production (Fash, 1991; Rue et al., 2002). By 300-400 CE, these small settlements had developed into a thriving city, marked by significant population growth and the emergence of diverse economic activities (Andrews and Fash, 2005). In 426/427 CE, *K'inich Yax K'uk' Mo'*, the dynastic founder of Copán, arrived according to epigraphic evidence. He commissioned the valley's first Classic Maya style monument, inscribed with hieroglyphic texts, and established the initial structure of the Acropolis. Over time, his successors expanded these into a monumental complex (Andrews and Fash, 2005; Fash, 2002).

The Acropolis, spanning three hectares, was the epicentre of Copán's political, economic, and religious life. It encapsulates the East and West Courts, temples, administrative buildings, the royal residence and a ball court (Martínez de Velasco Cortina and Vega Villalobos, 2015). High-relief sculptures adorned its facades, while numerous stelae and altars decorated its plazas (Andrews and Fash, 2005; Martínez de Velasco Cortina and Vega Villalobos, 2015). One of the Copán's most impressive monuments, the Hieroglyphic Stairway, was completed in 757 CE under the 15th ruler. It remains the longest known pre-Columbian inscription across the Americas (Fash, 2002). Throughout the history of the Classic Copán, the dynastic succession saw 16 rulers, under whom Copán flourished as a political, economic, and ceremonial centre, surrounded by elite and commoner residential structures. At its peak, around 800 CE, the population is estimated to have reached approximately 20,000 inhabitants (Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015; Webster et al., 1992).

However, the political regime ultimately collapsed around 820 CE, marking the end of the Late Classic period in Copán (Andrews and Fash, 2005).

Nomenclature

To date, more than 4,500 structures have been confirmed in the Copán Valley (Andrews and Fash, 2005), where a total area of approximately 24 km is organised into 500-metre grid squares in order to identify and register each of these structures. Each square block is designated by a combination of a number and a letter, indicating its vertical and horizontal axes respectively (*e.g.*, 10L, 9M, or 8N). Within a square block, different structures are numbered sequentially, starting from one. For example, a structure labelled as “10J-45” represents the 45-th structure within the block marked as 10J. This nomenclature makes it possible to locate any structure on the archaeological map of the Copán Valley, published in 1979 (Fash and Long, 1983).

Archaeological contexts of seven Classic Copán individual sequenced in this study

The structures, from which the Classic Copán individuals shotgun-sequenced were excavated, are registered as two distinct residential clusters. The first cluster lies approximately one kilometre west of the Acropolis, specifically within the 10J block, comprising Structures 10J-39, 10J-45, and 10J-71. The second cluster, located north of the Great Plaza, falls within the 9L block and consists of structures 9L-104 and 9L-105 (Figs. 2.1-2.3).

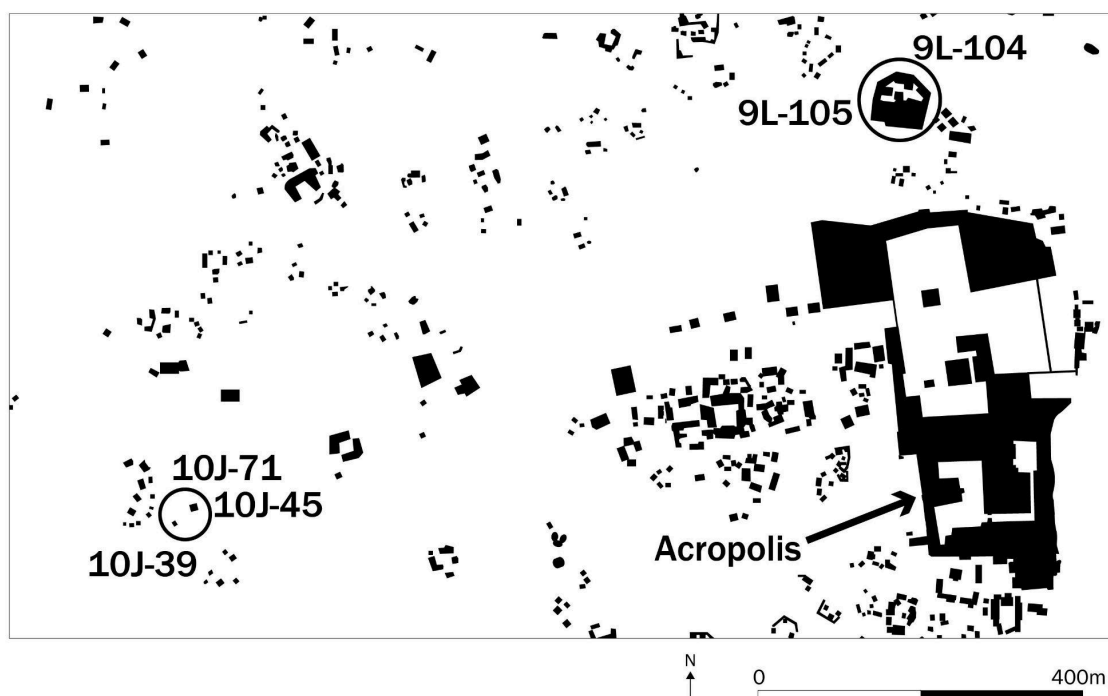


Figure 2.1. Locations of the 10J-39, 10J-45, 10J-71, 9L-104, and 9L-105 structures relative to the Acropolis.



Figure 2.2. Detailed locations of structures 10J-39, 10J-45, and 10J-71.

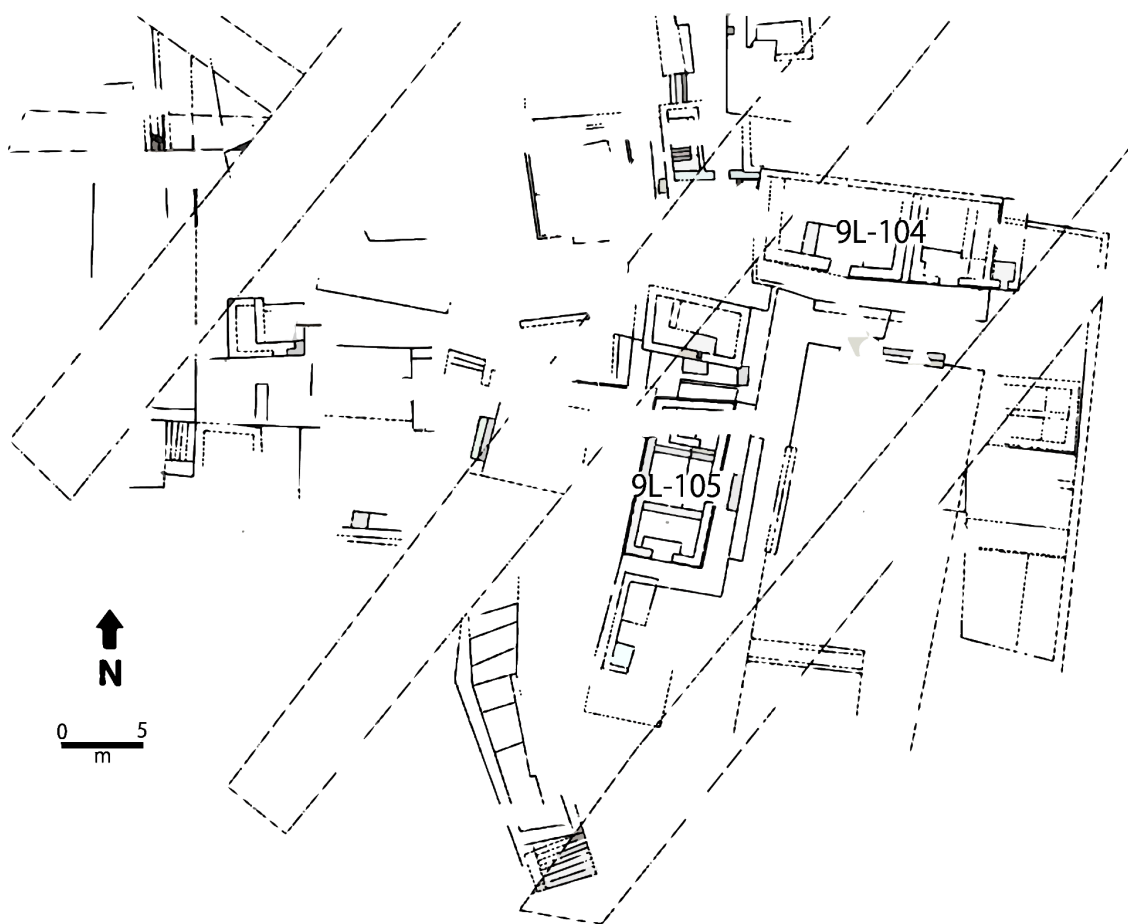


Figure 2.3. Detailed locations of structures 9L-104 and 9L-105.

A wealth designation was used to provide an approximate measure of social standing, with “0” representing the lowest and “5” the highest. This estimate was based on various markers including the form of the burial structure and the number, size, and complexity of accompanying burial offerings. The burial form, contexts, and contents, marked by the vaulted stone chamber with two stone burial slabs, as well as the presence of exquisite jade offerings inside and outside of the chamber provide compelling evidence that CpM13 held a position as one of Copán’s dynastic rulers (Cobos, 2004; Price et al., 2014). Notably, this individual was accompanied with two jade pectorals. One of these pectorals features intricate engravings of mat motifs, symbolising the rulership and divine power of the buried individual (Schele and Mathews, 1998). The second pectoral displays a carved motif of the Pax god, signifying the military prowess of the entombed individual. Adjacent to the burial site of CpM13 within the same structure (10J-45), CpM12 was found beneath the wall of the basement

at the funerary temple constructed for CpM13, without accompanying any offering; CpM12 is then identified as a sacrificial companion (Cobos, 2004; Price et al., 2014).

It is important to note that there is a discrepancy between the wealth designation scores assigned to individuals and their archaeological contexts. These two structural clusters are located in close proximity to the core of Copán, namely the Great Plaza and Acropolis, indicating their significance in the Classic Copán state. Therefore, the inhabitants of these structures can be regarded as part of the elite. However, it is worth mentioning that the wealth designation measures are based on the criteria outlined by Krejci and Culbert (Krejci and Culbert, 1995), which may not be fully applicable to burial contexts in the ancient Copán society.

CpM0203 and CpM0506

Burial 7-2000, located within the interior of the southeast corner of Structure 10J-39, was discovered in 2000, as part of the Integral Program for the Conservation of Copán Archaeological Park (Programa Integral de Conservación del Parque Arqueológico Copán, or its acronym in Spanish, PICPAC). This archaeological project was conducted by the Honduran Institute of Anthropology and History (IHAH), under the direction of Seiichi Nakamura with the assistance of Melvin Fuentes. There were human remains belonging to two distinct individuals; neither individuals were accompanied by any artefacts as offerings. Burial 7(1) was identified as a primary *in situ* burial during field work, whereas Burial 7(2) was subsequently recognized in a laboratory setting due to the presence of identical bone elements. The remains of Burial 7(1), designated as CpM0203, were well-preserved, comprising an almost complete adult skeleton. In contrast, the remains of Burial 7(2) (CpM0506) were fragmented. As evident in the photograph taken by Seiichi Nakamura (Fig. 2.4), only a portion of the skull from CpM0506 was found resting beside the waist of CpM0203. This burial arrangement suggests that CpM0506 was a secondary burial accompanying CpM0203. While this observation led us to hypothesise a potential kinship between these two individuals, our analysis demonstrates that there is no close genetic relationship between them, at least up to the second-degree level.



Figure 2.4. Burial 7-2000. The photo was taken by Seiichi Nakamura.

CpM11

Burial 27-2000 was uncovered during the PICPAC excavation in 2000. The remains of an individual, CpM11, were discovered *in situ* outside the southern wall of Structure 10J-71, at the base brick-line level (Fig. 2.5). While the remains are generally well-preserved, they were fragmented but identified as an adult. Due to the presence of a jade bead left as an offering, this individual was categorised as level 1 in terms of wealth designation.



Figure 2.5. Burial 27-2000. The photo was taken by Seiichi Nakamura.

CpM12

Burial 35-2000 was excavated as part of the PICPAC in 2000, exhibiting the well-preserved remains of an adult. The human remains, designated as CpM12, were discovered *in situ* directly beneath the first brick-line of the west-facing wall (front) of the foundation of 10J-45-2nd, approximately 20 to 40 cm below the surface, without any formal delineation for burial. Structure 10J-45-2nd was constructed as an adoration shrine following the presumed burial of a governor (CpM13 from Burial 36-2000). The

positioning of CpM12, with the right arm placed in front of the chest and legs crossed as if seated (Fig. 2.6), in close proximity to the burial chamber of Burial 36-2000, suggests that this individual was the sacrificed companion of the presumed governor, CpM13. Additionally, one of the jade pendants found within the offering box (5-2001), unearthed near Burial 35-2000, depicts a figure displaying a gesture of respect and subordination towards another individual, potentially the high-status individual 36-2000 (CpM13) (Fig. 2.7). This gesture is reminiscent of the burial posture observed in CpM12. Contrary to early studies that concluded CpM12 was female based on morphological characters, our genomic analysis clearly identifies this individual as male.



Figure 2.6. Burial 35-2000. The photo was taken by Seiichi Nakamura.

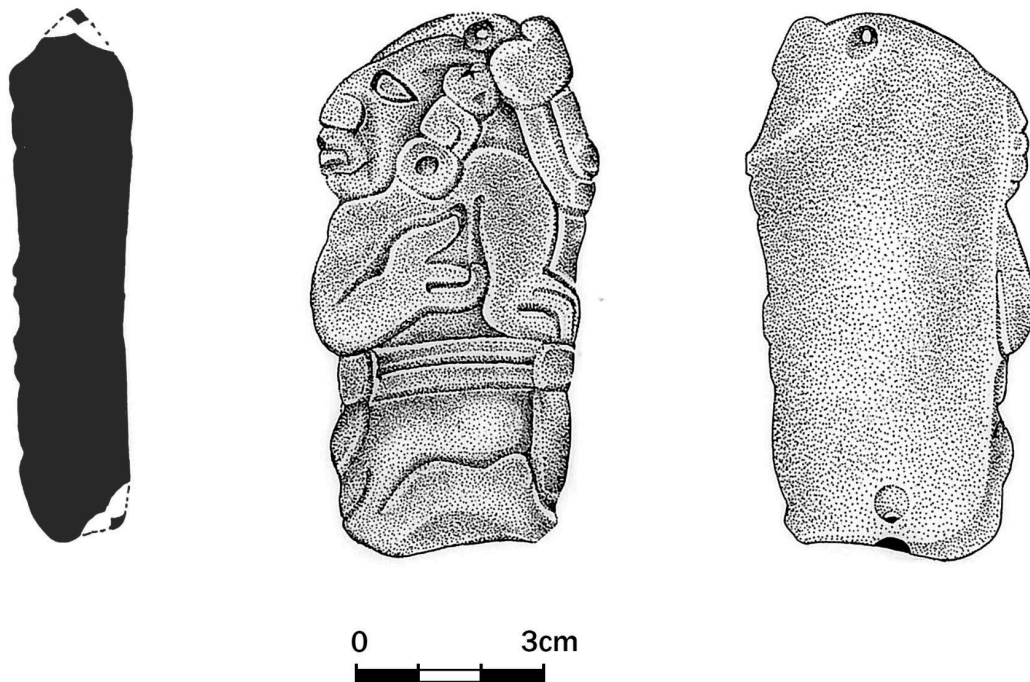


Figure 2.7. Jade pendants found within the offering box (5-2001), unearthed near Burial 35-2000. The picture was made by PICPAC.

CpM13

Burial 36-2000 represents one of the largest and most significant burials uncovered outside the Acropolis in Copán's 130-year history of archaeological investigation (Fig. 2.8). The excavation, conducted as part of the PICPAC project, was led by Seiichi Nakamura, with the assistance of Melvin Fuentes, in September 2000. The exceptional aspect of Burial 36-2000 lies in its contradiction of a commonly accepted hypothesis that the ruling class of Copán were exclusively buried within the Acropolis, the political, economic, and ritual hub of ancient Copán society.



Figure 2.8. Burial 36-2000 and the grave goods and interior of the chamber tomb. The photo was taken by Seiichi Nakamura.

The body remains of CpM13 are in varying states of preservation, ranging from regular to poor; they consist of fragmented and incomplete remains of an adult. Initially presumed to be a secondary burial, further analysis uncovered this individual to be the primary burial *in situ*. Red pigments were discovered on the surface of the bones. The tomb exhibited the characteristic vault of a royal tomb and contained two enormous jade pectorals (Fig. 2.9), one measuring 20 cm in length and the other 24 cm, adorned with

engraved designs symbolising the political and military authority of the individual. The archaeological evidence strongly suggests that this individual was one of the dynastic rulers of the 5th century.

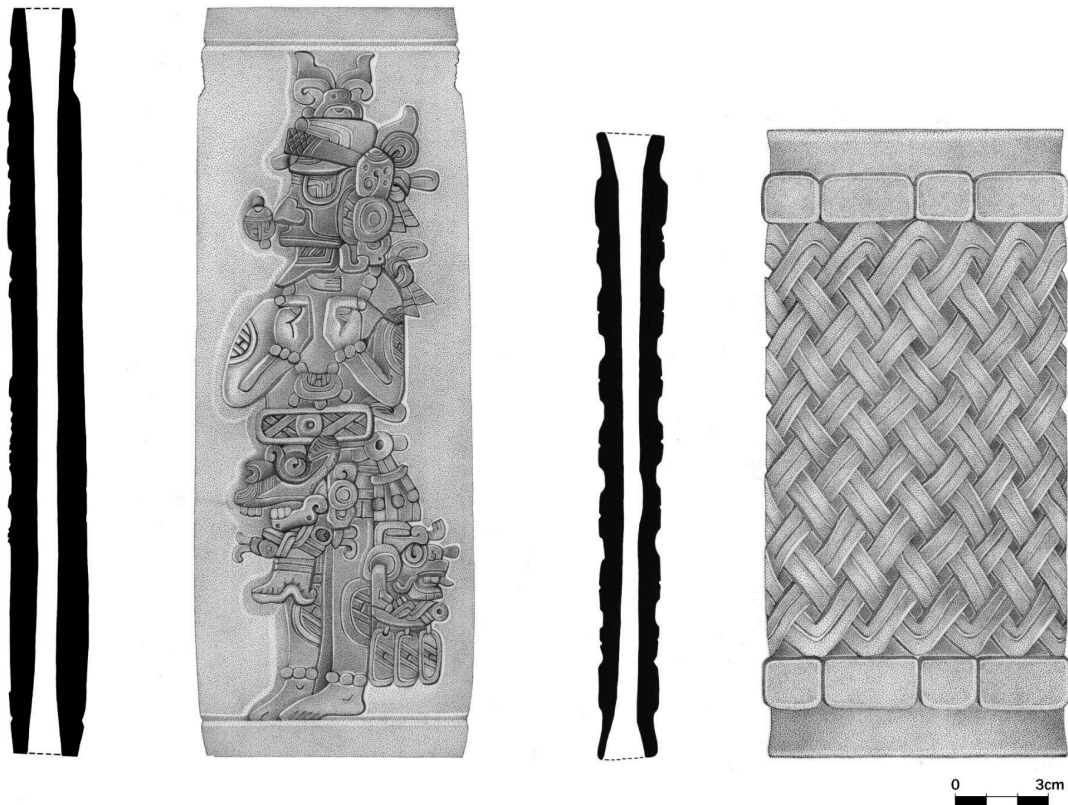


Figure 2.9. Jade pectorals from Burial 36-2000, with one measuring 24 cm in length and the other 20 cm, adorned with engraved designs. The picture was made by PICPAC.

CpM34 and CpM4243

Burial 94 and Burial 157, located in Structure 9L-104 and Structure 9L-105 respectively, were excavated in 2007 and 2016 as part of the new Copán Archaeological Project (known by its acronym in Spanish: PROARCO, Proyecto Arqueológico Copán). PROARCO was conducted by the Honduran Institute of Anthropology and History (IHAH), under the direction of Seiichi Nakamura with the assistance of Melvin Fuentes. These two structures were found in group 9L-23, also known as the Núñez-Chinchilla group, located 150 metres north of the Great Plaza.

Burial 94 contained the fragmented and incomplete remains of a girl (CpM34; Fig. 2.10), in a poorly preserved state. Archaeological evidence suggests that she may have been a sacrificed companion of another burial found in a box within the same structure, 9L-104. In contrast, Burial 157 consisted of the fragmented but well-conserved cranium of CpM4243 (Fig. 2.10), accompanied by a ceramic offering. However, there is a possibility that this offering might have belonged to a nearby burial registered as Burial 159, one of the burial groups found in Structure 9L-105. Therefore, CpM4243 may have been one of the companions of the main burial within Structure 9L-105. According to established ceramic chronology and C14 dating results, CpM4243 is contemporaneous with Burial 36-2000, despite the considerable distance between the two structures, approximately 1 km apart.



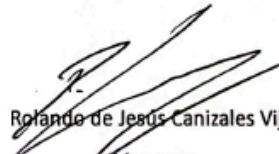
Figure 2.10. Burial 94 (left) and Burial 157 (right). The photos were taken by Seiichi Nakamura.

Ethics statement

All archaeological human remains included in this study were acquired under permission from the Honduran Institute of Anthropology and History, with authorization granted by the General Manager, Dr. Rolando Canizales (see [Fig. 2.11](#)). The findings of this study will be shared with the authorities at the Honduran Institute of Anthropology and History. Furthermore, the results will be disseminated among local communities through the institution.

Certificate

The undersigning Rolando de Jesús Canizales Vigil, Ph. D. Manager of the Honduran Institute of Anthropology and History (IHAH), CERTIFIES THAT the IHAH has authorized the scientific team lead by Prof. Seiichi Nakamura, Director of the Research Center for Next Generation Archaeological Studies at Komatsu University, Japan, to carry out paleogenomic studies in the laboratory facilities of Kanazawa University, Japan, using materials excavated from the burials at the Mayan site of Copán by PICPAC and PROARCO projects. Likewise, hereby states that permission has been granted to such team to present and publish the results of the above mentioned research in scientific journals.


Rolando de Jesús Canizales Vigil
Manager

Honduran Institute of Anthropology and History (IHAH)

Tegucigalpa, Honduras

February 29th, 2024


Villa Roy, Barrio Buenos Aires,
Tegucigalpa, M.D.C.
asistentegerencia.ihah@gmail.com
Tel.: (504) 2222-0079

Constancia

El infrascrito Dr. Rolando de Jesús Canizales Vijil, Gerente del Instituto Hondureño de Antropología e Historia (IHAIH), HACE CONSTAR QUE el IHAIH ha autorizado al equipo científico del Prof. Seiichi Nakamura, Director del Centro de Investigación Estudios Arqueológicos para Siguierte Generación de la Universidad de Komatsu, Japón, para que realice estudios paleogenómicos en el laboratorio de la Universidad de Kanazawa, Japón, utilizando materiales de los entierros del sitio Maya de Copán excavados por los proyectos PICPAC y PROARCO. Asimismo, hace constar que se le ha concedido el permiso para presentar y publicar los resultados de dichas investigaciones en revistas científicas.


Dr. Rolando de Jesús Canizales Vijil
Gerente
Instituto Hondureño de Antropología e Historia



Tegucigalpa, Honduras, 29 de febrero de 2024



Villa Roy, Barrio Buenos Aires,
Tegucigalpa, M.D.C.
asistentegerencia.ihah@gmail.com
Tel.: (504) 2222-0079

Figure 2.11. Letters confirming approval for the use of archaeological samples in DNA analysis. The original letter written in Spanish and its English translation are provided.

Chapter 3

Ancient genomes reveal demographic trajectories during the Classic Maya period

Abstract

The genetic history of the ancient Maya remains relatively unexplored, particularly in the Maya highlands. This chapter presents seven newly sequenced genomes from the city-state of Copan, dating back to the Classic Maya period, including a dynastic ruler and his sacrificial companion. Population structure analyses reveal strong genetic continuity in the Maya region from the Late Archaic period (~5,600 BP) to the present. However, this continuity was accompanied by genetic influx from highland Mexican populations. Time-stamped genomes indicate that this ancestry appeared in the Maya region as early as the Early-to-Middle Classic period, suggesting interactions between Maya and non-Maya populations that may have shaped the region's genetic landscape.

Contributions:

This work is truly collaborative and interdisciplinary, and I here clarify contributions from our team:

- **Sample Provision:** Rolando Canizales (Honduran Institute of Anthropology and History or HIAH) and Omar Talavera (HIAH) supplied the bone samples used in the research.
- **Archaeology and Anthropology:** Seiichi Nakamura (Komatsu University), Melvin Fuentes (HIAH), and Masahiro Ogawa (Komatsu University) were responsible for excavating the sites and gathering archaeological and anthropological information.
- **Excavation Support:** Salvador Varela (HIAH), Zoila Madrid (HIAH), and Carlos Carbajal (HIAH) provided logistical and practical support during the excavation process.
- **Radiocarbon Dating:** Minoru Yoneda (The University of Tokyo) performed the radiocarbon dating necessary to establish the age of the samples.

- **Sample processing, DNA extraction, and Library preparation:** Takashi Gakuhari conducted laboratory analysis of ancient DNA from the samples.
- **Genomic Data Processing and Analysis:** I processed and analyzed the genetic data, with the support from Lara M. Cassidy (Trinity College Dublin) and Shigeki Nakagome (Trinity College Dublin).

Introduction

The ancient city of Copán, located in the Maya highlands of western Honduras, is renowned for its well-preserved sculptures and the hieroglyphic stairway, reflecting its rich cultural and historical significance (Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015). By 1,000 BCE, the inhabitants already practised agriculture and constructed elevated stone platforms for their living quarters and for burying their dead underneath it. By 300-400 CE, the small settlement had become a vibrant city with economic opportunities for a growing population (Andrews and Fash, 2005).

Origins of royal families of the Classic Copán

Due to its geographical location at the southeastern edge of the Maya area (Fig. 3.1), it has been a subject of debate whether or not the inhabitants of Copán were ethnically Maya (Andrews and Fash, 2005; Suzuki et al., 2020). Early studies on mitochondrial⁵ haplogroups from Classic Copán populations identified frequencies differing from those observed in both present-day Maya and in ancient Maya from other city-states (Merriwether et al., 1997; Roca-Rada et al., 2020). Maya architecture, pottery and hieroglyphs appeared in Copán only after 400 CE, whereas Preclassic and Protoclassic pottery at Copán share closer affinities with those from central Honduras and El Salvador. This evidence suggests that the early inhabitants of the Copán Valley are unlikely to have been ethnically Maya. However, by 426/427 CE, Copán came under the rule of a Maya elite who spoke a Maya language (Andrews and Fash, 2005).

It has been posited that members of a widespread Maya elite network, potentially originating from the central Petén region of Guatemala—home to the city-state of Tikal—migrated to Copán. These elites may have intermarried with local residents, establishing themselves in the region. Hieroglyphic inscriptions at Copán identify its first dynastic founder, *K'inich Yax K'uk' Mo'*, as an outsider who assumed power in 426/427 CE. Although his name is distinctly Mayan, meaning “Great Sun Green Quetzal Macaw,” and early depictions present him in traditional Maya regalia, later portraits emphasise his foreign, Teotihuacán-style elements (Andrews and Fash,

⁵ Mitochondrial DNA is present in the cell mitochondria rather than in the nucleus. It is passed down from mother to child. It is present in multiple copies, facilitating its sequencing.

2005; Ferguson, 2007). On Altar Q, which features the 16 rulers of Copán, *K'inich Yax K'uk' Mo'* is the only figure explicitly depicted as a Teotihuacán warrior (Bell et al., 2004).

Archaeological and epigraphic evidence indicate that Teotihuacán, a large city-state in the valley of Mexico, established trade with Tikal in the 3rd century CE, using Tikal as a hub to extend political influence across Petén and beyond, reaching distant centres such as Copán, Quiriguá and Pusilhá. Prior to its collapse in 600 CE, Teotihuacán reached its maximum splendour much earlier, between 350-450 CE, coinciding with the arrival of the dynastic founder in Copán. Altar Q inscriptions state that, before acceding to the throne, *K'inich Yax K'uk' Mo'* undertook a 156-day journey to a distant region to receive symbols of rulerships. Whether this distant region was Teotihuacán, Tikal, or elsewhere remains uncertain. However, he brought two pyrite mirrors—similar to those excavated at Tikal—as gifts for his wife (Martínez de Velasco Cortina and Vega Villalobos, 2015).

Isotopic evidence from tooth enamel further supports the foreign origin of *K'inich Yax K'uk' Mo'*. Oxygen isotopic ratios, which reflect the rainfall of a person's childhood environment, and strontium isotopic ratios, which vary with geological substrates and enter the body through food and water, both indicate a non-local origin for *K'inich Yax K'uk' Mo'* and local origin for a royal woman presumed to be his wife (Price et al., 2010). Further isotopic analysis of Classic Copán populations reveals that nearly 50% of individuals—especially elites—were non-local, suggesting that many elites buried at Copán grew up elsewhere before relocating to the city (Suzuki et al., 2020).

Overall, it is unlikely that *K'inich Yax K'uk' Mo'* originated from Teotihuacán. Instead, his adoption of Teotihuacán-style imagery may have been a strategic effort to legitimise his authority and that of his successors. His exact birthplace remains uncertain, as isotopic data require regional comparisons for precision. Furthermore, the extent of social and cultural assimilation among these non-local elites, along with their potential genetic contributions, warrants further investigation.

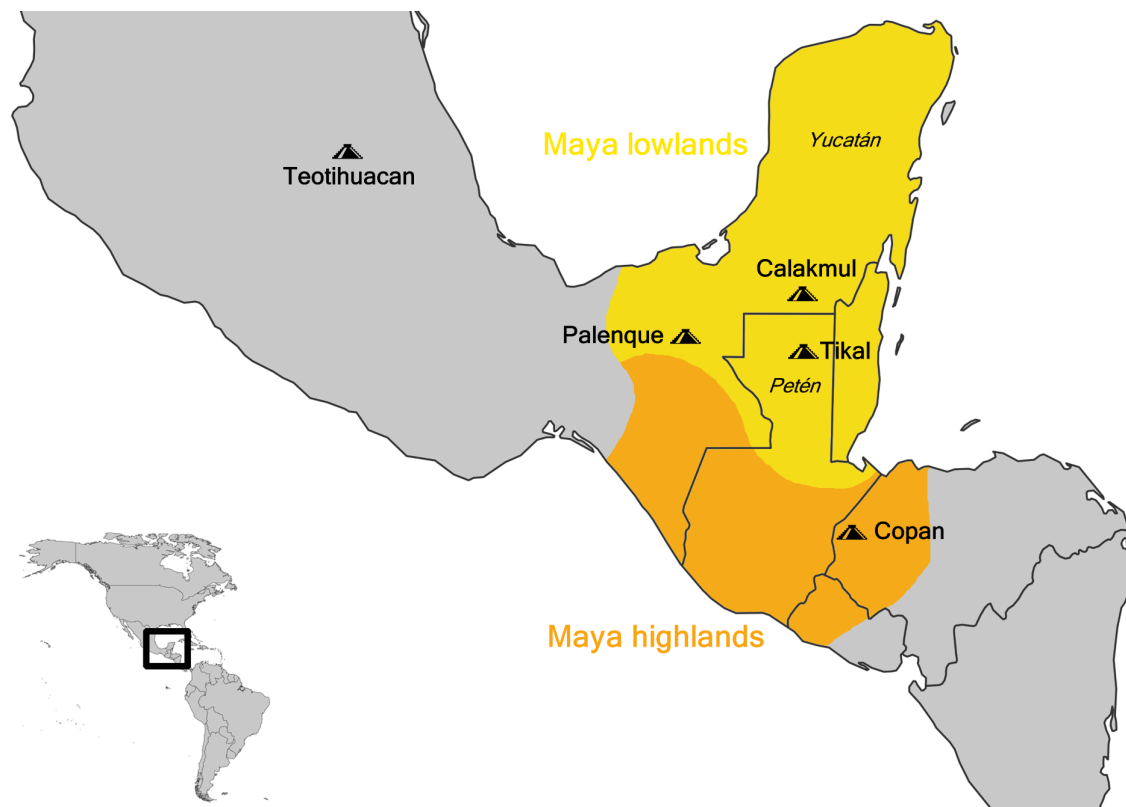


Figure 3.1. Major urban centres in Mesoamerica during the Classic period. The Maya area is highlighted in yellow and orange for the Maya lowlands and highlands respectively. City locations are pinpointed by a pyramid, and regions are in *italic*.

Archaeogenetics in the Maya region

Previous analyses of ancient genomes in the Maya region have uncovered the presence of a distinct genetic ancestry dating back to the Early Holocene, approximately 9,300 to 7,300 years before present (BP) (Posth et al., 2018), which constitutes a lineage likely diverged from the initial southward migrations through the Americas. By the Late Archaic period, around 5,600 to 3,700 years BP, the region saw gene flow from further south (Kennett et al., 2022). While a later influx of highland Mexican ancestry—represented by indigenous groups from Oaxaca, Mixe and Zapotec—has influenced the genetic makeup of present-day Maya populations (Kennett et al., 2022), the timing of its arrival remains elusive. Recent analysis of the Terminal Classic genomes from Chichén Itzá in Yucatán provides evidence of genetic continuity into a contemporary population from the same region (Barquera et al., 2024).

Aims of Chapter 3

This study presents seven newly sequenced ancient genomes from Copán, all dating to the Classic Maya period ([Table 3.1](#)). This dataset includes the genome of an enigmatic dynastic ruler buried outside the Acropolis (labelled as CpM13) and an individual discovered nearby, likely linked to a sacrificial burial (labelled as CpM12), and five commoners (Price et al., 2014). By integrating these Copán genomes with a broader ancient and modern genomic dataset spanning North-to-South Americas and Siberia, my work aims to enhance our understanding of the genetic ancestry of the Classic Copán, as well as other ancient and present-day populations from the Maya region.

Results

This section is organised into six subsections, each presenting results from distinct analyses as follows:

- I.** Describing the newly sequenced genomes from the Classic Copán, assessing their authenticity, and exploiting isotopic data.
- II.** Contextualising paternal and maternal haplogroups from the Classic Copán into variation in their compositions across ancient Maya populations.
- III.** Establishing and validating a masking strategy to exclude non-American genomic segments in present-day American individuals.
- IV.** Examining genetic structures of ancient and present-day American populations.
- V.** Exploring genetic affinities between ancient and present-day Maya populations, as well as between the Maya and other American populations.
- VI.** Modeling admixture scenarios in the populations from the Maya region.

I. Ancient genomes from the Classic Copán

I. 1. Genomic data generation of Classic Copán individuals

The initial screening focused on 16 ancient skeletal remains excavated from the sites of Copán, as part of two projects (PICPAC and PROARCO), both led by our archaeological collaborator, Seiichi Nakamura. Based on mapped reads after filtering and duplicate removal, endogenous⁶ human DNA contents above 1% were achieved in ten of these bone extracts (Table S3.1), which were then further shotgun-sequenced to higher coverage. Out of these ten extracts, three pairs were identified as genetically identical (CpM02-CpM03, CpM05-CpM06 and CpM42-CpM43) by the kinship program READ (Relationship Estimation from Ancient DNA) (Fig. S3.1) (Monroy Kuhn et al., 2018). Given that these pairs were found in the same burial tomb, they were subsequently merged into single genomes: CpM0203, CpM0506 and CpM4243 respectively. Genomic coverage⁷ for the seven final individuals range between 0.033× and 0.491× (Table 3.1).

Table 3.1. Summary of Classic Copán genomes and their associated profiles

Individual (Project)	Top: Structure Middle: Burial Bottom: Dating (median; 2σ)	Endo%; clonality	Coverage: whole (X; Y)	Molecular sex	Contamination; mtDNA coverage	Haplogroup		Wealth designation (0-5)	Isotopic data (⁸⁷ Sr/ ⁸⁶ Sr; δ ¹³ C; δ ¹⁸ O)
						mtDNA	Y-chr.		
CpM0203 (PICPAC)	10J-39 7(1)-2000 716; 657-774 CE	15.91%; 6.54%	0.4486 (0.2823; 0.0005)	Female	0.00%; 21.563	A2ad	-	0	0.7117; -4.96; -3.65
CpM0506 (PICPAC)	10J-39 7(2)-2000 717; 659-775 CE	19.31%; 7.35%	0.4908 (0.2987; 0.0005)	Female	0.48%; 13.284	A2+(64) +16129	-	0	-
CpM11 (PICPAC)	10J-71 27-2000 721; 667-774 CE	1.81%; 0.74%	0.0051 (0.002; 0.0004)	Male	0.00%; 0.125	NA	NA	1	-
CpM12 (PICPAC)	10J-45 35-2000 ND	30.8%; 11.25%	0.3076 (0.1203; 0.0236)	Male	0.00%; 13.900	C1c4	Q1b	0 (Sacrificed burial)	0.7065; -0.89; -4.39
CpM13 (PICPAC)	10J-45 36-2000 484; 425-542 CE	9.55%; 6.25%	0.0925 (0.0335; 0.0065)	Male	1.09%; 2.301	A2u1	Q1b	5 (Dynastic ruler)	0.7075; -3.81; -2.72
CpM34 (PROARCO)	9L-104 94 574; 543-605 CE	3.96%; 1.09%	0.0334 (0.029; 0.0001)	Female	0.00%; 2.047	A2	-	0	0.7064; -2.40; -3.50
CpM4243 (PROARCO)	9L-105 157 495; 431-558 CE	2.17%; 0.38%	0.0498 (0.0387; 0.0001)	Female	0.29%; 6.466	C1c	-	1	-

⁶ As ancient DNA is contaminated by DNA from the environment (ancient and present-day), the endogenous DNA refers to the human DNA belonging to the bone remains.

⁷ Genomic coverage is the mean number of sequencing reads for each genomic position. The higher the coverage, and the more certain is the DNA base read by the sequencing machine.

It is important to note that our team also attempted an enrichment approach using the myBaits Expert Human Affinities Kit (Daicel Arbor Biosciences) to recover DNA from the Classic Copán individuals. However, the enrichment was unsuccessful because the libraries failed to yield sufficient coverage at the targeted 1240k SNP positions, even for individuals with relatively high endogenous content such as CpM0506 (CpM05: 15.3%; CpM06: 16.8%) and CpM12 (30.76%). This outcome was likely due to the extreme fragmentation of their DNA with median fragment lengths ranging from 37 to 42 bp (Fig. 3.2). Given this limitation, we focused on generating as much genomic data as possible through whole-genome shotgun sequencing. This approach, coupled with UDG treatment of the DNA, is widely recognized as the best practice for maximising data yield and quality from highly degraded ancient DNA. By employing this strategy, we ensured that our methodology was robust and appropriate for the nature of the samples.

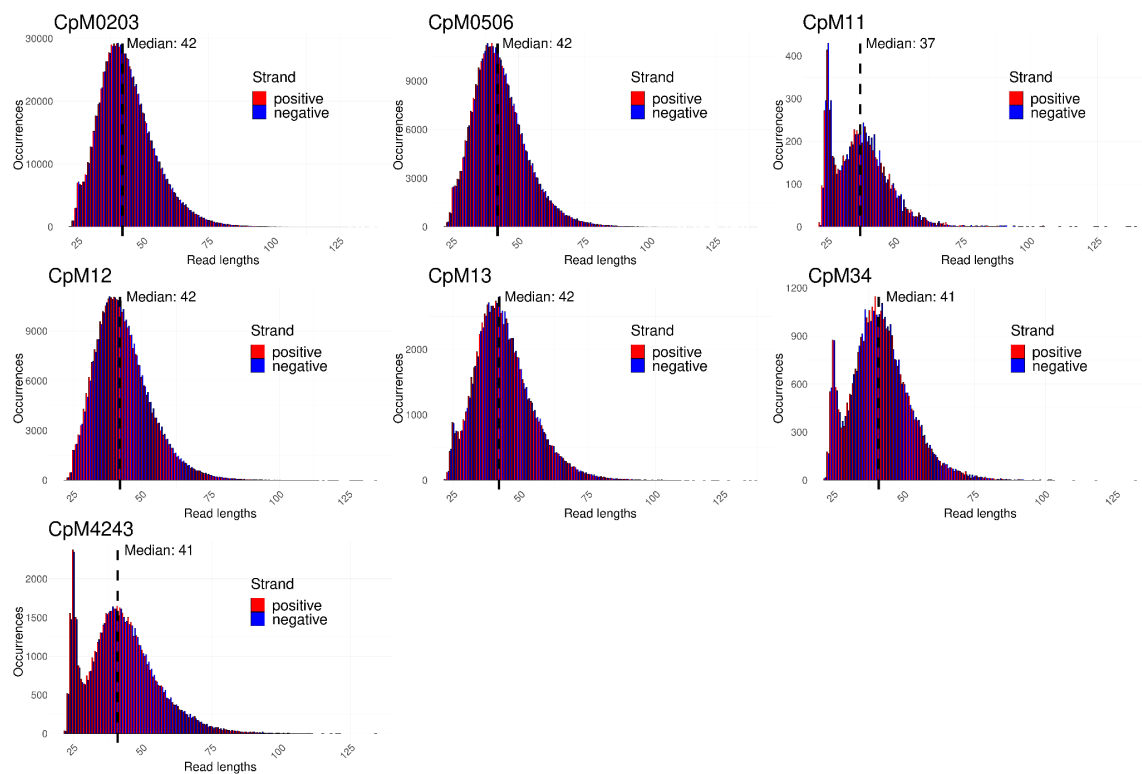


Figure 3.2. Fragment length distributions for newly sequenced Classic Copán genomes. Read length distributions are shown for the seven Classic Copán individuals, with forward strands in red and reverse strands in blue. The black dashed line indicates the median fragment length for each individual.

Radiocarbon dating conducted on collagen extracted from the corresponding bones reveals that all individuals align with the Classic Maya period, yet they are split into two temporal groups: either the Early-to-Middle (circa 250-550 CE) or the Middle-to-Late Classic (circa 550-830 CE) Maya phases.

I. 2. Authentication, kinships, and molecular sexing

To remove any first- or second-degree relatives, READ was re-ran on the seven Classic Copán individuals (Fig. S3.2). The kinship analysis indicated potential second-degree relationships between CpM11 and CpM34 or between CpM11 and CpM4243. However, due to large error bars, these kinship estimates may be unreliable. Consequently, CpM11 was excluded from all downstream analyses.

Deamination profiles at the terminal ends of DNA fragments were analysed to confirm the ancient origin of sequence reads generated from the DNA extracts of the Copán individuals. Ancient DNA accumulates damage patterns over time, with deamination being the most prevalent. This damage is observed as C-to-T or G-to-A misincorporation in ancient genomes (Briggs et al., 2007; Do and Dobrovic, 2012). All of the Copán individuals show clear evidence of postmortem damages (Fig. 3.3), along with short DNA fragment lengths (Fig. 3.2), providing strong support for their authenticity as ancient DNA.

These low levels of contamination were further supported by the X-chromosome data from males, which provides an additional method to assess and control for non-endogenous DNA contamination. Since males have only one X-chromosome (compared to two in females), a single allele is expected at each site, and any observed polymorphism likely results from contamination or sequencing errors (Rasmussen et al., 2011). Molecular sexing identified three males among the seven individuals analysed, including CpM12, a sacrificed victim previously classified as female based on their morphological characteristics (Mizoguchi and Nakamura, 2006; Price et al., 2014). Contamination estimates measured using Analysis of Next Generation Sequencing Data (ANGSD) are presented in Table 3.2. Both methods compare the number of minor reads to the number of major reads, however Method 1 assumes that sequencing errors are independent across all reads and sites, while Method 2 accounts for the fact that errors are often correlated, particularly in dataset filtered for coverage depth (Rasmussen et al.,

2011). While the results indicate low contamination levels, they are not considered statistically significant, likely due to limited genomic coverage.

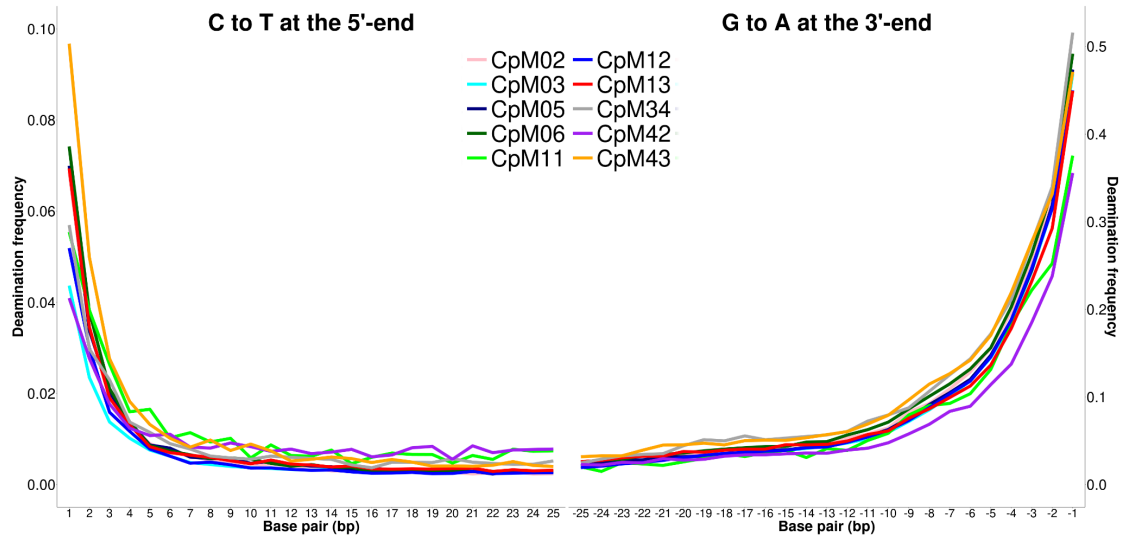


Figure 3.3. DNA damage plots for newly sequenced Classic Copán individuals. Damage patterns on the left plot show C>T misincorporations at the 5'-end, while those on the right plots show G>A misincorporations at the 3'-end. As sequencing libraries were constructed with the NEBNext Ultra II DNA Library Prep Kit that uses USER enzyme for exercising adaptors taking a hairpin loop structure, C-to-T substitutions observable at the 5'-end become less pronounced compared with G-to-A substitutions at the 3'-end.

Table 3.2. Contamination estimates in males based on X-chromosomes.

	CpM11		CpM12		CpM13	
	Method 1	Method 2	Method 1	Method 2	Method 1	Method 2
<i>p</i> -value	0.6101	1	0.6101	0.3758	0.6101	1
Contamination	0.009%	0.007%	0.009%	0.225%	0.009%	0.007%
log-likelihood	7.55	0.27	7.55	7.55	7.55	0.64
SE	0.0006	0	0.0006	0.0024	0.0006	0

Similar to quantifying X-chromosome mismatches in males, mitochondrial DNA can be used to estimate contamination in both males and females (Furtwängler et al., 2018). Unlike the Y-chromosome, which is transmitted only from father to son,

mitochondrial DNA is inherited from the mother by all offspring. Furthermore, mitochondrial DNA is present in greater quantities than nuclear DNA within a cell, resulting in higher sequencing coverage and facilitating its analysis (Merriwether et al., 1997). Contamination based on mtDNA is detected by measuring levels of heterozygosity at positions where endogenous ancient DNA is expected to be homogenous. Deviations from the expected homogeneity—such as the presence of multiple alleles at a single site—indicate contamination from exogenous sources, such as present-day human DNA (Furtwängler et al., 2018). In the Copán individuals, mitochondrial coverage ranged from 0.125× for CpM11 to 21.563× for CpM0203 (Table 3.1). Contamination levels ranged from 0% to 1.09%, confirming minimal modern human contamination.

I. 3. Isotopic profiles

In previous studies, isotopic data were collected from tooth enamel for four out of the seven Copán individuals (unavailable for CpM0506, CpM11, and CpM4243) (Price et al., 2014; Suzuki et al., 2020), which are summarised in Table 3.1. Isotope ratios from archeological remains can provide valuable insights, including human mobility. Strontium, which varies based on geology and rock types, is incorporated into bones and teeth through the food chain. Since tooth enamel forms during early childhood and undergoes minimal change afterward, strontium isotope ratios ($^{87}\text{Sr}/^{86}\text{Sr}$) can reveal childhood geographic origins when compared to baseline data from surrounding areas (Price et al., 2010). Previous studies of Classic Maya individuals from Copán identified nearly 50% of the individuals analysed as non-locals—specifically 2 in 10 in Price et al. (2010), 14 in 32 in Price et al. (2014), and 28 in 66 in Suzuki et al. (2020)—particularly among elite groups. This includes the “stranger-king” *K'inich Yax K'uk' Mo'* (Price et al., 2014, 2010; Suzuki et al., 2020). Using a general baseline measurement from across the Maya region for comparison (Suzuki et al., 2020), $^{87}\text{Sr}/^{86}\text{Sr}$ ratios from the Copán individuals suggest a local origin for CpM12 and CpM34 (within the Copán baseline range of 0.7060-0.7072). In contrast, CpM0203 and CpM13 are likely non-locals, with a potential origin from the Maya mountains of southern Belize for CpM0203, and a possible southern Petén origin for CpM13.

Another indicator of the place of birth is the isotopic ratio of oxygen. Clouds forming over tropical and coastal regions are typically enriched with heavier oxygen isotopes (^{18}O), which precipitate first, leaving polar and inland rains lighter with higher proportions of ^{16}O . This isotopic ratio becomes sealed in tooth enamel during childhood through the consumption of local drinking water, providing a signal of the water sources at the place of birth (Price et al., 2014, 2010; Suzuki et al., 2020). The oxygen isotopic ratios ($\delta^{18}\text{O}$) of CpM0203, CpM12, and CpM34 fall within the range typical of locales near Copán, unlike that of CpM13, which differ significantly (Suzuki et al., 2020). When considered alongside, these data indicate a local Copán origin for CpM12 and CpM34. However, the birthplace of CpM0203 and CpM13 remains uncertain. Since baseline measurements remain sparse throughout Mesoamerica, it is indeed challenging to accurately identify specific birth origins solely from isotopic data (Price et al., 2014). More importantly, isotopic analysis cannot identify non-local individuals beyond first generation immigrants. Genetic data is required to explore long-term interactions between local and non-local populations, for a better understanding of subsequent social and cultural change.

Isotopic ratios in tooth enamel also reflect dietary practices. Carbon isotopes ($\delta^{13}\text{C}$) are influenced by the types of plants consumed, while nitrogen isotopes ($\delta^{15}\text{N}$) distinguish between terrestrial and marine food sources (Kennett et al., 2022; Somerville et al., 2013). By comparing isotopic data from Copán individuals with those of Preclassic and Classic Maya individuals (Kennett et al., 2022; Somerville et al., 2013), it is evident that the Copán population aligns with other Classic Maya in their heavy reliance on maize as a primary dietary source (Fig. 3.4).

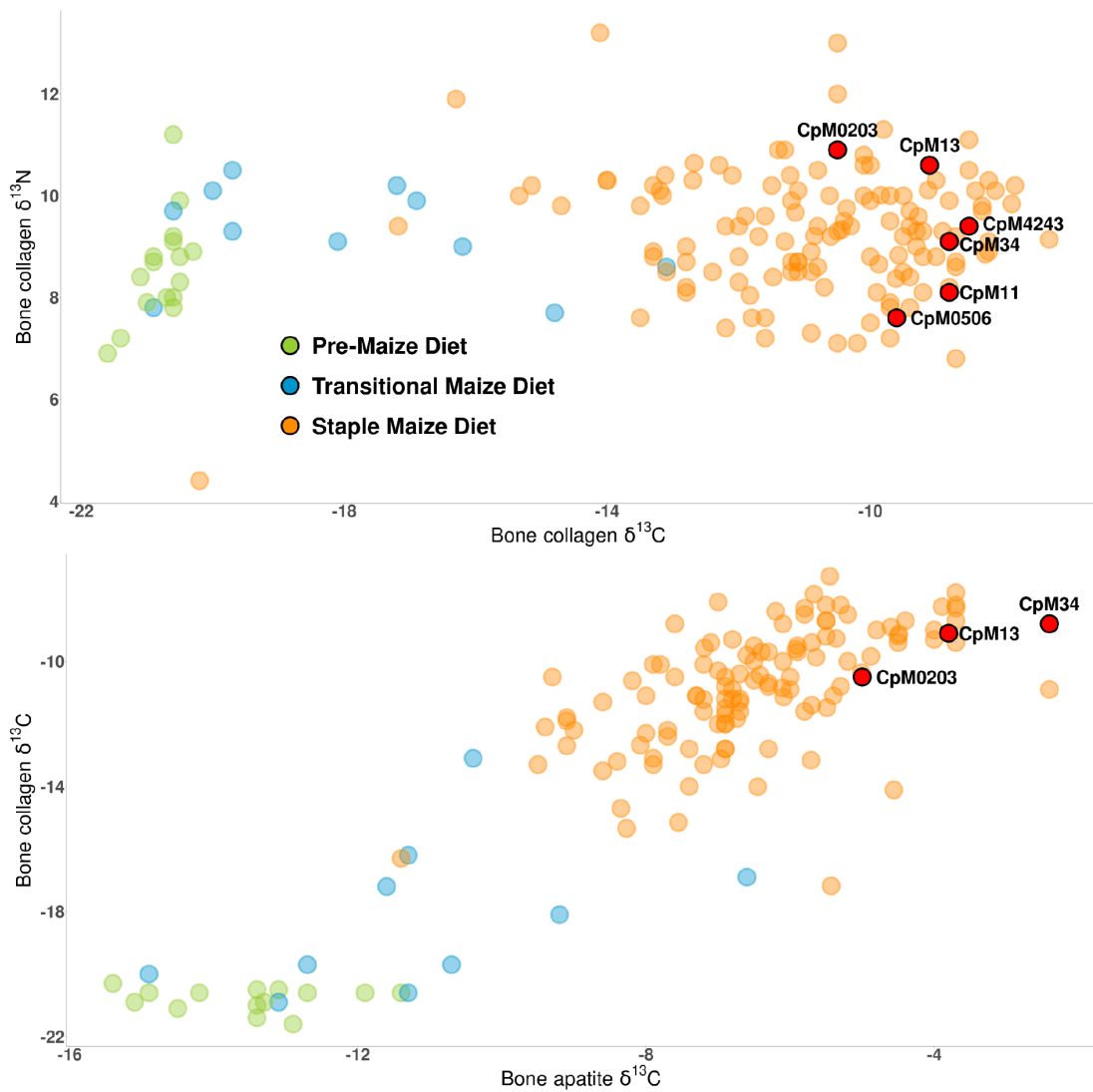


Figure 3.4. Isotopic profiles of the Classic Copán individuals (red) and ancient Maya individuals with pre-maize (green), transitional maize (blue), and staple maize diets (orange). The top and bottom plots show values of $\delta^{13}\text{C}_{\text{collagen}}$ (x-axis) and $\delta^{13}\text{N}_{\text{collagen}}$ (y-axis) or those of $\delta^{13}\text{C}_{\text{apatite}}$ (x-axis) and $\delta^{13}\text{C}_{\text{collagen}}$ (y-axis), respectively. Due to a lack of isotopic data, CpM12 or CpM0506, CpM11, and CpM4243 are not included in the top or bottom plot. All data, except for the Classic Copán individuals, are referred from previous studies (Kennett et al., 2022; Somerville et al., 2013).

II. Uniparental markers of the Classic Copán

II. 1. Identifying paternal markers from the Classic Copán

Two uniparental markers are widely studied in genetics: mitochondrial DNA (mtDNA) and the Y-chromosome. Mitochondrial DNA is maternally inherited and present in all individuals, while the Y-chromosome is paternally inherited and present exclusively in males. These distinctions in inheritance allow geneticists to trace maternal and paternal lineages separately. By analysing these markers, it is possible to reconstruct the origins and divergences of human ancestry.

Given a high degree of genetic similarity in maternally and paternally inherited markers among Indigenous Americans (*e.g.*, the Y-chromosome haplogroup Q shared by nearly all American males) (Grugni et al., 2019; Llamas et al., 2016), they possibly descended from a single migratory subset of a Eurasian gene pool that entered the Americas via Beringia between 20,000 and 15,000 years ago.

Out of the seven ancient individuals sequenced from Copán, three were identified as males: CpM11, CpM12 and CpM13 (Table 3.1). Genomic coverage varied among these individuals, affecting the resolution of Y-chromosome haplogroup assignments. For CpM11, low genomic coverage prevented the detection of Y-chromosome markers, making it impossible to identify its paternal haplogroup..

CpM13, sequenced at a nuclear genome coverage of 0.0925×, carried a limited number of Y-chromosome markers (Fig. 3.5). Specifically, 17 mutations aligned with the BT branch and 6 with the CT branch on the phylogenetic tree. However, these markers are located close to the root of the tree and provide limited insight. A sub-group within the D branch was detected, but this classification was excluded due to conflicts with four other markers diverging from the CT branch and not associated with the DE branch. These conflicting markers included one mutation from the GHIJK branch, which diverged from the F branch (supported by two markers), and one from the CF branch. Additionally, two informative markers, Z12681 and L57, were identified as belonging to sub-groups H3a2a2 and Q1b, respectively. Given the limited number of markers, my analysis classified the Y-chromosome haplogroup of CpM13 as either H3a2a2 or Q1b, with low certainty.

CpM12, sequenced at a higher coverage of at 0.3076×, provided more Y-chromosome markers (Fig. 3.5). Similar to CpM13, the most abundant markers aligned with the BT (39 hits) and CT (22 hits) branches. A subgroup E classification was detected, but this was excluded due to inconsistencies with other markers that did not belong to the DE branch. There were six markers aligned with the F branch, which diverges from the DE branch. Among the remaining 16 markers, the majority were associated with the Q branch: two markers for Q, four for Q1, and two for Q1b. These mutations—M1074, F1728, F1304, F1426, CTS2006, CTS5804, L57 and F2122—support the assignment of CpM12 to the Q1b branch.

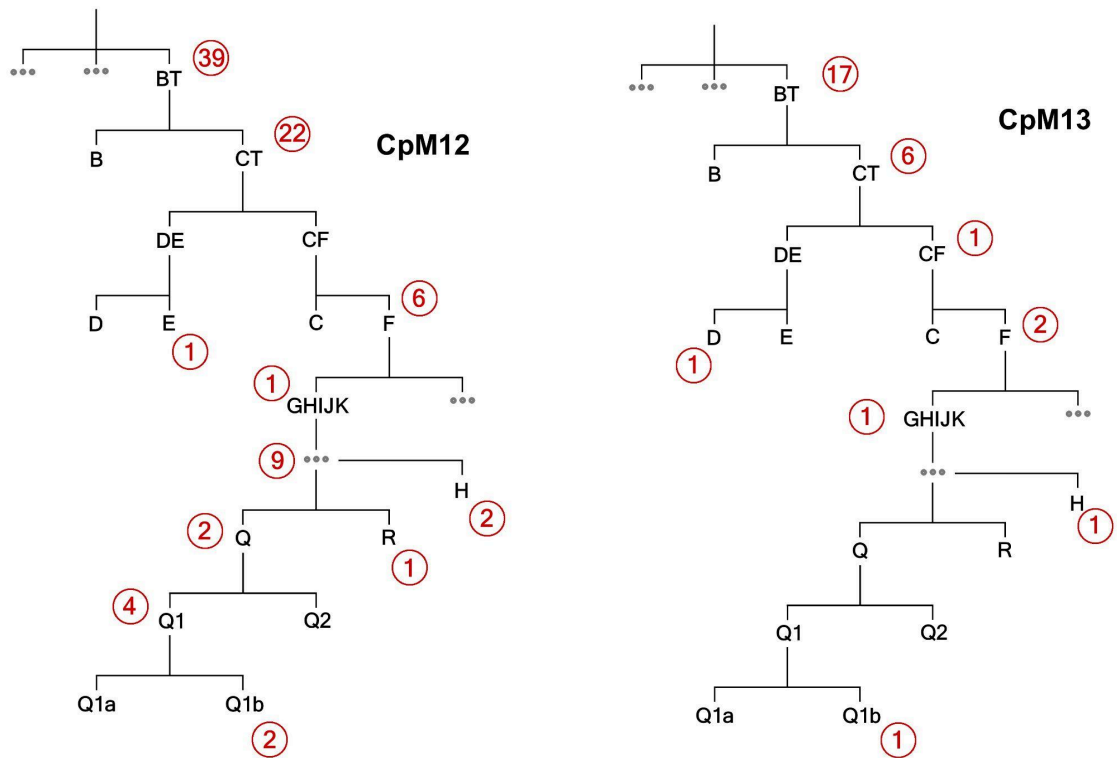


Figure 3.5. Y-chromosome hits for individuals CpM12 and CpM13. Number of hits are in red circles for individuals CpM12 (left) and CpM13 (right), overlaid on a simplified version of the Y-chromosome tree by collapsing branches which are symbolised by three grey dots. The Y-chromosome tree is based on <https://www.yfull.com/tree/>.

II. 2. Interpreting paternal haplogroups from the Classic Copán

When counting the number of hits alone, CpM13 is equally likely to be classified as H3a2a2 or Q1b, with one hit for each sub-group. However, considering that most of published ancient American males belong to the Q1 branch, it is more probable

that CpM13 also belongs to the Q1 branch rather than the H branch (Campelo dos Santos et al., 2022; Capodiferro et al., 2021; de la Fuente et al., 2018; Kennett et al., 2022; Nägele et al., 2020; Nakatsuka et al., 2023, 2020a, 2020b; Popović et al., 2021; Posth et al., 2018; Scheib et al., 2018; Severson et al., 2022; Villa-Islas et al., 2023).

CpM12 is confidently classified in Q1b. Considering that the Maya practiced patrilineal succession, with power typically passed from father to son (Martínez de Velasco Cortina and Vega Villalobos, 2015; Verdugo et al., 2020), the paternal haplogroup of CpM13 offers potential insights into the lineage of the royal family of Copán. The discovery of CpM12—a sacrificial individual—within CpM13’s royal burial tomb raises intriguing questions about their relationship. Could CpM12 have been a war captive from a lower social class, or was an elite individual, possibly a war captive from a rival city-state, sacrificed to honour the deceased king?

Sacrifice played a central role in Maya civilisation. Royal rulers, regarded as semi-divine mediators between the natural and the supernatural worlds (Martínez de Velasco Cortina and Vega Villalobos, 2015), were believed to possess blood imbued with special power. This belief may have led to the preference for elite individuals during sacrificial rituals. However, due to the limited sample size and insufficient Y-chromosome reads, we are unable to definitively address these questions with the current data.

II. 3. Identifying maternal haplogroups from the Classic Copán

Among Indigenous Americans, six mtDNA haplogroups have been identified: A2, B2, C1, D1, X2a and more recently identified and rarely found D4h3 (Bandelt et al., 2003; González-Martín et al., 2015; Kumar et al., 2011). Five haplogroups—A2, B2, C1, D1, D4h3—are of Asian origin. Their coalescence time and geographic distributions support the theory of a single migration from East Asia, followed by a period of isolation in Beringia, and subsequent expansion into the Americas. Although the evolutionary history of X2a is debated due to its West Eurasian origin, its presence in early Holocene individuals such as Kennewick Man (~8500 BP) suggest it may have arrived with the initial peopling of the Americas, pointing to a more complex migration history than previously assumed (Bandelt et al., 2003; Kumar et al., 2011; Rasmussen et al., 2015; Sandoval et al., 2009).

In the ancient Copán individuals, mtDNA analysis identified haplogroups A2 and C1 (Table 3.1), supporting their Indigenous American origin. CpM0203 belongs to haplogroup A2ad, while CpM13 is classified as A2u1, and without additional markers, CpM34 is assigned to A2. CpM0506 also harbours the markers for the A2 haplogroup but with additional mutations that prevent precise subgroup assignment. Its mtDNA may represent an undefined branch of A2. Similar undefined subgroups, such as “A2+(64),” have been identified in ancient individuals from the southern cone (Nakatsuka et al., 2020b), the Caribbeans (Fernandes et al., 2021; Nägele et al., 2020), New Brunswick in Canada (Raghavan et al., 2015), and Belize in Mesoamerica (Kennett et al., 2022). The subgroup “A2+(64)” may represent extinct lineages found only in ancient Americans.

CpM12 and CpM4243 fall within the C1 haplogroup, with CpM12 classified as C1c4 and CpM4243 as C1c (“PhyloTree,” 2016; van Oven and Kayser, 2009). Lastly, CpM11 could not be confidently typed due to low mitochondrial coverage (<0.1× compared to >2× for the other individuals).

II. 4. Comparing Classic Copán maternal haplogroups to present-day Mesoamericans

Consistent with studies on present-day Central American populations, the haplogroup X2a was not detected in any of the Classic Copán individuals, further supporting its restriction to North American populations (Gorostiza et al., 2012; Sandoval et al., 2009). In present-day Mesoamericans, the previously reported frequencies of haplogroups A2, B2, C1, D1 and D4h3a based on 520 individuals are as follows: 49.4%, 24.2%, 19.4%, 4.4% and 2.3% ($n=520$) (González-Martín et al., 2015). Among the Maya, who inhabited the southeastern region of Mesoamerica, haplogroup A2 is the most prevalent at 79.6%, followed by C1 (9.1%), B2 (6.8%) and D1 (4.5%) ($n=44$) (Gorostiza et al., 2012). The haplogroup frequencies observed in Classic Copán individuals align with those of present-day Maya, with A2 as the most common haplogroup, followed by C1, though the proportions differ. Among the Classic Copán individuals, four individuals (CpM0203, CpM0506, CpM13 and CpM34), which are corresponding to 67%, belong to haplogroup A2, compared to 79.6% in present-day Maya. The remaining two individuals (CpM12, and CpM4243; 33%) fall within haplogroup C1, compared to 9.1% in present-day Maya. These differences may reflect

the influence of small sample size in this study, but also potentially population heterogeneity among the Maya.

Heterogeneity within present-day Maya groups has been previously suggested by González-Martin and colleagues (González-Martín et al., 2015). The study found that F_{ST} values become significant only when the Maya were divided into individuals from the Yucatán Peninsula and those from other regions. Furthermore, the Maya from the Yucatán Peninsula exhibited the highest genetic diversity of all Mesoamerican groups. These results support a potential population structure within the Maya, which may result from differences in past trade, politics and migration routes between Classic Maya states.

II. 5. Contextualising the Classic Copán with ancient Maya

Ancient DNA can elucidate when and how genetic heterogeneity in the Maya region emerged. Given the abundance of mtDNA relative to nuclear DNA—approximately 1,500 times more copies per cell—mtDNA offers a significant advantage for such studies (Merriwether et al., 1997). This abundance made mtDNA the primary focus of ancient DNA research before the advent of next-generation sequencing technologies enabled more comprehensive analyses of autosomal DNA.

Mitochondrial DNA has been widely sequenced in the context of ancient Maya, which are summarised in [Figure 3.6](#). Ochoa-Lugo and colleagues sequenced the mitochondrial genomes of 38 individuals from various sites across the Maya region (Ochoa-Lugo et al., 2016). Gonzáles-Oliver *et al.* typed 24 individuals from Xcarat in Mexico; Verdugo *et al.* studied 17 individuals from Midnight Terror Cave in Belize; and Merriwether *et al.* examined 9 individuals from the Copán site in Honduras (González-Oliver et al., 2001; Merriwether et al., 1997; Verdugo et al., 2020).



Figure 3.6. Overview of mitochondrial haplogroups in ancient Maya from published studies. Arrows correspond to the geographic location of ancient Maya individuals (the Maya region is highlighted in a darker shade of green). Pie charts represent the frequencies of each mitochondrial group found. For visual clarity, mitochondrial groups belonging to A, B, C and D are coloured in orange, yellow, blue and green respectively. Original study, sample sizes and archeological dates are displayed next to corresponding pie charts.

All studies including the data from newly sequenced Classic Copán, except for those from Merriwether and colleagues, support A as the most common haplogroup found across ancient Maya cities, followed by haplogroup C (González-Olivier *et al.*, 2001; Ochoa-Lugo *et al.*, 2016; Verdugo *et al.*, 2020). Nine Copán individuals sequenced by Merriwether and colleagues show a markedly different haplogroup distribution, with eight individuals typed as C and one as D. Based on this distinct composition, the study suggested that Copán may have been genetically isolated from other Maya cities. In contrast, my analysis reveals a different haplogroup distribution (four individuals as A and two as C), aligning more closely with other ancient Maya populations.

While both studies are limited in sample size in representing true frequencies (Merriwether *et al.*, 1997), several factors may explain the discrepancies. First, the unusually high frequency of haplogroup C in Merriwether *et al.*, (1997) may reflect consanguinity among elite individuals (González-Olivier *et al.*, 2018). Second, concerns about contamination are not adequately addressed in Merriwether *et al.*, (1997). The study took a PCR-based approach to amplify and sequence mtDNA, which did not leave any signature of *post-mortem* damages typically observed in ancient DNA when data

are generated by next-generation sequencing (Roca-Rada et al., 2020). By contrast, in my study, contamination estimates were negligible (<2%) (Table 3.1), *post-mortem* damage patterns were confirmed (Fig. 3.3), and familial relationships were rigorously screened (Fig. S3.2).

The haplogroup frequencies observed in the Classic Copán (67% typed A2 with $se=19.2\%$, and 33% typed C1 with $se=19.2\%$) closely match those reported for other ancient Maya cities (73% typed A2 with $se=5.0\%$, and 20% typed C1 with $se=4.5\%$). This alignment challenges earlier interpretations of genetic isolation at Copán, particularly on the maternal side, and instead suggests connections to other Maya kingdoms. These findings are consistent with inscriptions and archeological records, which highlight a cohesive Maya culture characterised by long-distance trade, inter-city marriages, and frequent conflicts (Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015; Sharer, 2009; Söchtig et al., 2015).

II. 6. Population structures of ancient Maya through mtDNA

Based on mtDNA variation observed between present-day northern and southern Maya groups, possibly reflecting past demographic events (González-Martín et al., 2015), I investigated population structure in ancient Maya. Individuals from this study, as well as from previous studies, were categorised geographically into two regions: the Yucatán Peninsula lowlands (north) and the highlands (south). I also analysed temporal variation, dividing individuals into two periods: the Classic period (250-900 CE) and the Postclassic period (900-1,521 CE), which marks the rapid decline and abandonment of many Maya cities (Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015). Finally, given the Maya's highly stratified society, I compared mtDNA haplogroups between elites and commoners (Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015; Sharer, 2009). Social status was inferred through archeological contexts, though often with uncertainty. Only CpM13 from the Classic Copán could be confidently classified as a royal elite, and a limited number of individuals as commoners, based on their tombs' locations in working quarters (Ochoa-Lugo et al., 2016). Results of these analyses are presented in Figure 3.7. However, due to the limited representation of elites ($n=1$), social class-related structure could not be reliably assessed.

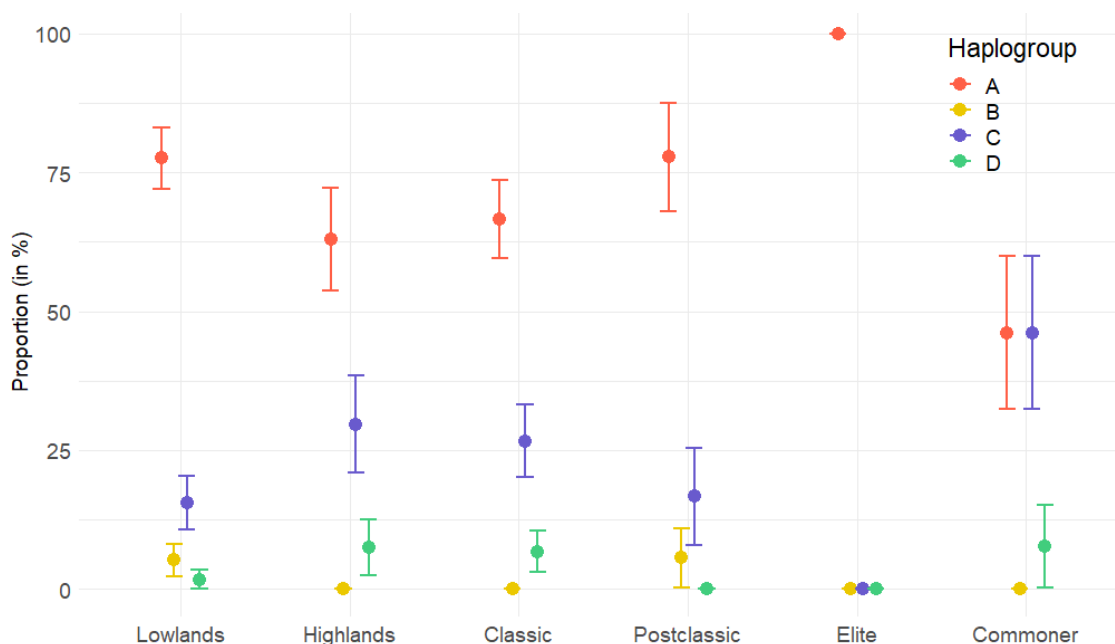


Figure 3.7. A comparison of published ancient Maya mitochondrial frequencies. The error bars show the proportion and standard error of the four haplogroups (A, B, C and D in red, yellow, blue and green respectively), across geography (the lowlands versus the highlands), periods (Classic versus Postclassic) and social status (elite versus commoner). Standard error was calculated as follows: $SE = \sqrt{\frac{p(1-p)}{n}}$. The categories lowlands, highlands, Classic, Postclassic, elite and commoner consist of 58, 27, 45, 18, 1 and 13 individuals respectively.

Geographically, the highlands showed higher frequencies of haplogroups C and D and lower frequencies of haplogroups of A and B, compared to the lowlands. Temporally, the Classic period exhibited similar patterns, with higher frequencies of haplogroups C and D and lower frequencies of haplogroups A and B compared to the Postclassic period (Fig. 3.7). It is important to note that, while the lowlanders are equally divided between the temporal periods (18 individuals each), all highlanders are associated with the Classic period, complicating the distinction between geographic and temporal factors. Still, considering the overlaps in the standard errors across regions and periods, ancient Maya were likely a genetically homogenous population, without sub-structure between north and south or between pre- and post-collapse.

The observed diversity among present-day Maya groups may have arisen during the Colonial period rather than earlier eras. The arrival of Spanish conquistadors had devastating effects, introducing new diseases, waging war with advanced weaponry, and

implementing the *encomienda*⁸ system, which enslaved Indigenous populations. These factors caused severe population bottlenecks, territorial reorganisation, and subsequent isolation, likely driving the genetic differentiation observed in the maternal line of contemporary Maya populations (Livi-Bacci, 2006; Martínez de Velasco Cortina and Vega Villalobos, 2015).

While additional mtDNA analyses could further elucidate maternal genealogical histories, they remain limited to the maternal lineage. In this project, I leverage the autosomal shotgun genomes from the Classic Copán to conduct genome-wide analyses. This approach enables the application of population genetics methods unsuitable for mtDNA data, offering new insights into demographic trajectories during the Classic Maya period.

⁸ *encomienda*: Spanish labour system that rewarded *encomenderos* with Indigenous people to use as slave

III. Masking genomic segments derived from non-American ancestry in present-day Americans

III. 1. Non-American proportions in present-day Maya genomes from HOA

Before analysing genome-wide data, I identified potential non-American genomic segments in the present-day American populations. These segments, originating from European and/or African admixture during post-Columbian time, are absent in ancient or pre-Columbian Americans, and could bias downstream analyses when analysing ancient and present-day Americans together. Specifically, European and African ancestries exhibit significant divergence from American ancestries, artificially amplifying genetic differentiation between ancient and contemporary populations.

To address this problem, I inferred local ancestry in present-day American genomes to identify and mask segments of non-Indigenous (i.e., African and European) origin. Several tools exist for local ancestry inference—such as LAMP (Sankararaman et al., 2008), HapMix (Price et al., 2009), ELAI (Guan, 2014), RFMix (Maples et al., 2013), and MOSAIC (Salter-Townshend and Myers, 2019)—each with different assumptions and capabilities. Earlier tools like LAMP and HapMix are limited to two-way admixture models and are therefore not suited for complex three-way admixture typical of American populations. More recent methods, such as ELAI and RFMix support multi-way admixture but require a priori specification of source populations, meaning the user must define in advance which ancestries to infer. In contrast, MOSAIC does not require prior specification of ancestral groups. Instead, it uses a two-layer Hidden Markov Model (HMM) and linkage disequilibrium (LD) patterns to infer ancestry directly from the genomic data. This makes MOSAIC particularly suitable for masking, as some present-day individuals might carry exclusively Indigenous American ancestry and thus require no masking. Differentiating Indigenous Americans from Siberians remains challenging due to their shared evolutionary history; however, the focus here is on masking segments introduced through recent gene flow—specifically African and European ancestry—that postdates the Maya period and would confound analyses.

Global populations from the Human Origin Array (HOA) dataset were classified into seven regions based on their geographic origins: Africa, America (representing

Indigenous American ancestry), Siberia, East Asia, Middle East, South Asia and West Eurasia (representing European ancestry). Local ancestry inference was first applied to the present-day Maya within the HOA dataset, with the assumption of three distinct ancestral sources. These three sources identified in the Maya are coloured by blue, green, and orange in [Figure 3.8](#).

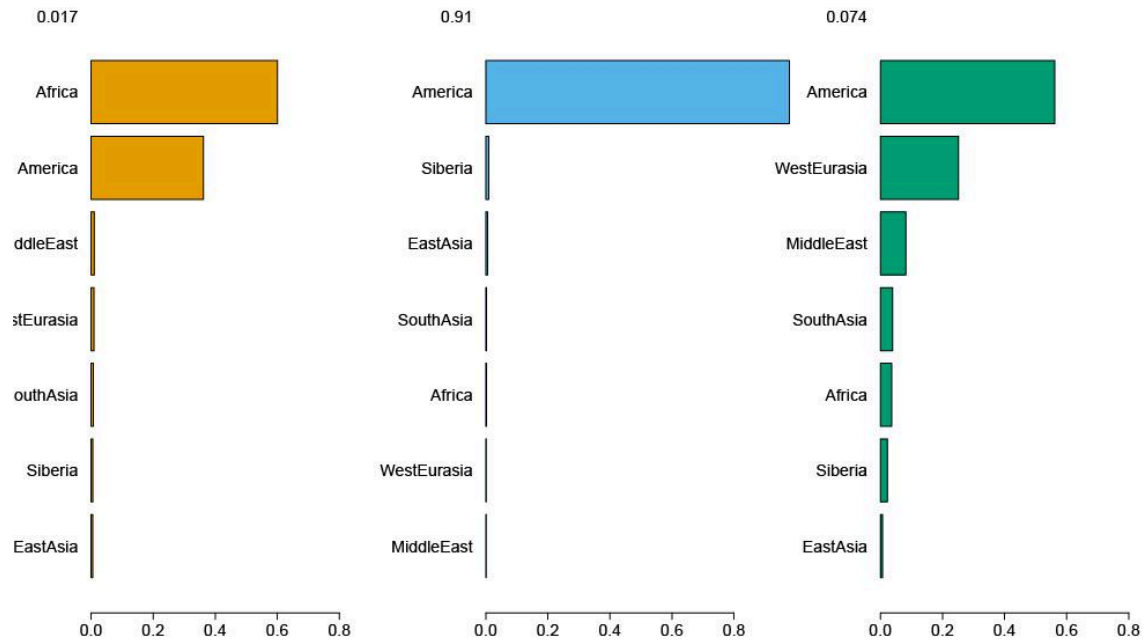


Figure 3.8. Three-way admixture proportion in present-day Maya from HOA as inferred by MOSAIC. Each colour paints the make-up of an ancestry inferred by MOSAIC. Values on top are the admixture proportions from each source ancestry.

West Eurasian and African segments identified in the Maya are also present in the reference panel of Americans, yet these segments exhibit greater genetic similarity to the reference panel of West Eurasians or Africans based on F_{st} values ([Fig. 3.9](#)). American ancestry constitutes the majority of the Maya genomes (91%), with the remaining segments attributed to West Eurasian (7.4%) and African (1.7%) ancestries ([Fig. 3.8](#)).

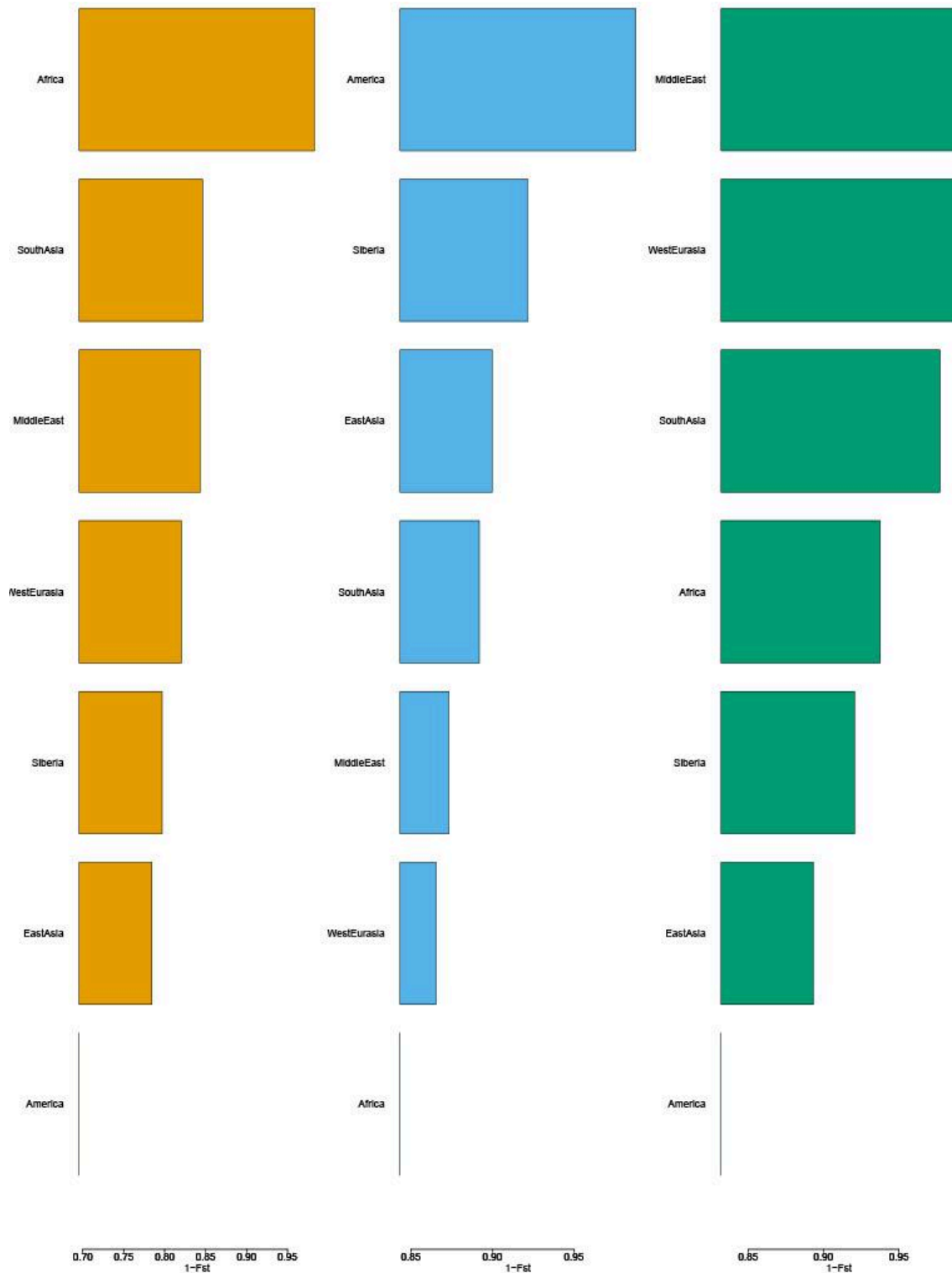


Figure 3.9. Similarities of the orange, blue, and green ancestry detected by MOSAIC in present-day Maya from HOA relative to the other continents. Here, the orange, blue and green components represent African, American and European ancestries respectively. As F_{ST} defines genetic divergence, the x-axis of $1-F_{ST}$ shows here genetic similarity.

The minor West Eurasian and African contributions in the present-day Maya genomes are consistent with the historical events of the 15th-century colonisation, when Europeans arrived in Central America and subsequently began transporting Africans to work in the fields. This admixture likely occurred between 400 and 500 years ago, approximately 20 generations ago, assuming a generation time of 20-25 years.

MOSAIC can also estimate admixture timing by fitting curves that measure the probability of making pairs of ancestry-specific haplotypes at varying genetic distances (Salter-Townshend and Myers, 2019). For the Maya, the estimated admixture date is approximately 8 generations ago (Fig 3.10). This estimate is more recent than the initial contact period, which occurred around 20 generations ago (Livi-Bacci, 2006). However, it is reasonable to assume that admixture did not begin immediately after contact, as the arrival of more Europeans and Africans and the onset of gene flow likely occurred gradually. Furthermore, admixture is not a single discrete event but rather a process that unfolds over time. Because MOSAIC is designed to capture the most recent major admixture signals, an estimate of 8 generations is consistent with historical expectations of admixture occurring some time after initial contact.

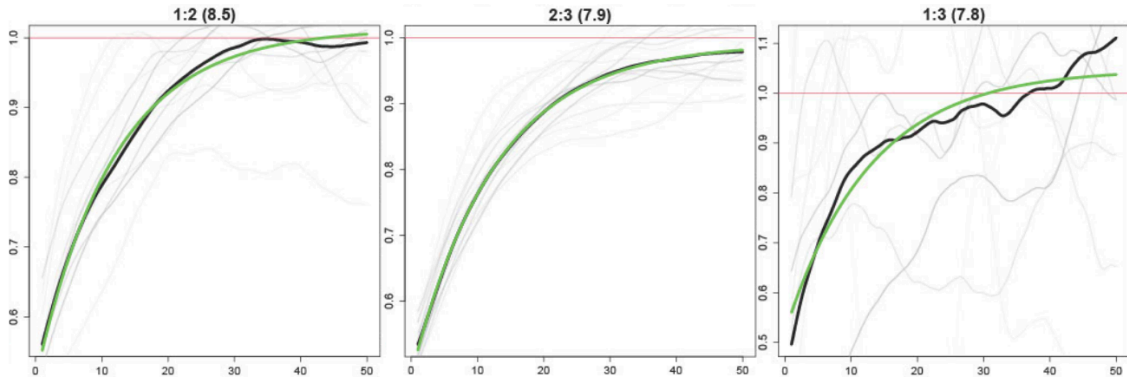


Figure 3.10. Three-way co-ancestry curves in present-day Maya from HOA as estimated by MOSAIC. Values on top correspond to source ancestries from Fig. 3.8 and Fig. 3.9: 1 is orange, 2 is blue, and 3 is green. So the first plot is the co-ancestry between the orange and blue ancestries, the second plot is the co-ancestry between the blue and green ancestries, and the last plot is the co-ancestry between the orange and green ancestries. The black lines are empirical across all individuals, the light grey are per individual, the green is the fitted single-event co-ancestry curve. The co-ancestry plots use exponential decay for the ratio of probabilities of pairs of local ancestries (y-axis) as a function with the genetic distance (x-axis) to estimate the decay parameter which corresponds to admixture date in generations (value in bracket on top).

III. 2. Non-American proportions in present-day Maya genomes from SGDP

To validate the local ancestry inference performed with MOSAIC, I tested whether the results obtained for the 18 Maya individuals from the HOA data could be replicated in the two Maya individuals from the SGDP data ([Table S3.2](#)). Both SGDP Maya individuals are included in the HOA database, making them directly comparable. As not all SGDP individuals are included in HOA ([Table S3.3](#)), this replication analysis is critical to confirm the robustness of local ancestry inference and to evaluate the performance of MOSAIC when applied to datasets with small numbers of individuals per population (*e.g.*, 1-3 individuals).

Both datasets aim to capture the diversity of worldwide present-day populations, with a focus on Americans: HOA includes 98 American individuals (Patterson et al., 2012), while SGDP includes 22 (Mallick et al., 2016) ([Table S3.2](#)). However, the two datasets differ markedly in their genomic coverage. The HOA references 594,896 genomic sites, while the SGDP includes 3,867,550 sites after filtering for transversions and a 1% minor allele frequency. Although HOA includes 4.5 times more American individuals than SGDP, SGDP records 6.5 times more variants.

Worldwide populations from the SGDP were classified into different groups based on their geographic origins. A three-way admixture model was applied to two individuals of the present-day Maya from SGDP. MOSAIC identified two American sources (blue and green) and one West Eurasian source (in orange) ([Fig. 3.11](#)). When the two American sources were combined, the genome-wide ancestry estimates for these individuals were 97.8% American and 2.2% West Eurasian ancestry, without any African ancestry. These results differ from the genome-wide estimates for the Maya individuals in the HOA, which identified 91% American, 7.4% West Eurasian, and 1.7% African ancestries ([Fig. 3.8](#)). The observed differences may be due to averaging genome-wide ancestries with other HOA individuals that are not included in the SGDP.

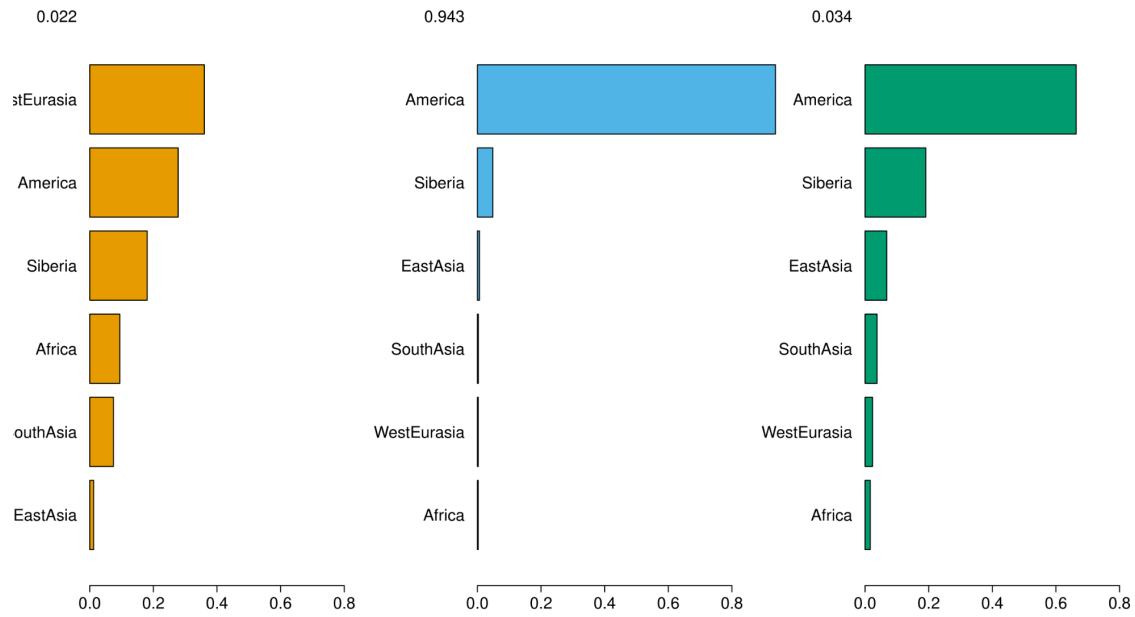


Figure 3.11. Three-way admixture proportion in present-day Maya from SGDP as inferred by MOSAIC. Each colour paints the make-up of an ancestry inferred by MOSAIC. Values on top are the admixture proportions from each source ancestry.

III. 3. Non-American segments in present-day Maya from HOA and SGDP

To investigate further, I directly compared non-American segments between the two Maya individuals present in both datasets. Global ancestry represents the mean ancestry across the entire genome, while local ancestry identifies the ancestry source specific to a genomic position. Local ancestry profiles along chromosomes in the Maya are presented in [Figure 3.12](#), with the HOA and SGDP at the top and bottom respectively. African ancestry, coloured by orange, is negligible, with very short fragments identified only from the HOA. The lengths and locations of European segments (green for the HOA and orange for the SGDP) are comparable across both datasets, such as those on chromosome 17 in S_Mayan-1 and chromosome 16 in S_Mayan-2. The segments identified as African origin in the HOA (orange) are considered as West Eurasian origin in the SGDP dataset. It is also important to note that both blue and green components define American ancestry in the SGDP data.

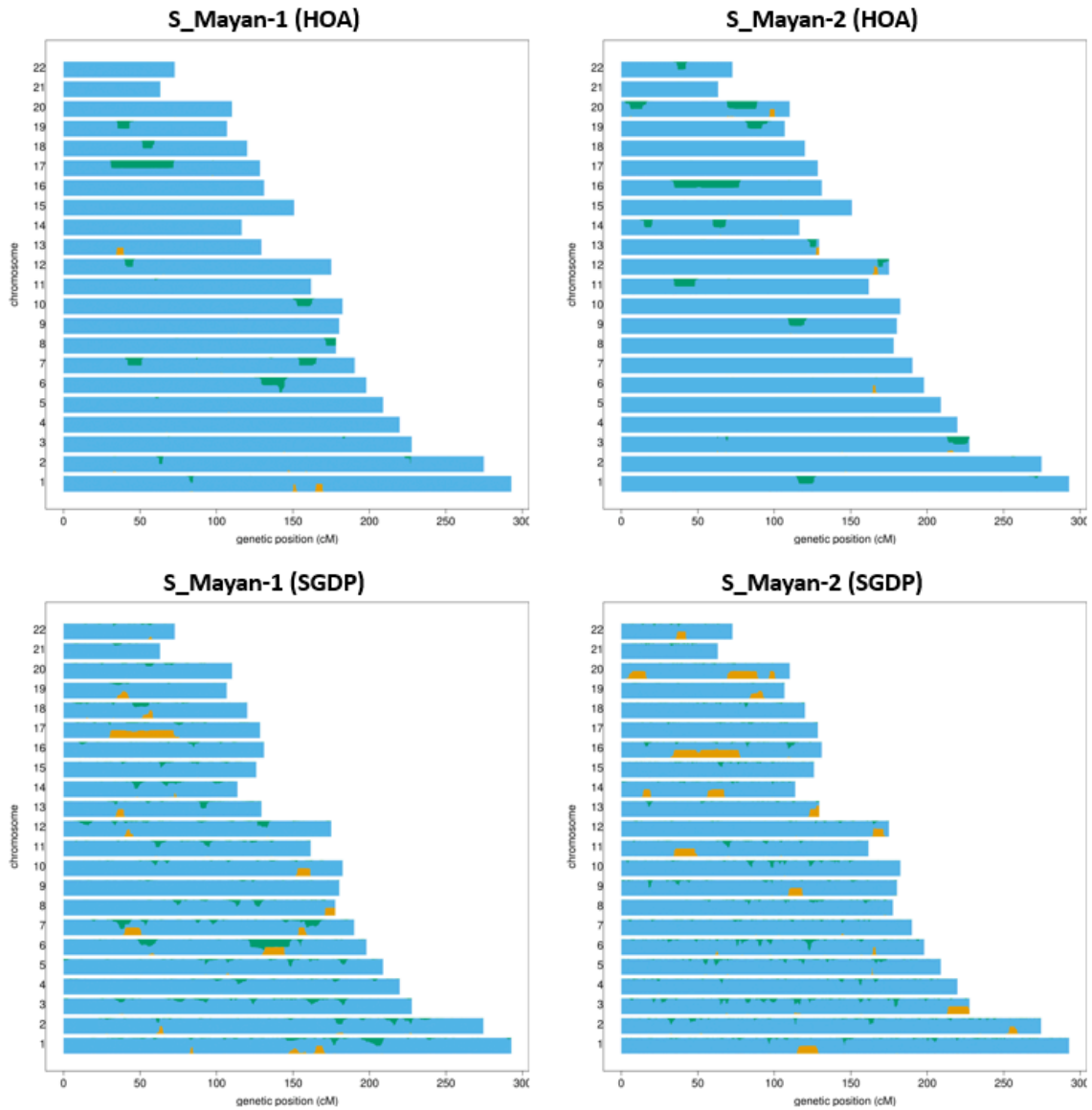


Figure 3.12. Local-ancestries for present-day Maya from HOA and SGDP. The two Maya individuals are present twice, either from the HOA dataset (top) or from the SGDP dataset (bottom). Local-ancestry corresponds to American, European and African in blue, green and orange respectively for HOA, and to American and European in blue/green and orange for SGDP.

Overall, masking non-American segments in the present-day Maya from the HOA and SGDP datasets resulted in the almost same genomic regions being excluded. This finding confirms that MOSAIC accurately infers local ancestry in the SGDP dataset, even when the number of individuals per population is small.

The two individuals of the Maya from the SGDP data exhibit higher American ancestry on average (97.8%) than the genome-wide average (91%) from 18 Maya

individuals from the HOA. This may suggest these two individuals were specifically selected based on their minimum amount of non-American ancestry. Ancestry proportions of these 18 individuals (36 chromosomes) are summarised in [Figure 3.13](#). Those chromosomes derived from the Maya individuals who are also present in the SGDP data are highlighted as gold triangles. Their American ancestries are consistently high across chromosomes, with a couple of exceptions in chromosomes 16 and 17. Still, their genome-wide proportions of American ancestry are among the highest.

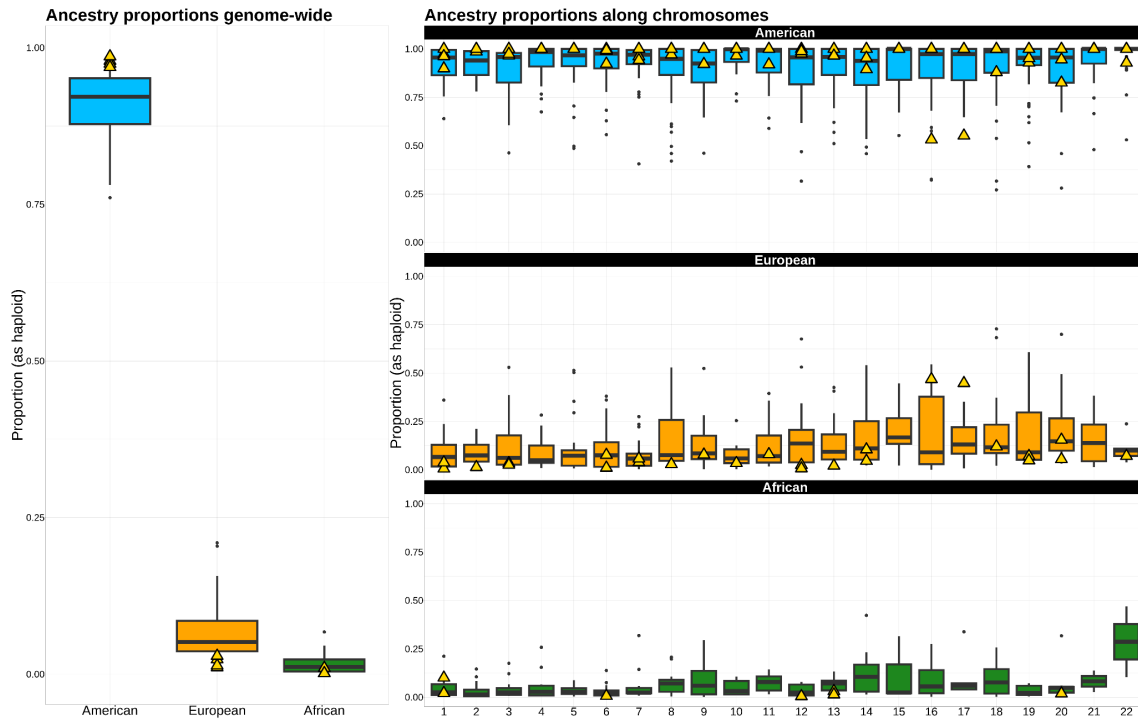


Figure 3.13. Ancestry proportion in present-day Maya from HOA. The left panel summarises ancestry proportion genome-wide, while the right panel details ancestry proportion along the chromosome with each ancestry shown separately. Proportions were estimated on Maya individuals from HOA as haploids, and gold triangles highlight individuals also present in SGDP. Throughout, American, European and African sources ancestries are colour coded in blue, orange and green respectively.

III. 4. Non-American proportions in present-day Americans

I applied MOSAIC to the rest of present-day American populations in the HOA and SGDP datasets. In the HOA dataset, the proportions of American ancestry are on average higher than 91% (Fig. 3.14), with stronger contribution from European ancestry than that from African ancestry, except for Mixe or Pima that have either African or European impact. None of Karitiana, Piapoco and Surui show any detectable level of non-American ancestry. The dates estimated for those admixed populations support the events less than 16 generations ago, consistent with colonial history. Genome-wide proportions of American ancestry for every individual from HOA and SGDP are detailed in Table S3.3.

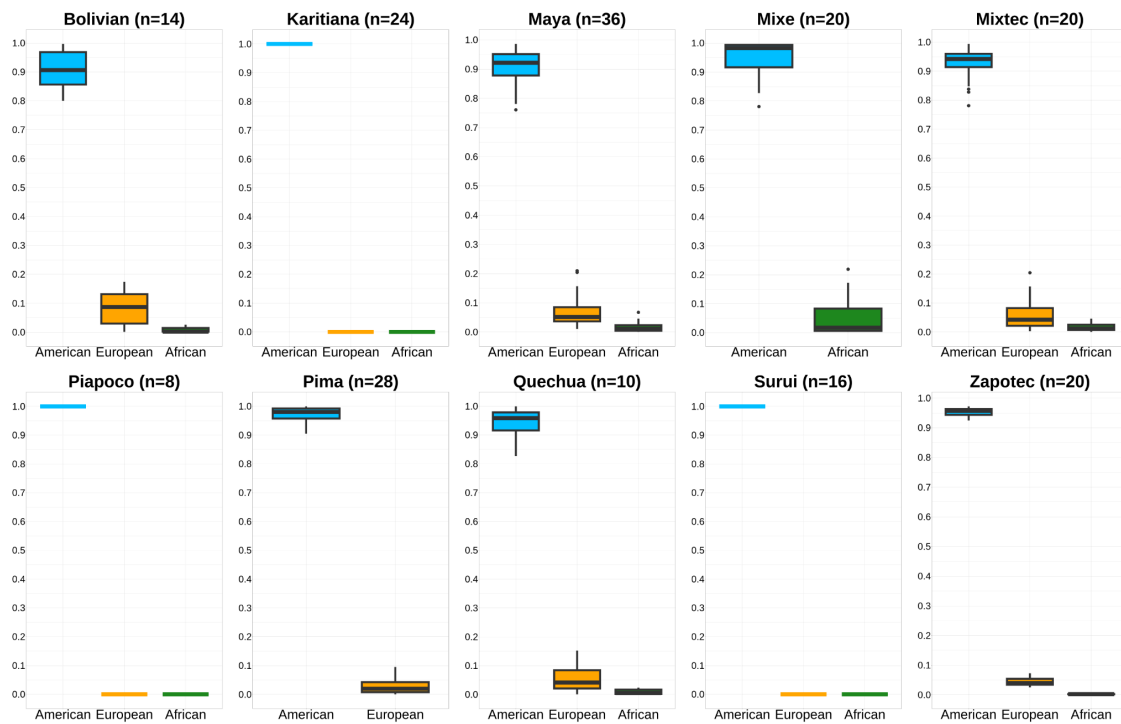


Figure 3.14. Ancestry proportions in present-day Americans from HOA. Top label names American population from HOA, with n as the number of haploids. So given that $2n$ is a diploid, if $n=36$ haploids then $2n=18$ diploids. Throughout, American, European and African sources ancestries are colour coded in blue, orange and green respectively.

IV. Population structure within the Americas

IV. 1. Genetic structure in ancient and present-day Americans

To contextualise newly-sequenced Copán individuals within the broader landscape of American demography, the Classic Copán genomes were combined with genomic data from previously published studies. This created a dataset comprising 674 ancient genomes representing 192 populations, after removing individuals previously marked as either contaminated or related (Fig. 3.15 and Table S3.4). This large set of ancient genomes was further merged with present-day populations from the HOA or SGDP datasets (Table S3.2).

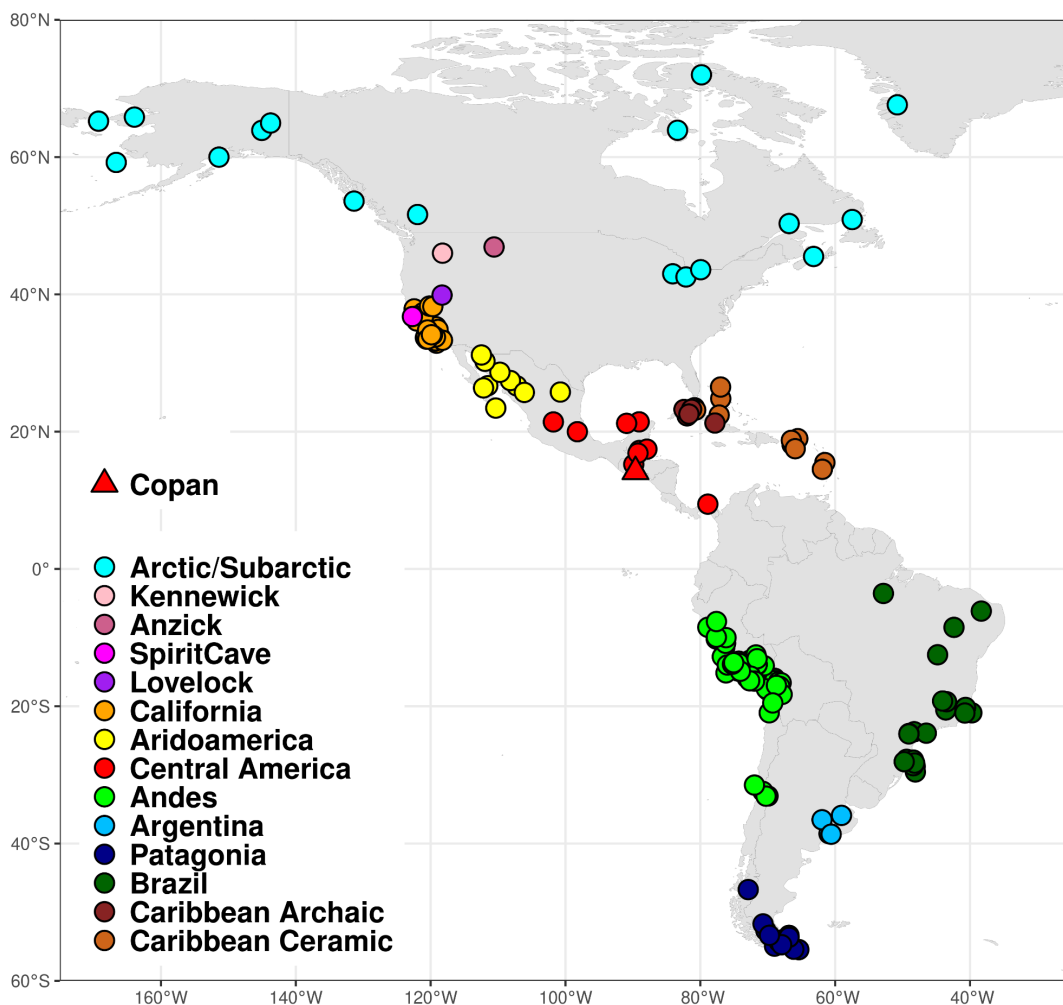


Figure 3.15. Geographic locations of the Classic Copán and publicly available ancient genomes analysed in this thesis. The map shows the locations of the Classic Copán marked with a red triangle, along with other ancient genomes analysed. Three ancient populations from Siberia (AG2, MA1, and Birnirk), included in the analysis, are not plotted.

The genome-wide affinities of ancient and present-day Indigenous American populations were explored using Principal Component Analysis (PCA), where ancient individuals were projected onto genomic variation among present-day individuals. The HOA dataset was chosen over SGDP as a reference of present-day populations to maximise the number of individuals contributing to eigenvector calculation, as it contains 98 individuals compared to 22 individuals from the SGDP dataset ([Table 3.3](#)).

In [Figure 3.16](#), the PCA clearly differentiates two present-day Amazonian populations of Surui and Karitiana (dark green triangles) from the rest of populations, likely resulting from strong genetic drift reported in previous studies (Castro E Silva et al., 2021; Skoglund et al., 2015). Pima, a present-day population from Aridoamerica, forms a separate cluster on the third PC (yellow triangles). The other present-day populations from Central and South America cluster around the centre of the plots, along with the majority of ancient individuals. Within this large cluster, ancient individuals from Aridoamerica (yellow circles) tend to spread toward the present-day Pima.

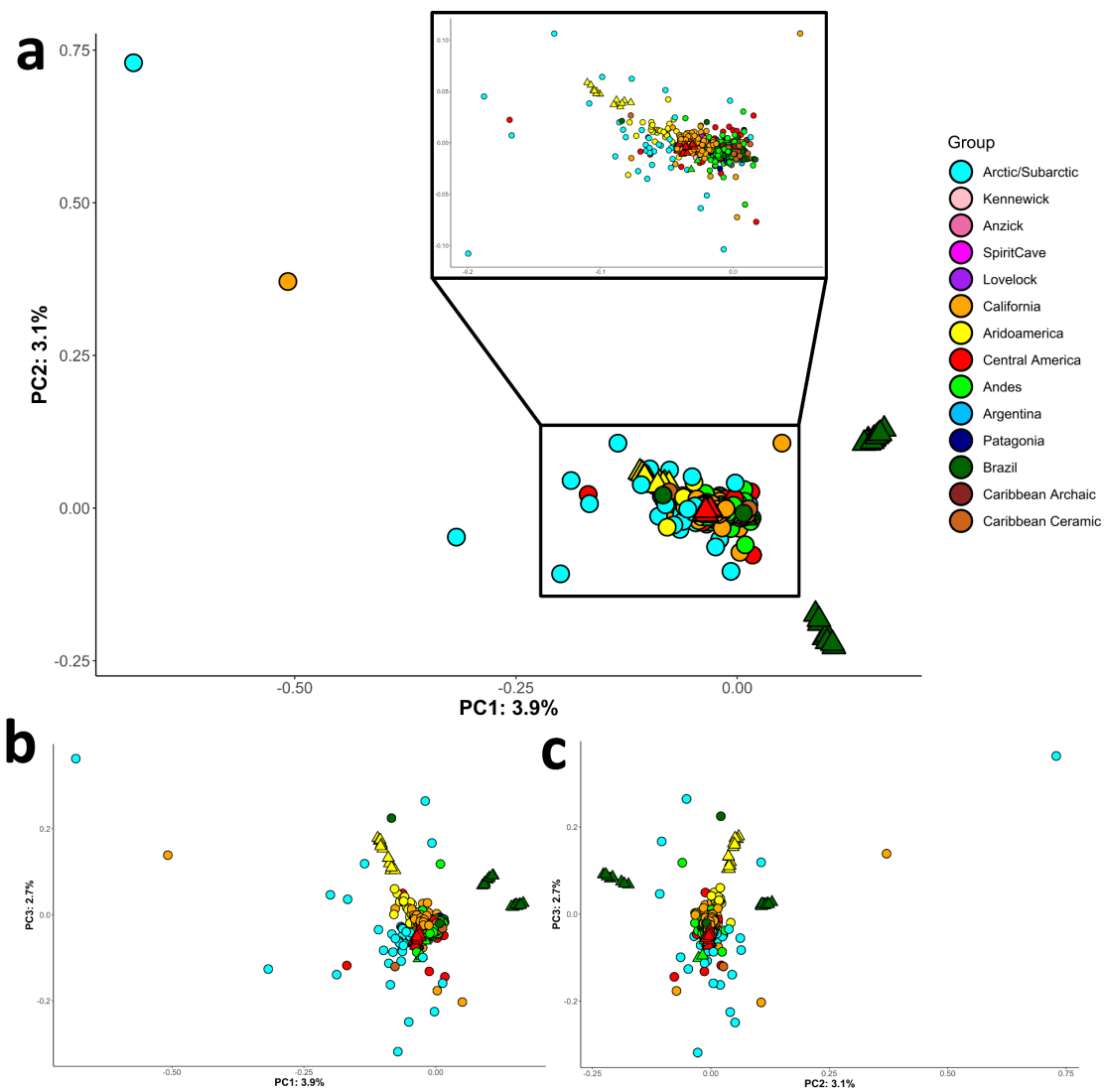


Figure 3.16. Population structure in the Americas as revealed by PCA. Principal Components Analysis (PCA) visualising ancient American individuals (as circles) projected onto present-day American individuals from the HOA panel (as triangles). Both ancient and present-day individuals are colour-coded according to geographic location (see Fig. 3.15; Table S3.4). Different principal components are shown, with PC1 versus PC2 in (a) and a zoomed view of the centre outlined by a square, PC1 versus PC3 in (b), and PC2 versus PC3 in (c).

Projecting ancient genomes onto modern genetic variation through PCA relies on the assumption that past genetic diversity is adequately represented by present-day populations. However, this assumption may not hold true for Indigenous American populations, whose ancestors underwent significant loss of genetic diversity during European colonisation (Livi-Bacci, 2006). Consequently, contemporary populations in

the Americas may not serve as an ideal reference panel for capturing the genetic diversity of the past, at least in the context of PCA.

To further investigate population structures of ancient Americans, I employed multidimensional scaling (MDS) based on a matrix of outgroup- f_3 statistics measured from all pairs of ancient populations and individuals. Unlike PCA, MDS does not require projection, making it possible to directly compare genetic affinities between ancient Americans without relying on present-day populations. I chose to base the MDS transformation on outgroup f_3 -statistics, which measure the amount of shared genetic drift between pairs of populations since their divergence from an outgroup. Compared to methods based on direct pairwise genetic difference, this approach is more robust to issues common in ancient DNA, such as low coverage and *post-mortem* damage. Although outgroup f_3 -statistics are not affected by genetic drift that is unique to either population being compared, if a population has experienced substantial drift in isolation, its f_3 values with other groups will be relatively low because f_3 -statistics fails to detect much shared drift. As a result, this can visually exaggerate distances, potentially distorting the MDS plot's spatial interpretation.

This analysis effectively separates ancient populations into distinct clusters that mirror their geographic origins, such as the Arctic/Subarctic (cyan), California (orange), Aridoamerica (yellow), the Caribbean specifically from the Archaic periods (dark brown), and the Andes (green) (Fig. 3.17; Fig. S3.3). Notably, the ancient populations from Aridoamerica are positioned between the ancient Californians and Central Americans, reflecting their intermediate geographic location. Although the Arctic and Subarctic populations might be expected to cluster far from other American groups due to their extreme geographic location, they instead appear near the center of the MDS plot, possibly because most individuals within these populations are dated to the early Holocene, thus suggesting they retain ancestral genetic signals from which other populations have since drifted. The Caribbean populations are divided into two temporal groups: the earlier Archaic period (8,000-2,800 years ago) and the later Ceramic period (after 2,800 years ago). Populations from the Ceramic period show genetic affinities with South American populations, consistent with previous findings (Nägele et al., 2020). Central American populations (red), along with those from Brazil

(dark green), Argentina (sky blue), Patagonia (dark blue), and the Caribbean during the Ceramic period (light brown), form a central cluster in the plot.

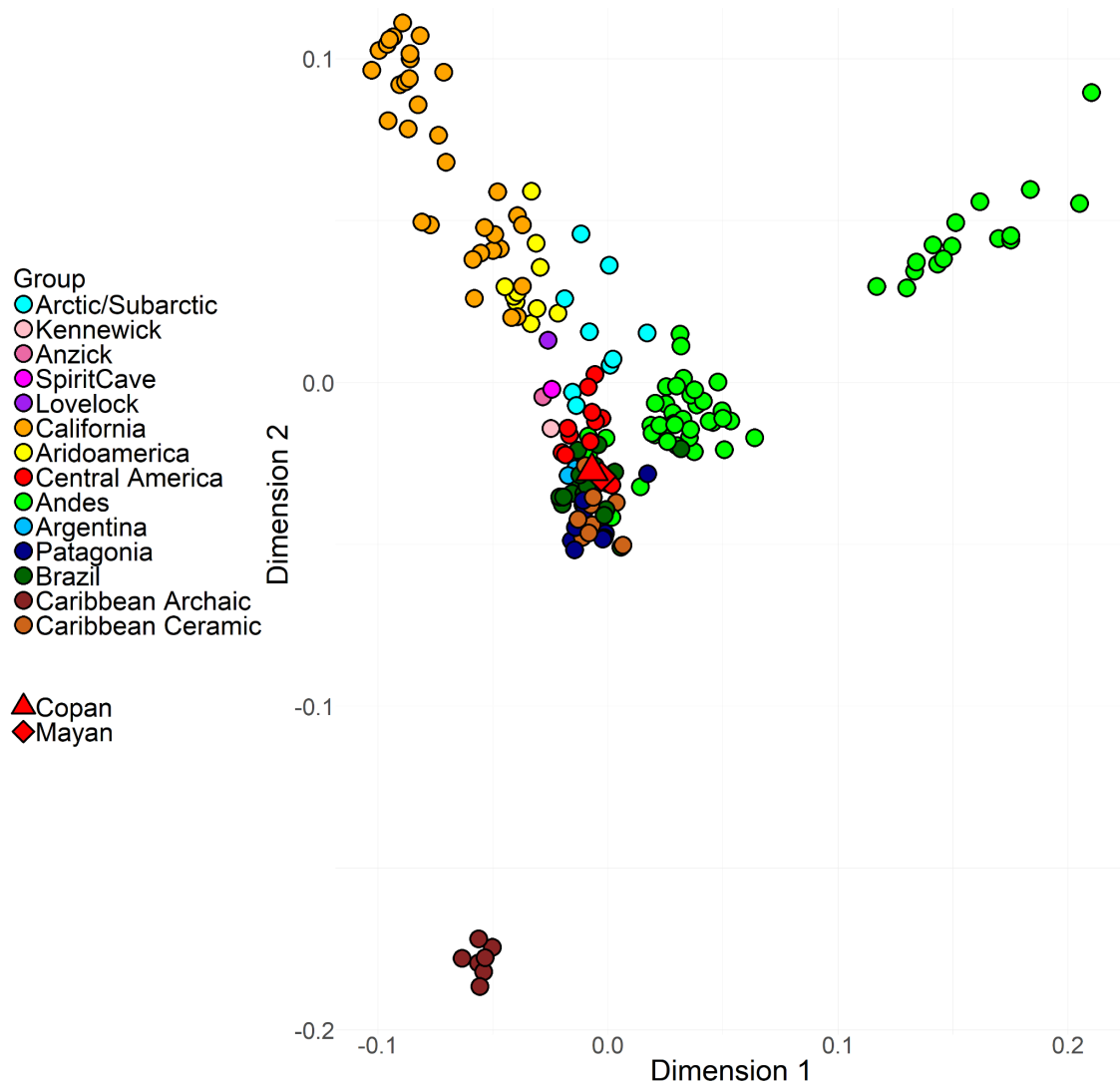


Figure 3.17. Population structure in the Americas as revealed by MDS. Multidimensional scaling (MDS) plot for the matrix of outgroup f_3 -statistics for all pairs of ancient and present-day populations. Populations are colour-coded based on their geographic location (see Fig. 3.15; Table S3.4), with the Classic Copán represented by a red triangle and the present-day Maya by a red diamond.

While some individuals from Patagonia show distinct positioning along Dimension 3, the central cluster—comprising populations from diverse geographic regions—remains consistent across multiple dimensions (Fig. S3.3). This pattern suggests potential genetic links among these regions. Within this central cluster, the

present-day Maya (red diamond) are the closest to the Classic Maya from Copán (red triangle).

IV. 2. Genomic profiles of ancient and present-day Americans

The ADMIXTURE analysis provides a detailed picture of genetic relationships among ancient populations and individuals (Fig. 3.18). At $K = 9$ (Fig. 3.19; Fig. S3.4; Fig. S3.6), the red component is prevalent across several southern populations including those from Central America, the Ceramic Caribbean, the earlier Andes, Brazil, and Argentina. This shared ancestry mirrors their clustering in the MDS plots and supports a genetic link between those regions. This red component is consistently high in the ancient Central American individuals, particularly maximised in the Early Holocene (9,300-7,400 years old) and LateArchaic (5,600-3,700 years old) individuals (Fig. 3.19b).

The other ancestral components uniquely or exclusively characterise different geographic populations. The genetic component identified at $K = 2$ corresponds to the Archaic Caribbean (brown), the most distinctive population on the MDS plots (Fig. 3.18). Individuals from Aridoamerica exhibit the yellow ancestor, which is also prevalent in Central Americans, including the present-day Mixe, Mixtec, and Zapotec at around 50%, and the present-day Maya at around 30%. This component is also observable in the Classic Copán, Chichén Itzá, and Mexico Colonial but not in earlier populations, the Early Holocene (9,300-7,400 years old) and Late Archaic (5,600-3,700 years old) from Belize.

Populations from the Andes are genetically divided into two groups: light green and magenta. This pattern is also observable from the MDS plots (Fig. 3.18), which is possibly defined by a temporal dimension around 550-520 years BP, corresponding to the emergence of the Inca empire (Cartwright, 2014).

Individuals from Patagonia (dark blue), who form a cluster in Dimension 3 of the MDS plots (Fig. S3.3a-b), become unique in their genetic ancestry at $K = 7$ (Fig. 3.19a; Fig. 3.18), supporting the isolation from the rest of South American populations (Posth et al., 2018; Silva et al., 2022; Willerslev and Meltzer, 2021).

Most individuals from California are characterised by either the orange component, unique to the Channels Islands, or the purple one, representing the rest of

the region. Some ancient individuals from California exhibit mixed ancestries of the purple, red, and yellow components, including the oldest individuals dated 3,400 years old (dating for San_Francisco_Bay was not available) (Fig. 3.19a; Fig. S3.6). Given that the yellow and red components are predominant in Aridoamerica and Mesoamerica, respectively, these mixed ancestries may reflect gene flow from Northwest Mexico into California starting around 5,200-4,900 BP, as previously reported (Nakatsuka et al., 2023).

The cyan component, which appears at $K=8$, defines the Early Holocene individuals from North America, including Anzick (Rasmussen et al., 2014) and Kennewick Man (Rasmussen et al., 2015), and Brazil (*i.e.*, Sumidouro) (Moreno-Mayar et al., 2018b), possibly representing the ancestor who made the earlier migration wave into South America (Moreno-Mayar et al., 2018b).

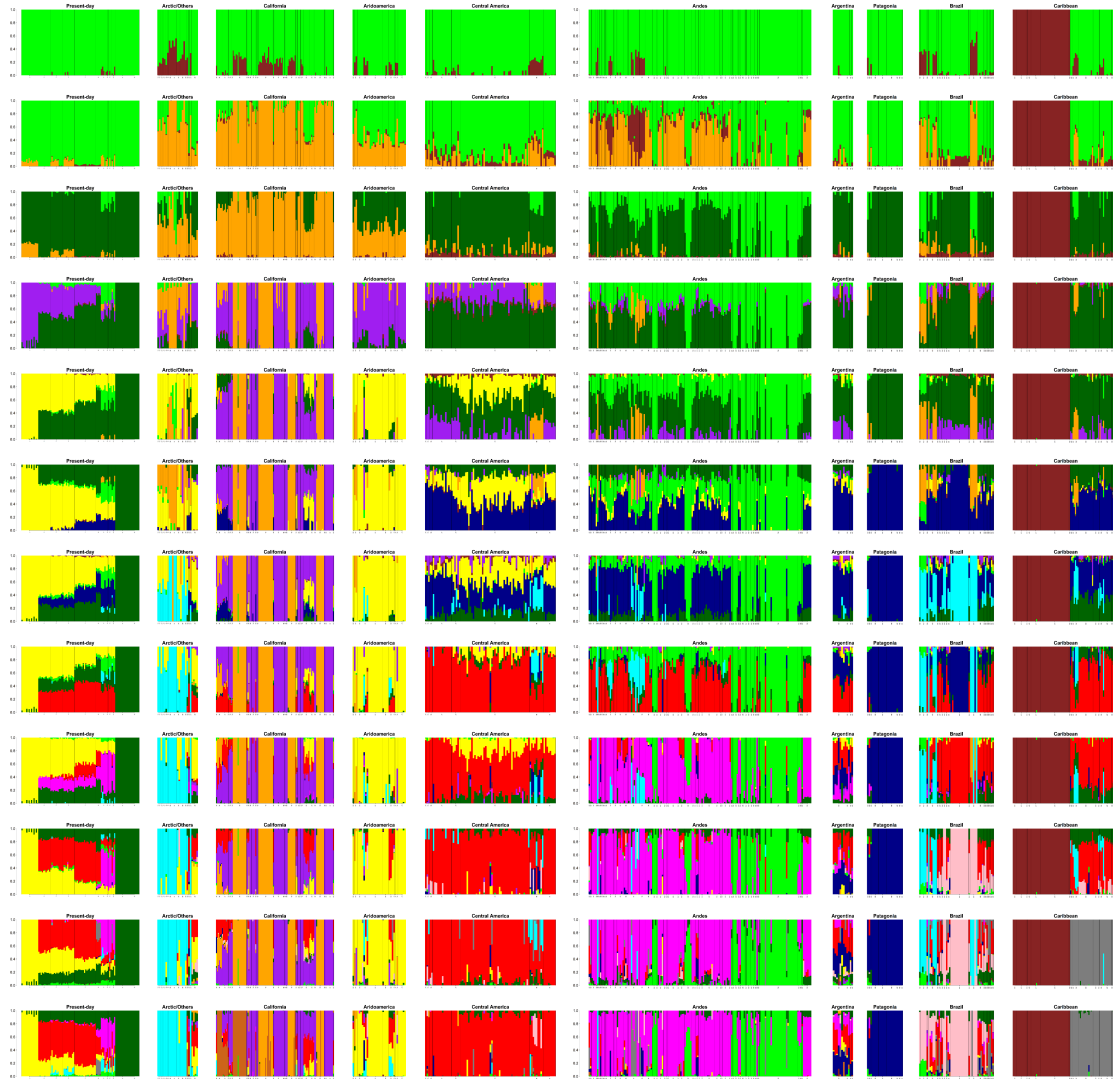


Figure 3.18 ADMIXTURE results from ancestral clusters from $K = 2$ (top) to 13 (bottom) in present-day (far left) and ancient populations from the Americas. Ancient individuals are grouped based on their geographic contexts and sorted by their associated dates within each geographic group, with the oldest on the left and the most recent on the right. Present-day individuals are sorted based on latitude, with the most north on the left and the most south on the right. A detailed list of the populations and the number of individuals included in each population is provided in [Table S3.4](#). At $K = 2$, populations split into Archaic Caribbean (brown) and non-Archaic Caribbean (green) ancestry. At $K = 3$, an orange component appears, which is prevalent across ancient populations and particularly defines the genetic ancestry of populations from the California Channel Islands. At $K = 4$, a dark green component separates the green populations, distinguishing present-day populations from the Amazon region in Brazil. At $K = 5$, a purple component appears, shared by the present-day Pima and ancient populations from California and Aridoamerica. At $K = 6$, a yellow component defines the ancient Aridoamerican and present-day Mesoamerican populations. Additional ancestral components appear at higher K values, including a dark blue component for ancient Patagonian populations at $K = 7$, cyan for the Arctic at $K = 8$, red for Central America at $K = 9$, magenta for the Andes at $K = 10$, pink for Brazil at $K = 11$, grey for

the Ceramic Caribbean at $K = 12$, and chocolate for the California Channel Islands at $K = 13$.

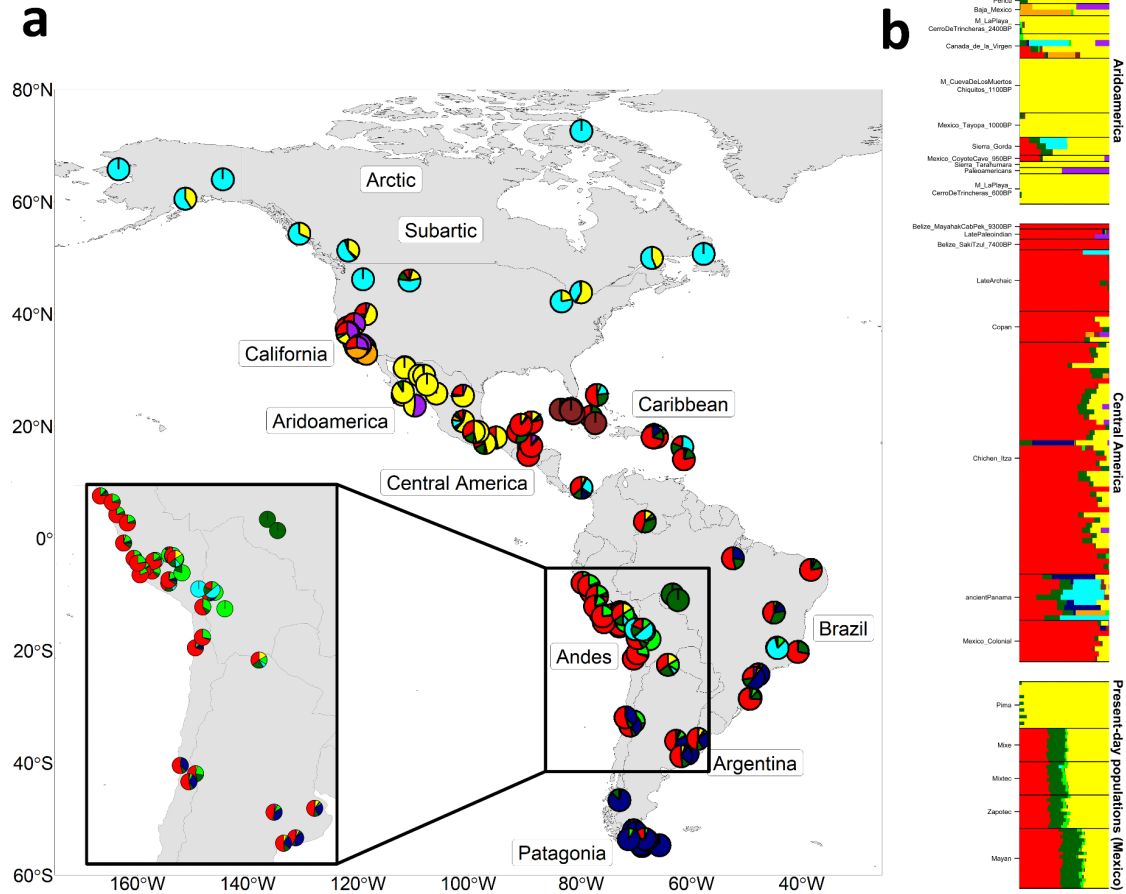


Figure 3.19. Genomic profile of ancient and present-day American populations. Geographic distribution of distinct ancestral components identified from ADMIXTURE analysis with $K = 9$. (a) Pie charts represent the composition of different ancestors in ancient and present-day American populations, with a zoom on the densely sampled region of the Andes. (b) Ancestral composition at the individual level of individuals from the Maya region and Aridoamerica ($K = 9$; results from $K = 2$ to $K = 13$ are presented in Fig. 3.18). Ancient individuals are grouped based on their geographic contexts and sorted by their associated dates within each geographic group, with the oldest on the left and the most recent on the right. Present-day individuals are sorted based on latitude, with the most north on the left and the most south on the right. A detailed list of the populations and the number of individuals included in each population is provided in Table S3.4.

IV. 3. Population migration in the Americas

To better understand the population structure of ancient Americans, I employed Estimated Effective Migration Surface (EEMS) to model migration rates in the context of geographical barriers (Fig. 3.20).

At the southernmost tip of the Americas, Patagonia appears highly isolated from the rest of South America, while showing substantial internal migration within the Patagonian peninsula. At the northernmost end of the Americas, Alaska shows low migration for individuals dated to the Early and Middle Holocene. These samples potentially represent descendants of the population that remained isolated in Beringia during the LGM. The limited migration in the region suggests geographical barriers due to the harsh cold climatic conditions.

In the Andes, the EEMS results reveal significant isolation of Andean individuals from other regions of South America, consistent with the rugged geography of the region (Gnecchi-Ruscone et al., 2019). However, there is a clear hub of migration around the Titicaca basin, suggesting this area served as a point of connectivity for populations in the central Andes.

In the Caribbean, migration patterns appear fragmented, possibly reflecting the mixture in the dataset of individuals from the Archaic and Ceramic phase of the Caribbean. These two groups, previously shown to cohabit without evidence of admixture (Nägele et al., 2020; Schroeder et al., 2018; Willerslev and Meltzer, 2021), may contribute to the observed geographic barriers in EEMS.

On a continental scale, EEMS highlights a discernible migration corridor from North America to South America, with reduced migration in desertic regions, such as of Aridoamerica and Patagonia.

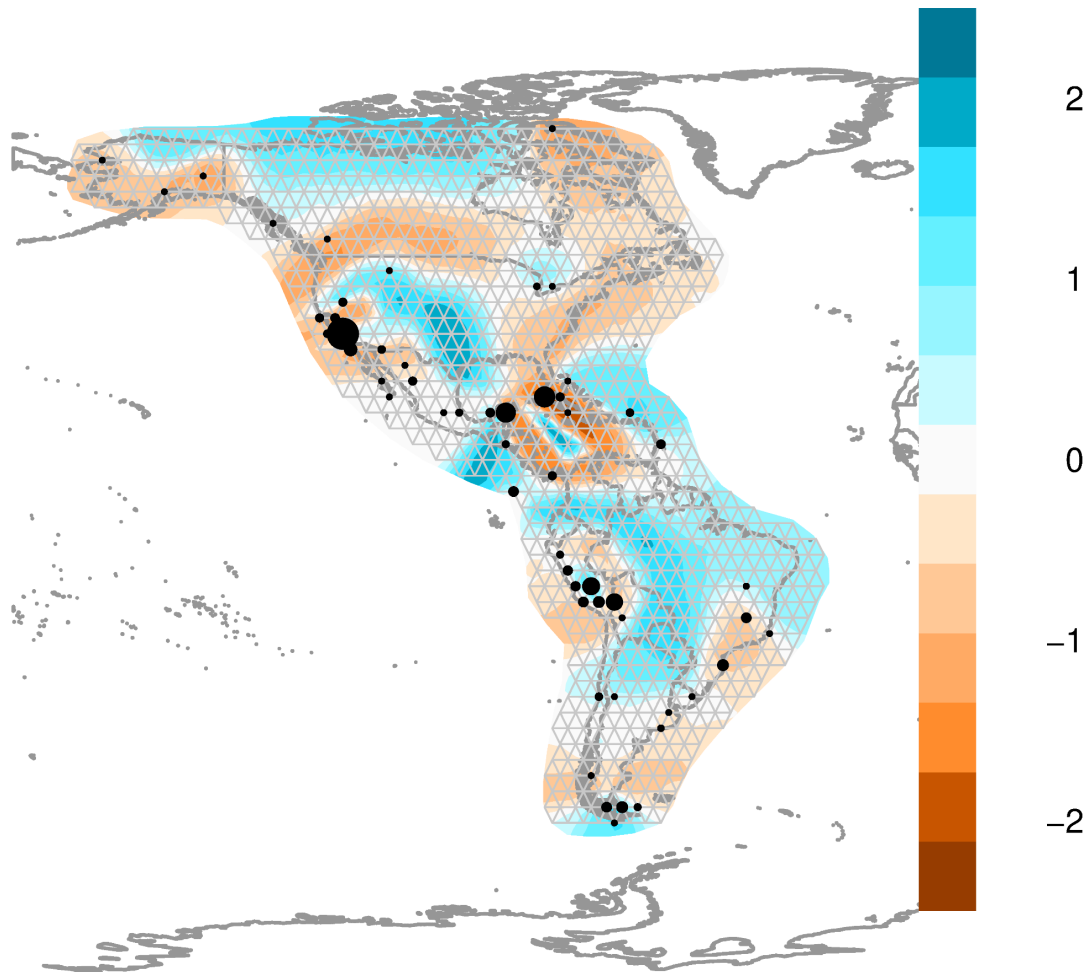


Figure 3.20. Migration rates in ancient America. Effective migration rates are estimated in the log scale with normalisation so that 0 corresponds to the overall mean migration rate. As such, blue corresponds to migration corridors, while orange highlights geographical barriers.

One important caveat of the EEMS analysis is the mixing of samples from different time periods, which may obscure temporal shifts in migration patterns. Migration dynamics likely evolved over the 15,000 years since the initial peopling of the Americas. To address this, I divided the dataset into three archeological periods: Early Holocene (before 8,000 BP), Middle Holocene (8,000-2,500 BP), and Late Holocene (after 2,500 BP). The results, visualised in the supplementary materials ([Fig. S3.8](#)), show that insufficient sampling density for the Early and Middle Holocene periods limit resolution for interpretation. For the Late Holocene, the patterns were mostly consistent with results obtained using all samples. Large regions remain unsampled, complicating EEMS's ability to detect fine-scale migration patterns.

V. Genetic affinities in Maya

V. 1. Genetic affinities to Classic Copán

To investigate genetic relationships between the Classic Copán and populations from Siberia and the Americas, I computed f_4 -statistics in the form of $f_4(\text{Mbuti}, X; Y, \text{Classic Copán})$. This analysis leveraged the SGDP panel, filtered for transversions and 1% minor allele frequencies in global populations to minimise the effects of *post-mortem* damage—a critical consideration in ancient DNA studies. The computation iteratively evaluated each of the 200 populations as X , with the remaining populations designated as Y . A Z -score above 3.0 is considered statistically significant in the genetic affinity of population X to the Copán relative to population Y . Populations X were ranked based on the frequency with which they demonstrated greater genetic affinity to the Copán compared to each of the other 199 populations, with higher ranks representing stronger affinities to the Copán ([Fig. 3.21](#)).

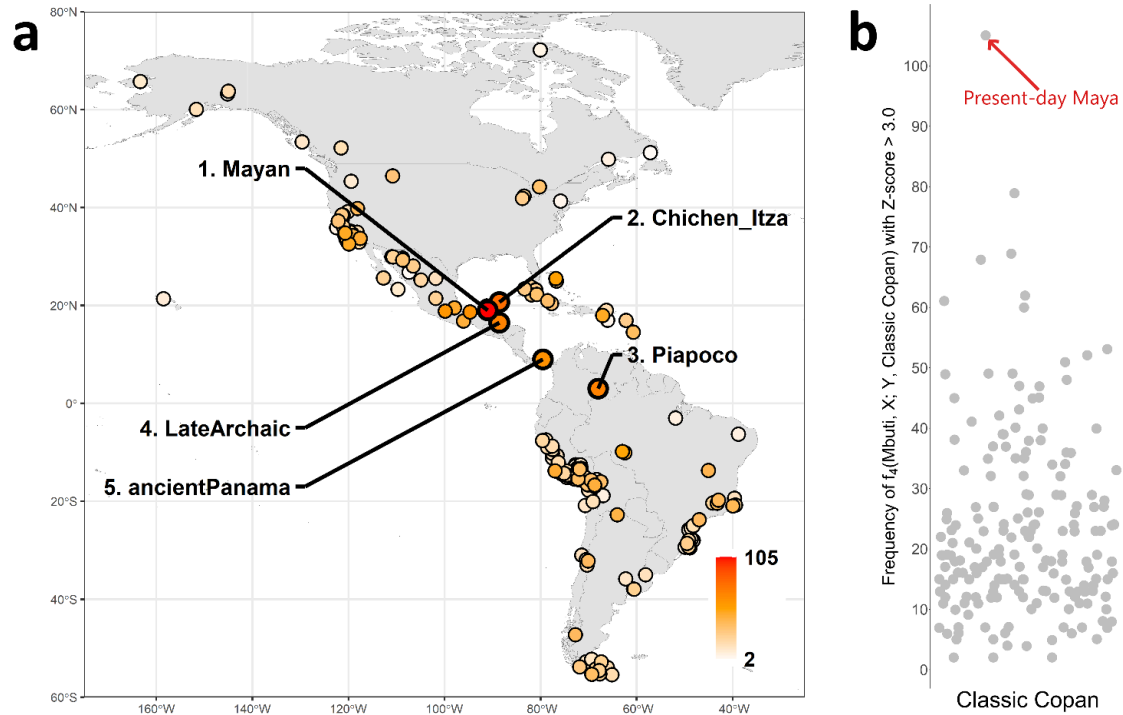


Figure 3.21. Ancient and present-day populations ranked by significant genetic affinities to the Classic Copán in the form of $f_4(\text{Mbuti}, X; Y, \text{Classic Copán})$. (a) A geographical map illustrates a population X who exhibits an excess genetic affinity to the Classic Copán compared to a population Y , with a Z-score > 3.0. The colour gradation reflects the frequency with which a population stands out in the f_4 -tests. The top five populations are labelled on the map, with their ranking. Three populations from Siberia (AG2, MA1, and Tlingit) are included in the tests, but are not shown on the map, all of which are positioned with lower rankings in the test. (b) The plot represents the frequencies and rankings of all populations mapped in (a). The present-day Maya population emerges as the most prevalent outlier.

The results highlight the present-day Maya as the top population, consistent with the findings from the MDS analysis (Fig. 3.17). This corresponds to the maximum number of significant hits, 105 out of the 199 possible combinations, with most of the remaining 94 tests involving fewer than 5000 SNPs, which likely limited their statistical power. The other Classic Maya population from Chichén Itzá is ranked as the second from the top, supporting a certain level of genetic continuity within the Maya region (Fig. 3.21a). These patterns are also supported when this analysis was performed for each of the Classic Copán individuals separately (Fig. 3.22).

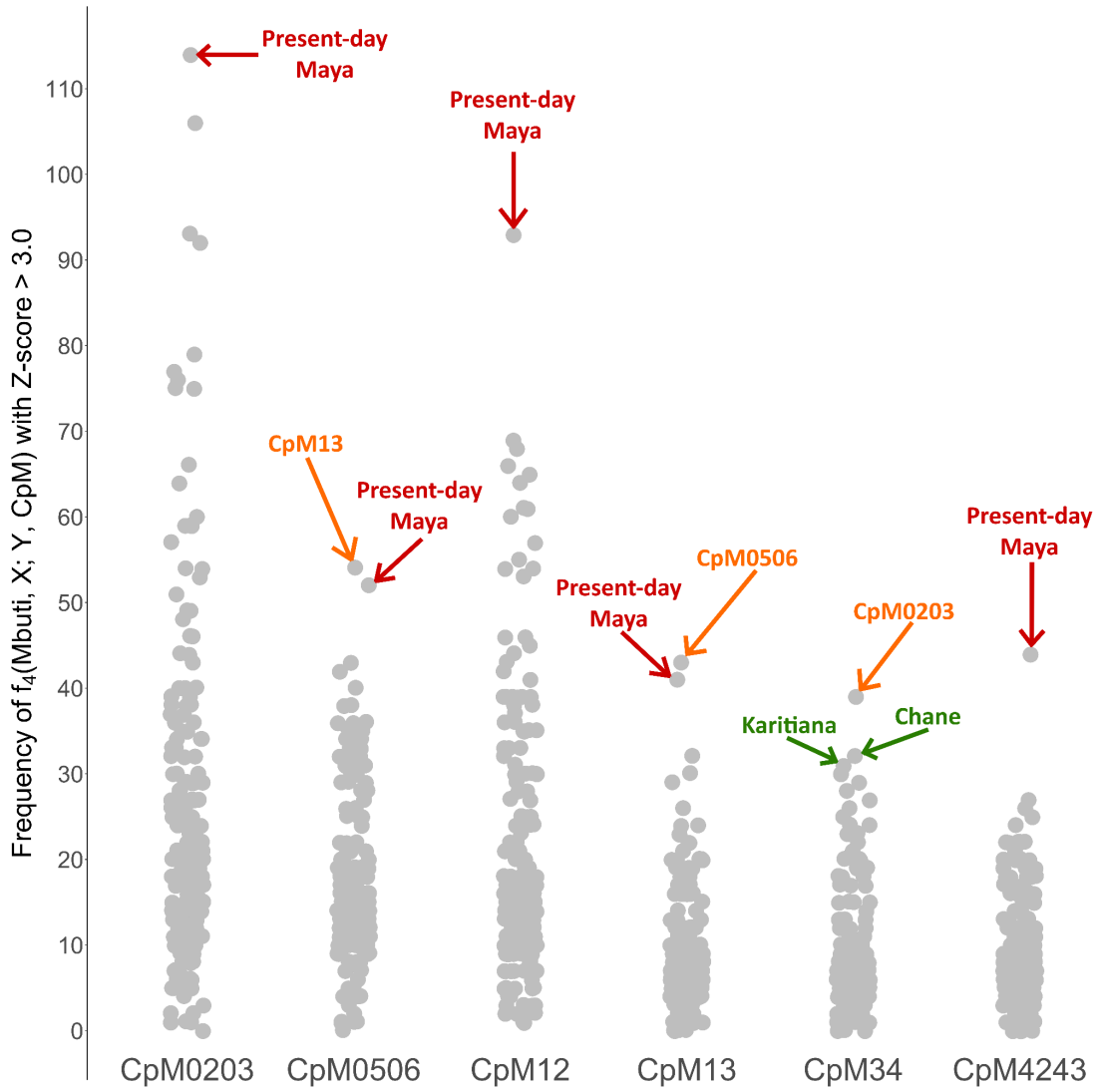


Figure 3.22. Distributions of populations ranked by significant genetic affinities in $f_4(\text{Mbuti}, X; Y, \text{CpM})$. The plot illustrates the frequency with which a population stands out in the f_4 -tests. Each plot represents a population X who exhibits an excess genetic affinity to a Copán individual compared to a population Y , with a Z -score > 3.0 .

V. 2. Genetic continuity in the Maya region

I further tested whether these genetic affinities are reciprocally supported across a time-transect of ancient genomes in the Maya region. The target population, Classic Copán, was replaced with each of the following populations: the individuals from the Late Archaic period, around 5,600 to 3,700 years BP, in Belize, those associated with the Terminal Classic period from Chichén Itzá in the Yucatán Peninsula, the Colonial genomes from Campeche in the Yucatán Peninsula, and the present-day Maya (Fig. 3.23). These analyses reveal Maya groups as the top populations, providing robust evidence of genetic connections across time. In particular, the geographic and temporal ties between the Chichén Itzá and Colonial Mexico are mutually supported from $f_4(\text{Mbuti}, X; Y, \text{Chichen_Itza})$ and $f_4(\text{Mbuti}, X; Y, \text{Mexico_Colonial})$ (Figs. 3.23b-c). The Early Holocene individuals from Belize, dating back to 9,300 years BP and 7,400 years BP, exhibit significant affinities to the Late Archaic (Fig. 3.23a), as reported in a previous study (Kennett et al., 2022). The present-day Maya consistently emerges as the population with strong affinities to ancient Maya populations from the Classic period onward (Fig. 3.21a; Figs. 3.23b-c).

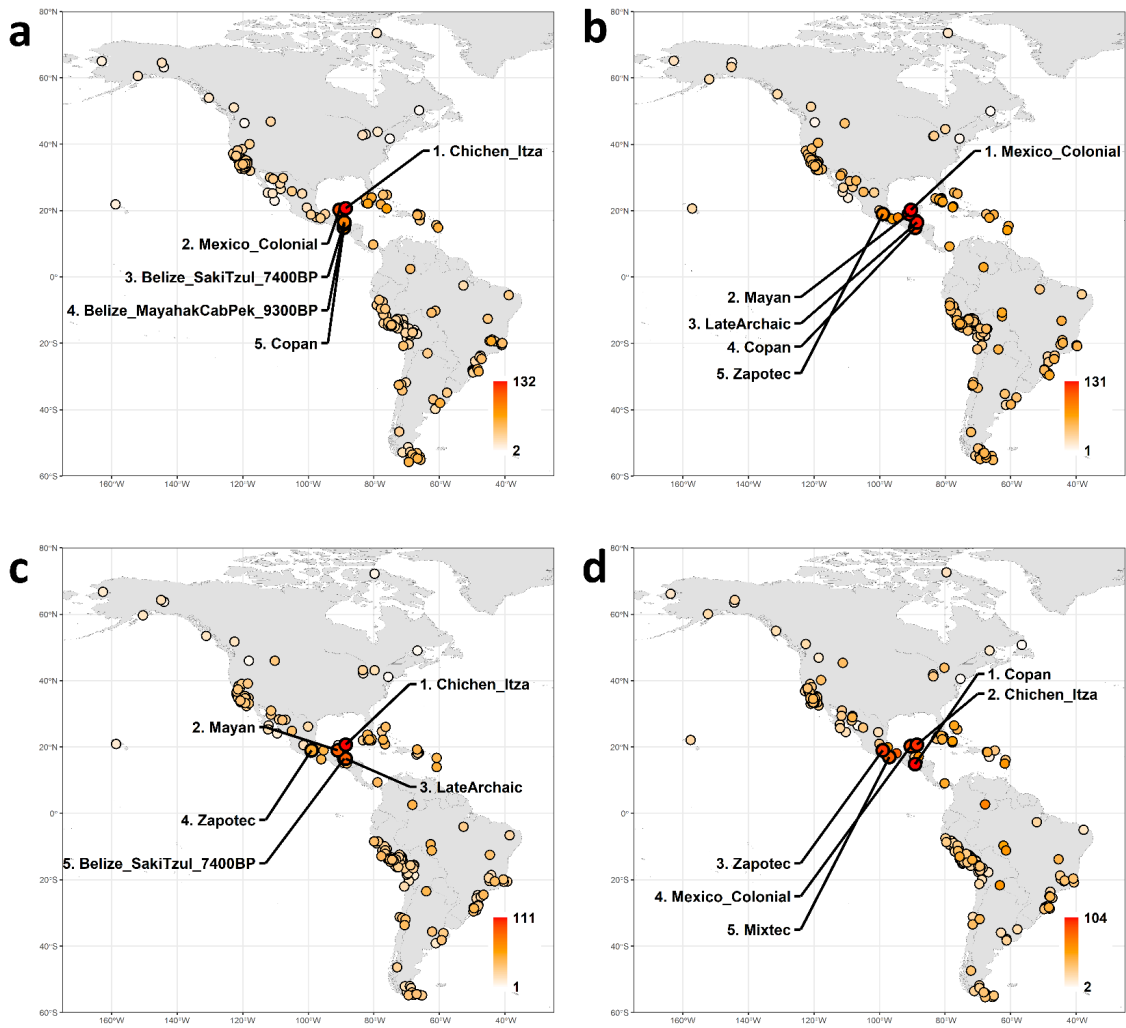


Figure 3.23. Ranking of ancient and present-day populations based on their genetic affinities to populations in the Maya region. The genetic affinities are tested using $f_4(\text{Mbuti}, X; Y, \text{Target})$, where the Target is (a) Late Archaic, (b) Chichén Itzá, (c) Colonial Mexico, and (d) present-day Maya. A geographical map illustrates a population X who exhibits an excess genetic affinity to the target population compared to a population Y , with a Z -score > 3.0 . The colour gradation reflects the frequency with which a population stands out in the f_4 -test. The top five populations are labelled in the map, with their ranking. Three populations from Siberia (AG2, MA1, and Tlingit) are included in the tests, but are not shown on the map, all of which are positioned with lower rankings in the tests.

VI. Genetic ancestries in ancient Maya

VI. 1. Admixture modelling of ancient and present-day Maya populations

To investigate the genetic ancestries contributing to populations in the Maya region over time, the genetic affinities of each pair of the Maya groups were compared to those of all other American populations, using $f_4(\text{Mbuti}, X; Y, Z)$. In this form, Y and Z represent pairs of populations from the Maya region, while X is any other populations in the dataset (Table 3.3). Compared to the Late Archaic, the later Maya populations (Chichén Itzá, Colonial Mexico, and present-day Maya) share greater affinity to the Sierra Gorda or other Mesomaerican populations, such as the Zapotec, Mixe, and Mixtec. The Sierra Gorda, while excavated in a southern region of Aridoamerica, lived in an area that was part of Mesoamerica before the severe droughts of 1,000-1,200 CE, which reshaped the ecological landscape (Villa-Islas et al., 2023). In contrast, several Early Holocene individuals from Belize, Chile, and Argentina exhibit greater genetic affinities with the Late Archaic than with the later Maya populations.

Table 3.3. Comparison of genetic affinities between pairs of populations from the Maya region

Form of f_4 -tests	Population Y	Population X with $Z < -3$	Population X with $Z > 3$
$f_4(\text{Mbuti}; X; Y, \text{LateArchaic})$	Copan	-	-
	Chichen_Itza	Sierra_Gorda Mixe Zapotec	LatePaleoindian Belize_MayahakCabPek_9300BP LocaDoSuin_9100BP
	Mexico_Colonial	Zapotec	LatePaleoindian Chile_CaletaHuelen_1100BP_MH NorthTierradelFuego_LaArcillosa2_5800BP Playa_del_Mango
	Mayan	Sierra_Gorda Mixe Mixtec Zapotec	LatePaleoindian Argentina_ArroyoSeco2_7700BP Argentina_LagunaChica_6800BP Belize_SakiTzul_7400BP Chile_LosRieles_10900BP USA_CA_Carpinteria_7000BP
$f_4(\text{Mbuti}; X; Y, \text{Copan})$	Chichen_Itza	-	-
	Mexico_Colonial	-	-
	Mayan	-	Kaillachuro LSCI LSN Mainland_Chumash Rio_Uncallane Sumidouro & Sumidouro_exclude TIW
$f_4(\text{Mbuti}; X; Y, \text{Chichen_Itza})$	Mexico_Colonial	-	Copan
	Mayan	-	LateArchaic Argentina_LagunaChica_6800BP USA_CA_Goleta_1200BP USA_CA_Goleta_4700BP USA_CA_LakeCachuma_400BP
$f_4(\text{Mbuti}; X; Y, \text{Mexico_Colonial})$	Mayan	-	Argentina_LagunaChica_6800BP

Using the Late Archaic population as a source of pre-existing ancestry in the Maya region, I modelled potential gene flow from populations with greater affinities to the later Maya populations (*e.g.* Zapotec, Mixe, and Sierra Gorda) into ancient and present-day populations (Table S3.5). The admixture modelling shows that the Sierra Gorda is an unlikely source of gene flow due to tail probabilities below 5% or unrealistic admixture proportions across the Maya populations, with the exception of the Classic Copán. In contrast, when the Zapotec or Mixe are used as a source, the admixture model is supported for all Maya populations, with a substantially better fit observed for the Zapotec.

Using the Zapotec as the second source of ancestry, admixture proportions for Late Archaic ancestry estimated are as follows: 93.9% ($\pm 2.6\%$) in the Classic Copán,

48.3% ($\pm 8.0\%$) in the Chichén Itzá, 48.0% ($\pm 11.9\%$) in the Colonial Mexico, and 76.5% ($\pm 9.4\%$) in the present-day Maya (Fig. 3.24). These results indicate that Late Archaic ancestry was dominant in the Classic Copán, but declined by the Terminal Classic period, as seen in the Chichén Itzá, where the two ancestral sources contributed nearly equally to the genetic makeup. This balance between the two ancestries remained stable during the Colonial period in Mexico, but shifted back toward a predominance of Late Archaic ancestry in the present-day Maya.

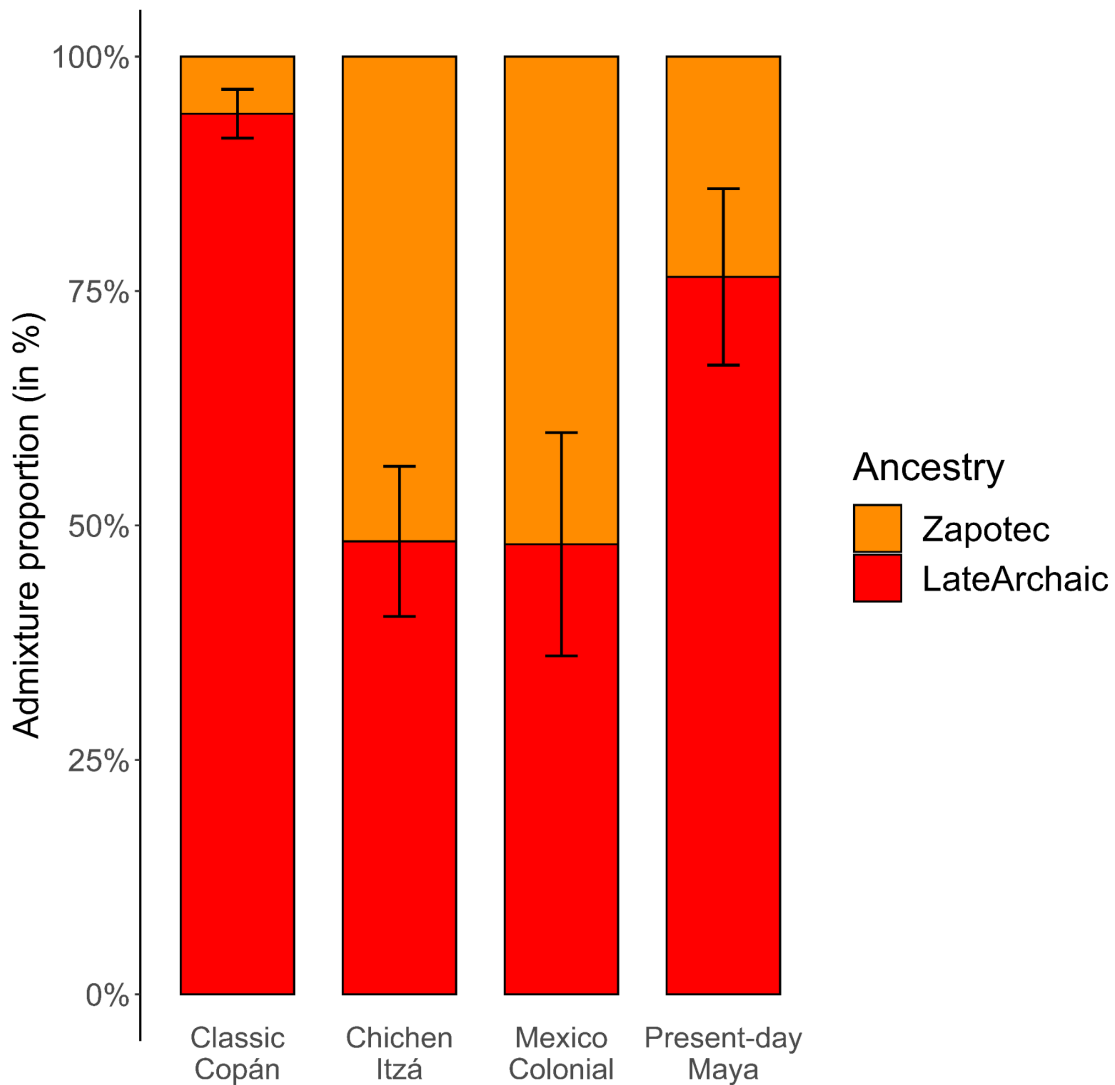


Figure 3.24. Genetic ancestry of ancient and present-day Maya populations. Bar plots show proportions of genetic ancestors in the Classic Copán, Chichén Itzá, Mexico Colonial, and present-day Maya populations. Colours represent different sources for admixture modelling: pre-existing ancestry in the Maya region represented by LateArchaic in red, and highland Mexico represented by Zapotec in orange. Vertical bars depict ± 1 standard error estimated by *qpAdm*.

VI. 2. Potential variation within the Classic Copán

The six individuals from the Classic Copán are divided into two temporal groups to examine temporal variation in genetic ancestry: the Early-to-Middle Classic period (CpM13, CpM12, and CpM4243) and the Middle-to-Late Classic period (CpM34, CpM0203, and CpM0506). The two-way admixture model of Late Archaic and Zapotec ancestries fits both temporal groups ([Table S3.6](#)). Admixture proportions are very similar between the groups, with 93.1% ($\pm 3.2\%$) Late Archaic ancestry in the Early-to-Middle group and 94.3% ($\pm 3.9\%$) in the Middle-to-Late Classic group. This indicates no temporal differences in ancestry proportions within the Classic Copán.

Given the hierarchical structure of ancient Maya society and isotopic evidence suggesting that CpM13 originated from outside Copán (Martínez de Velasco Cortina and Vega Villalobos, 2015; Price et al., 2010; Schele, 1992), I tested whether CpM13 exhibits distinct genetic ancestry compared to the commoners of the Classic Copán, based on $f_4(\text{Mbuti}, X; \text{CpM13}, \text{Commoners})$, where X represents the present-day Maya or potential sources of gene flow from Mesoamerica, including the Zapotec, Mixtec, and Mixe. None of the f_4 -statistics support significant differences between CpM13 and the commoners, with Maya ($Z = 0.666$), Zapotec ($Z = 0.051$), Mixtec ($Z = -0.145$), and Mixe ($Z = 0.949$).

To further investigate the influx of Zapotec-like ancestry at individual levels, I applied the two-way admixture model to each individual from the Classic Copán ([Table 3.4](#)). Most individuals from the Classic Copán are explained solely by Late Archaic ancestry, with two exceptions: CpM13, the dynastic ruler, and CpM12, his sacrificial companion.

Table 3.4. Testing fittings of the two-way admixture model between LateArchaic and highland Mexican ancestry to each individuals from the Classic Copán

Individual	Admixture proportion for source 1	Admixture proportion for source 2	Tail prob for dual model	Tail prob for source 1 (LateArchaic)	Tail prob for source 2 (Zapotec)	Nested <i>p</i>-value against source 1
CpM0203	95.5% ($\pm 8.2\%$)	4.5% ($\pm 8.2\%$)	0.048	0.096	0.003	-nan
CpM0506	98.4% ($\pm 2.9\%$)	1.6% ($\pm 2.9\%$)	0.950	0.973	1.253e-04	-nan
CpM12	93.0% ($\pm 4.9\%$)	7.0% ($\pm 4.9\%$)	0.371	0.250	1.204e-04	0.086
CpM13	84.3% ($\pm 14.4\%$)	15.7% ($\pm 14.4\%$)	0.414	0.263	0.097	0.071
CpM34	95.9% ($\pm 8.0\%$)	4.1% ($\pm 8.0\%$)	0.838	0.872	0.034	0.591
CpM4243	96.3% ($\pm 5.7\%$)	3.7% ($\pm 5.7\%$)	0.832	0.857	0.001	0.513

Discussion

This integrated analysis of newly generated Classic Copán genomes, in comparison with published ancient and present-day American genomes, uncovers genetic changes during the Classic Maya period. The population structure analyses of MDS, ADMIXTURE, and f_4 -statistics (Figs. 3.17, 3.19, 3.21 and 3.23) strongly support genetic continuity in the Maya region from the Late Archaic period to the present. However, this continuity was not exclusive, as time-stamped genomes from the region reveal the emergence of genetic ancestry associated with highland Mexican populations represented by the Zapotec, potentially from a region such as Oaxaca, beginning in the Early-to-Middle Classic period.

Our analysis shows that this gene flow extended to Copán (Fig. 3.24), located at the geographic periphery of the Classic Maya states, but was diluted relative to more northerly populations, such as that at Chichén Itzá. These findings highlight the role of mobility, potentially connected to the Maya elite network, in the formation of the Copán kingdom, as indicated from archeological evidence (Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015; Sharer, 2009; Söchtig et al., 2015). The strong genetic ties from the Late Archaic period through to the present support the amalgamation of local populations with incoming members of a greater Maya network as a strategy for expansion of Maya polities, as proposed by archaeological studies (Andrews and Fash, 2005).

Although the confidence intervals in Figure 3.24 suggest that Classic Copán could be modeled entirely with Late Archaic ancestry, the dual-ancestry model shows a much stronger statistical fit in Table S3.5. This significant dual-model, although largely of Late Archaic ancestry, may reflect heterogeneity within the Copán population. Indeed, the dual-model is well supported at the individual level for CpM12 and CpM13, and while three other individuals also show significant dual models, the single-source model is more strongly supported in these cases. With the current data, it is difficult to determine whether this individual-level variation reflects real biological structure or limitations due to low-coverage. However, the dual-model being clearly supported for the enigmatic ruler, CpM13, with a higher proportion of highland Mexican ancestry ($15.7\% \pm 14.4\%$) compared to commoners, the result possibly reflects social

stratification within the Maya society, implying the genetic origins of the Copán's royal lineage rooted outside the Maya region.

A foreign origin for the elite Maya of Copán is further supported by archaeological and linguistic evidence. Before the arrival of the first dynastic ruler around 400 CE, there is no clear sign of Maya culture in the Copán Valley: neither Maya-style architecture nor hieroglyphs inscriptions are present, and pottery from the Late Preclassic and Protoclassic periods resemble those from non-Maya regions of central Honduras and eastern El Salvador (Andrews and Fash, 2005). The abrupt appearance of Maya writing and material culture coincides with the reign of K'inich Yax K'uk' Mo', suggesting an elite intrusion from the Maya heartland. His names, and those of his successors, are Mayan, and writing itself was prestigious and restricted to the nobility, with scribes (often called "royal librarians") holding high status. Kings were sometimes even depicted with the distinctive headband of scribes, a bundle of pens tied to the forehead, signaling their training in the sacred art of writing (Sharer, 2009). This suggests that it was the elite who introduced and maintained Maya language and ideology in Copán, further reinforcing their distinct origins and sociopolitical role.

This cultural context potentially aligns with the dual-source ancestry found in CpM13, an early ruler of Copán and likely descendant of K'inich Yax K'uk' Mo' (Schele, 1992). CpM12 also supports a dual-source ancestry, although his status as a sacrificial companion leaves open the question of his social origin. The Maya may have practiced sacrifices on elite captives from rival city-states, believing their blood held greater ritual power. For example, the king's blood as a semi-divine mediator between the natural and supernatural worlds, practiced bloodletting ceremonies on himself (Martínez de Velasco Cortina and Vega Villalobos, 2015).

Interestingly, strontium isotopic ratios from CpM13 indicate non-local origin in his childhood, as was the case for the dynastic founder *K'inich Yax K'uk' Mo'*, both possibly coming from the Petén region (Price et al., 2014). Given that ruling power was passed from father to son (Martínez de Velasco Cortina and Vega Villalobos, 2015; Schele, 1992; Sharer, 2009), this potentially suggests that boys from higher social classes, including heir to the thrones, were sent outside of Copán, possibly to Tikal in the Petén region which was one of the major cities of the Classic Maya period (Fig. 3.1) and has been a historical ally of Copán (Andrews and Fash, 2005), as foster or wardship

children to strengthen political alliances. This strategy might explain why so many elite individuals, around 50%, were found to have grown up outside of Copán (Suzuki et al., 2020).

While the increase of highland Mexican ancestry in the Chichén Itzá compared with the Classic Copán may reflect temporal shifts in this ancestry, it could also indicate geographic variation. In particular, Chichén Itzá and Colonial Mexico, both from the Yucatán Peninsula in Mexico, exhibit similar proportions of highland Mexican ancestry. A key limitation of this study is the small sample size—only seven Classic Copán individuals, many of whom with low genomic coverage—which constrains the statistical power of admixture modelling (Harney et al., 2021). Moreover, the non-random nature of the sampling, focused on individuals from a specific residential or patio group, further limits the scope of inference. Denser spatial and temporal sampling from the Maya region, as well as Mesoamerica, is crucial to uncover genetic origins and demography of Classic Maya populations.

Materials and Methods

I. Individual processing, haplogroup determination and contamination validation

I. 1. Sampling and archeological contexts

Petrous bone fragments from a total of 16 fossils excavated from the sites of Copán in western Honduras were sampled and brought for processing at Kanazawa University in Japan. CpM0203, CpM0506, CpM11, CpM12 and CpM13 were excavated in the year 2000 as part of PICPAC (Programa Integral de Conservación del Parque Arqueológico Copán), and CpM34 and CpM4243 were excavated in 2007 and 2016 respectively as part of PROARCO (Proyecto Arqueológico Copán).

I. 2. DNA extraction

Bone sample processing was carried out in a dedicated ancient DNA facility at Kanazawa University where all the precautions for ancient DNA processing were followed as described in a previous study (Llamas et al., 2017). Bone samples were photographed extensively before further processing. All bones were exposed to ultraviolet light for 15 min on either side to remove surface contaminants. Further cleaning of their surface was performed with a drill before extraction. The otic capsule region of the petrous temporal bone was targeted for sampling. An aliquot of 40-100 mg of the bone powder was subjected to a silica column method with an initial washing step by predigestion solution (Damgaard et al., 2015). DNA extracts were purified with MinElute silica columns (QIAGEN) and eluted at a volume of 30 µl of the elution buffer.

I. 3. Sequencing and damage assessment

The initial screening of each individual and blank controls was performed by constructing a double-stranded DNA next-generation sequencing library using a NEBNext Ultra II DNA Library Prep Kit for Illumina with bead-based size selection. The 16 initial bone extracts were screened on a MiSeq Illumina platform. The authenticity of ancient DNA was determined in non-UDG-treated libraries by looking at

degradation patterns at 5'- and 3'-end using mapDamage2.0 (Fig. 3.3) (Jónsson et al., 2013).

The 10 extracts with endogenous contents between 1% and 17% were selected for high-coverage sequencing and subsequently treated with uracil-DNA-glycosylase (UDG) before a second library construction to reduce the impact of DNA degradation and damage (Rohland et al., 2015). High-coverage sequencing for UDG-treated libraries was carried out on a NovaSeq 6000 Illumina platform (150-bp paired-end reads) at Macrogen (Republic of Korea).

I. 4. Sequence data processing

Adapters were trimmed using *AdapterRemoval* version 2.2.2 (Schubert et al., 2016), with the following parameters: `--collapse --minadapteroverlap 1 --minlength 25 --minquality 25 --trimqualities`. Trimmed reads were aligned to the hs37d5 human reference genome using *BWA* (Burrows-Wheeler Aligner) version 0.7.5 (Li and Durbin, 2009), with relaxed parameters, `"bwa aln -l 16500 -n 0.01 -o 2."` We used *picard-tools* version 2.22.1 (<https://broadinstitute.github.io/picard/>) to add reads groups, remove PCR duplicates, and merge all different libraries sequenced on MiSeq and NovaSeq platforms into a single BAM per individual. The reads were further filtered for length longer than 34 bp and mapping quality higher than 25. Two final processing steps were carried out on the merged files: indels were realigned using Genome Analysis Toolkit (GATK) version 3.7-0 (McKenna et al., 2010), and the quality of the two bases at the end of each read was manually reduced to a score of 2 ("softclipped"), to avoid calling genotype from these damage-prone regions. Genomic coverage and depth were measured using *qualimap* version 2.2.1 (Okonechnikov et al., 2016).

I. 5. Relatedness with READ

Kinship using READ (Monroy Kuhn et al., 2018) was estimated for the ten samples retained for high-coverage sequencing. Three pairs of samples were identified as genetically identical (CpM02-CpM03, CpM05-CpM06, CpM42-CpM43; Fig. S3.1) and were subsequently merged into single individuals as CpM0203, CpM0506 and CpM4243 respectively with *picard-tools* version 2.22.1. READ was then re-ran for the seven Classic Copán individuals to control for first- and second-degree relatives, which

led to CpM11 being removed from some analysis as a possible relative to CpM34 and CpM4243 (Fig. S3.2).

I. 6. Molecular sex determination

Molecular sex was determined after filtering reads with a minimum mapping quality of 30, and calculating the ratio of Y-chromosome reads to the total number of sex chromosomes (R_y). Following a previous study, the individual was assigned female if $R_y < 0.016$ or male if $R_y > 0.075$ (Skoglund et al., 2013).

I. 7. Contamination of individuals based on Y-chromosome

Because males possess a Y-chromosome, they only have one X-chromosome with a single allele at each site, so any polymorphisms would translate to reads derived from sequencing errors or contamination. There are two methods for calculations: they both compare the number of major reads (allele with the highest frequency for a given site) to the number of minor reads (all other alleles different to the highest frequency allele for a given site), but the first assumes that the error rate for each site is independent while the second removes this assumption (Rasmussen et al., 2011). This approach is integrated into ANGSD (Analysis of Next Generation Sequencing Data) version 0.928 (Korneliussen et al., 2014; popgen.dk/angsd/index.php/Contamination) which produced results in Table 3.2.

I. 8. Contamination of individuals based on mitochondrial DNA

Contamination rates based on mitochondrial DNA were estimated as described in previous methods (Cassidy et al., 2016; Sánchez-Quinto et al., 2012). Briefly, contamination was based on the rate of secondary bases that did not match the consensus sequence at haplotype-defining and private mutations defined by *HaploGrep*.

I. 9. Y-chromosome haplogroup determination

All newly sequenced individuals determined to be male were piled up to the 2019 version of the International Society of Genetic Genealogy database (ISOGG) using GATK version 3.7-0 (McKenna et al., 2010), with a mapping quality of 20 and a base quality of 30.

I. 10. Mitochondrial haplogroup determination

To assign mitochondrial (mt) haplogroups, sequence reads uniquely mapped to mtDNA were extracted from BAM files using *SAMtools* (version 1.7) (Danecek et al., 2021) to avoid biases from nuclear-mitochondrial segments (NUMTs), and re-aligned to the revised Cambridge reference sequence using *BWA* (version 7.5a) (Li and Durbin, 2009). Consensus sequences were determined to mapping quality higher than 25, a minimum depth of consensus of 5, and a base quality of 30 using *mpileup* and *bcftools* within the *SAMtools* package. A specific haplotype was then assigned to each consensus sequence using *HaploGrep* (version 2.2.9) (Weissensteiner et al., 2016), listed in [Table 3.1](#).

II. Dataset preparation: pileup and masking

II. 1. Reads trimming

The processing pipeline for newly-sequenced Copán individuals was applied for published ancient genomes with the exception of using cutadapt version 1.15 (Martin, 2011) instead of AdapterRemoval. Published sequences were downloaded from the European Nucleotide Archive (ENA) as FASTQ files. If FASTQ files were not available, BAM files were realigned to the reference genome and then put through the same pipeline.

Adapters were detected in only two of the published studies included in this analysis: one from 2014 (Raghavan et al., 2014a), and one from 2015 (Raghavan et al., 2015) affecting only some files for Chinchorro individual. The adapter sequence was obtained from supplementary materials and adapters were subsequently trimmed as follow:

```
cutadapt --adapter CAAGCAGAAGACGGCATACGAGAT \  
  --quality-base 33 \  
  --quality-cutoff 2 \  
  --minimum-length 25 \  
  --format fastq ${individual[i]}.fastq.gz \  
  --output ${individual[i]}.trimmed25bp
```

Quality assessment of raw data performed using FastQC v0.11.4 (Andrews, 2010) confirmed the absence of adapters in individuals from the 2014 study (Fig. 3.25), whereas adapters were still detected in the Chinchorro individual. Additional trimming was carried out using AdapterRemoval version 2.3.1 (Schubert et al., 2016), however, adapters remained present (Fig. 3.26).

```
AdapterRemoval --adapter1 CAAGCAGAAGACGGCATACGAGAT \  
  --qualitybase 33 \  
  --minquality 2 \  
  --minlength 25 \  
  --trimns --trimqualities \  
  --file1 ${individual[i]}.fastq.gz --gzip \  
  --basename ${individual[i]}.trimmed25bp
```

Since BAM files were not publicly available for this study, Chinchorro individual for population MummyMaderasEnco2 was generated by merging FASTQ files that were pre-trimmed, and ignoring FASTQ files that still contained the adapters.

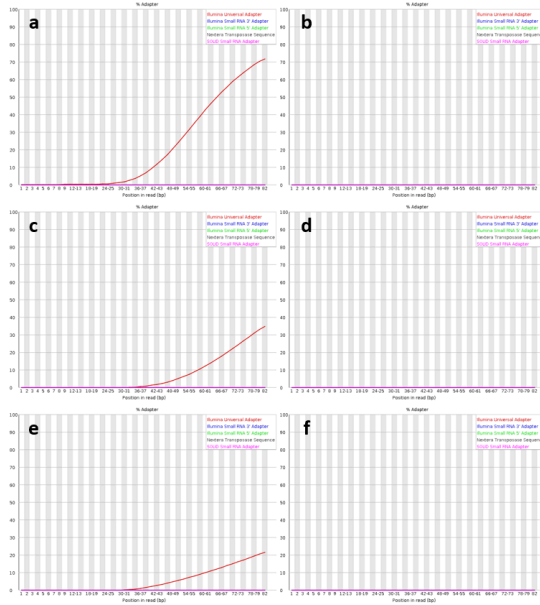


Figure 3.25. Examples of FastQC reports of adapter contents for individuals from Raghaven *et al.*, 2014 (a, c and e) Before cutadapt and (b, d and f) after cutadapt was applied. Raw files of Alarnerk3_lib1 (a-b), MARC1491_lib2 (c-d) and XIV_H126_lib4 (e-f) as examples.

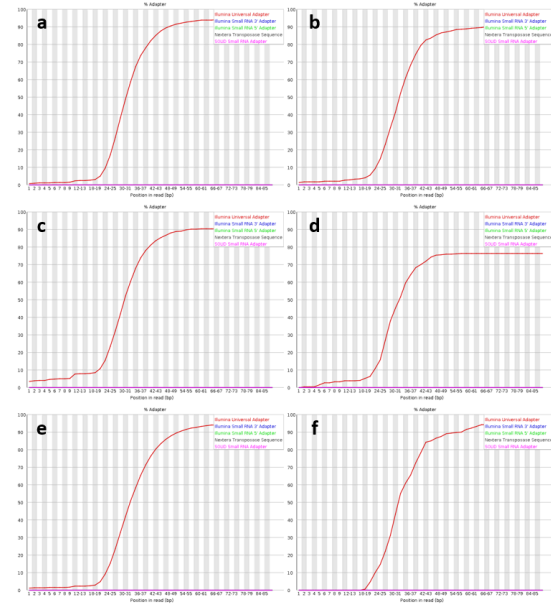


Figure 3.26. Examples of FastQC reports of adapter contents for Chinchorro from Raghaven *et al.*, 2015 (a, c and e) After cutadapt and (b, d and f) after AdapterRemoval were applied. Raw files Chinchorro_13 (a-b), Chinchorro_17 (c-d) and Chinchorro_20 (e-f) as examples.

II. 2. Pileup of ancient with present-day genomes

I created a large panel that includes newly sequenced Classic Copán genomes and publicly available ancient genomes from Siberia and the Americas (Fig. 3.15 and Table S3.4). Individuals previously reported as related or suspected of contamination were removed. All ancient individuals were genotyped for the biallelic SNP sites present on two different reference panels, to which they were subsequently merged: (i) the Human Origin Array (HOA) consisting of 594,896 SNP sites for 1,963 present-day, ancient and reference genomes (Lazaridis *et al.*, 2014), and (ii) the Simon Genome Diversity Project (SGDP) consisting of 278 present-day, ancient and reference genomes (Mallick *et al.*, 2016) filtered beforehand for autosomal transversion-only SNPs with a

minor allele frequency of 1%, which left 3,867,550 SNP sites in the merged data. For each SNP site, a high-quality single base (bq30) was randomly called per position to create a pseudo-haploid genotype using GATK version 3.7-0 (McKenna et al., 2010).

II. 3. Masking with MOSAIC

MOSAIC, or MOSAIC Organises Segments of Ancestry In Chromosome, was developed by Michael Salter-Townshend and Simon Myers in 2019. For each target haplotype, the software employs hidden Markov models to create a matrix of probabilities of possible ancestries (Salter-Townshend and Myers, 2019). MOSAIC was applied to present-day populations from the HOA or SGDP datasets. While MOSAIC does not require prior knowledge on source ancestry, it relies on a predefined set of reference populations (Salter-Townshend and Myers, 2019). In the original, worldwide populations from the SGDP dataset were divided into 6 regions, but I decided to further divide “West Eurasia” between West Eurasia and Middle East as this group would have otherwise consisted of at least twice the number of populations from the other regions. Additionally, I divided “East Asia” into East Asia and Siberia, following the genetic clade identified by the authors (Mallick et al., 2016). Worldwide present-day populations were therefore divided into African, American, Siberian, East Asian, Oceanian, South Asian, West Eurasian and Middle Eastern groups based on geography: the Middle East span over the Arabian peninsula with the frontier between Egypt and Israel marking its border with Africa, the Bosphorus Strait marking its border with West Eurasia, the Caucasus mountains marking its border with Siberia, and the frontier between Afghanistan and Pakistan marking its border with South Asia; while East Asia differentiated from South Asia with the frontier between India and Myanmar, from Siberia with the frontier between China and Mongolia, and from Oceania with the Banda sea. With this border outline, the Dusun and Igorot people from Malaysia and the Philippines were therefore grouped as East Asians rather than Oceanians, in accordance with their genetic clustering (Mallick et al., 2016). Two exceptions were made: although Tlingit originate from Alaska they were grouped as Siberians instead of Americans for their genetic grouping with other Siberian populations (Mallick et al., 2016), and African-Americans from Denver were grouped as Africans instead of Americans.

Three ancestries were assumed in present-day Americans, expecting a main American ancestry, and two minor West Eurasian and African ancestries originating from the XVth century colonisation. MOSAIC was executed separately for each target American population using an iterative approach, where non-target American populations were designated as “American,” with the following command:

```
Rscript MOSAIC/mosaic.R \
    ${american_pop} mosaic_input \
    --ancestries 3 \
    --number ${nb_ind_in_pop} \
    --panels "Africa America Siberia EastAsia SouthAsia
WestEurasia MiddleEast"
```

For running MOSAIC on SGDP populations, only SNPs intersecting with sites in the HOA dataset were utilised, resulting in 548,131 genomic sites for SGDP and 594,896 for HOA.

To obtain the physical coordinates of non-American segments for masking, grid points based on recombination distances were converted to genomic positions with the “grid_to_pos()” function in the MOSAIC package in R (Salter-Townshend and Myers, 2019). These coordinates were masked in each corresponding individual using *plink* (version v1.90b6.22), with the options *--keep* to extract the individual, *--zero-cluster* and *--within* to set the non-American sites to missing, and *--merge-list* to pileup all Americans back to the dataset (Purcell et al., 2007). To classify as an American segment, both chromosomes of the individual had to have been identified as American, otherwise it was masked.

III. Genome-wide analysis

III. 1. Principal Components Analysis (PCA)

PCA was conducted using smartpca (v18140) from the EIGENSOFT package (v8.0.0) (Patterson et al., 2006). All ancient Siberians and Americans were projected onto present-day Americans in the HOA dataset (Pima, Mixe, Mixtec, Zapotec, Maya, Bolivian, Quechua, Piapoco, Karitiana, and Surui). The options specified were as follows to adjust for the projection of ancient data:

- “killr2: NO” to eliminate SNPs in linkage disequilibrium with nearby SNPs
- “numoutlieriter: 0” to prevent the removal of ancient individuals
- “lsqproject: YES” to help the projection of data with high missingness
- “autoshrink: YES” to stretch the axes of computed compared to non-computed samples.

III. 2. Multidimensional scaling (MDS)

MDS was conducted based on outgroup f_3 -statistics measured for all combinations of ancients and present-day American populations from the SGDP panel, with Mbuti as the outgroup, using AdmixTools v6 package (Patterson et al., 2012). Populations were excluded from MDS analysis if fewer than 100 SNPs were included in any of the tests performed for a particular population. Additionally, to improve the resolution of population structure in the Americas in MDS, specific ancient individuals from Siberia and the Arctic were excluded from the analysis (AG2, MA1, Late Dorset, and Middle Dorset).

III. 3. ADMIXTURE

Unsupervised genetic clustering was conducted by ADMIXTURE version 1.23 (Alexander et al., 2009) for ancient and present-day American individuals with at least 10,000 SNPs in the HOA panel. SNPs were pruned for sites in linkage disequilibrium using PLINK v1.90b4.4 (Purcell et al., 2007), with a sliding window size of 50 variants, a step size of 5 variants, and an r^2 threshold of 0.2 (--indep-pairwise 50 5 0.2), leaving 184,489 SNPs for analysis. Ten replicates with random seeds were run for each of the clusters (K) from 2 to 13. The run with the highest likelihood across replicates was

plotted. Results from $K = 9$ were plotted on a geographical map using `mapmixture` implemented in R (Jenkins, 2024).

III. 4. Estimated effective migration surface (EEMS)

To visualise population migration in the context of geographical barriers in the Americas, I used the software EEMS, which models genetic differentiation in space by highlighting regions of higher-than-average and lower-than-average gene flow on a grid (Petkova et al., 2016). The surface of the American continent was captured as a regular grid, setting demes to 1000, where each sample is assigned to the closest demes. Pairwise dissimilarity matrix was generated with the program `bed2diffs v1`. EEMS was run with the version *run_eems_snps* specific for SNP data, with tuned parameters so that the proposals are not accepted too rarely (less than 10%) or too often (more than 40%): $q_{\text{EffectProposalS2}} = 0.013$, $q_{\text{SeedsProposalS2}} = 1.3$, $m_{\text{EffectProposalS2}} = 2$, $m_{\text{SeedsProposalS2}} = 0.2$, and $m_{\text{rateMuProposalS2}} = 0.005$. For each run, the convergence of the MCMC chain was checked for convergence, which was the case after 8,000,000 iterations and 7,000,000 burn-in. The run was repeated three times, with random seeds 123, 456, and 789 respectively (Fig. S3.7). The run with the highest likelihood was selected for the EEMS figure in the main text, corresponding to run number 1.

III. 5. F -statistics

I used `qpDstat` (version 900) on the merged dataset of ancient individuals and the SGDP panel of present-day populations, with the f_4 mode in the `AdmixTools v6` package (Patterson et al., 2012). The African population, Mbuti, was specified as an outgroup, and standard errors in f_4 -statistics were calculated from the default jackknife approach. Statistical significance in genetic affinity between populations was defined as Z -scores greater than 3.0, with more than 1,000 SNPs used for the test.

III. 6. Admixture modelling by `qpAdm`

To model admixture events in the Classic Copán and other ancient and present-day Maya populations, I applied `qpAdm v1000` in the `AdmixTools v6` package (Patterson et al., 2012) to the merged data of ancient and SGDP populations, coupled

with the option of “allsnps: YES”. A set of 13 ancient and present-day populations was used as an outgroup in the modelling: Han ($n=3$), Pima ($n=2$), Mixe ($n=3$), Karitiana ($n=3$) (Mallick et al., 2016), MA1 ($n=1$) (Raghavan et al., 2014b), USR1 ($n=1$) (Moreno-Mayar et al., 2018a), ASO_4000 ($n=1$) (Scheib et al., 2018), Sumidouro5 ($n=1$), Ayayema ($n=1$), PuntaSantaAna ($n=1$), SpiritCave ($n=1$) (Moreno-Mayar et al., 2018b), Anzick ($n=1$) (Rasmussen et al., 2014), and Taino ($n=1$) (Schroeder et al., 2018). When Mixe was tested as a source ancestry, they were replaced by Zapotec ($n=2$). An admixture model with a tail probability greater than 0.05 was considered as a fit.

Chapter 4

Maximising genomic insights from low-coverage ancient genomes: Performance of GLIMPSE and the promise of double-imputation pipelines

Abstract

Accurate genotype imputation is critical for maximising the information retrieved from low-coverage ancient genomes. This chapter evaluates the performance of GLIMPSE, a state-of-the-art imputation tool designed for low-coverage data, on ancient American genomes. Using two high-coverage individuals, USR1 (11,500 years old, Alaska) and Sumidouro5 (10,400 years old, Brazil), imputation was assessed across genomic coverages ranging from $0.01\times$ to $5.0\times$ using metrics of concordance, accuracy, and SNP recovery.

GLIMPSE consistently achieved concordance and accuracy rates exceeding 99% at homozygous sites across all coverages. At heterozygous sites, accuracy significantly improved between $0.01\times$ and $0.05\times$, reaching a plateau at approximately 99% from $0.1\times$ onwards. However, transitions in Sumidouro5 exhibited lower concordance and accuracy, likely due to *post-mortem* DNA damage. SNP recovery demonstrated a steady increase across all site categories until plateauing at $0.5\times$ coverage, with heterozygous sites showing a threefold increase in recovery between $0.1\times$ and $0.5\times$. Building on these findings, I propose a novel double-imputation pipeline that integrates GLIMPSE with Beagle5, aiming to further enhance SNP recovery and imputation performance.

Contributions:

The double-imputation pipeline combining GLIMPSE and Beagle5, explored in this project, was originally developed by Dr. Lara Cassidy. This approach has been applied to ancient genomes from Europe (Cassidy et al., 2025). My research extends its application to ancient American genomes, demonstrating its broader potential for studying diverse ancient populations.

Introduction

The study of ancient DNA (aDNA) provides a powerful lens into the genetic history of past populations, shedding light on migration patterns, population structures, and evolutionary trajectories. This knowledge influx has been a direct result of the integration of next-generation sequencing technologies with improved methods for recovering DNA from ancient samples (Nägele et al., 2022). However, this field faces technical and biological challenges due to the fragile nature of DNA preserved in ancient remains. Over time, aDNA undergoes significant *postmortem* degradation, characterised by fragmentation, contamination from environmental sources, and deamination damage (Nägele et al., 2022; Orlando et al., 2015).

Deamination replaces the amino group on cytosine with a keto group, converting it to uracil. During amplification in library preparation, DNA polymerase pairs uracil with adenine instead of guanine, which results in C-to-T or G-to-A transition mutation in data. This error can be experimentally addressed by using a Uracil-Specific Excision Reagent (USER) Enzyme, as applied to the Classic Copán genomes. However, this correction is not uniformly effective across the genome. Consequently, transition sites are often excluded in ancient DNA studies to avoid spurious results, which in turn significantly reduces the number of sites available for analysis, as transitions account for between half and two-thirds of the genome (Do and Dobrovic, 2012; Orlando et al., 2015).

Additionally, contamination—whether from modern human sources or environmental microbes—further reduces the proportion of endogenous DNA retrievable from ancient remains. Combined with extensive fragmentation and deamination, these factors pose significant challenges to recovering high-quality genomic data. As a result, the majority of ancient genomes are sequenced at low coverage (Garrido Marques et al., 2024; Sousa Da Mota et al., 2023). This limitation significantly hinders confident genotype calling and the ability to accurately infer heterozygosity, a key measure for understanding genetic diversity and evolutionary history. Due to these challenges, ancient DNA studies often rely on pseudo-haploid data, where one allele per variant site is randomly sampled to create a simplified representation of the genome. While this approach has been proven to be effective in

understanding the human past, it precludes the use of tools designed for diploid data, which leverage heterozygosity and haplotypic information to explore demographic histories, kinships, and local selection pressures (Kousathanas et al., 2017; Sousa Da Mota et al., 2023). While sequencing at higher coverage certainly mitigates these issues, it comes with two major drawbacks: increased financial costs and the need to destroy more of the precious archaeological material. Another solution is to computationally increase genomic coverage through imputation, a method that is cost-effective and preserves valuable archeological materials by maximising low-coverage ancient genomes.

Genotype imputation is a technique that infers missing genotypes in low-coverage genomes or partially genotyped datasets by leveraging linkage disequilibrium (LD) and haplotype sharing within a reference panel of fully sequenced genomes (De Marino et al., 2022; Sousa Da Mota et al., 2023). By comparing the observed genotypes of a study sample, containing many missing sites, to a reference panel of fully sequenced genomes, shared haplotypes can be identified. This enables the reconstruction of haplotypes in the sample as mosaics of short segments from the reference panel, allowing the inference of unobserved genotypic sites and the re-creation of diploid data (Das et al., 2018; De Marino et al., 2022). Most imputation tools employ the Li and Stephen's hidden Markov Model (HMM), which underpins this process by estimating the probability of observing a haplotype as a combination of haplotype segments from the reference panel. This model accounts for recombination events and potential mutations, ensuring accurate reconstruction of missing genetic information (Das et al., 2018; Sousa Da Mota et al., 2023).

Initially, imputation was applied in genome-wide association studies (GWAS) to inexpensively approximate whole-genome sequence data using genome-wide single-nucleotide polymorphism (SNP) array data (Das et al., 2018). However, it has been relatively recently adapted for ancient genomics to reconstruct diploid genomes from low-coverage data (Clemente et al., 2021; Gamba et al., 2014; Jones et al., 2015; Margaryan et al., 2020; Martiniano et al., 2017; Saupe et al., 2021).

To maximise the inference of missing haplotypes in low-coverage ancient genomes, a two-step imputation has been proposed. This approach involves estimating genotype likelihoods by Beagle4, removing sites with lower genotype probabilities, and

re-calling those sites with Beagle5, followed by final genotype probability filtering (Hui et al., 2020). In the next chapter, I utilise a double-imputation pipeline combining GLIMPSE and Beagle5. GLIMPSE is specifically designed for use with low-coverage genomes (Rubinacci et al., 2021).

A key challenge in the use of imputation for ancient genomes is the reliance on reference panels derived from present-day populations, highlighting potential concerns about the representation of ancient populations, in particular those populations that do not leave their genetic legacy in modern descendants. However, previous studies have demonstrated that ancient humans, including Indigenous Americans, can be accurately imputed (Sousa Da Mota et al., 2023). Here, my research focuses on the validation of imputation by GLIMPSE, which constitutes the first step of the double-imputation pipeline in the context of ancient genomes from the Americas, an underrepresented continent in reference panels.

Results

To assess the performance of GLIMPSE in reconstructing diploid genotypes, genotypes called from original high-coverage genomes are compared with those imputed from downsampled low-coverage genomes. Performance was evaluated with two different metrics of concordance and accuracy, previously employed in a study (Cooke et al., 2024), where I was involved as a co-author. Concordance measures the proportion of sites in the original diploid data successfully recovered through imputation, while accuracy assesses the proportion of imputed sites that are validated against the original diploid data. These metrics were measured for mutation types (*i.e.*, transitions and transversions) and genotypes (*i.e.*, homozygous reference, homozygous alternative, and heterozygous sites).

The imputation performance was tested on two high-coverage ancient genomes: USR1, a 11,500-year-old individual from Alaska (Moreno-Mayar et al., 2018a), and Sumidouro5, a 10,400-year-old individual from Brazil (Moreno-Mayar et al., 2018b). Given USR1, a representative of the Ancient Beringians—the earliest peoples of the Americas—, and Sumidouro5, as a descendant of the early migration wave to South America, these are well-suited for assessing imputation performance in the context of ancient American genomes.

Both individuals displayed high concordance for homozygous reference and homozygous alternative sites across varying coverages, achieving over 99.85% concordance for both transitions and transversions (Fig. 4.1). However, concordance in heterozygous sites varied with coverages and mutation types.

For USR1, heterozygous concordance at 0.01× coverage was 56.1% for transversions and 34.6% for transitions. At 0.05×, this rate significantly increased to 96.6% for transversions and 90.1% for transitions, which continued to improve with the increase of coverage and plateaued at 99.85% for transversions and 98.7% for transitions.

For Sumidouro5, the difference in the concordance between transitions and transversions was more pronounced compared with that observed in USR1. Heterozygous concordance started lower, at 68.5% for transversions and 4.6% for transitions at 0.01× coverage. At 0.05×, concordance improved to 96.7% for

transversions and 40.8% for transitions, and further to 98.8% for transversions and 62.8% for transitions at 0.1 $\times$. A plateau was reached at approximately 99.8% for transversions and 86.9% for transitions at 2 $\times$ coverage.

Concordance for heterozygous transversions was consistently higher than for transitions, achieving over 95% from as low as 0.05 $\times$ coverage. By contrast, heterozygous transitions reached 90% concordance for USR1 but only 40% for Sumidouro5 at the same coverage.

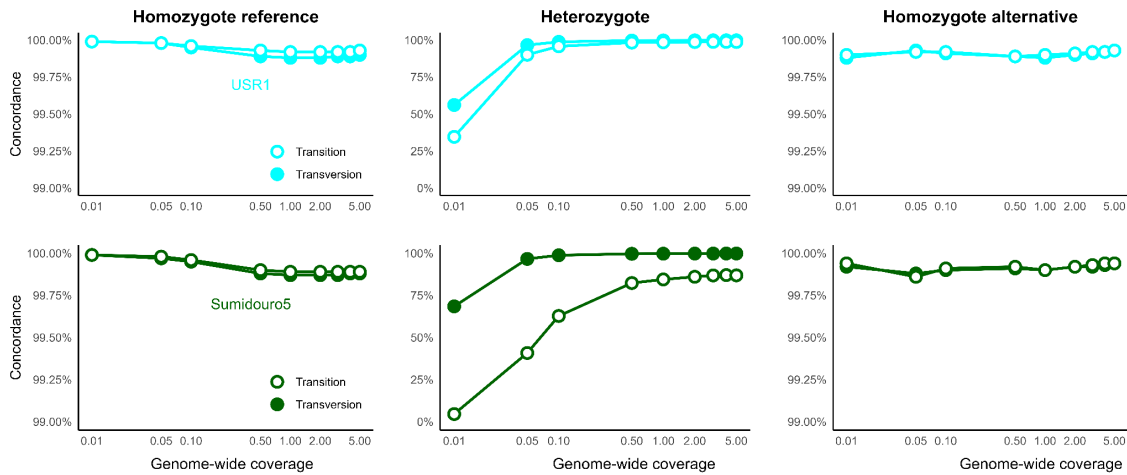


Figure 4.1. Assessing the performance of GLIMPSE imputation through concordance. The ability to impute ancient Americans was assessed by downsampling two individuals, USR1 in blue (top) from Alaska and Sumidouro5 in dark green (bottom) from Brazil, to 0.01 $\times$, 0.05 $\times$, 0.1 $\times$, 0.5 $\times$, 1 $\times$, 2 $\times$, 3 $\times$, 4 $\times$ and 5 $\times$ and imputing them to higher coverage. Performance was assessed through concordance rate, which was defined as rate at which diploid sequence data was correctly re-created through imputation. Results were divided into variant site categories (homozygous reference, heterozygous, and homozygous alternative), and by mutation type (transition, as represented by open circles, and transversion, represented by closed circles). The scale for y-axis is adjusted to site categories, but do not vary within individuals to facilitate interpretation.

Accuracy is overall higher than concordance, supporting the idea that imputed genotypes with genotype probabilities over 99% are likely to be correct. USR1 exhibited consistently high performance across coverages and genotypes, with accuracy over 99% from as low as 0.01 $\times$ coverage (Fig. 4.2). It is important to note that, different from the concordance results, accuracy is comparable between transitions and transversions.

In Sumidouro5, accuracy for homozygotes exceeded 94.6%, with transversions achieving 99.8%. Similarly, heterozygous sites exhibited high accuracy, but with no noticeable differences between transitions and transversions. At 0.05 $\times$, transitions achieved an accuracy of 98.8%, slightly outperforming transversions at 98.5%. By 0.1 $\times$, accuracy for transitions and transversions reached a plateau of approximately 99.1% and 98.8%, respectively.

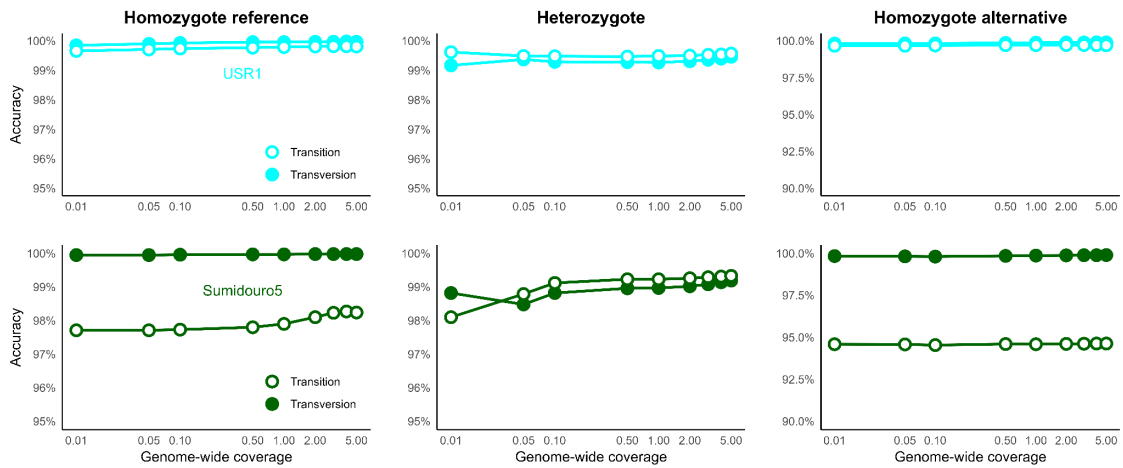


Figure 4.2. Assessing the performance of GLIMPSE imputation through accuracy. The ability to impute ancient Americans was assessed by downsampling two individuals, USR1 in blue (top) from Alaska and Sumidouro5 in dark green (bottom) from Brazil, to 0.01 $\times$, 0.05 $\times$, 0.1 $\times$, 0.5 $\times$, 1 $\times$, 2 $\times$, 3 $\times$, 4 $\times$ and 5 $\times$ and imputing them to higher coverage. Performance was assessed through accuracy, which was defined as how often imputed data accurately matched the sequence data. Results were divided into variant site categories (homozygous reference, heterozygous, and homozygous alternative), and by mutation type (transition, as represented by open circles, and transversion, represented by closed circles). The scale for y-axis is adjusted to site categories, but do not vary within individuals to facilitate interpretation.

Whilst concordance and accuracy were high, it is equally important to consider the number of sites successfully recovered through imputation. To evaluate this, the number of SNPs matching the original diploid data and the imputed downsampled data was counted across coverage levels for USR1 and Sumidouro5 (Fig. 4.3). As expected, the number of transition sites recovered was significantly higher than for transversion sites.

For homozygous reference sites, the number of recovered SNPs increased steadily until plateauing at approximately 2.1 million for transversions and 4.3 million for transitions in USR1, and 1.6 million and 2.9 million in Sumidouro5. In homozygous

alternative sites, a sharp increase was observed up to $0.5\times$ coverage, after which SNP recovery plateaued at around 3.2 million for transitions and 6.7 million for transitions in USR1, and at 1.8 million and 3.1 million in Sumidouro5. For heterozygous sites, the number of recovered SNPs increased approximately threefold between $0.1\times$ and $0.5\times$ coverage, before reaching a plateau at 3.8 million for transversions and 7.6 million for transitions in USR1, and 2.3 million and 4.6 million in Sumidouro5. Taken together, these results indicate that while imputation performed with high concordance and accuracy, particularly from as low as $0.05\times$ coverage, the number of recovered sites is fully maximised when the coverage is equal to or higher than $0.5\times$.

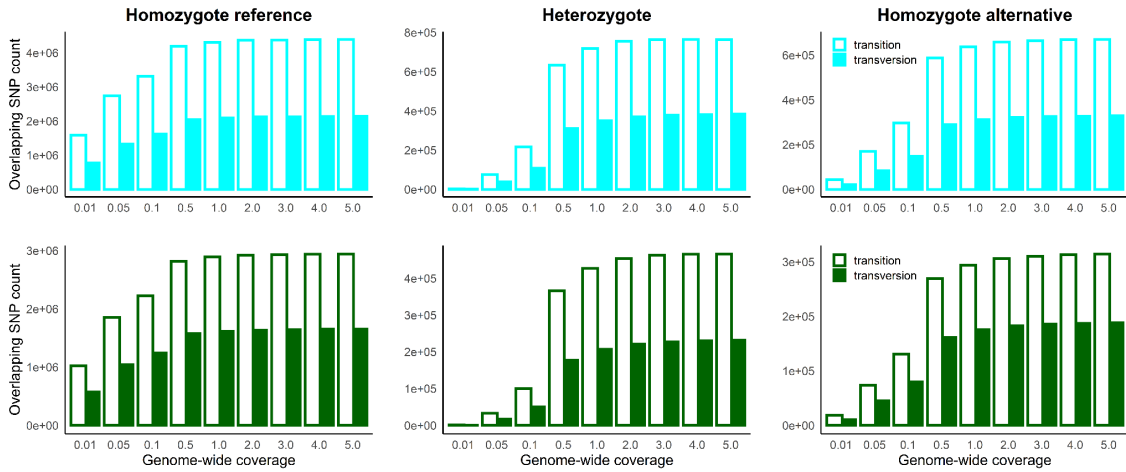


Figure 4.3. Overlapping SNP sites produced by imputation. A bar plot representing the total number of SNP sites, divided by variant site categories (homozygous reference, heterozygous, and homozygous alternative), and by mutation type (white fill for transition and coloured fill for transversion). Imputation was assessed by downsampling two individuals, USR1 from Alaska in blue and Sumidouro5 from Brazil in dark green, to $0.01\times$, $0.05\times$, $0.1\times$, $0.5\times$, $1\times$, $2\times$, $3\times$, $4\times$ and $5\times$ and imputing them to higher coverage.

Discussion

This chapter demonstrates the robust performance of imputation using GLIMPSE for ancient American genomes, even at extremely low genomic coverage. Concordance and accuracy in homozygous sites consistently achieved 99% across all genomic coverages as low as 0.01×, with the exception of transitions in Sumidouro5 (Fig. 4.1; Fig. 4.2). However, heterozygous sites, which are critical for understanding genetic diversity and evolutionary history, revealed more nuanced patterns. Concordance in heterozygous sites improved markedly between 0.01× and 0.05× and reached a plateau from 0.1× onwards, except for transitions in Sumidouro5. Accuracy in heterozygous sites similarly plateaued at approximately 99% from 0.1×. Similar patterns were found in ancient Asian individuals, with a significant improvement in concordance between 0.01× to 0.05× going from below 20% to over 80% for heterozygous sites and plateauing around 99% from 0.1×. Likewise, accuracy in heterozygous sites plateaued from 0.1× coverage, around 99% (Cooke et al., 2024). These findings assure the power of GLIMPSE to recover missing genotypes from low-coverage ancient genomes.

The number of SNP sites recovered by imputation steadily increases until plateauing at 0.5× coverage across all genotypes, with heterozygous sites exhibiting a threefold improvement and homozygous alternative sites a twofold improvement between 0.1× and 0.5× (Fig. 4.3). This indicates that the reliability of imputed genotypes is confidently high even at coverages as low as 0.1×. However, a coverage of at least 0.5× is required to recover a comprehensive set of data.

The performance for transitions in Sumidouro5 was noticeably lower than that in USR1, with reduced concordance in heterozygous sites (Fig. 4.1) and reduced accuracy in homozygous sites (Fig. 4.2). Similar patterns have been reported in previous studies of imputation in ancient American genomes, with Sumidouro5 consistently showing lower performance for transitions (Garrido Marques et al., 2024; Sousa Da Mota et al., 2023). This can be possibly attributed to *post-mortem* DNA damage. Sumidouro5, in particular, is known to be a heavily damaged genome, with a deamination rate of approximately 40% at read termini (Moreno-Mayar et al., 2018b). This high rate of

damages likely caused an excess of heterozygous calls in the high-coverage reference data, even after quality filtering. In such cases, the observed genotypes used as comparison for imputation may themselves be affected by damage and thus incorrect. The differences in transitions for Sumidouro5 suggest that imputation might have successfully inferred the correct genotypes, even when the observed calls were compromised by damages. This may highlight the potential of imputation to mitigate the effects of *post-mortem* DNA damage, correcting erroneous calls and enhancing the reliability of downstream analyses (Garrido Marques et al., 2024; Sousa Da Mota et al., 2023).

Although minor, a small decrease in concordance at higher coverage is observed for both USR1 and Sumidouro5 in homozygous reference sites (even though the y-axis in Fig. 4.1 is zoomed between 99% and 100%). This trend is somewhat unexpected, as denser SNP datasets typically improve imputation performance. Moreover, at higher coverage, concordance is higher for transitions than for transversions.

One possible explanation is that higher coverage allows for the recovery of more low-frequency variants, including those with low minor allele frequency (MAF) in the reference panel (Hui et al., 2020). While this increased sensitivity improves resolution, it can also introduce noise when recovering rare alleles, occasionally leading to miscalls by the imputation algorithm. However, this does not fully explain why transversions exhibit lower concordance than transitions.

This discrepancy may be further explained by the fact that transitions are more frequent in the genome and better represented in reference panels, allowing for more confident imputation. In contrast, transversions are rarer and may lack the same level of statistical support.

Although transversions show higher concordance at heterozygous sites, particularly at lower coverage, this is likely due to the down-weighting of transition calls during genotype likelihood estimation to minimize the effects of post-mortem deamination. In contrast, at homozygous reference sites, where around 10 times more SNPs are assessed than at heterozygous sites (Fig. 4.3), the overall impact of low-frequency SNP recovery may be more noticeable. This could explain why transitions outperform transversions in homozygous reference concordance at higher

coverage.

To further explore this pattern, it would be useful in future work to stratify imputation concordance by MAF bins, helping to test whether attempts to impute rare alleles underlie the observed dip in performance.

This chapter has demonstrated the robust performance of GLIMPSE imputation in ancient American genomes, with high concordance and accuracy achieved from $0.1\times$ coverage. However, despite the reliable performance at low coverages, filtering for 99% genotype certainty results in many sites falling below the threshold at coverages under $0.5\times$, limiting the number of SNPs recovered. Most SNPs are recovered only from $0.5\times$ coverage, emphasizing the need for strategies that maximise SNP recovery without compromising reliability.

Previous research has highlighted the advantages of a second imputation step, which refine the initial imputation results and increase the number of recovered SNPs (Hui et al., 2020). While earlier studies employed Beagle4 followed by Beagle5, our collaborator, Dr. Lara Cassidy, has proposed a novel double-imputation pipeline that combines GLIMPSE with Beagle5. Given that Beagle has already shown high imputation accuracy, even in ancient American populations (Sousa Da Mota et al., 2023), this double-imputation pipeline offers a promising strategy for improving SNP recovery in low-coverage ancient genomes.

Beyond improving SNP recovery, the diploid data generated by imputation open new avenues for analyses that are not possible with pseudo-haploid data. For example, studies of identity-by-descent (IBD) segments allow for the detection of runs of homozygosity (ROH), estimation of kinship between individuals, and inference of effective population size. These analyses provide critical insights into population demography, shedding light on inbreeding, population structure, and historical changes in population size. By enabling such studies, imputation enhances the scope of research into ancient human populations, allowing for a more comprehensive understanding of their genetic diversity and evolutionary history.

Methods

1. Processing and downsampling of ancient genomes

Ancient genomes of USR1, a 11,500 years old individuals from Alaska (Moreno-Mayar et al., 2018a), and Sumidouro5, a 10,400 years old individuals from Brazil (Moreno-Mayar et al., 2018b), were aligned and processed using the same pipeline as Chapter 3. The high coverage individuals, 13.35× and 13.42× respectively, were independently downsampled to 0.01×, 0.05×, 0.1×, 0.5×, 1×, 2×, 3×, 4× and 5× with DownsampleSam from Picard-tools version 1.101 (available at <http://broadinstitute.github.io/picard/>) with random seed “1234” prior to imputation. Downsampling is the process of randomly discarding reads to achieve the desired reduced genomic coverage.

2. Step 1 of double-imputation pipeline: GLIMPSE

The downsampled data were imputed with GLIMPSE (Rubinacci et al., 2021) coupled with the use of 1000 Genome Phase 3 as a reference panel (The 1000 Genomes Project Consortium et al., 2015). To minimise the effects of DNA damage, which is mostly occurring in transition sites (Orlando et al., 2015), genotype likelihoods for transitions were set to zero to exclude those sites from imputation, as suggested from a previous study (Hui et al., 2020). Imputed SNPs with 99% certainty only were retained for subsequent analysis. Imputed sites were classified in the three following categories: homozygote reference (1/1), homozygote alternative (0/0) and heterozygote (1/0).

3. Validation of imputation

Diploid variants were called from processed BAM files using HaplotypeCaller, which is part of GATK version 3.7-0 (McKenna et al., 2010). Direct comparisons were made between sites common to the downsampled imputed and the diploid genomes. Performance was measured under two metrics, as detailed in a previous study (Cooke et al., 2024): a) concordance, which was defined as the rate at which diploid sequence data was correctly re-created through imputation (i.e., overlapping sites divided by the sites present in the diploid data), and accuracy, which was defined as how often imputed data accurately matched the sequence data (i.e., overlapping sites divided by the number of

imputed sites). In other words, the number of sites from the diploid found in the imputed were either, divided by the total number of sites from the diploid to measure concordance, or divided by the total number of sites from the imputed to measure accuracy.

Chapter 5

Tracing population demography in the Americas using identical-by-descent (IBD) segments

Abstract

The transition from hunting-gathering to sedentary farming has driven profound demographic, cultural, and technological shifts throughout human history. While well-documented in other continents, this impact in the Americas remains less thoroughly explored. In this chapter, identical-by-descent (IBD) segments identified from imputed genomes of ancient Americans are leveraged to reconstruct demographic patterns and population dynamics. Patterns of runs-of-homozygosity (ROH) confirms bottlenecks and small founding population sizes during the initial settlement of the Americas. As expected, hunter-gatherer populations exhibit longer and more numerous ROH segments compared to sedentary farming populations, reflecting differences in lifestyle and their impacts on population demography. Effective population size estimates further reveal significant growth during key periods of agricultural innovation, such as the incorporation of fully-domesticated maize as a staple crop in the Maya region and the likely influence of the Tiwanaku Empire's agricultural advancements in the Andes. Regional analyses also uncover distinct patterns of connectivity and structure. Lastly, in California, a maritime route between San Nicolas and San Clemente Islands may have developed within the last millennium, possibly absent during earlier periods (4,500 BP).

Introduction

During the Pleistocene, human populations across all continents subsisted as hunter-gatherers. However, the Holocene marked a transformative era during which several communities independently domesticated plants and transitioned to sedentary farming (Diamond and Bellwood, 2003; Stoneking et al., 2023). The domestication of crops and subsequent agricultural practices facilitated increased food production, enabling larger population sizes and the accumulation of food surpluses. This shift laid the foundation for the development of complex technologies, social stratification, centralised states, and urbanisation (Diamond and Bellwood, 2003; Stoneking et al., 2023). Transitions from a hunting-gathering lifestyle to sedentary farming occurred on every continent, as illustrated by the following non-exhaustive list of examples: (a) in Europe, originating from the Levent, the Neolithic transition is characterised by the spread of farming, domesticated animals, but also ceramics reaching regions as far as the Iberian peninsula, the British Isles and southern Scandinavia (Günther and Jakobsson, 2016), (b) in Africa, Neolithic farmers from Europe introduced wheat, barley and lentils (Llorente et al., 2015), (c) in south Asia, with Neolithic farmers from Pakistan moving into the Indus Valley, leading to the emergence of villages and towns (Misra, 2001), (d) in east Asia where hunter-gatherers mixed with incoming farmers associated with the Neolithic spreading from South China (McColl et al., 2018), and (e) in Oceania, the agriculturists from Taiwan expanded into Southeast Asia and Oceania with the development of sophisticated sea navigation technology and spreading Austronesian languages, leading to maritime trade networks (Oliveira et al., 2022).

In the Americas, however, the narrative is less straightforward. While plants were domesticated in various regions, major dispersal events often appear disconnected from the spread of farming (Stoneking et al., 2023). For instance, maize, originally domesticated in Mesoamerica or Isthmo-Colombian Area, reached the southwestern United States (Merrill et al., 2009), possibly pushed by population expansion from Mexico around 5,000 years ago (Villa-Islas et al., 2023). Yet, although the dispersal routes and genetic adaptation of maize have been relatively studied (Da Fonseca et al., 2015; Ramos-Madrigal et al., 2025), the impact of agriculture on human

societies—particularly on mobility, kinship patterns, and demographic dynamics—remains largely unexplored.

This chapter leverages shared genomic segments analysis to address these questions. For instance, identical-by-descent (IBD) segments—genomic tracts shared between individuals and inherited from a common ancestor—provide key insights into kinships and population demography. The length of IBD segments reflects the time to the most recent shared ancestor: longer segments indicate recent shared ancestry, such as within-family relationships, while shorter segments accumulate in smaller populations over generations, reflecting past bottlenecks and enabling estimates of effective population size (Browning and Browning, 2015, 2012). A notable example of IBD analysis comes from a study on ancient Rapa Nui. Researchers identified third- to fourth- relatives using IBD segments analysis (Moreno-Mayar et al., 2024), which surpassed the detection limits of READ, a tool for pseudo-haploid data that identifies kinship only up to second-degree relationships (Monroy Kuhn et al., 2018). The same study also applied runs-of-homozygosity (ROH) analysis, which provides complementary insights: longer ROH point to consanguinity or inbreeding, while an accumulation of shorter ROH reflects small population sizes or historical bottlenecks. The results revealed a high proportion of short ROH in ancient Rapa Nui individuals, which together with population size estimates, suggest a stable population size over the past 1,000 years. These findings challenge the ecocide hypothesis, which argues that deforestation and human activity caused a population collapse in the 1600s. Instead, the study concluded that the Rapa Nui population remained resilient despite environmental changes (Moreno-Mayar et al., 2024).

In studies of ancient Americans, ROH analyses have similarly provided valuable insights into population size and structure. For example, small effective population sizes have been identified in eastern South America (Ferraz et al., 2023) and Patagonia (de la Fuente et al., 2018). Conversely, analysis of an ancient Ceramic Caribbean individual from the Bahamas suggested an unexpectedly large effective population size, far exceeding what the island could sustain. This discrepancy points to high mobility and interconnectedness between Caribbean islands, a conclusion further supported by archeological evidence (Schroeder et al., 2018). Additionally, ROH-based estimates of effective population size have been interpreted to suggest insights into subsistence

strategies, with the authors proposing that a large population could only be sustained through farming. They further hypothesise south-to-north migration from Mexico to California linked to maize agriculture (Nakatsuka et al., 2023).

In this study, I compile genomic data from previously published ancient individuals across the Americas, including both shotgun and capture sequencing data, and apply imputation to augment coverage depth. Using this dataset, I infer IBD segments to address two key questions: (1) what communities can be identified from kinship patterns; and (2) how effective population sizes changed over time in the Americas.

Results

I. Societal insights into ancient American communities

I. 1. Different kinship results in the Classic Maya of Chichén Itzá

IBD segments were inferred in 330 ancient individuals spanning across the Americas. Details of IBD sharing between each pair of individuals, including their corresponding degree of relatedness, are summarised in [Table 5.1](#).

The closest relationship identified was between two individuals from the Chichén Itzá population, YCH033 and YCH054, who shared 191 IBD segments totalling to 2006.82cM ([Table 5.1](#)). While the total length of these IBD segments indicates a relationship of second-degree relatives, it falls short of the level of IBD sharing typically associated with identical twins, as defined in a previous study (Barquera et al., 2024). This is possibly due to the focus of my analysis on IBD1, rather than including both IBD1 and IBD2, potentially leading to an underestimation of very close kinship. Similarly, my analysis could not confirm the previously reported second-degree relationship between YCH34 and YCH41 from the Chichén Itzá (Barquera et al., 2024), as the observed IBD sharing did not indicate any significant kinship between these individuals. While the Barquera analysis used the method BREADR (Rohrlach et al., 2023) to estimate the pairwise-mismatch rate (PMR) based on pseudo-haploid data, my analysis relied on detecting IBD segments from imputed diploid genotypes. A possible explanation for the discrepancy is that BREADR may be more sensitive to close kinship in low-coverage data, as it operated directly on mismatch rates without relying on imputation, whereas IBD detection may miss close relationships when imputation quality is low or when phasing errors obscure long shared segments.

Table 5.1. Kinship in the ancient Americas according to IBD sharing

Pedigree	Individual 1*	Individual 2*	Number of IBD segments shared	Sum of IBD length shared (in cM)	Kinship degree
Figure 5.2	ESN_SN17_merged	ESN_SN48	56	377.756	4th
	ESN_SN17_merged	ESN_SN44_merged	84	717.551	3rd
	ESN_SN44_merged	ESN_SN48	34	216.513	5th
	ESN_SN40	ESN_SN44_merged	40	272.092	5th
-	USA_CA_SantaRosa_7400BP_I11557	USA_CA_SantaRosa_7400BP_I11558	46	359.852	4th
-	LSN_SN03	LSN_SN50	21	144.565	5th
	LSN_SN50	LSN_SN51	48	293.813	4th
Figure 5.3	LSN_SN10	LSN_SN11_merged	26	172.046	5th
	LSN_SN13_merged	LSN_SN53	41	229.965	5th
	USA_CA_SanClemente_500BP_I0751	LSN_SN13_merged	20	154.007	5th
	LSCI_SC03_merged	LSN_SN11_merged	27	168.76	5th
	USA_CA_SanClemente_500BP_I0751	LSN_SN11_merged	23	145.4	5th
	LSN_SN11_merged	LSN_SN53	20	143.39	5th
	LSN_SN11_merged	LSN_SN13_merged	48	377.485	4th
	LSCI_SC05_merged	LSN_SN13_merged	25	153.669	5th
	LSN_SN13_merged	LSN_SN52	24	158.058	5th
	LSCI_SC05_merged	LSN_SN11_merged	26	189.309	5th
	LSCI_CT01_merged	USA_CA_SanClemente_500BP_I0751	24	145.751	5th
	USA_CA_SanClemente_500BP_I0750	LSCI_SC03_merged	19	154.35	5th
	USA_CA_SanClemente_500BP_I0760	LSCI_SC05_merged	23	137.513	5th
	USA_CA_SanClemente_500BP_I0751	LSCI_SC03_merged	35	230.821	5th
	LSCI_CT01_merged	LSCI_SC05_merged	27	139.222	5th
	LSCI_SC01	LSCI_SC04	22	138.22	5th
	USA_CA_SanClemente_500BP_I0751	LSCI_SC05_merged	31	186.605	5th
	USA_CA_SanClemente_500BP_I0756	LSCI_SC03_merged	36	236.01	5th
	USA_CA_SanClemente_500BP_I0751	LSCI_SC04	30	166.144	5th
	USA_CA_SanClemente_500BP_I0760	LSCI_SC03_merged	40	293.246	4th
	USA_CA_SanClemente_500BP_I0756	USA_CA_SanClemente_500BP_I0760	32	231.024	5th
	USA_CA_SanClemente_500BP_I0750	USA_CA_SanClemente_500BP_I0751	33	214.63	5th
	USA_CA_SanClemente_500BP_I0751	USA_CA_SanClemente_500BP_I0758	20	139.247	5th
	USA_CA_SanClemente_500BP_I0750	USA_CA_SanClemente_500BP_I0760	20	148.963	5th
	USA_CA_SanClemente_500BP_I0751	USA_CA_SanClemente_500BP_I0756	30	234.204	5th
	USA_CA_SanClemente	USA_CA_SanClemente_500BP_I0760	62	513.105	4th

	_500BP_I0751	0			
-	Anzick-1_Anzick-1	Chile_LosRieles_10900BP_I11974	27	199.975	5th
-	Canimar_Abajo_CAO011	Canimar_Abajo_CAO028	28	222.733	5th
-	Canimar_Abajo_CAO029	Canimar_Abajo_CAO022026	19	158.931	5th
-	Cueva_de_los_Esqueletos_CDE002	Cueva_de_los_Esqueletos_CDE004	36	201.026	5th
-	Sierra_Gorda_2417J	Sierra_Gorda_2417Q	18	141.842	5th
	Sierra_Gorda_333B	Sierra_Gorda_2417J	47	423.592	4th
-	Chichen_Itza_YCH033	Chichen_Itza_YCH054	191	2006.82	2nd
-	Rio_Uncallane_IL3	Rio_Uncallane_IL7	17	161.618	5th
	Rio_Uncallane_IL3	Rio_Uncallane_IL5	23	200.353	5th
	Rio_Uncallane_IL4	Rio_Uncallane_IL7	20	160.951	5th
Figure 5.1	Sumidouro_Sumidouro4	Sumidouro_Sumidouro6	34	346.087	4th
	Sumidouro_Sumidouro5	Sumidouro_Sumidouro7	22	181.136	5th
	Sumidouro_Sumidouro6	Sumidouro_Sumidouro7	44	400.747	4th
	Sumidouro_Sumidouro5	Sumidouro_Sumidouro6	25	260.068	5th
	Sumidouro_Sumidouro4	Sumidouro_Sumidouro7	14	144.864	5th
	Sumidouro_Sumidouro4	Sumidouro_Sumidouro5	32	280.811	4th
-	IPK12_IPK12	IPK13_IPK13	30	160.413	5th
-	NorthTierradelFuego_Grouped_500BP_I12354	MA577_MA577	31	205.509	5th
Figure 5.4	Fuego-Patagonian_894	BeagleChannel_Grouped_1500-100BP_I12943	24	142.521	5th
	Fuego-Patagonian_890	Fuego-Patagonian_894	25	172.998	5th
	BeagleChannel_Grouped_1500-100BP_I12941	BeagleChannel_Grouped_1500-100BP_I12943	28	192.734	5th
	BeagleChannel_Grouped_1500-100BP_I12941	IPY10_IPY10	34	195.16	5th
	BeagleChannel_Grouped_1500-100BP_I12943	IPY08_IPY08	31	186.661	5th
	BeagleChannel_Grouped_1500-100BP_I12941	IPY08_IPY08	27	137.519	5th
	Fuego-Patagonian_890	IPY08_IPY08	24	155.066	5th
	Fuego-Patagonian_895	IPY10_IPY10	43	270.916	5th
	Fuego-Patagonian_894	IPY08_IPY08	40	301.136	4th
	Fuego-Patagonian_894	Fuego-Patagonian_895	37	242.6	5th
	Fuego-Patagonian_895	IPY08_IPY08	64	450.029	4th
	IPY08_IPY08	IPY10_IPY10	52	363.896	4th
	Fuego-Patagonian_894	IPY10_IPY10	29	181.502	5th
	BeagleChannel_Grouped_1500-100BP_I12943	IPY10_IPY10	29	167.682	5th

* pairs of individuals are labelled as “[population]_[individual]”

I. 2. A possible patriarchal society for the hunter-gatherer Sumidouro of Brazil

When kinships are identified within a community, constructing a pedigree linking these individuals can offer valuable insights into societal and funerary practices. This approach has proven particularly revealing for the Sumidouro population from Brazil, where six pairs of relatives were detected among four individuals. These relationships consist of 4th- and 5th-degree connections. A proposed pedigree, illustrated in [Figure 5.1](#), respects the inheritance of the Y-chromosome haplogroup. Sumidouro5, Sumidouro6, and Sumidouro7 are linked through the paternal line, all sharing the haplogroup Q-M3 (M848), while Sumidouro4 is connected through the maternal line, having inherited the haplogroup Q-M3(xY4308) from an unknown parent. Maternal haplogroups were not incorporated into the pedigree since the D1 haplogroup, shared by Sumidouro4, Sumidouro6, and Sumidouro7, could simply reflect its prevalence within the population, with accounts of haplogroup D4h3a found in Sumidouro5.

Sumidouro5, Sumidouro7, Sumidouro6, and Sumidouro4 are dated to approximately 10,405, 10,271, 10,083, and 10,096 years BP. The pedigree inferred from IBD sharing indicates that the burial site was used over five generations, suggesting that the Sumidouro hunter-gatherer population returned to this site across generations to inter their family members.

One female individual, Sumidouro8, dated to be younger (9,849 years BP) than the others, shows no familial relationships. While the sample size is small, these findings may suggest that the Sumidouro community was patriarchal, with females being buried alongside their husband's family.

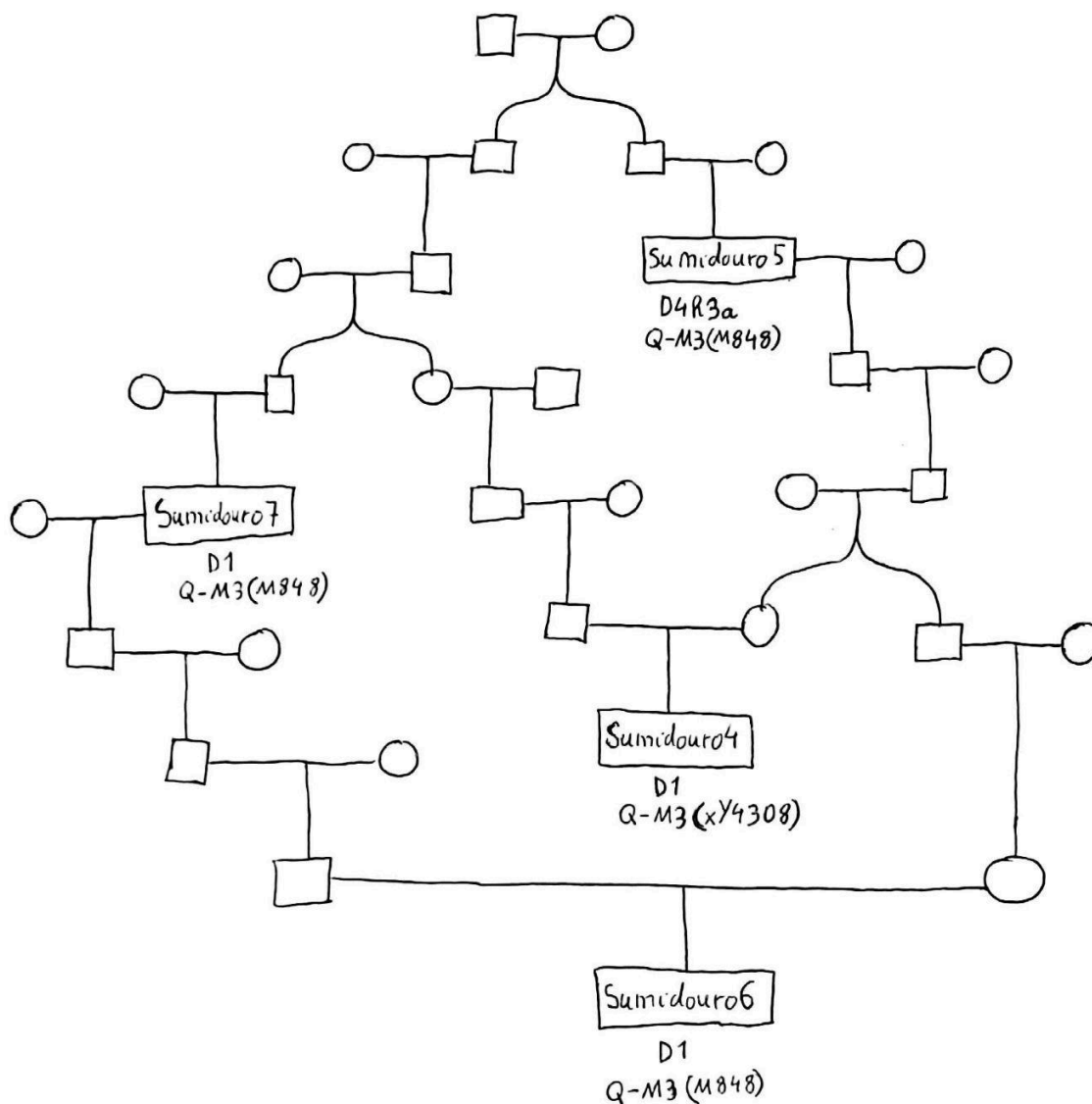


Figure 5.1. Pedigree of Sumidouro according to IBD sharing. Males are represented as squares and females as circles. Offsprings are depicted as descending from the union of two parents, with siblings branching off from the same parental link. Mitochondrial haplogroup is indicated below each individual, followed by Y-chromosome haplogroup if the individual is male.

I. 3. A family pedigree for early individuals from San Nicolas Island

In addition to the Sumidouro population, other communities also exhibit familial connections. For instance, in the ancient Andean population of Rio_Uncallane, fifth-degree relations were found for three pairs, IL3-IL7, IL3-IL5, and IL4-IL7 ([Table 5.1](#)). Similarly, in the Sierra_Gorda population from Mexico, a 4th-degree relationship was identified between individuals 333B and 2417J, alongside a fifth-degree relationship between 2417J and 2417Q.

Larger kinship networks were observed within populations from California, with a clear distinction between earlier (Early San Nicolas or ESN) and later (Late San Nicolas or LSN) groups, despite both groups being excavated from the same site on San Nicolas Island.

Kinship analysis among four individuals from the ESN population identified one 3rd-degree relationship between SN17 and SN44, one 4th-degree relationship between SN17 and SN48, and two 5th-degree relationships between SN40 and SN44 and between SN44 and SN48. Taking account of these relationships, along with inheritance patterns of mtDNA and Y-chromosome haplogroups, a potential pedigree is inferred in [Figure 5.2](#). SN17 and SN44 both share the same Y-chromosome haplogroup of Q1a*, which places them on the same paternal lineage. In contrast, SN40 and SN44 are linked through their maternal line, as both share the mitochondrial haplogroup A2cb. SN48 or SN17 carries haplogroup A2 or A2ca, respectively. Radiocarbon dating is available only for SN44 (4,647 years BP) and SN17 (4,517 years BP), positioning SN44 closer to the root of the pedigree compared to SN17.

As with the Sumidouro population, these results suggest that this hunter-gatherer community of ESN possibly utilised the same burial site over several generations for extended family members. Unlike Sumidouro, however, there is no apparent differentiation in burial customs between males and females.

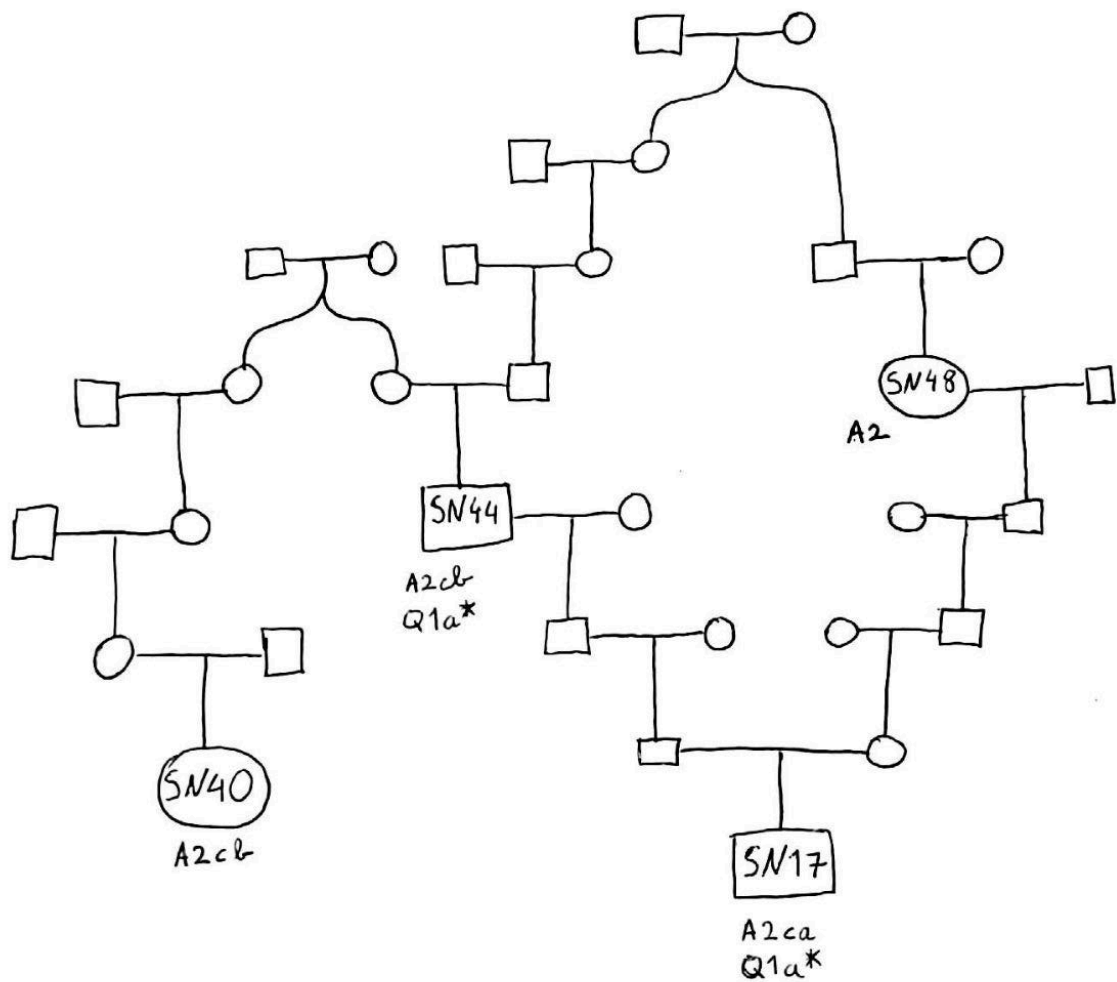


Figure 5.2. Pedigree of Early San Nicolas (ESN) individuals according to IBD sharing. Males are represented as squares and females as circles. Offsprings are depicted as descending from the union of two parents, with siblings branching off from the same parental link. Mitochondrial haplogroup is indicated below each individual, followed by Y-chromosome haplogroup if the individual is male.

I. 4. An intricate connection between maritime hunter-gatherers of San Nicolas, San Clemente, and Santa Catalina islands

The most intricate pedigree is observed among the LSN individuals. For instance, SN13 has kinship ties with five relatives, including I0751, an individual from San Clemente Island. I0751, in turn, shares familial connections with ten individuals, including CT01, who is associated with the LSCI population. CT01, in turn, has five relatives ([Table 5.1](#) and [Fig. 5.3](#)). Collectively, this extended pedigree links individuals from different islands within the southern Channel Islands (Nakatsuka et al., 2023; Scheib et al., 2018).

While CT01 from the LSCI was excavated from Santa Catalina Island, the other LSCI individuals (SC01, SC03, SC04, and SC05) were recovered from San Clemente Island. San Clemente Island, Santa Catalina Island, and San Nicolas Island—the origin of the LSN individuals—are all situated off the coast of California, within 100 kilometers of each other. The familial connections revealed by the pedigree suggest frequent migration among these islands, likely facilitated by intermarriage and regular travel. This is consistent with the populations' demonstrated seafaring abilities, as supported by archaeological evidence of extensive shell middens, indicating a marine-focused diet that remained consistent over the 8,000 years of occupation of these islands (Scheib et al., 2018).

Radiocarbon dating further supports overlapping timeframes for these populations: individuals from San Nicolas Island (LSN population) date between 1,172 and 840 years BP (median: 947 years BP), CT01 from Santa Catalina Island dates to 387 years BP, other LSCI individuals from San Clemente Island date between 1,101 and 884 years BP (median: 1,055 years BP), and individuals from the USA_CA_SanClemente_500BP population date between 825 and 759 years BP (median: 825 years BP) (Nakatsuka et al., 2023; Scheib et al., 2018). No kinship was identified between individuals from the LSN or LSCI and those from the ESN, as their occupations are temporarily separated by roughly 3,000 years.

Interestingly, no kinship was detected with individuals from the USA_CA_SantaRosa_300BP (Santa Rosa Island) or Island_Chumash group (San Miguel Island), even though these groups are separated only by 80-150 kilometers from the groups from San Nicolas Island or San Clemente Island. However, their temporal

dimensions are not exactly the same. The two Island_Chumash individuals from San Miguel Island are dated to 826 and 641 years BP, while the five individuals from Santa Rosa are dated between 940 and 595 years BP (median: 600 years BP). Therefore, this lack of kinship may be due to the difference in temporal contexts between individuals from the LSN or LSCI and those from Santa Rosa Island or San Miguel Island.

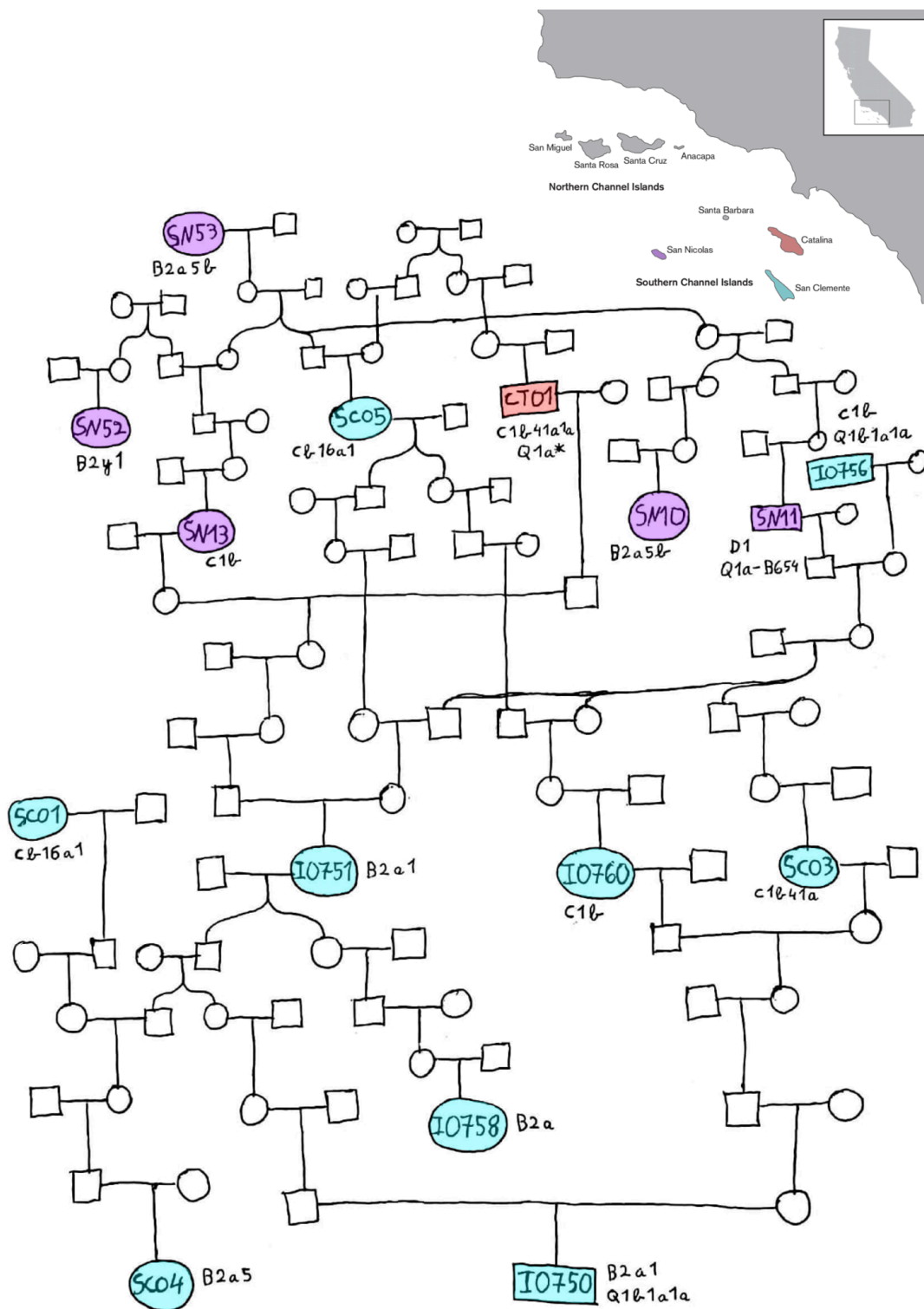


Figure 5.3. Pedigree of Islanders from Late San Nicolas (LSN), Santa Catalina, and San Clemente according to IBD sharing. Males are represented as squares and females as circles. Offsprings are depicted as descending from the union of two parents, with siblings branching off from the same parental link. Mitochondrial haplogroup is indicated below each individual, followed by Y-chromosome haplogroup if the individual is male. In the top right corner, a map of the Channel Islands illustrates the geographical origin of individuals.

I. 5. Cultural biased pedigree in Patagonia

The pedigree inferred from IBD sharing links individuals from three populations in Patagonia: BeagleChannel_Grouped_1500-100BP consisting of individuals excavated from sites near the Beagle Channel dating from approximately 1,500 to 100 BP, Fuego-Patagonian also excavated from the shores of the Beagle Channel and dating to around 150 BP, and maritimePatagonia dating to around 1,000 years BP and comprising of two individuals associated with the Yaghan culture and two others with the Kawéskar culture (Fig. 5.4). For instance, the individual I12943 from the BeagleChannel is not only a 5th-degree relative of I12941 from the BeagleChannel, but also a 5th-degree relative of 894 from the Fuego-Patagonian and a 5th-degree relative of IPY08 and IPY10, both from the maritimePatagonia (Table 5.1). All individuals have been carbon-dated to a similar time period, ranging between 190 and 150 years BP, except for IPY10, which dates back to 810 years BP (de la Fuente et al., 2018; Nakatsuka et al., 2020b; Raghavan et al., 2015). Their possible intricate familial relationships are summarised in the pedigree (Fig. 5.4). Notably, individuals 890 and 894 share a maternal great-grandmother, making them 5th-degree relatives and passing on the mitochondrial haplogroup D4h3a. Similarly, IPY08 and IPY10 are connected through a great-grandfather for the former and a grandfather for the latter on their paternal side, making them 4th-degree relatives, with both carrying the Y-chromosome haplogroup Q1a2a1a1. Additionally, IPY10, I12941, and I12943 are all 5th-degree relatives of one another, inheriting the mitochondrial haplogroup D4h3a from a common maternal ancestor, whether a great-great-great-great-grandmother, grandmother, or great-great-grandmother.

All three populations were excavated within Tierra del Fuego Archipelago at the southernmost edge of Patagonia, near the Beagle Channel strait (de la Fuente et al., 2018; Nakatsuka et al., 2020b; Raghavan et al., 2015). The sites are located within 200 kilometers of each other. While the region's irregular terrain, including glaciers, dense forests, and plateaus, could have hindered overland travel, the Beagle Channel strait likely served as a natural waterway connecting the populations.

Interestingly, among the four maritimePatagonia individuals, IBD was only detected in the two associated with the Yaghan culture (or Yámana), not in the two associated with the Kawéskar culture. Similarly, among the four Fuego-Patagonian

individuals analysed, IBD sharing segments were detected only in the three associated with the Yaghan culture and not in the individual linked to the Selk'nam culture. The Beagle Channel strait intersects the territories of the four main cultural groups of the Tierra del Fuego Archipelago: the Yaghan in the south, the Kawéskar in the west, the Selk'nam in the north, and the Haush in the east. The Yaghan and Kawéskar cultures shared a strong reliance on marine resources, which they accessed via canoes. In contrast, the Selk'nam focused on hunting guanaco and birds and gathering seashells along the shoreline, while the Haush hunted both terrestrial and marine prey. While some researchers suggest that intermarriage was common in boundary areas between cultural groups, others argue that unions were rare (de la Fuente et al., 2018; Nakatsuka et al., 2020b). The results here suggest that geographical proximity alone did not determine marital connections. Intermarriage may have been facilitated by shared cultural practices, as exemplified by the Yaghan, or, conversely, the spread of the Yaghan culture could have been driven by intermarriage.

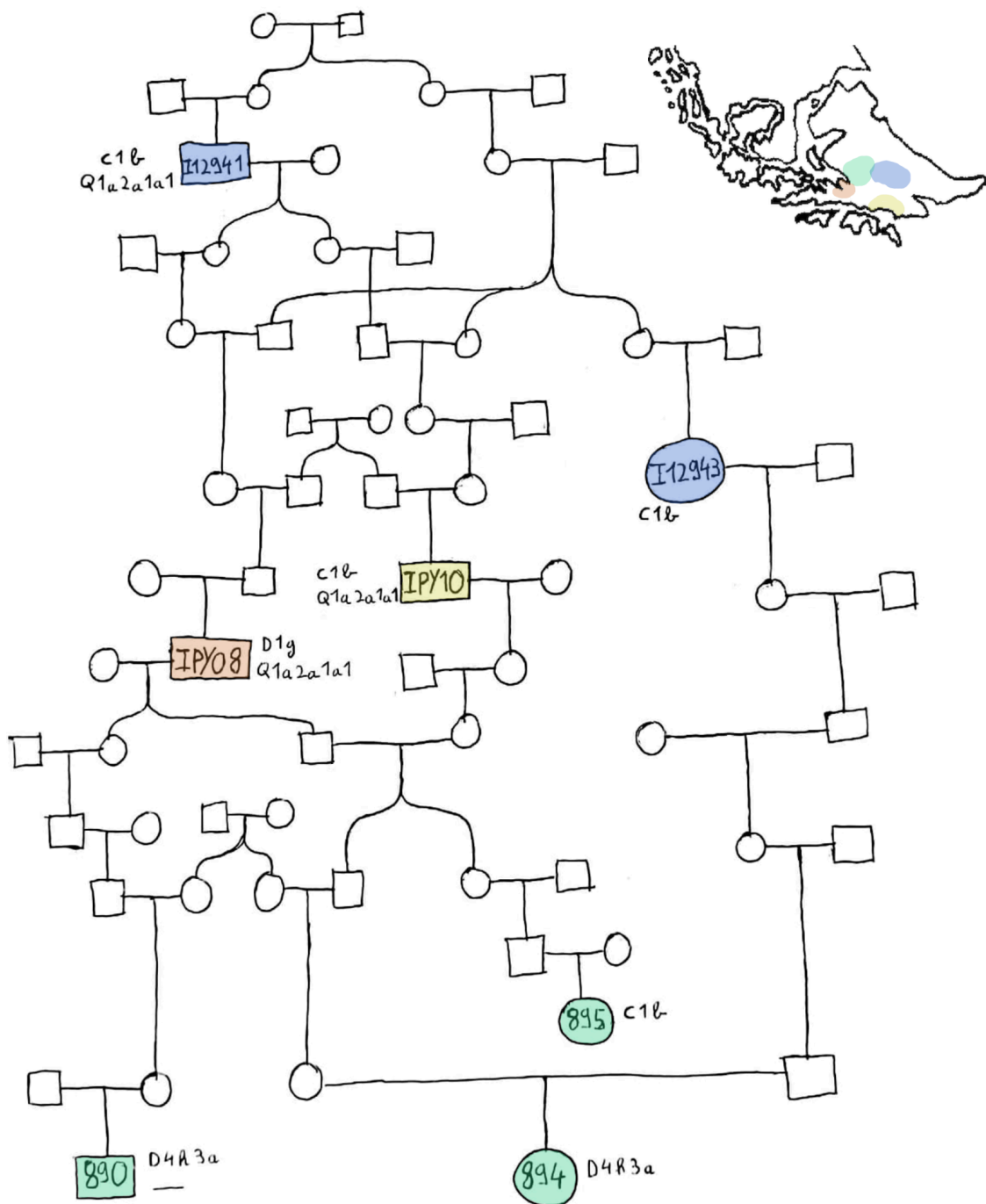


Figure 5.4. Pedigree of individuals from Patagonia according to IBD sharing. Males are represented as squares and females as circles. Offsprings are depicted as descending from the union of two parents, with siblings branching off from the same parental link. Mitochondrial haplogroup is indicated below each individual, followed by Y-chromosome haplogroup if the individual is male. In the top right corner, a map of southern Patagonia illustrates the geographical origin of individuals.

I. 6. A surprising link between Anzick and an individual from Chile

Lastly, a surprising 5th-degree relationship was detected between Anzick, an individual from Montana, and I11974, an individual from Chile ([Table 5.1](#)). Although both date to similar periods—12,656 years BP for Anzick and 10,935 years BP for I11974—they were excavated from sites separated by a vast distance. Anzick is associated with the Clovis culture, the earliest widespread archeological complex in North America (Rasmussen et al., 2014). In contrast, I11974 was excavated in central Chile and is associated with a culture characterised by extensive, stratified shell middens (Posth et al., 2018).

Given that a small founder population expanded rapidly throughout the American continents, the observed relationship between Anzick and I11974 may reflect a more complex interplay of ancient populations structure and migration patterns rather than a direct kinship. For instance, the 10,400-year-old Sumidouro individuals from Brazil and the 5,100-year-old Ayayema individual from Patagonia share a common ancestor associated with the initial migration into South America. This lineage, traced back to the ANC-A related to Anzick, highlights the complexity of ancient human movements (Moreno-Mayar et al., 2018b). In the following section, I will examine whether these potential kinships arise from demographic processes rather than familial relationships.

II. Demographic impact of ancient American population

II. 1. Small IBD segments in ancient Americans

To assess the accuracy of the kinship signals, the data was compared against simulated datasets generated by Ariano *et al.* (Ariano et al., 2022). These simulations tested three scenarios—second-degree, third-degree, and fourth-degree relatives—with the results from ancient Americans overlaid on the plot (Fig. 5.5). On this plot, the second-degree relationship between two individuals from Chichén Itzá (YCH033 and YCH054) and the third-degree relationship between the early individuals from San Nicolas Island (SN17 and SN44) (Table 5.1) are highlighted by red and green circles. The other pairs of individuals have 4th- or 5th-degree relationships, in blue and purple on the plot. However, while the total length of IBD segments for ancient Americans aligns with the simulated data for the same degree of kinship, the number of IBD segments exceeds for each scenario. This discrepancy indicates that ancient Americans accumulate IBD segments of comparable total length, but those segments are generally shorter.

Given the separation of two groups on the simulated data for 2nd-degree relationships, I considered an effect driven by pairing-specific factors, such as differences between granddaughter-grandmother and niece-aunt relationships, both examples of second-degree kin. To rule out such an effect, I examined variation in the coefficient of relatedness across different pairing cases within each degree of kinship (Fig. S5.1), but observed no significant differences. This suggests that the observed separation is not due to the specific relationship type within each degree of kinship.

The method used to infer degrees of relatedness relies on dividing the total lengths of IBD segments by twice the genome's recombination map length. This approach focuses exclusively on segment length, but not on the number of segments. The shorter IBD segment sizes observed in ancient Americans are likely a reflection of demographic factors, specific to the settling history of the Americas, such as the small effective population size of the initial founder population (Llamas et al., 2016). To further investigate demographic effects, the next section delves into Runs-Of-Homozygosity (ROH) analysis, which provides additional insights into the bottlenecks and inbreeding patterns.

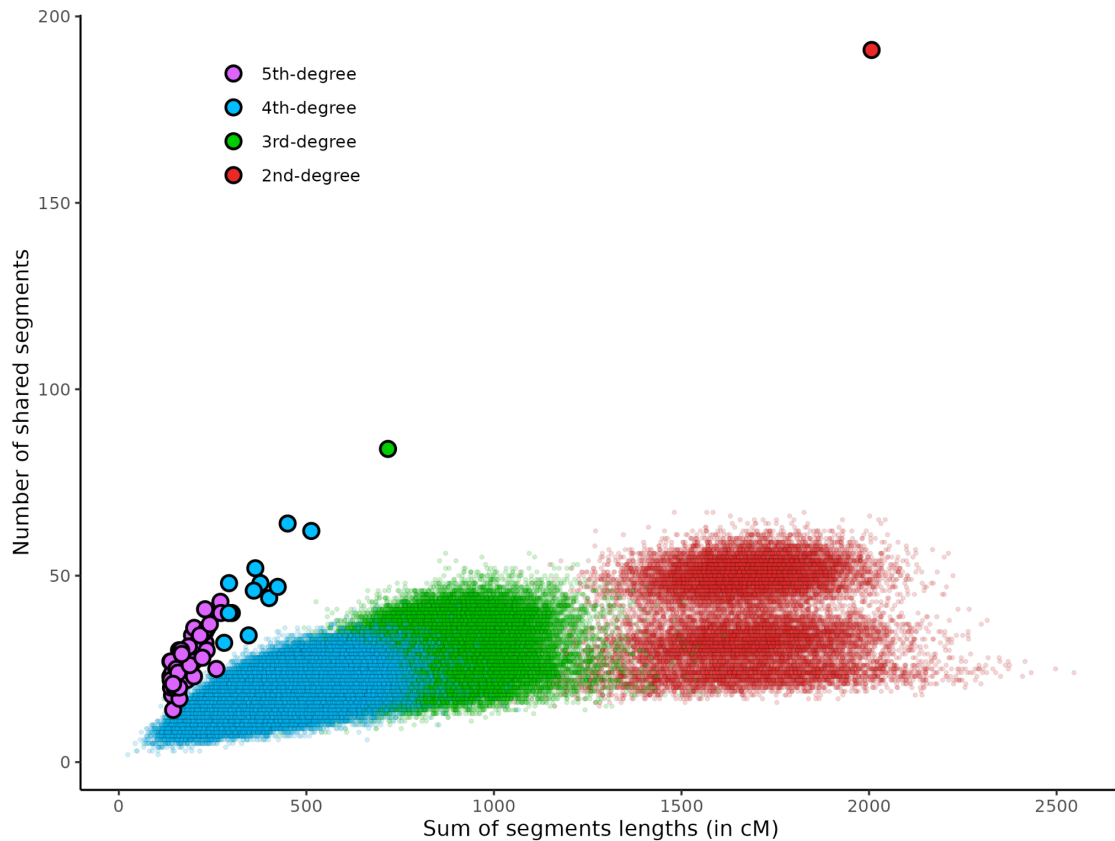


Figure 5.5. Kinship results for ancient Americans compared to simulated data. Kinship results for ancient Americans (larger points and bold outline) are overlaid onto simulated data 2nd-, 3rd-, and 4th-degree relationships (smaller points and in high transparency). The x -axis shows the sum of shared segment lengths (in centiMorgans), while the y -axis represents the number of shared segments. Degrees of relatedness are represented in different colours: red, green, blue and purple for 2nd-, 3rd-, 4th- and 5th-degree relationships respectively.

II. 2. Runs-Of-Homozygosity analysis

Building on the patterns observed in IBD segment sizes, the analysis of Runs-Of-Homozygosity (ROH) provides a complementary perspective on the demographic history of ancient populations. While IBD reveals shared ancestry between individuals, ROH patterns offer insights into population diversity and structure within individual genomes. Specifically, ROH analysis enables the detection of extended stretches of homozygosity, which are shaped by demographic events such as bottlenecks, founder effects, and inbreeding. These patterns help to disentangle the genetic impact of the small effective population sizes and prolonged isolation suggested by the fragmented IBD segments in ancient Americans.

In this section, ROH segments in an individual are categorised by different segment lengths: very long ROH ($> 20\text{cM}$ in red), long ROH ($12\text{-}20\text{cM}$ in yellow), medium ROH ($8\text{-}12\text{cM}$ in light blue), and short ROH ($4\text{-}8\text{cM}$ in dark blue) (Fig. 5.6). Given that those who were distantly related primarily represent hunter-gatherer or marine hunter-gatherer groups (Figs. 5.1-4), I first focus on populations where no evidence of a sedentary lifestyle or farming practices was found.

The highest cumulative ROH sum is observed in SN37, an early individual from San Nicolas Island, with twice the ROH sum of the next individual (Fig. 5.6). However, no kinship relationships were identified between SN37 and other individuals in the dataset (Table 5.1 and Fig. 5.2). The preponderance of short ROH segments ($4\text{-}8\text{cM}$) in SN37 suggests that these segments were inherited from distant rather than close ancestors. Similarly, individuals with the second, third, fifth, sixth, seventh, and ninth highest ROH sums—SN44, SN17, SN32, SN40, SN48, and SN54—all belong to the same early San Nicolas Island population and exhibit greater accumulations of short ROH compared with long ROH. Considering the multiple levels of distant kinship within this community (Fig. 5.2), the absence of evidence for inbreeding, and the excess number of IBD segments (Fig. 5.5), these results support the scenario of a small founding population followed by prolonged isolation.

The fourth highest ROH sum is observed in I0748 from San Clemente Island (Fig. 5.6). This individual exhibits particularly high amounts of long ROH segments ($> 20\text{cM}$), indicative of recent inbreeding. No direct kinship was detected for I0748, despite the San Clemente population presenting the most intricate pedigree (Fig. 5.3). The later inhabitants of both San Nicolas and San Clemente islands exhibit lower ROH sums and shorter ROH lengths compared to earlier populations, possibly reflecting increased genetic diversity through interactions with other seafaring groups.

The Sumidouro individuals from Brazil, dated to approximately 10,000 years ago likely represent some of the oldest individuals in South America (Llamas et al., 2016; Moreno-Mayar et al., 2018b; Raghavan et al., 2015; Rasmussen et al., 2014; Willerslev and Meltzer, 2021). Sumidouro4, Sumidouro5, Sumidouro6, and Sumidouro7 show low overall ROH sums (119cM , 83cM , 75cM , and 120cM , respectively), with no clear evidence of very long ROH segments (Fig. 5.6). While Sumidouro8 displays slightly higher ROH levels (307cM), these are mostly dominated

by short and medium lengths. These results suggest a relatively large population size during the Early Holocene in South America.

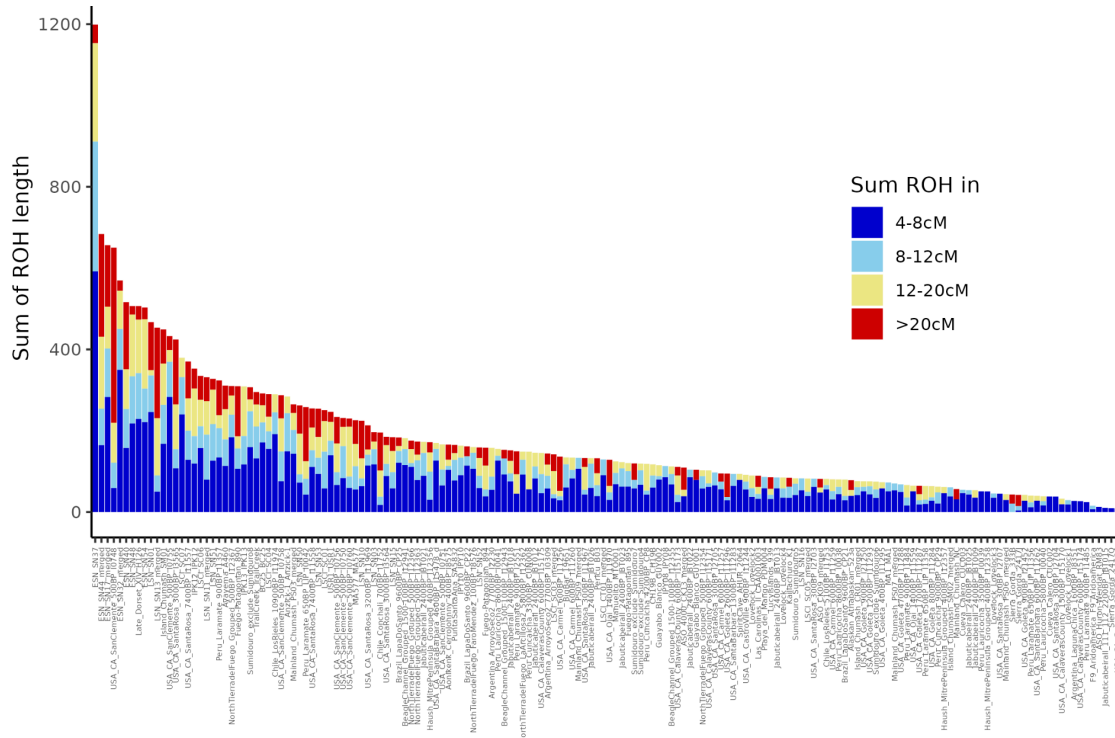


Figure 5.6. Runs-Of-Homozygosity for imputed non-sedentary non-farmers from ancient America. Each histogram depicts the accumulated sum of ROH segments for an individual with no evidence for a sedentary settlement. ROH sum is represented in four colours for four length categories: very long ROH of $> 20\text{cM}$ in red, long ROH for segments between 12 and 20cM in yellow, medium ROH for segments between 8 and 12cM in light blue, and short ROH for segments between 4 and 8cM in dark blue.

Below is the ROH plot for individuals associated with evidence of a sedentary lifestyle or farming practices (Fig. 5.7). To facilitate direct comparison with hunter-gatherer populations analysed earlier, the y-axis scale has been maintained. The dataset includes comparable numbers of individuals from both groups, with 166 hunter-gatherers and 138 sedentary farmers. However, the patterns displayed in the two figures reveal striking differences.

Among hunter-gatherers, 47% (78 individuals) exhibit very long ROH ($> 20\text{cM}$), whereas this proportion is more than halved in sedentary farmers, with only 21% (29 individuals). Additionally, the cumulative sum of ROH, averaged across all individuals, shows a nearly 2.5-fold difference, with 174cM for hunter-gatherers

compared to just 72cM for sedentary farmers. These differences align with expectations based on subsistence strategies and their demographic implications. Sedentary farmers, by producing more food and supporting larger communities, likely experienced population growth and admixture, leading to lower levels of ROH overall. In contrast, the smaller populations of hunter-gatherers would have been more prone to isolation, reflected in their higher levels of ROH.

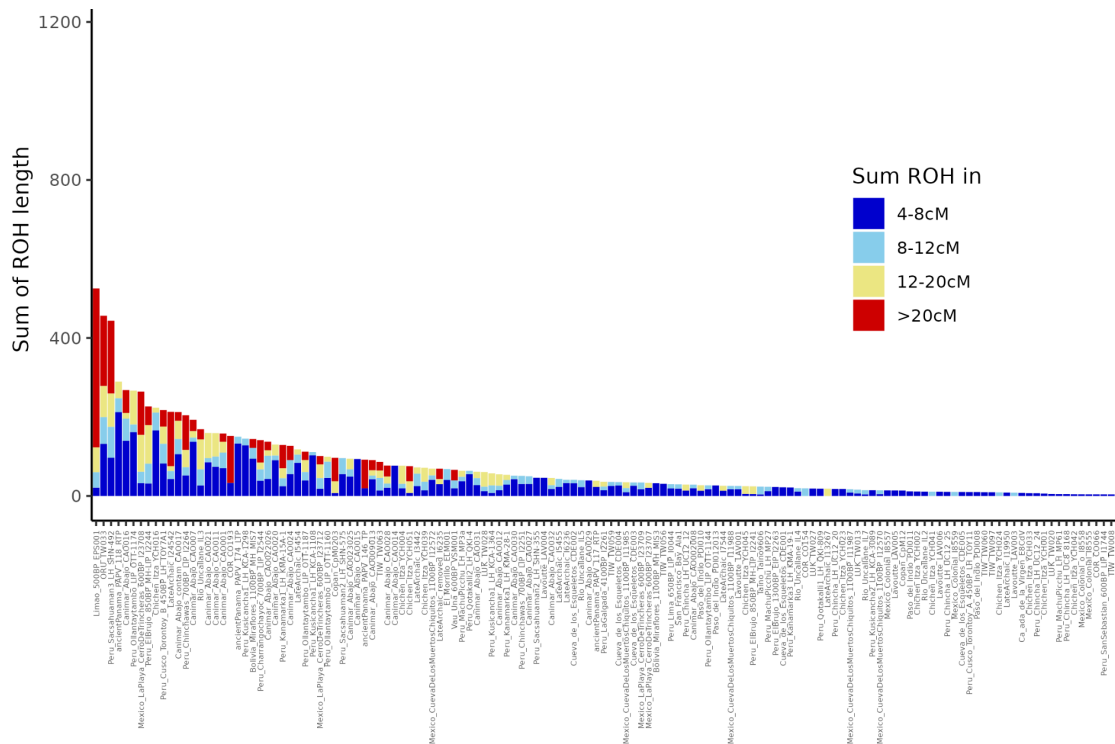


Figure 5.7. Runs-Of-Homozygosity for imputed sedentary farmers from ancient America. Each histogram depicts the accumulated sum of ROH segments for an individual with evidence for a sedentary settlement. ROH sum is represented in four colours for four length categories: very long ROH of > 20cM in red, long ROH for segments between 12 and 20cM in yellow, medium ROH for segments between 8 and 12cM in light blue, and short ROH for segments between 4 and 8cM in dark blue. The y-axis scale is set to match the previous plot of non-sedentary Americans for comparison.

II. 3. Effective population size in the ancient Americas

To further investigate the demographic factors underlying the genetic patterns revealed in the ROH analysis, this section explores temporal changes in effective population sizes across different regions. Populations were grouped by geographical regions following the classification from Chapter 3 (Fig. 3.15 and Table S3.4), with one

exception for the ancient Panamanian as they are the only Central Americans to fall outside of the Maya area. Given sufficient data for the Maya region and the geographical diversity of Central America, this analysis focuses on the Maya area instead of Central America. Considering the samples where carbon dating was available (Fig. S5.2), population effective sizes were estimated for regions of California, Aridoamerica, the Maya area, the Andean mountains, Brazil and Patagonia.

In California, the earlier analyses in this chapter (Fig. 5.3) suggested that individuals from San Nicolas and San Clemente Islands formed a closely connected community. This raised the possibility of dividing the region into subgroups, separating for example, the islanders from mainland Californians. However, the lack of a clear definition for these subgroups posed a challenge. To avoid unnecessary fragmentation and potential noise, all Californians were treated as a single group, assuming that any demographic shifts would be diluted but still visible within the larger dataset.

The largest population effective size was found in the Maya region, reflecting the history of large civilisations in the area (Regueiro et al., 2013). Overall, the estimation of effective population size reveals several periods of significant population growth across the Americas, often overlapping: in the Maya area between 4,700 and 1,100 years BP, in the Andes between 1,800 and 800 years BP, in the Caribbean between 1,300 and 900 years BP, and in Patagonia between 1,000 and 400 years BP (Fig. 5.8). Some of these increases correspond to key historical and agricultural transitions. In the Maya area, this period coincides with the establishment of maize as a fully domesticated and staple crop (Kennett et al., 2022; Regueiro et al., 2013). In the Andes, the population growth aligns with the golden age of the Tiwanaku empire, which revolutionised high-altitude maize farming: raised-fields increased crop yield and protected them from night frost by releasing the heat accumulated in the water canals during the day (*Tiwanaku Part 1: The City*, 2021). These agricultural innovations likely drove increased food production, supporting a larger population.

Conversely, population declines are observed in the Maya area between 1,100 and 300 years BP, Aridoamerica between 1,200 and 600 years BP, and Patagonia between 400 and 100 years BP (Fig. 5.8). In the Maya region, this sharp decline corresponds to the well-documented period of collapse, marking the transition from the Classic to the Postclassic period (Fash, 1991; Martínez de Velasco Cortina and Vega

Villalobos, 2015). In Patagonia, the decline in the most recent period likely reflects the arrival of Europeans, who introduced diseases that decimated local populations (Livi-Bacci, 2006).

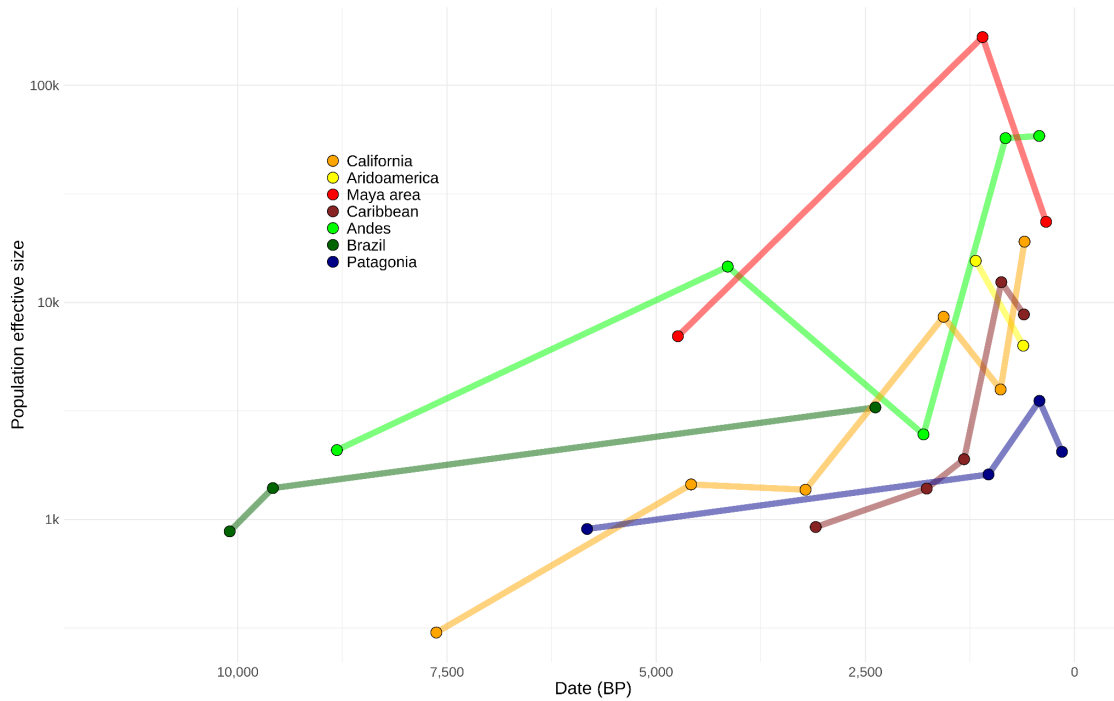


Figure 5.8. Temporal changes in population size across different geographic populations in the Americas. Circles indicate population effective size for data points, across time before present, and lines trace those data points for each geographical region. Individuals were divided into regions following the classification from Chapter 3 (Fig. 3.15 and Table S3.4).

Discussion

In this chapter, IBD segments were leveraged to investigate kinships and demography among ancient American individuals. Pedigrees were reconstructed for different geographic and temporal populations, including those for the Sumidouro individuals from Lagoa Santa in Brazil (approximately 10,184 years BP) (Fig. 5.1), early San Nicolas individuals from California (approximately 4,500 years BP) (Fig. 5.2), later San Nicolas and San Clemente individuals from California (approximately 1,000 years BP) (Fig. 5.3), and for individuals associated with the Yaghan culture in Patagonia (Fig. 5.4).

While IBD-based methods are powerful tools for detecting shared ancestry and kinships, their interpretation must take into account the demographic history of the populations involved. In principle, the likelihood of detecting distant relatives increases as effective population size decreases, due to higher background relatedness and increased genetic drift. Several populations in this study—particularly those from Patagonia and California—show evidence of long-term isolation and small population sizes, which could inflate the detection of distant IBD sharing. In such cases, some of the IBD segments may not reflect true genealogical kinship, but rather the genetic signature of small, drifted populations. An example is the detection of fifth-degree relatedness between Anzick, a ~12,600 year-old individual from Montana, and an individual from Los Rieles in Chile, dated to ~10,900 years ago. Despite their similar ages, the vast geographic distance between them raises the possibility that small ancestral population size, rather than recent shared family, is responsible for this inferred relationship. These results underscore the need for cautious interpretation of IBD sharing in ancient populations, especially in light of demographic complexity. Future studies incorporating explicit simulations of ancient populations structure and range expansion will be crucial to better understand how such factors shape IBD segment distribution and the reliability of inferred kinship degrees in ancient genomic datasets.

An additional consideration is the minimum IBD segment length used to infer relatedness. In this study, a threshold of 3 cM was applied, which allowed for the identification of many distant kin relationships, particularly at fourth- and fifth-degree

level. However, more conservative thresholds (such as 8 cM) greatly reduced the number of inferred relatives, with most kinship signals disappearing in ancient Americans under this criterion. This pattern supports the hypothesis that a portion of the observed IBD sharing may reflect the effects of small effective population sizes or drift, rather than true recent genealogical ties. Notably, the literature does not yet agree on a universal threshold for ancient DNA, likely because the optimal cutoff may vary across populations depending on their demographic history. This further underscores the need for future simulation-based work tailored to ancient population structures, in order to determine robust thresholds for IBD-based relatedness inference and to improve the reliability of kinship estimates.

Despite the presence of numerous IBD segments in ancient Americans, these segments were relatively short ([Fig. 5.5](#)), likely reflecting underlying demographic processes. However, this pattern does not appear to result from the bottleneck effect associated with the initial peopling of the Americas. Early Holocene individuals, such as the Sumidouro population, did not exhibit excessive ROH, suggesting that post-glacial populations expanded rapidly following isolation in Beringia. This aligns with mitochondrial DNA evidence indicating a sudden diversification between 16,000 and 13,000 years BP (Llamas et al., 2016).

In contrast, early San Nicolas individuals exhibited high levels of ROH, primarily consisting of short ROH segments, indicative of a persistently small population rather than recent inbreeding. Later individuals from the same region displayed reduced ROH and increasing kinship connections with individuals from neighbouring islands of San Clemente and Santa Catalina. This suggests that increased interactions with seafaring groups contributed to greater genetic diversity. These genetic interactions were observed among populations from the southern Channel Islands, with no evidence from the northern Channel Islands or the mainland populations. This suggests different population histories between the southern Channel Islands and the rest of California, consistent with ADMIXTURE results from Chapter 3 where the southern Channel Islands define a unique genetic component ([Fig. S3.6](#)).

This genetic divergence between southern and northern Channel Islanders aligns with differences in ancestry and language history. Between 5,000 and 2,000 years BP, southern Channel Islanders experienced increasing genetic contributions from

populations in Mexico, whereas this Mexican-related ancestry only appeared in the northern Channel Islands around 2,000 BP (Nakatsuka et al., 2023). This genetic pattern coincides with linguistic distinctions; during the Colonial period, southern Channel Islanders spoke an Uto-Aztecan language with origins in Mexico, whereas northern Channel Islanders spoke a language from an unrelated language family.

The influence of Mexican-related ancestry into southern California may explain the increases in estimated population size observed, potentially around 5,000 years BP and another possibly around 2,000 years BP (Fig. 5.8). The previous study estimated effective size for distinct regions as follows: Northern Channel Islands (388 ± 42), Southern Channel Islands (175 ± 13), Southern California mainland (519 ± 49), and Central California (418 ± 39) (Nakatsuka et al., 2023), while I account for population effective size over multiple time points to capture a trajectory. While differences may reflect different methods used, hapROH for the initial study and IBDNe for myself, the higher population size values in my analysis are expected, as I grouped all Californian populations together.

I refrain from discussing population sizes in absolute numbers due to limitations in estimations. Confidence intervals were not included in Figure 5.8 as they would clutter the visualisation and limit interpretability. These wide intervals reflect substantial uncertainty, primarily due to the small number of individuals contributing to each estimate. As a result, the inferred trends in effective population size should be interpreted with caution. The recommended number of individuals for each data point in IBDNe is 50 (Browning and Browning, 2015), whereas in some cases, I used as few as two individuals. While this low sample size is a common limitation in archaeogenetics, it nonetheless allowed for exploratory inference of population size changes over time, though these should be regarded as tentative and interpreted conservatively. It is also critical to distinguish between census population size (N_c) and effective population size (N_e). While census size reflects the total population, effective size only includes the breeding individuals genetically contributing to the next generation (Frankham, 1995). For humans, the census population size is generally no more than three times the effective population size (Felsenstein, 1971). Being mindful of the low sample size and the distinction between N_e (effective population size) and N_c (census population size), I refrained from discussing population sizes in absolute numbers. Instead, I focused on

the relative evolution of effective population size over time, comparing trends across different regions calculated using the same methodology.

Across the Americas, hunter-gatherer populations consistently exhibited higher levels of ROH than sedentary farming populations (Figs. 5.6-7), reinforcing the demographic contrast between these subsistence strategies. Significant increases in effective population size coincides with two major events (Fig. 5.8): the incorporation of fully domesticated maize as a staple crop in the Maya region (Kennett et al., 2022; Regueiro et al., 2013), and the likely agricultural advancements of the Tiwanaku empire in the Andes (*Tiwanaku Part I: The City*, 2021).

The Central Andes functioned as a major center of cultural and political power, home to the Tiwanaku and Inca civilisations. The Tiwanaku Empire (1,840-750 BP) pioneered high-altitude agriculture, developing raised-field systems that improved crop resilience and productivity. These fields, constructed with stone and gravel foundations and irrigated via adjacent canals, created microclimates that mitigated frost damage while sustaining aquatic ecosystems (*Tiwanaku Part I: The City*, 2021). These agricultural innovations facilitated population expansion, as reflected in the IBDNe estimates.

Lastly, a surprising population decline is observed in Aridoamerica during the period coinciding to the rise of the Aztec Empire and its capital, Tenochtitlán. As a thriving center of trade and culture, Tenochtitlán attracted merchants and visitors from across the Aridoamerica (*Children of the sun: the Aztecs*, 2019; Hodge et al., 1993; Van Dop Dejesus, 2015). However, my dataset lacks Aztec individuals, meaning the observed decline may reflect demographic shifts in other populations rather than a direct observation of Aztec dominance.

Methods

1. Data preparation and genotype imputation

Ancient genomes from the Americas, both previously published and newly-sequenced Copán individuals, were aligned and processed using the same pipeline as Chapter 3. In short, adapters were trimmed either with *AdapterRemoval* version 2.2.2 (Schubert et al., 2016) or with *cutadapt* version 1.15 (Martin, 2011), reads were aligned to the hs37d5 human reference genome using *BWA* (Burrows-Wheeler Aligner) version 0.7.5 (Li and Durbin, 2009), and filtered for the length longer than 34 bp and mapping quality higher than 25 before being softclipped by 2. Diploid data was generated through genotype imputation using a double-imputation pipeline, as introduced in Chapter 4.

All bi-allelic variants were called from ancient data, and the corresponding genotype likelihoods were calculated using *bcftools* v1.11 (Li, 2011) with the arguments “-mpileup” and “-call” respectively, and subsequently indexed with “-index”. The genotype likelihood score for all transition sites was manually lowered to “0” before imputation in order to account for potential errors caused by damaged ancient DNA that would influence the imputation output.

Ancient genomes were individually imputed using *GLIMPSE* v1.0.1 (Rubinacci et al., 2021). Input genomes were split into chunks of 2 Mb windows with a buffer size of 200 kb using *GLIMPSE_chunk* and subsequently imputed with *GLIMPSE_phase*, coupled with the 1000 Genomes Phase 3 as the reference panel for imputation (The 1000 Genomes Project Consortium et al., 2015), consisting of 75,880,357 SNP sites. The chunks were then ligated using *GLIMPSE_ligate*. Output SNPs were further filtered based on their output genotype likelihood score, with only SNPs reaching 99% considered in further analysis.

To ensure high genotype imputation performance, shotgun-sequenced genomes below $0.1\times$ were filtered out. To ensure a similar starting quality in capture genomes, a 6% missing rate threshold was applied in the *GLIMPSE*-imputed genomes (Fig. S5.3). All autosomal SNPs present on the 1240K capture panel were extracted from the *GLIMPSE*-imputed genomes for the second round of imputation.

These genomes were then subjected together for imputation using BEAGLE v5.4 (Browning et al., 2021) based on the reference panel of the 1000 Genome Phase 3 (The 1000 Genomes Project Consortium et al., 2015). This batch imputation leveraged markers from other ancient individuals to improve the overall imputation quality. The 1240K sites were then re-extracted from BEAGLE-imputed genomes. Finally, indels, multiallelic and sites below 1% frequency were removed for IBD analysis.

2. Inferring IBD segments

Identical-by-descent (IBD) segments across the double-imputed genomes of 330 ancient individuals were identified using refinedIBD (Browning and Browning, 2013). To correct for genotyping and haplotype-phase switching errors, which could split true IBD segments into shorter adjacent segments, neighbouring breaks were joined if gaps contained at most one discordant heterozygote and were less than 0.6 cM in length, using the command *merge-ibd-segments* from refinedIBD.

3. Kinship analysis

To enhance power in identifying shared IBD segments, results from three independent runs of refinedIBD with different random seed numbers were merged (Browning and Browning, 2013). To minimise false positives, only segments longer than 3cM and with high LOD score (> 10) were retained. Focusing on distant kinship, only IBD1 segments were retained for analysis, ignoring IBD2 segments—representing segments where the two individuals share both haplotypes—which are particularly common in closer relatives such as siblings (Browning et al., 2021). Pairs of individuals which shared more than 3 segments totalising to over 24cM in length were considered for kinship analysis. Coefficient of relatedness calculated by dividing the total sum of IBD segment length shared (in centiMorgans) by twice the length of the genome's recombination map, approximated to 6200cM. A coefficient of relatedness exceeding thresholds of 0.3536, 0.1786, 0.0884, 0.0442, and 0.022 correspond to 1st degree, 2nd degree, 3rd degree, 4th degree, and 5th degree relatives respectively, where, for example, 1st degree can be either a sibling or parent-offspring relation, 2nd degree is either or uncle-niece/aunt-nephew or grandparent-grandchild relation, and so on.

4. Pedigree validation with simulation

To assess the accuracy of the kinship signals, the data was compared against simulated data generated by Ariano (Ariano et al., 2022). The simulations were based on 21 Irish imputed individuals published by Cassidy (Cassidy et al., 2020), free from confounding factors such as inbreeding, relatedness, population structure, or recent changes in population sizes. This dataset served as the input for the software ped-sim (Caballero et al., 2019; Campbell et al., 2015). Three inbreeding scenarios were tested: second-degree, third-degree, and fourth-degree relatives. Each scenario was simulated 400 times using randomly sampled founders.

5. Runs-Of-Homozygosity analysis

Homozygous-by-descent segments across the double-imputed genomes of 330 ancient individuals were identified using a single run of refinedIBD (Browning and Browning, 2013). To correct for genotyping and haplotype-phase switching errors, which could split true homozygous-by-descent segments into shorter adjacent segments, neighbouring breaks were joined if gaps contained at most one discordant heterozygote and were less than 0.6 cM in length, using the command *merge-ibd-segments* from refinedIBD. Only segments with LOD score > 3 were considered for ROH analysis, and classified by length: > 20 cM, 12-20cM, 8-12cM, and 4-8cM, following categories from previous studies in the Americas (Ferraz et al., 2023; Nakatsuka et al., 2023).

6. Population effective size estimation

IBD segments were inferred from a single run of refinedIBD (Browning and Browning, 2013), with LOD score > 3 and length > 2 cM. Populations were grouped by geographical regions following the classification from Chapter 3 (Fig. 3.15 and Table S3.4). For each geographical region, dated individuals were clustered into time points based on their temporal proximity (Fig. S5.2), with the median date of the individuals within each cluster serving as the representative time point. Population effective size estimates were obtained using IBDNe (Browning and Browning, 2015) with default parameters. The effective size for each time point was calculated as the harmonic mean of the estimated population sizes across the 5th and 50th generations.

Chapter 6

Conclusions and future directions

Some of the most enduring mysteries in Maya archaeology include the origins of Maya royal dynasties and the sudden collapse of the Classic Maya civilisation. The influx of highland Mexican ancestry into Maya populations from the Classic period onwards offers a new perspective, suggesting the existence of broader mobility and social networks beyond the conventional boundaries of the Maya territory. Genomic insights can be integrated into the widely accepted archaeological concept of the dichotomy between local and non-local or between Maya and non-Maya—commonly acknowledged in the understanding of the origins of people in the Classic Maya states.

The collapse of the Classic Maya civilisation, occurring around 900 CE, has long been a subject of research. While the term “collapse” is often used to describe the breakdown of political and economic structures, it is important to recognise that the Maya people and their cultural traditions persisted beyond this period and continue to thrive today (Martínez de Velasco Cortina and Vega Villalobos, 2015; Masson, 2012). In this context, the Maya collapse refers specifically to the decline of large, interconnected city-states and the complex sociopolitical systems that once defined the region.

The collapse was most pronounced in the southern Maya highlands, where major centers like Copán were completely abandoned. Copán provides a direct archaeological record of this decline, with its last inscription on the unfinished Altar L dated to 822 CE, marking the end of political authority in the city (Andrews and Fash, 2005). Multiple hypothesis have been proposed to explain this collapse, including internal conflict, as suggested by the imagery of warfare on Stela 11 dated to 820 CE (Andrews and Fash, 2005), political instability, evidenced by the rebellion of Quiriguá in the early 9th century when it was subservient to Copan since 436 CE (*Ancient Metropolis: The true scale of Mayan Cities | Treasure Tombs of the Ancient Maya*, 2022; Andrews and Fash, 2005), overpopulation and soil depletion (Andrews and Fash, 2005; Webster et al., 1992), and prolonged droughts (Folan et al., 1983; Haug et al., 2003; Hodell et al., 1995), leading to famine.

Genomic data support the increase in population size during the Classic period, which has been implicated from the archaeological records of dense and interconnected

city-states (Regueiro et al., 2013), as well as extensive trade and communication networks (*Ancient Metropolis: The true scale of Mayan Cities | Treasure Tombs of the Ancient Maya*, 2022; Fash, 1991; Martínez de Velasco Cortina and Vega Villalobos, 2015). However, the population growth was followed by a significant decline in effective population size observed in the dataset, potentially reflecting the collapse of the Maya sociopolitical system.

The genetic approach used in this thesis parallels recent studies of other ancient societies, demonstrating how genomic data can refine narratives of social collapse. A notable example is Rapa Nui, a remote island in the South Pacific, where demographic history has been reconstructed through ancient DNA (Moreno-Mayar et al., 2024). The long-standing ecological suicide hypothesis suggests that the Rapanui deforested their island, leading to famine, warfare, and ultimately a population collapse by the 1600s. However, genomic evidence contradicts this, revealing a consistent population size and a steady population increase rather than a sharp collapse. This findings underscores the importance of integrating genetic data with archeological and environmental evidence to reassess narratives of societal decline. In the case of the Maya, the demographic collapse detected through IBDNe supports existing hypotheses of a significant population decline but also raises questions about post-collapse population dynamics and potential migration patterns.

While the collapse of major city-states in the southern Maya highlands led to population declines, evidence suggests that some regions experienced an influx of people during this period. Simulations of population size in Copán indicate that its rapid growth between 600 and 750 CE was largely driven by immigration to the Copán Valley, though the origins of these migrants remain unknown (Webster et al., 1992). Following the political fragmentation of Copán, other centers, such Chichén Itzá, rose to prominence between 800 and 925 CE, attracting people from across the region as it expanded its influence over the eastern and western coasts of the Yucatán Peninsula, likely controlling coastal trade during the Postclassic period (Masson, 2012; Olga Hernández Espinoza and Márquez Morfín, 2015). The contemporaneous decline of Copán and the expansion of Chichén Itzá raise the possibility that at least some former inhabitant of Copán relocated to emerging political centers, seeking new economic opportunities.

This pattern of demographic shifts and possible migrations in response to societal collapse is not unique to the Maya region. Beyond Mesoamerica, similar processes have been observed in the Central Andes, in the Tiwanaku empire, around 750 years BP, when successive drought conditions prompted migrations from the Tiwanaku capital in the Titicaca basin to coastal regions with more favourable climates (Fehren-Schmitz et al., 2014).

Future research can address these gaps by expanding genomic sampling across more regions and time periods, improving sequencing methods for higher-resolution data, and fostering collaborations between geneticists, archaeologists, historians, and paleoclimatologists. By building on this work, future studies can better integrate genetic, cultural, historical, environmental perspectives to deepen our understanding of ancient societies and their interconnected histories.

Throughout this thesis, I have made a sustained effort to contextualise the genetic results within the broader framework of Maya archaeology and cultural history. This includes careful attention to terminology—for example, updating outdated population labels such as “Eskimo” to “Inuit”, following Indigenous preferences (Kaplan, n.d.). I hope that this research is respectful and of value to Maya communities today, especially to those in Honduras who are descendants of the ancient Copán civilisation. According to the 2013 census, there are 33,256 Maya descendants in Honduras (Minority Rights Group, 2018). The Minority Rights Group supports the rights of these communities—including language preservation, land rights, and cultural practices—and has highlighted the importance of scientific publications and media coverage relevant to Indigenous groups. In line with this goal, most of the work presented in this thesis has contributed to a publication (Murray et al., 2025), and the archeological team led by Seiichi Nakamura has participated in a documentary broadcasted in several countries and available online (*Ancient Metropolis: The true scale of Mayan Cities | Treasure Tombs of the Ancient Maya*, 2022). I have also engaged with viewers through the comments section on the YouTube page. While I was unable to personally travel to Honduras or visit local museums due to the constraints of my doctoral program, I have relied on the collaboration of archaeologists and my supervisor, who maintain contact with local institutions and communities (Fig. 2.11). I acknowledge my own limitations—I do not speak Spanish or any Mayan language, and

I have never been to Central America—but I remain committed to supporting equitable and inclusive research practices.

Ancient DNA research also raises important ethical concerns. The extraction of DNA from archaeological materials is inherently destructive, and there is growing pressure from museums and curators to preserve these rare and valuable artefacts for future generations. Recent calls for non-destructive or minimally invasive techniques highlight the urgent need for methodological innovation in this area (Dalal et al., 2023). In conducting this research, I have valued the biological materials entrusted to our care and aimed to maximise data output. Looking forward, I believe the field of population genomics must do more to centre the voices and interests of Indigenous peoples. Archaeologists and geneticists should work together not only to co-develop better conservation practices but also to ensure that research questions, study designs, and the interpretation of results are shaped by dialogue with the communities concerned. Wherever possible, reporting should be collaborative and inclusive, involving local stakeholders—whether community leaders, activists, or researchers—to strengthen partnerships with regional institutions and honour the cultural and historical significance of the materials under study.

Acknowledgements

I would like to thank my supervisor Dr Shigeki Nakagome who handed me this opportunity and entrusted me with such a great project. Shigeki always replied quickly to my emails and found time for meetings. He was also very invested in the project, guiding my every step, and shadowing the entire analysis.

Yet, the project was only possible due to the quality archeological work of Dr Seiichi Nakamura. Seiichi excavated the valuable bones of ancient Copán individuals, believed in this collaboration and entrusted us with the raw data.

Dr Niall Cooke, the only other member of the Nakagome lab at the start, played a crucial role in my progress. Early in my PhD, he generously shared his scripts and expertise, taking time out of his own research to guide mine and acting as my first mentor.

Scripts from Dr Lara Cassidy also kickstarted some analysis, particularly in imputation and IBD analysis. Lara consistently answered my questions with clarity and friendliness, and her deep knowledge and thoughtful criticism were invaluable throughout the process. Additionally, I thank Prof Dan Bradley for advertising the PhD program at the very beginning. I had emailed him with a spontaneous application, and he redirected me to this funded opportunity.

In the Bradley lab in Trinity College Dublin, I especially thank Dr Marco Rosaria Capodiferro for closely monitoring the publication of ancient genomes from the Americas. Marco also sat with me on multiple occasions to answer my questions and teach me about mitochondrial haplogroups. His comments, with those of Dr Kevin Daly and Dr Victoria Mullin during lab meetings, were an important source of feedback in validating my analysis and anticipating potential comments from journals during the publication process.

Naturally, I send my thanks to all my other colleagues within the Bradley, Cassidy and Huerta-Sanchez labs at Trinity College Dublin (in no particular order): Dr Valeria

Mattiangeli, Dr Bruno Ariano, Dr Andrew Hare, Dr Anahit Hovhannisyan, Dr Emily Breslin, Dr Conor Rossi, Dr Iseult Jackson, Dr Linda Ongaro, Aine Halpin, Deborah Diquelou, Louis l'Hôte, Harry Little, Bruno González, Ciaran O'Connor, Corey Alwell and Catherine Butt, as well as Dr Ross Byrne. They always made it a safe and friendly environment to work in.

I would also like to send a big thank you to Dr Karsten Hokamp, Dr Iseult Jackson and Dr Marco Rosario Capodiferro for putting time and effort into maintaining the server, rebooting it when necessary, and answering questions on pre-installed software.

Beyond my colleagues, I am deeply grateful to my peers from the Centre for Research Training (CRT) program cohort. Starting a PhD in a new country during the COVID pandemic was a considerable challenge, and my cohort provided invaluable support during that isolating time. The program brought together individuals of similar age and position, creating a unique sense of camaraderie. Together, we navigated the difficulties of cancelled flights, shifting regulations, and vaccine protocols. I hope I was as much a source of support for them as they were for me.

I would like to particularly thank newly graduated Dr Rasha Shraim from Trinity College Dublin, Dr Tembi Fashina and Dr Ashwini Shanmugan from Royal College of Surgeons in Ireland (RSCI), who were part of my CRT cohort. Our writing sessions in quiet, often obscure rooms turned the daunting thesis process into a shared and lighter experience.

I also want to acknowledge Dr Stéphane Petrousek, whose similar thesis deadline made him an invaluable companion during the final sprint. Together, we shared progress, setbacks, and advice, which gave me a sense of solidarity throughout the writing process. Finally, my heartfelt thanks go to my partner and family, who provided meals, encouragement, and unwavering support during those intense final months.

As part of the program organised by the CRT, we had to carry out a placement of a minimum of three months, outside of the Republic of Ireland. Dr Leo Speidel welcomed me in his group at the University College of London (UCL) in the United Kingdom. During the three months placement, Leo was very reachable and his office door was always open. He even continued to supervise my progress through online meetings set

every two weeks several after the end of my placement and until the end of my PhD. I thank him for his time and dedication.

At UCL, I was also warmly greeted by Dr Nancy Bird and Prof Garrett Hellenthal. And I was lucky to have the chance to meet the ancient genomics team at Francis Crick Institute under the group leader Dr Pontus Skoglund. He opened his lab to me, and I was able to attend journal clubs and seminars.

The additional funding provided by the CRT was invaluable in supporting my placement in London, especially given the high cost of living.

Finally, I send my thanks to the Training Coordinator Dr Simone Coughlan and the Programme Manager Dr Sandra Healy from the CRT for the smooth running of the program. They followed our progress throughout, answered our questions, listened to feedback and worked hard to improve our experience.

This research was funded by Science Foundation Ireland through the SFI Centre for Research Training in Genomics Data Science under Grant number 18/CRT/6214. This research was supported in part by the EU's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant H2020-MSCA-COFUND-2019-945385.



Trinity College Dublin

The University of Dublin

Centre for
Research
Training





As a concluding remark, I thank Dr Víctor Moreno-Mayar and Prof Ross McManus, respectively the external and internal examiners from the University of Copenhagen and Trinity College Dublin, for making me at ease during the viva. They provided for a comfortable environment which allowed for an interesting and enjoyable discussion.

Supplementary figures

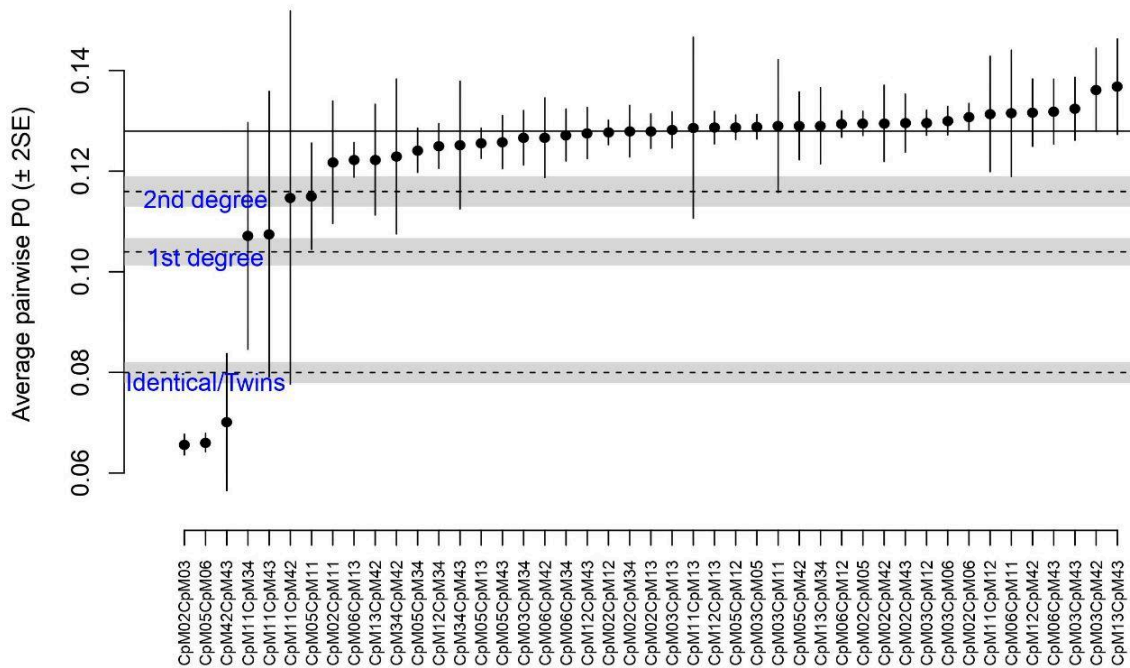


Figure S3.1. Results of kinship analysis by READ on the initial 16 bone extracts from Copán. Pairs of extracts tested are listed on the *x*-axis, while the *y*-axis represents average pairwise P0 scores. Error bars show the two standard errors of the mean, and dashed lines show the cut-off for the identical twins and 1st and 2nd degree relatives.

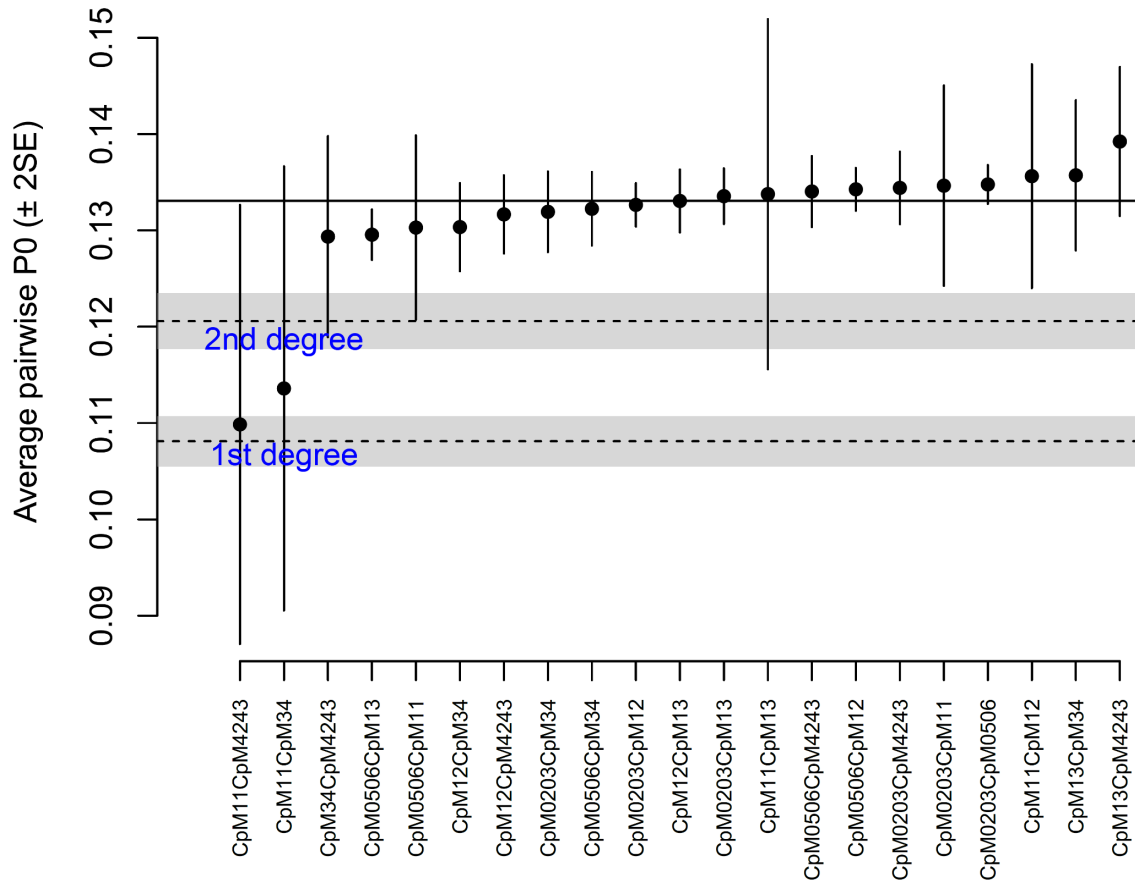


Figure S3.2. Results of kinship analysis by READ for the final Classic Copán individuals. Pairs of individuals tested are listed on the *x*-axis, while the *y*-axis represents average pairwise P0 scores. Error bars show the two standard errors of the mean, and dashed lines show the cut-off for the 1st and 2nd degree relatives.

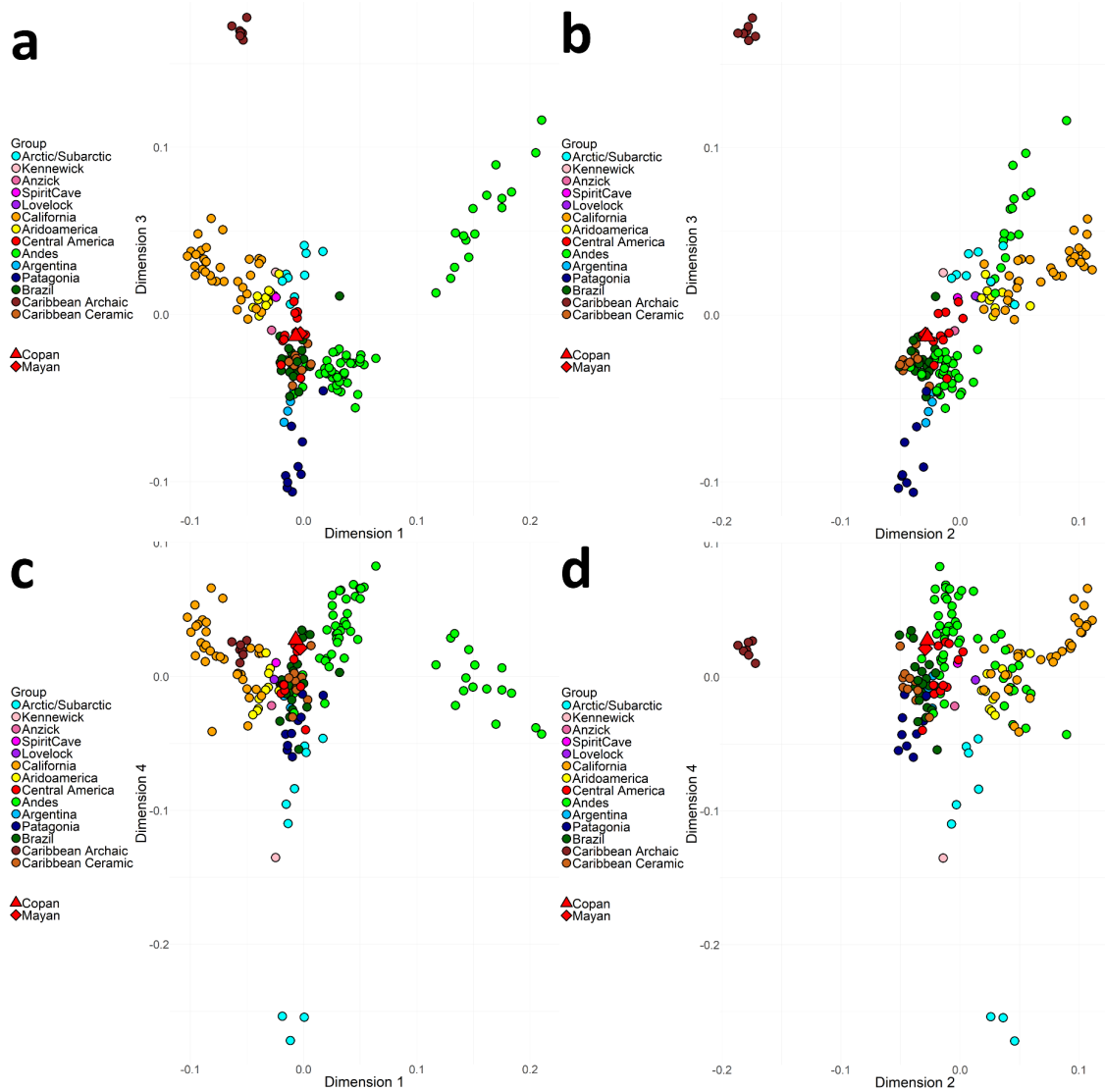


Figure S3.3. Population structure in the Americas as revealed by MDS. Multidimensional scaling (MDS) plot for the matrix of outgroup f_3 -statistics for all pairs of ancient and present-day populations. Populations are colour-coded based on their geographic location (see Fig. 3.15; Table S3.4), with the Classic Copan represented by a red triangle and the present-day Maya by a red diamond. Different dimensions are compared as follows: (a) Dimension 1 versus Dimension 3, (b) Dimension 2 versus Dimension 3, (c) Dimension 1 versus Dimension 4, and (d) Dimension 2 versus Dimension 4.

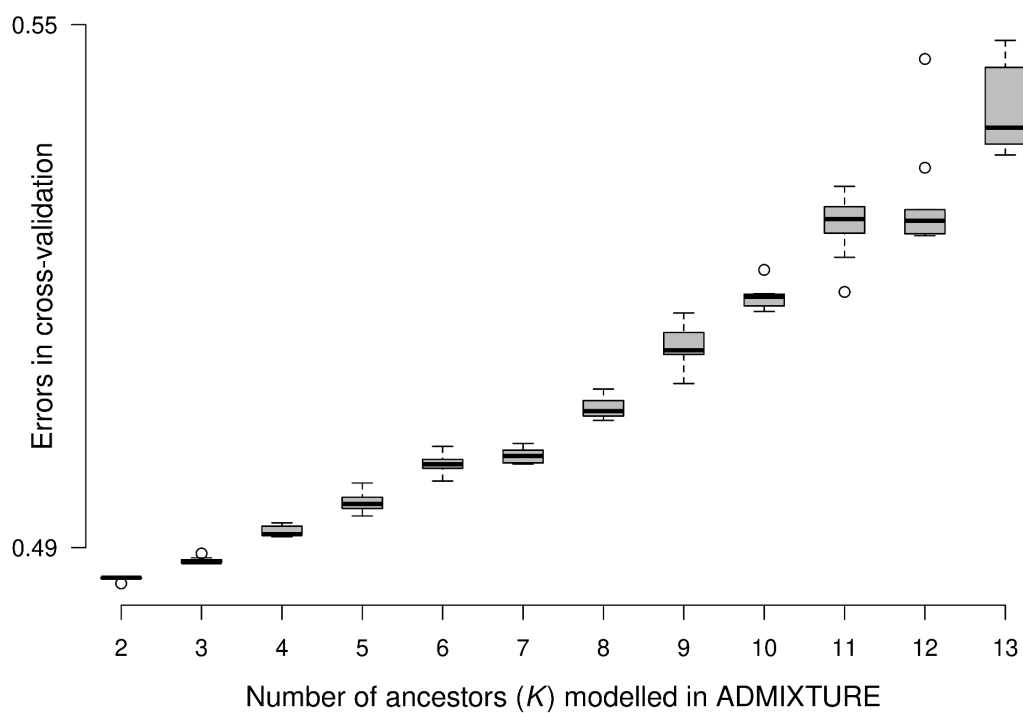


Figure S3.4. Cross validation (CV) errors used in ADMIXTURE analysis. Each K component includes 10 replicates.

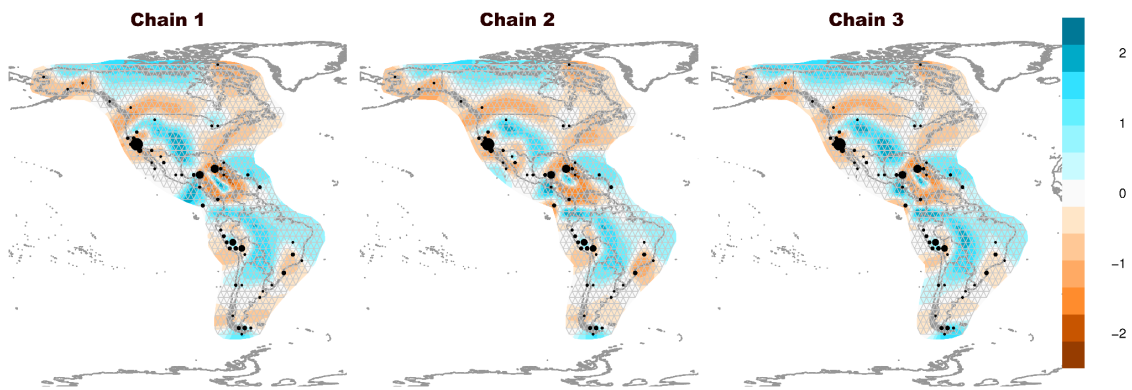


Figure S3.7. Three independent random runs of EEMS. The EEMS run was repeated three times, titled “Chain 1”, “Chain 2”, and “Chain 3”, with random seeds 123, 456, and 789 respectively. The run with the highest likelihood was selected for the EEMS main ([Fig. 3.20](#)), corresponding to run number 1. Estimated effective migration rates in the log scale with normalisation so that 0 corresponds to the overall mean migration rate. As such, blue corresponds to migration corridors, while orange highlights geographical barriers.

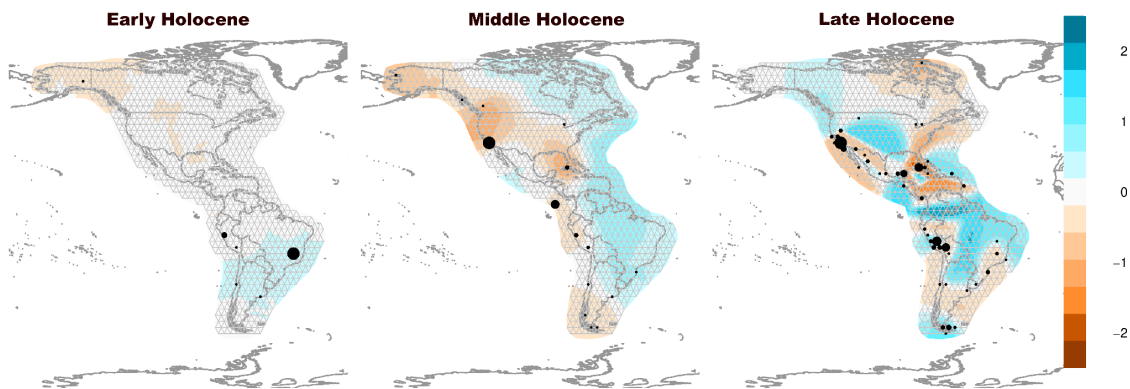


Figure S3.8. Migration rates for the Early Holocene, Middle Holocene, and Late Holocene periods. The Early Holocene was set as older than 8,000 years ago, the Middle Holocene as dating between 8,000 and 2,500 years ago, and the Late Holocene as more recent than 2,500 years ago. Estimated effective migration rates in the log scale with normalisation so that 0 corresponds to the overall mean migration rate. As such, blue corresponds to migration corridors, while orange highlights geographical barriers.

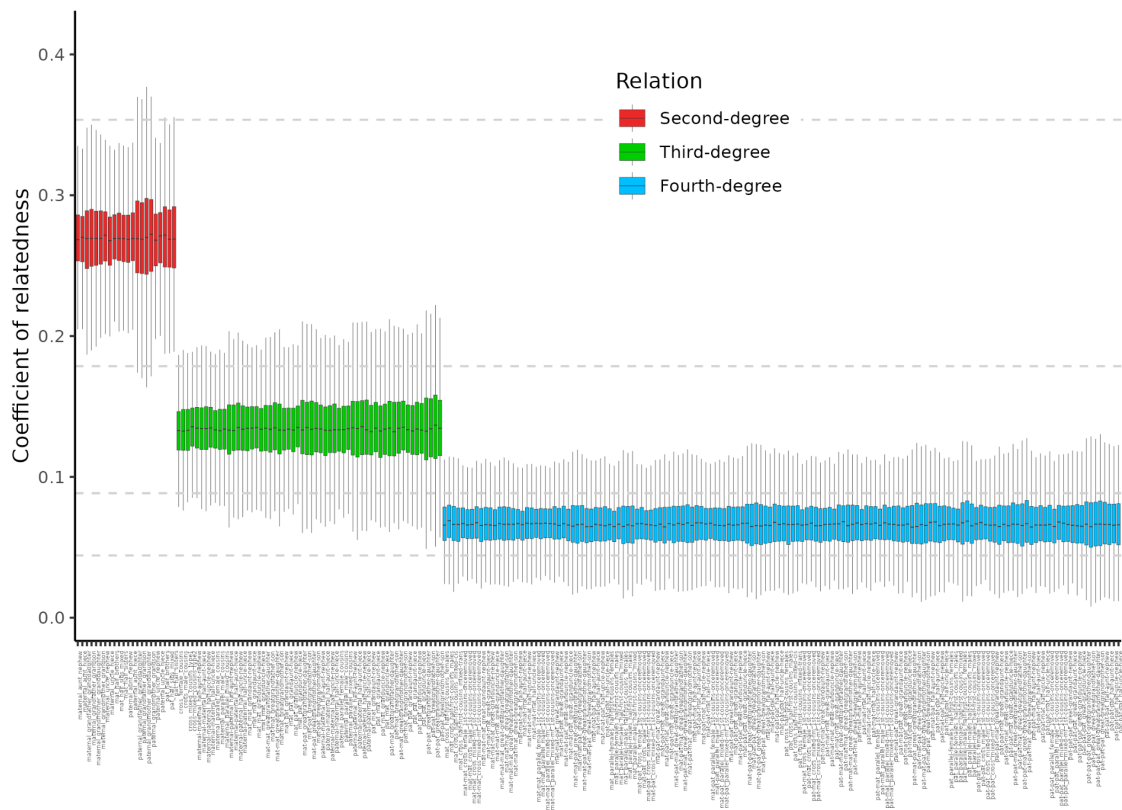


Figure S5.1. Coefficient of relatedness variation. Simulated data of 2nd-, 3rd-, and 4th-degree relationships are depicted in different colours (red, green, and blue respectively), with a boxplot for each pairing type, showing standard error. The dashed horizontal grey lines indicate the cutoffs applied for inferring degrees of relatedness: 4th-degree for coefficients between 0.0442 and 0.0884, 3rd-degree for coefficients between 0.0884 and 0.1786, and 2nd-degree for coefficients between 0.0884 and 0.3536.

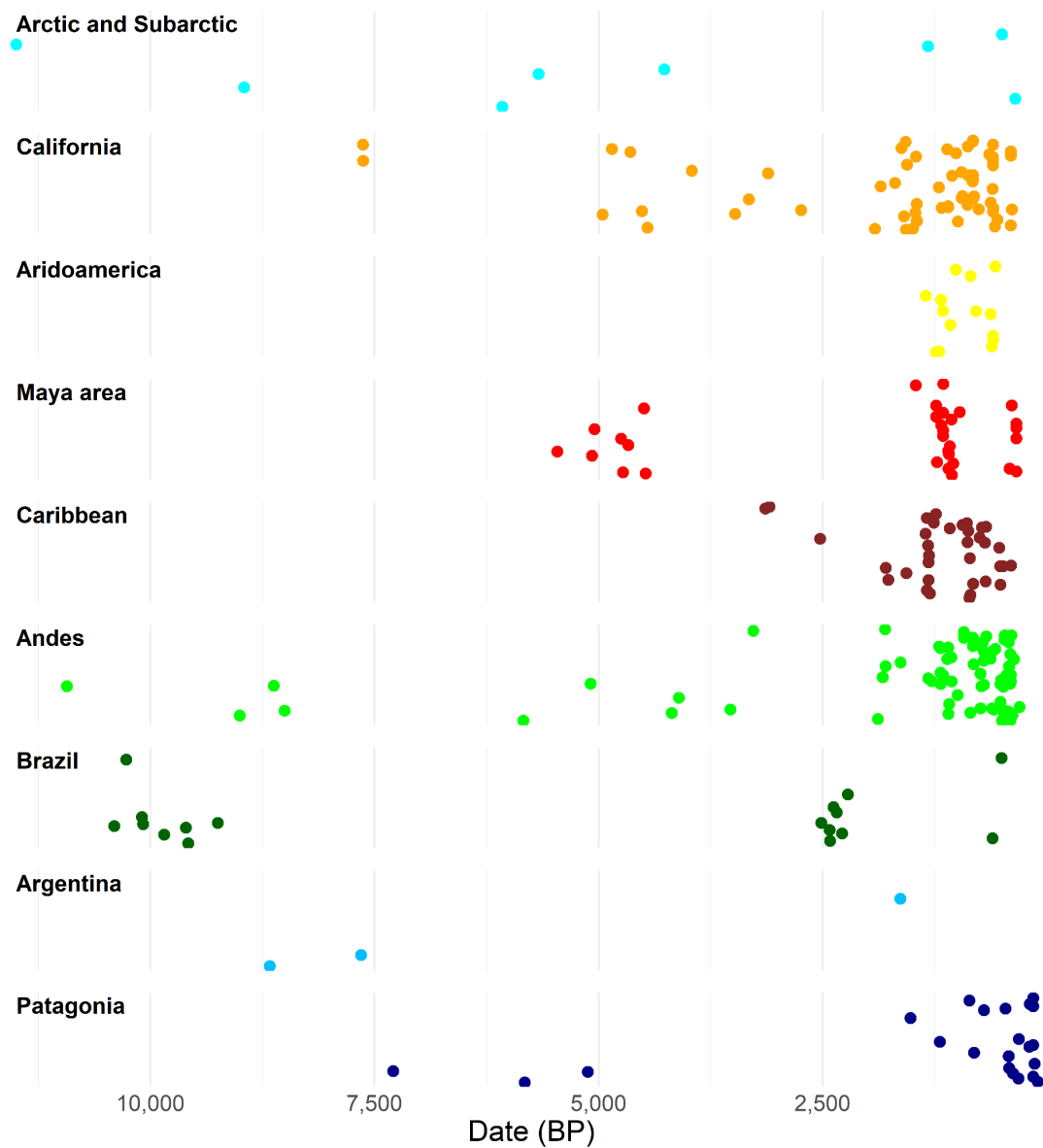


Figure S5.2. Temporal distribution of imputed individuals. Ancient Americas are separated according to geographical regions: cyan for the Arctic, orange for California, yellow for Aridoamerica, red for the Maya area, brown for the Caribbean, green for the Andes, dark green for Brazil, sky blue for Argentina, and dark blue for Patagonia (see [Fig. 3.15](#)). Each point is an ancient American individual jittered across its carbon datation in years before present.

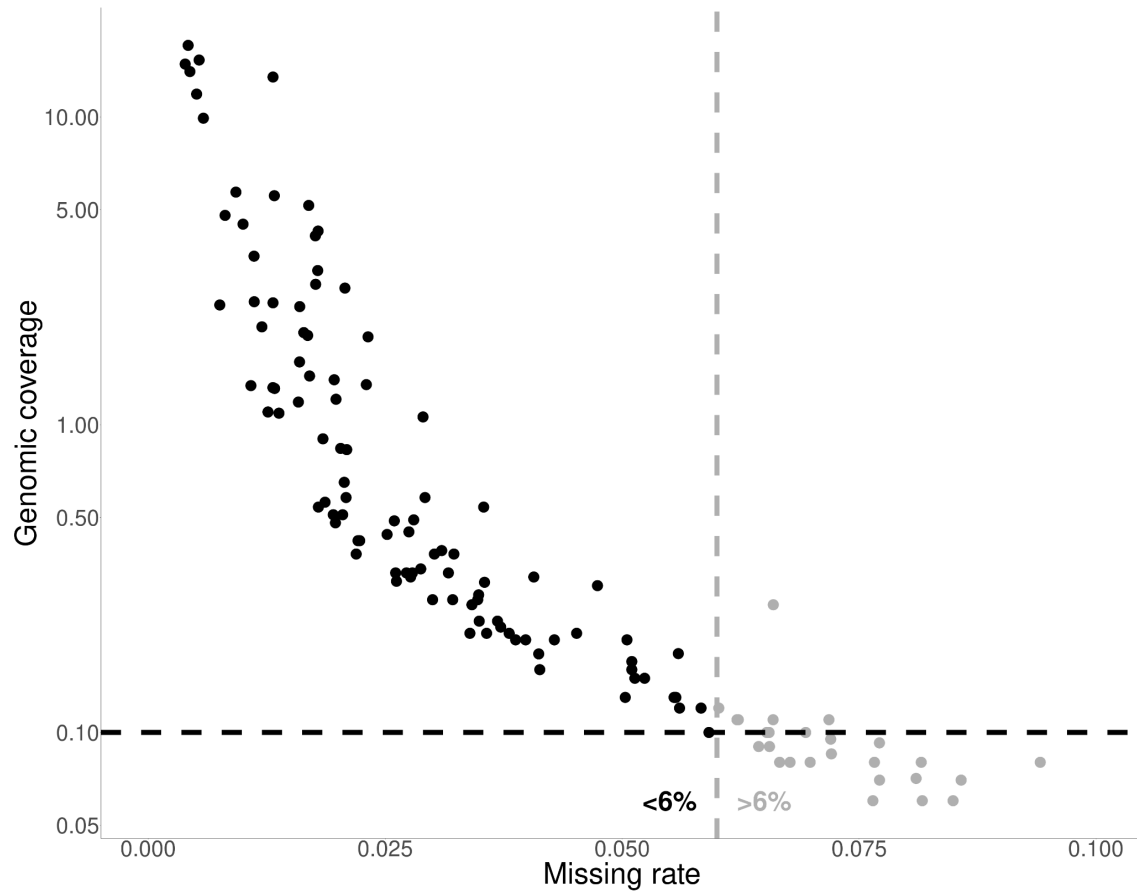


Figure S5.3. A relationship between missing rates in GLIMPSE-imputed genomes and their coverage in pre-imputed data. Each dot is an ancient individual who was originally whole-genome shotgun-sequenced. A dashed grey line indicates a 6% missingness cut-off, which corresponds to genomic coverage of $>0.1\times$, marked by a dashed black line.

Supplementary tables

Table S3.1. Endogenous content estimations in newly sequenced bone extracts

ID	Endogenous content ^a	Extract coverage ^b	Final coverage ^c
CpM02	17.31%	0.2363	0.4486
CpM03	7.17%	0.2123	
CpM05	15.25%	0.2843	0.4908
CpM06	16.83%	0.2064	
CpM08	0.44%		
CpM11	1.2%	0.0051	0.0051
CpM12	30.76%	0.3076	0.3076
CpM13	8.17%	0.0925	0.0925
CpM15	0.45%		
CpM17	0.62%		
CpM26	0.47%		
CpM30	0.63%		
CpM34	3.21%	0.0334	0.0334
CpM42	2.13%	0.0157	0.0498
CpM43	3.78%	0.0341	
CpM45	0.45%		

^aestimated from reads produced by MiSeq

^bestimated from reads produced by NovaSeq

^cfinal estimation from merged reads produced by MiSeq and NovaSeq

Table S3.2. Present-day American genomes count in HOA and SGDP

	HOA	SGDP	region
Bolivian	7	-	Argentina
Chane	-	1	Bolivia
Karitiana	12	3	Amazonia
Maya	18	2	Mexico
Mixe	10	3	Mexico
Mixtec	10	2	Mexico
Piapoco	4	2	Colombia
Pima	14	2	Arizona
Quechua	5	3	Peru
Surui	8	2	Amazonia
Zapotec	10	2	Mexico
<i>total</i>	<i>98</i>	<i>22</i>	

Table S3.3. American ancestry proportion in present-day Americans

Population	Individual ID	HOA		SGDP	
		American	Non-American	American	Non-American
Chane		NA		97.36%	2.64%
Bolivian	Bolivian81	79.93%	20.07%	NA	
	Bolivian39	98.77%	1.23%	NA	
	Bolivian41	82.31%	17.69%	NA	
	Bolivian46	96.24%	3.76%	NA	
	Bolivian62	86.89%	13.11%	NA	
	Bolivian71	67.65%	32.35%	NA	
	Bolivian72	66.22%	33.78%	NA	
Karitiana	HGDP00995	100%	0%	NA	
	HGDP00999	100%	0%	NA	
	HGDP01001	100%	0%	NA	
	HGDP01003	100%	0%	NA	
	HGDP01006	100%	0%	NA	
	HGDP01010	100%	0%	NA	
	HGDP01012	100%	0%	100%	0%
	HGDP01013	100%	0%	NA	
	HGDP01014	100%	0%	NA	
	HGDP01015	100%	0%	100%	0%
	HGDP01018	100%	0%	100%	0%
	HGDP01019	100%	0%	NA	
Maya	HGDP00854	83.03%	16.97%	NA	
	HGDP00855	95.37%	4.63%	95.64%	4.36%
	HGDP00856	89.33%	10.67%	NA	
	HGDP00857	95.49%	4.51%	95.37%	4.63%
	HGDP00858	88.91%	11.09%	NA	
	HGDP00859	85.76%	14.24%	NA	
	HGDP00861	63.25%	36.75%	NA	

	HGDP00862	79.41%	20.59%	NA	
	HGDP00863	83.24%	16.76%	NA	
	HGDP00864	90.05%	9.95%	NA	
	HGDP00865	90.93%	9.07%	NA	
	HGDP00868	73.25%	26.75%	NA	
	HGDP00869	64.65%	35.35%	NA	
	HGDP00870	87.38%	12.62%	NA	
	HGDP00872	86.60%	13.40%	NA	
	HGDP00873	91.14%	8.86%	NA	
	HGDP00875	78.23%	21.77%	NA	
	HGDP00877	76.11%	23.89%	NA	
Mixe	mixe0002	NA		95.74%	4.26%
	mixe0007	97.83%	2.17%	95.33%	4.67%
	mixe0009	97.48%	2.52%	NA	
	mixe0015	99.49%	0.51%	NA	
	mixe0018	99.13%	0.87%	NA	
	mixe0026	98.76%	1.24%	NA	
	mixe0027	98.23%	1.77%	NA	
	mixe0028	98.36%	1.64%	NA	
	mixe0029	99.45%	0.55%	NA	
	mixe0030	96.86%	3.14%	NA	
	mixe0035	98.09%	1.91%	NA	
	mixe0042	NA		98.14%	1.86%
Mixtec	Mixa0001	93.77%	6.23%	NA	
	Mixa0002	81.73%	18.27%	NA	
	Mixa0016	93.62%	6.38%	NA	
	Mixa0018	93.01%	6.99%	NA	
	Mixa0024	90.81%	9.19%	NA	
	Mixa0049	93.32%	6.68%	NA	
	Mixa0061	91.54%	8.46%	NA	

	Mixa0078	92.47%	7.53%	NA	
	Mixa0087	87.33%	12.67%	NA	
	Mixa0093	87.22%	12.78%	NA	
	Mixa0099	NA		86.89%	13.11%
	Mixa0105	NA		52.92%	47.08%
Piapoco	HGDP00702	100%	0%	100%	0%
	HGDP00704	100%	0%	NA	
	HGDP00706	100%	0%	100%	0%
	HGDP00970	100%	0%	NA	
Pima	HGDP01037	90.89%	9.11%	NA	
	HGDP01041	97.76%	2.24%	NA	
	HGDP01043	99.74%	0.26%	NA	
	HGDP01044	98.90%	1.10%	98.97%	1.03%
	HGDP01047	98.76%	1.24%	99.76%	0.24%
	HGDP01050	89.49%	10.51%	NA	
	HGDP01051	92.56%	7.44%	NA	
	HGDP01053	98.14%	1.86%	NA	
	HGDP01055	88.36%	11.64%	NA	
	HGDP01056	99.51%	0.46%	NA	
	HGDP01057	97.76%	2.24%	NA	
	HGDP01058	85.36%	14.64%	NA	
	HGDP01059	96.14%	3.86%	NA	
	HGDP01060	89.45%	10.55%	NA	
Quechua	NA11197	77.42%	22.58%	NA	
	NA11198	69.57%	30.43%	NA	
	NA11199	72.49%	27.51%	NA	
	NA11200	95.55%	4.45%	95.33%	4.67%
	NA11201	85.25%	14.75%	95.42%	4.58%
	Peru60	NA		85.13%	14.87%
	HGDP00832	100%	0%	NA	

Surui	HGDP00837	100%	0%	NA	
	HGDP00838	100%	0%	NA	
	HGDP00843	100%	0%	NA	
	HGDP00845	100%	0%	NA	
	HGDP00846	100%	0%	100%	0%
	HGDP00849	100%	0%	NA	
	HGDP00852	100%	0%	100%	0%
Zapotec	Zapo0046	92.18%	7.82%	NA	
	Zapo0059	91.88%	8.12%	NA	
	Zapo0064	88.48%	11.52%	NA	
	Zapo0072	88.98%	11.02%	NA	
	Zapo0076	89.40%	10.60%	NA	
	Zapo0082	94.06%	5.94%	NA	
	Zapo0083	92.24%	7.76%	NA	
	Zapo0091	92.33%	7.67%	NA	
	Zapo0095	89.10%	10.90%	NA	
	Zapo0097	90.33%	9.67%	NA	
	Zapo0098	NA		92.79%	7.21%
	Zapo0099	NA		88.51%	11.46%

Table S3.4. A summary of all ancient data analysed in Chapters 3 and 5, including newly generated Copán and previously published genomes.

Population label	#Individuals (processed)	#Individuals ^a (analysed)	Geography	Reference	Index ^b
AG2	1	1	Siberia	(Raghavan et al., 2014b)	-
MA1	1	1	Siberia	(Raghavan et al., 2014b)	-
Birnirk	2	2	Siberia	(Raghavan et al., 2014a)	-
USR1	1	1	Arctic/Subarctic	(Moreno-Mayar et al., 2018a)	13
USR2	1	1	Arctic/Subarctic	(Moreno-Mayar et al., 2018a)	14
Alaskan_Athabaskan	1	1	Arctic/Subarctic	(Scheib et al., 2018)	15
TrailCreek	1	1	Arctic/Subarctic	(Moreno-Mayar et al., 2018b)	16
Norton	1	1	Arctic/Subarctic	(Raghavan et al., 2014a)	-
LucyIslands	1	1	Arctic/Subarctic	(Raghavan et al., 2015)	17
BigBar	1	1	Arctic/Subarctic	(Moreno-Mayar et al., 2018b)	18
Saqqaq	2	2	Arctic/Subarctic	(Raghavan et al., 2014a)	-
ASO_4000	1	1	Arctic/Subarctic	(Scheib et al., 2018)	19
Pre_Dorset	1	1	Arctic/Subarctic	(Raghavan et al., 2014a)	-
Late_Dorset	2	2	Arctic/Subarctic	(Raghavan et al., 2014a)	20
Middle_Dorset	14	14	Arctic/Subarctic	(Raghavan et al., 2014a)	21
ASO_Huron-Wendat	2	2	Arctic/Subarctic	(Scheib et al., 2018)	22
Thule	4	4	Arctic/Subarctic	(Raghavan et al., 2014a)	-
ASO	12	12	Arctic/Subarctic	(Scheib et al., 2018)	23
OldMissionPoint	1	1	Arctic/Subarctic	(Raghavan et al., 2015)	24
Colonist	1	1	Arctic/Subarctic	(Scheib et al., 2018)	-
Anzick	1	1	Other	(Rasmussen et al., 2014)	25
SpiritCave	1	1	Other	(Moreno-Mayar et al., 2018b)	26
Kennewick	1	1	Other	(Rasmussen et al., 2015)	27
Lovelock	4	4	Other	(Moreno-Mayar et al., 2018b)	28
San_Francisco_Bay	1	1	California	(Scheib et al., 2018)	29
USA_CA_SantaRosa_7400BP	5	5	California	(Nakatsuka et al., 2023)	30
USA_CA_Carpinteria_7000BP	6	6	California	(Nakatsuka et al., 2023)	31
USA_CA_SantaRosa_4900BP	2	1	California	(Nakatsuka et al., 2023)	32
USA_CA_SantaBarbara_4900BP	1	0	California	(Nakatsuka et al., 2023)	-
USA_CA_Goleta_4700BP	3	3	California	(Nakatsuka et al., 2023)	33
USA_CA_SantaBarbara_4600BP	1	1	California	(Nakatsuka et al., 2023)	34
ESN	20	20	California	(Scheib et al., 2018)	35
USA_CA_Carpinteria_4000BP	1	1	California	(Nakatsuka et al., 2023)	36
USA_CA_Malibu_3400BP	4	2	California	(Nakatsuka et al., 2023)	37
USA_CA_SantaRosa_3200BP	2	2	California	(Nakatsuka et al., 2023)	38
USA_CA_SantaRosa_3000BP	5	4	California	(Nakatsuka et al., 2023)	39
USA_CA_Goleta_3000BP	1	1	California	(Nakatsuka et al., 2023)	40
USA_CA_Goleta_2600BP	1	1	California	(Nakatsuka et al., 2023)	41
USA_CA_Goleta_1700BP	1	1	California	(Nakatsuka et al., 2023)	42
LSN	15	15	California	(Scheib et al., 2018)	43
USA_CA_CalaverasCounty_1500BP	1	1	California	(Nakatsuka et al., 2023)	44
Mainland_Chumash	17	17	California	(Scheib et al., 2018)	45
USA_CA_Ojai_1400BP	2	2	California	(Nakatsuka et al., 2023)	46
USA_CA_SantaBarbara_1300BP	1	1	California	(Nakatsuka et al., 2023)	47
USA_CA_Goleta_1200BP	2	2	California	(Nakatsuka et al., 2023)	48
LSCI	9	9	California	(Scheib et al., 2018)	49
USA_CA_Monterey_1000BP	1	1	California	(Nakatsuka et al., 2023)	50
USA_CA_CastroVile_900BP	1	1	California	(Nakatsuka et al., 2023)	51
USA_CA_Goleta_900BP	4	3	California	(Nakatsuka et al., 2023)	52
USA_CA_Goleta_800BP	1	1	California	(Nakatsuka et al., 2023)	53
Island_Chumash	3	3	California	(Scheib et al., 2018)	54
USA_CA_CalaverasCounty_600BP	6	6	California	(Nakatsuka et al., 2023)	55
USA_CA_Carmel_600BP	5	5	California	(Nakatsuka et al., 2023)	56

USA_CA_SantaBarbara_600BP	2	2	California	(Nakatsuka et al., 2023)	57
USA_CA_SanClemente_500BP	8	8	California	(Nakatsuka et al., 2023)	58
USA_CA_PebbleBeach_500BP	1	0	California	(Nakatsuka et al., 2023)	-
USA_CA_Goleta_400BP	1	1	California	(Nakatsuka et al., 2023)	59
USA_CA_LakeCachuma_400BP	2	2	California	(Nakatsuka et al., 2023)	60
USA_CA_SantaRosa_300BP	6	6	California	(Nakatsuka et al., 2023)	61
USA_CA_PacificGrove_200BP	1	1	California	(Nakatsuka et al., 2023)	62
Pericu	1	1	Aridoamerica	(Scheib et al., 2018)	63
Baja_Mexico	2	2	Aridoamerica	(Scheib et al., 2018)	64
M_CoyoteCave_2700BP	1	0	Aridoamerica	(Nakatsuka et al., 2023)	-
M_LaPlaya_CerroDeTrincheras_2400BP	3	3	Aridoamerica	(Nakatsuka et al., 2023)	65
Canada_de_la_Virgen	4	4	Aridoamerica	(Villa-Islas et al., 2023)	66
M_CuevaDeLosMuertosChiquitos_1100BP	9	9	Aridoamerica	(Nakatsuka et al., 2023)	67
Mexico_Tayopa_1000BP	7	4	Aridoamerica	(Nakatsuka et al., 2023)	68
Sierra_Gorda	4	4	Aridoamerica	(Villa-Islas et al., 2023)	69
Mexico_CoyoteCave_950BP	1	1	Aridoamerica	(Nakatsuka et al., 2023)	70
Sierra_Tarahumara	1	1	Aridoamerica	(Villa-Islas et al., 2023)	71
M_SanLorenzo_700BP	1	0	Aridoamerica	(Nakatsuka et al., 2023)	-
Paleoamericans	4	4	Aridoamerica	(Raghavan et al., 2015)	72
M_LaPlaya_CerroDeTrincheras_600BP	6	6	Aridoamerica	(Nakatsuka et al., 2023)	73
Pre-Columbian	2	2	Aridoamerica	(Raghavan et al., 2015)	-
PreLatePaleoindian	1	0	Central America	(Kennett et al., 2022)	-
Belize_MayahakCabPek_9300BP	1	1	Central America	(Posth et al., 2018)	74
LatePaleoindian	3	2	Central America	(Kennett et al., 2022)	75
Belize_SakiTzul_7400BP	2	2	Central America	(Posth et al., 2018)	76
EarlyArchaic	1	0	Central America	(Kennett et al., 2022)	-
LateArchaic	15	12	Central America	(Kennett et al., 2022)	77
Copán	7	6	Central America	This study	78
Chichen_Itza	61	52	Central America	(Barquera et al., 2024)	79
ancientPanama	9	9	Central America	(Capodiferro et al., 2021)	80
colonialPanama	1	1	Central America	(Capodiferro et al., 2021)	-
Mexico_Colonial	10	10	Central America	(Tiesler et al., 2022)	81
Mexico_Colonial_African	1	1	Central America	(Tiesler et al., 2022)	-
Mexico_Colonial_o	1	1	Central America	(Tiesler et al., 2022)	-
Chile_LosRieles_10900BP	1	1	Andes	(Posth et al., 2018)	82
Peru_Cuncaicha_9000BP	1	1	Andes	(Posth et al., 2018)	83
Peru_Lauricocha_8600BP	3	3	Andes	(Posth et al., 2018)	84
Kaillachuro	1	1	Andes	(Lindo et al., 2018)	85
SMP	1	1	Andes	(Lindo et al., 2018)	86
MummyMaderasEncoC2	1	1	Andes	(Raghavan et al., 2015)	-
Peru_Lauricocha_5800BP	1	1	Andes	(Posth et al., 2018)	87
Chile_LosRieles_5100BP	1	1	Andes	(Posth et al., 2018)	88
Peru_Cuncaicha_4200BP	1	1	Andes	(Posth et al., 2018)	89
Peru_LaGalgada_4100BP	1	1	Andes	(Posth et al., 2018)	90
Peru_Lauricocha_3500BP	1	1	Andes	(Posth et al., 2018)	91
Peru_Cuncaicha_3300BP	1	1	Andes	(Posth et al., 2018)	92
Rio_Uncallane	5	5	Andes	(Lindo et al., 2018)	93
Peru_Ullujaya_1350BP_EIP	2	2	Andes	(Nakatsuka et al., 2020a)	94
Peru_ElBrujo_1300BP_EIP	5	5	Andes	(Nakatsuka et al., 2020a)	95
Chile_CaletaHuelen_1100BP_MH	3	3	Andes	(Nakatsuka et al., 2020a)	96
TIW	8	8	Andes	(Popović et al., 2021)	97
LUK	4	4	Andes	(Popović et al., 2021)	98
Peru_Lima_1450-1050BP_EIP-MH	5	5	Andes	(Nakatsuka et al., 2020a)	99
Bolivia_Miraflores_1100BP_MH	4	3	Andes	(Nakatsuka et al., 2020a)	100
Bolivia_Iroco_1050BP_MH	2	1	Andes	(Nakatsuka et al., 2020a)	101
Bolivia_Tiwanaku_1000BP_MH	4	4	Andes	(Nakatsuka et al., 2020a)	102
Peru_Campanayuq_1000BP_MH	3	3	Andes	(Nakatsuka et al., 2020a)	103
Peru_Chinchawas_1000BP_MH	2	2	Andes	(Nakatsuka et al., 2020a)	104
Peru_Ullujaya_950BP_MH	1	1	Andes	(Nakatsuka et al., 2020a)	105

Peru_Laramate_900BP	6	6	Andes	(Posth et al., 2018)	106
Peru_Palpa_950BP_MH	1	1	Andes	(Nakatsuka et al., 2020a)	107
Peru_ElBrujo_850BP_MH-LIP	4	4	Andes	(Nakatsuka et al., 2020a)	108
Peru_Ollantaytambo_LIP	5	5	Andes	(Salazar et al., 2023)	109
Chile_Pukara_700BP_LIP	2	2	Andes	(Nakatsuka et al., 2020a)	110
Chile_PicaOcho_700BP	1	1	Andes	(Posth et al., 2018)	111
Chile_Conchali_700BP	2	2	Andes	(Posth et al., 2018)	112
Peru_Charrangochayoc_700BP_LIP	1	1	Andes	(Nakatsuka et al., 2020a)	113
Peru_Chinchawas_700BP_LIP	3	3	Andes	(Nakatsuka et al., 2020a)	114
Peru_Lima_650BP_LIP	7	7	Andes	(Nakatsuka et al., 2020a)	115
Peru_Laramate_650BP_LIP	4	4	Andes	(Nakatsuka et al., 2020a)	116
Peru_SanSebastian_600BP_LIP	2	1	Andes	(Nakatsuka et al., 2020a)	117
ORU	1	1	Andes	(Popović et al., 2021)	118
COR	4	4	Andes	(Popović et al., 2021)	119
Peru_Palpa_550BP_LIP	2	2	Andes	(Nakatsuka et al., 2020a)	120
Peru_CasaConcha2_LH	1	1	Andes	(Salazar et al., 2023)	121
Peru_Qotakalli2_LH	1	1	Andes	(Salazar et al., 2023)	122
Peru_Qotakalli1_LH	3	3	Andes	(Salazar et al., 2023)	123
Peru_SanSebastian_500BP_LH	1	1	Andes	(Nakatsuka et al., 2020a)	124
Aconcagua	1	1	Andes	(Moreno-Mayar et al., 2018b)	125
Peru_Kusicancha1_LH	3	3	Andes	(Salazar et al., 2023)	126
Peru_Sacsahuaman1_LH	1	1	Andes	(Salazar et al., 2023)	127
Peru_Sacsahuaman2_LH	3	3	Andes	(Salazar et al., 2023)	128
Peru_Sacsahuaman3_LH	1	1	Andes	(Salazar et al., 2023)	129
Peru_Cusco_Torontoy_A_450BP_LH	2	2	Andes	(Nakatsuka et al., 2020a)	130
Peru_Cusco_Torontoy_B_450BP_LH	1	1	Andes	(Nakatsuka et al., 2020a)	131
Peru_Mesayocpata_450BP_LH	1	1	Andes	(Nakatsuka et al., 2020a)	132
Peru_Kusicancha2_LH	1	1	Andes	(Salazar et al., 2023)	133
Peru_MachuPicchu_LH	33	33	Andes	(Salazar et al., 2023)	134
Peru_Kanamarka1_LH	2	2	Andes	(Salazar et al., 2023)	135
Peru_Kanamarka2_LH	1	1	Andes	(Salazar et al., 2023)	136
Peru_Kanamarka3_LH	1	1	Andes	(Salazar et al., 2023)	137
Peru_Chincha_LH	6	6	Andes	(Bongers et al., 2020)	138
Argentina_ArroyoSeco2_7700BP	8	7	Argentina	(Posth et al., 2018)	139
Argentina_LagunaChica_6800BP	3	3	Argentina	(Nakatsuka et al., 2020a)	140
Argentina_LagunaToro_2400BP	1	1	Argentina	(Nakatsuka et al., 2020b)	141
Argentina_LagunaChica_1600BP	1	1	Argentina	(Posth et al., 2018)	142
PuntaSantaAna	1	1	Patagonia	(Moreno-Mayar et al., 2018b)	143
NorthTierradelFuego_LaArcillosa2_5800BP	1	1	Patagonia	(Nakatsuka et al., 2020b)	144
Ayayema	1	1	Patagonia	(Moreno-Mayar et al., 2018b)	145
BeagleChannel_Grouped_1500-100BP	4	4	Patagonia	(Nakatsuka et al., 2020b)	146
NorthTierradelFuego_Grouped_500BP	6	5	Patagonia	(Nakatsuka et al., 2020b)	147
Haush_MitrePeninsula_Grouped_400BP	6	6	Patagonia	(Nakatsuka et al., 2020b)	148
Aonikenk_CerroJohnny_400BP	1	1	Patagonia	(Nakatsuka et al., 2020b)	149
Fuego-Patagonian	4	4	Patagonia	(Raghavan et al., 2015)	150
NorthTierradelFuego_FaroMendez_100BP	1	1	Patagonia	(Nakatsuka et al., 2020b)	151
Sumidouro5	1	1	Brazil	(Moreno-Mayar et al., 2018b)	152
Sumidouro	4	4	Brazil	(Moreno-Mayar et al., 2018b)	153
Capelinha_10400BP	1	1	Brazil	(Ferraz et al., 2023)	154
Brazil_LapaDoSanto_9600BP	7	6	Brazil	(Posth et al., 2018)	155
LocaDoSuin_9100BP	1	1	Brazil	(Ferraz et al., 2023)	156
Brazil_Laranjal_6700BP	2	2	Brazil	(Posth et al., 2018)	157
Brazil_Moraes_5800BP	1	1	Brazil	(Posth et al., 2018)	158
SerradaCapivara	1	1	Brazil	(Raghavan et al., 2015)	-
Cabecuda_3200BP	2	2	Brazil	(Ferraz et al., 2023)	159
Limao_2700BP	1	1	Brazil	(Ferraz et al., 2023)	160
CubataoI_2700BP	2	2	Brazil	(Ferraz et al., 2023)	161
JabuticabeiraII_2400BP	14	12	Brazil	(Ferraz et al., 2023)	162
JabuticabeiraII_111_112_2200BP	2	1	Brazil	(Ferraz et al., 2023)	163

Brazil_Jabuticabeira2_2000BP	5	5	Brazil	(Posth et al., 2018)	164
Lima_1900BP	4	4	Brazil	(Ferraz et al., 2023)	165
JabuticabeiraII_102_1300BP	1	1	Brazil	(Ferraz et al., 2023)	166
GalhetaIV_1200BP	1	1	Brazil	(Ferraz et al., 2023)	167
Vau_Una_600BP	1	1	Brazil	(Ferraz et al., 2023)	168
Lima_500BP	1	1	Brazil	(Ferraz et al., 2023)	169
Palmeiras	1	1	Brazil	(Ferraz et al., 2023)	170
Kaingang	1	1	Brazil	(Ferraz et al., 2023)	171
PedraDoAlexandre_2_undated	1	1	Brazil	(Ferraz et al., 2023)	172
Guayabo_Blanco	3	3	Caribbean	(Nägele et al., 2020)	173
Playa_del_Mango	8	8	Caribbean	(Nägele et al., 2020)	174
Las_Carolinas_II	1	1	Caribbean	(Nägele et al., 2020)	175
Cueva_Calero	4	3	Caribbean	(Nägele et al., 2020)	176
Cueva_del_Perico	8	8	Caribbean	(Nägele et al., 2020)	177
Canimar_Abajo	26	23	Caribbean	(Nägele et al., 2020)	178
Manuelito	1	1	Caribbean	(Nägele et al., 2020)	179
Tibes	2	2	Caribbean	(Nägele et al., 2020)	180
Taino	1	1	Caribbean	(Schroeder et al., 2018)	181
Anse_a_la_Gourde	4	4	Caribbean	(Nägele et al., 2020)	182
Lavoutte	12	12	Caribbean	(Nägele et al., 2020)	183
Preachers_Cave	4	4	Caribbean	(Nägele et al., 2020)	184
Los_Indios	1	1	Caribbean	(Nägele et al., 2020)	185
Punta_Candelero	5	4	Caribbean	(Nägele et al., 2020)	186
Paso_del_Indio	7	7	Caribbean	(Nägele et al., 2020)	187
El_Morrillo	1	1	Caribbean	(Nägele et al., 2020)	188
Cueva_de_los_Esqueletos	5	5	Caribbean	(Nägele et al., 2020)	189

^aThe column represents the number of individuals remaining after excluding those identified as related, contaminated, or low-coverage in previous studies. A count of zero in this column indicates that this population has been excluded from all analyses.

^bThe indices correspond to the numbers labelled in the ADMIXTURE plots of [Fig. 3.18](#). Populations marked with a hyphen are excluded from the analysis due to having fewer than 10,000 SNPs available after LD pruning. Indices 1 through 12 represent present-day American populations.

Table S3.5. Testing fittings of two-way admixture models between LateArchaic and Zapotec, Mixe, or Sierra_Gorda ancestries to the genetic ancestry of ancient and present-day populations from the Maya region.

Model parameters	Ancient Maya populations									Present-day Maya		
	Classic Copán			Chichén Itzá			Colonial Mexico					
Model	Two-way admixture											
DoF	11	11	11	11	11	11	11	11	11	11	11	11
χ^2	9.308	10.682	16.671	9.663	19.052	30.722	7.791	21.570	27.361	2.881	14.933	23.634
Tail prob	0.594	0.470	0.118	0.561	0.060	0.001	0.732	0.028	0.004	0.992	0.186	0.014
Admixture proportions	Source 1:											
	Late Archaic											
	93.9% (±2.6%)	97.1% (±2.2%)	98.0% (±3.5%)	48.3% (±8.0%)	58.2% (±9.9%)	22.5% (±28.6%)	48.0% (±11.9%)	55.0% (±14.1%)	50.6% (±41.9%)	76.5% (±9.4%)	79.2% (±15.9%)	108.9% (±7.8%)
	Source 2:											
	Zapotec	Mixe*	Sierra Gorda	Zapotec	Mixe*	Sierra Gorda	Zapotec	Mixe*	Sierra Gorda	Zapotec	Mixe*	Sierra Gorda
	6.1% (±2.6%)	2.9% (±2.2%)	2.0% (±3.5%)	51.7% (±8.0%)	41.8% (±9.9%)	77.5% (±28.6%)	52.0% (±11.9%)	45.0% (±14.1%)	49.4% (±41.9%)	23.5% (±9.4%)	20.8% (±15.9%)	-8.9% (±7.9%)
Model	Single ancestry with Source 1											
DoF	12	12	12	12	12	12	12	12	12	12	12	12
χ^2	14.282	11.878	15.824	56.130	44.876	56.294	34.971	39.010	34.989	22.618	25.807	23.740
Tail prob	0.283	0.456	0.199	1.134e-07	1.082e-05	1.060e-07	4.732e-04	1.048e-04	4.702e-04	0.031	0.011	0.022
Model	Single ancestry with Source 2											
DoF	12	12	12	12	12	12	12	12	12	12	12	12
χ^2	57.998	80.311	63.930	25.615	44.111	31.811	18.688	36.663	31.412	9.578	19.926	32.108
Tail prob	5.216e-08	3.600e-12	4.296e-09	0.012	1.462e-05	1.480e-03	0.096	2.522e-04	1.704e-03	0.653	0.068	0.001
Model selection	Testing dual ancestry vs. single ancestry scenarios											
Nested p -value against Source 1	0.026	0.274	-nan	0.000	5.560e-07	0.297	0.000	1.023e-04	0.044	0.000	0.025	0.744
Nested p -value against Source 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.000	0.000

*In the set of right populations, Mixe was replaced by Zapotec

Table S3.6. Testing fittings of the two-way admixture model between LateArchaic and highland Mexican ancestry to the genetic ancestry of each of the two temporal groups in the Classic Copán

Model parameters	Classic Copán	
	Early-to-Middle	Middle-to-Late
Model	Two-way admixture	
DoF	11	11
χ^2	5.095	15.065
Tail prob	0.926	0.18
Admixture proportions	Source 1: Late Archaic	
	93.1% ($\pm 3.2\%$)	94.3% ($\pm 3.9\%$)
	Source 2: Highland Mexican (represented by Zapotec)	
	6.9% ($\pm 3.2\%$)	5.7% ($\pm 3.9\%$)
Model	Single ancestry with Source 1	
DoF	12	12
χ^2	10.482	16.575
Tail prob	0.574	0.166
Model	Single ancestry with Source 2	
DoF	12	12
χ^2	48.09	44.024
Tail prob	3.014e-06	1.513e-05
Model selection	Testing dual ancestry vs. single ancestry scenarios	
Nested <i>p</i> -value against Source 1	0.020	0.219
Nested <i>p</i> -value against Source 2	0.000	0.000

References

- Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Amplify Core Knowledge Language Arts. Grade 5, [2022 edition]. ed, 2022. . Amplify Core Knowledge Language Arts, Brooklyn, New York.
- Ancient Metropolis: The true scale of Mayan Cities | Treasure Tombs of the Ancient Maya, 2022. . tvf international.
- Andermann, T., Torres Jiménez, M.F., Matos-Maraví, P., Batista, R., Blanco-Pastor, J.L., Gustafsson, A.L.S., Kistler, L., Liberal, I.M., Oxelman, B., Bacon, C.D., Antonelli, A., 2020. A Guide to Carrying Out a Phylogenomic Target Sequence Capture Project. *Front. Genet.* 10, 1407. <https://doi.org/10.3389/fgene.2019.01407>
- Andrews, 2010. A quality control tool for high throughput sequence data.
- Andrews, E.W., Fash, W.L. (Eds.), 2005. Copán: the history of an ancient Maya kingdom, 1st ed. ed, School of American Research advanced seminar series. School of American Research : James Currey, Santa Fe, N.M. : Oxford.
- Ariano, B., Mattiangeli, V., Breslin, E.M., Parkinson, E.W., McLaughlin, T.R., Thompson, J.E., Power, R.K., Stock, J.T., Mercieca-Spiteri, B., Stoddart, S., Malone, C., Gopalakrishnan, S., Cassidy, L.M., Bradley, D.G., 2022. Ancient Maltese genomes and the genetic geography of Neolithic Europe. *Current Biology* 32, 2668-2680.e6. <https://doi.org/10.1016/j.cub.2022.04.069>
- Ávila-Arcos, M.C., De La Fuente Castro, C., Nieves-Colón, M.A., Raghavan, M., 2022. Recommendations for Sustainable Ancient DNA Research in the Global South: Voices From a New Generation of Paleogenomicists. *Front. Genet.* 13, 880170. <https://doi.org/10.3389/fgene.2022.880170>
- Bandelt, H.-J., Herrnstadt, C., Yao, Y.-G., Kong, Q.-P., Kivisild, T., Rengo, C., Scozzari, R., Richards, M., Villems, R., Macaulay, V., Howell, N., Torroni, A., Zhang, Y.-P., 2003. Identification of Native American Founder mtDNAs Through the Analysis of Complete mtDNA Sequences: Some Caveats. *Ann Human Genet* 67, 512–524. <https://doi.org/10.1046/j.1469-1809.2003.00049.x>
- Barquera, R., Del Castillo-Chávez, O., Nägele, K., Pérez-Ramallo, P., Hernández-Zaragoza, D.I., Szolek, A., Rohrlach, A.B., Librado, P., Childebayeva, A., Bianco, R.A., Penman, B.S., Acuña-Alonzo, V., Lucas, M., Lara-Riegos, J.C., Moo-Mezeta, M.E., Torres-Romero, J.C., Roberts, P., Kohlbacher, O., Warinner, C., Krause, J., 2024. Ancient genomes reveal insights into ritual life at Chichén Itzá. *Nature* 630, 912–919. <https://doi.org/10.1038/s41586-024-07509-7>
- Bell, E.E., Canuto, M.A., Sharer, R.J., Society for American Archaeology (Eds.), 2004. Understanding early classic Copan, 1st ed. ed. University Of Pennsylvania Museum of Archaeology And Anthropology, Philadelphia.
- Bergsland, K., n.d. Encyclopaedia Britannica. Eskimo-Aleut languages. URL <https://www.britannica.com/topic/Eskimo-Aleut-languages>
- Bongers, J.L., Nakatsuka, N., O’Shea, C., Harper, T.K., Tantaleán, H., Stanish, C., Fehren-Schmitz, L., 2020. Integration of ancient DNA with transdisciplinary dataset finds strong support for Inca resettlement in the south Peruvian coast. *Proc. Natl. Acad. Sci. U.S.A.* 117, 18359–18368. <https://doi.org/10.1073/pnas.2005965117>
- Briggs, A.W., Stenzel, U., Johnson, P.L.F., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., Pääbo, S., 2007. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14616–14621. <https://doi.org/10.1073/pnas.0704665104>

- Brookfield, J., 1996. Population genetics. *Current Biology* 6, 354–356.
[https://doi.org/10.1016/S0960-9822\(02\)00493-1](https://doi.org/10.1016/S0960-9822(02)00493-1)
- Browning, B.L., Browning, S.R., 2013. Improving the Accuracy and Efficiency of Identity-by-Descent Detection in Population Data. *Genetics* 194, 459–471.
<https://doi.org/10.1534/genetics.113.150029>
- Browning, B.L., Tian, X., Zhou, Y., Browning, S.R., 2021. Fast two-stage phasing of large-scale sequence data. *The American Journal of Human Genetics* 108, 1880–1890. <https://doi.org/10.1016/j.ajhg.2021.08.005>
- Browning, S.R., Browning, B.L., 2015. Accurate Non-parametric Estimation of Recent Effective Population Size from Segments of Identity by Descent. *The American Journal of Human Genetics* 97, 404–418.
<https://doi.org/10.1016/j.ajhg.2015.07.012>
- Browning, S.R., Browning, B.L., 2012. Identity by Descent Between Distant Relatives: Detection and Applications. *Annu. Rev. Genet.* 46, 617–633.
<https://doi.org/10.1146/annurev-genet-110711-155534>
- Burger, J., Hummel, S., Herrmann, B., Henke, W., 1999. DNA preservation: A microsatellite-DNA study on ancient skeletal remains. *Electrophoresis* 20, 1722–1728.
[https://doi.org/10.1002/\(SICI\)1522-2683\(19990101\)20:8<1722::AID-ELPS1722>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1522-2683(19990101)20:8<1722::AID-ELPS1722>3.0.CO;2-4)
- Caballero, M., Seidman, D.N., Qiao, Y., Sannerud, J., Dyer, T.D., Lehman, D.M., Curran, J.E., Duggirala, R., Blangero, J., Carmi, S., Williams, A.L., 2019. Crossover interference and sex-specific genetic maps shape identical by descent sharing in close relatives. *PLoS Genet* 15, e1007979.
<https://doi.org/10.1371/journal.pgen.1007979>
- Campbell, C.L., Furlotte, N.A., Eriksson, N., Hinds, D., Auton, A., 2015. Escape from crossover interference increases with maternal age. *Nat Commun* 6, 6260.
<https://doi.org/10.1038/ncomms7260>
- Campelo dos Santos, A.L., Owings, A., Sullasi, H.S.L., Gokcumen, O., DeGiorgio, M., Lindo, J., 2022. Genomic evidence for ancient human migration routes along South America's Atlantic coast. *Proc. R. Soc. B.* 289, 20221078.
<https://doi.org/10.1098/rspb.2022.1078>
- Capodiferro, M.R., Aram, B., Raveane, A., Rambaldi Migliore, N., Colombo, G., Ongaro, L., Rivera, J., Mendizábal, T., Hernández-Mora, I., Tribaldos, M., Perego, U.A., Li, H., Scheib, C.L., Modi, A., Gómez-Carballea, A., Grugni, V., Lombardo, G., Hellenthal, G., Pascale, J.M., Bertolini, F., Grieco, G.S., Cereda, C., Lari, M., Caramelli, D., Pagani, L., Metspalu, M., Friedrich, R., Knipper, C., Olivieri, A., Salas, A., Cooke, R., Montinaro, F., Motta, J., Torroni, A., Martín, J.G., Semino, O., Malhi, R.S., Achilli, A., 2021. Archaeogenomic distinctiveness of the Isthmo-Colombian area. *Cell* 184, 1706–1723.e24.
<https://doi.org/10.1016/j.cell.2021.02.040>
- Cartwright, M., 2014. Inca Civilization. *World History Encyclopedia*.
- Cassidy, L.M., Maoldúin, R.Ó., Kador, T., Lynch, A., Jones, C., Woodman, P.C., Murphy, E., Ramsey, G., Dowd, M., Noonan, A., Campbell, C., Jones, E.R., Mattiangeli, V., Bradley, D.G., 2020. A dynastic elite in monumental Neolithic society. *Nature* 582, 384–388. <https://doi.org/10.1038/s41586-020-2378-6>
- Cassidy, L.M., Martiniano, R., Murphy, E.M., Teasdale, M.D., Mallory, J., Hartwell, B., Bradley, D.G., 2016. Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proc. Natl. Acad. Sci. U.S.A.* 113, 368–373. <https://doi.org/10.1073/pnas.1518445113>
- Cassidy, L.M., Russell, M., Smith, M., Delbarre, G., Cheetham, P., Manley, H., Mattiangeli, V., Breslin, E.M., Jackson, I., McCann, M., Little, H., O'Connor, C.G., Heaslip, B., Lawson, D., Endicott, P., Bradley, D.G., 2025. Continental

- influx and pervasive matrilocality in Iron Age Britain. *Nature*.
<https://doi.org/10.1038/s41586-024-08409-6>
- Castro E Silva, M.A., Ferraz, T., Bortolini, M.C., Comas, D., Hünemeier, T., 2021. Deep genetic affinity between coastal Pacific and Amazonian natives evidenced by Australasian ancestry. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2025739118.
<https://doi.org/10.1073/pnas.2025739118>
- Cavalli-Sforza, L.L., 2005. The Human Genome Diversity Project: past, present and future. *Nat Rev Genet* 6, 333–340. <https://doi.org/10.1038/nrg1596>
- Children of the sun: the Aztecs, 2019.
- Children of the sun: the Incas, 2019.
- Children of the sun: the Maya, 2019.
- Clemente, F., Unterländer, M., Dolgova, O., Amorim, C.E.G., Corrado-Santos, F., Neuenschwander, S., Ganiatsou, E., Cruz Dávalos, D.I., Anchieri, L., Michaud, F., Winkelbach, L., Blöcher, J., Arizmendi Cárdenas, Y.O., Sousa Da Mota, B., Kalliga, E., Souleles, A., Kontopoulos, I., Karamitrou-Mentessidi, G., Philaniotou, O., Sampson, A., Theodorou, D., Tsipopoulou, M., Akamatis, I., Halstead, P., Kotsakis, K., Urem-Kotsou, D., Panagiotopoulos, D., Ziota, C., Triantaphyllou, S., Delaneau, O., Jensen, J.D., Moreno-Mayar, J.V., Burger, J., Sousa, V.C., Lao, O., Malaspinas, A.-S., Papageorgopoulou, C., 2021. The genomic history of the Aegean palatial civilizations. *Cell* 184, 2565-2586.e21.
<https://doi.org/10.1016/j.cell.2021.03.039>
- Cobos, R. (Ed.), 2004. *Culto funerario en la sociedad maya: memoria de la Cuarta Mesa Redonda de Palenque*, 1. ed. ed. Presented at the Mesa Redonda de Palenque, Instituto Nacional de Antropología e Historia, México, D.F.
- Cooke, N.P., Murray, M., Cassidy, L.M., Mattiangeli, V., Okazaki, K., Kasai, K., Gakuhari, T., Bradley, D.G., Nakagome, S., 2024. Genomic imputation of ancient Asian populations contrasts local adaptation in pre- and post-agricultural Japan. *iScience* 27, 110050.
<https://doi.org/10.1016/j.isci.2024.110050>
- Da Fonseca, R.R., Smith, B.D., Wales, N., Cappellini, E., Skoglund, P., Fumagalli, M., Samaniego, J.A., Carøe, C., Ávila-Arcos, M.C., Hufnagel, D.E., Korneliussen, T.S., Vieira, F.G., Jakobsson, M., Arriaza, B., Willerslev, E., Nielsen, R., Hufford, M.B., Albrechtsen, A., Ross-Ibarra, J., Gilbert, M.T.P., 2015. The origin and evolution of maize in the Southwestern United States. *Nature Plants* 1, 14003.
<https://doi.org/10.1038/nplants.2014.3>
- Dalal, V., Pasupuleti, N., Chaubey, G., Rai, N., Shinde, V., 2023. Advancements and Challenges in Ancient DNA Research: Bridging the Global North–South Divide. *Genes* 14, 479. <https://doi.org/10.3390/genes14020479>
- Damgaard, P.B., Margaryan, A., Schroeder, H., Orlando, L., Willerslev, E., Allentoft, M.E., 2015. Improving access to endogenous DNA in ancient bones and teeth. *Sci Rep* 5, 11184. <https://doi.org/10.1038/srep11184>
- Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. *GigaScience* 10, giab008.
<https://doi.org/10.1093/gigascience/giab008>
- Das, S., Abecasis, G.R., Browning, B.L., 2018. Genotype Imputation from Large Reference Panels. *Annu. Rev. Genom. Hum. Genet.* 19, 73–96.
<https://doi.org/10.1146/annurev-genom-083117-021602>
- Davies, D., 2016. *Maya: a journey through the Maya world*. Millgate House, Sandbach.
- Davies, D., 2012. *The Maya Calendar Explained*. Educational Resources on the Maya by Dr Diane Davies. URL
<https://www.mayaarchaeologist.co.uk/public-resources/maya-world/maya-calendar-system/>

- de la Fuente, C., Ávila-Arcos, M.C., Galimany, J., Carpenter, M.L., Homburger, J.R., Blanco, A., Contreras, P., Cruz Dávalos, D., Reyes, O., San Roman, M., Moreno-Estrada, A., Campos, P.F., Eng, C., Huntsman, S., Burchard, E.G., Malaspinas, A.-S., Bustamante, C.D., Willerslev, E., Llop, E., Verdugo, R.A., Moraga, M., 2018. Genomic insights into the origin and diversification of late maritime hunter-gatherers from the Chilean Patagonia. *Proc. Natl. Acad. Sci. U.S.A.* 115. <https://doi.org/10.1073/pnas.1715688115>
- De La Fuente Castro, C., Cortés, C., Raghavan, M., Castillo, D., Castro, M., Verdugo, R.A., Moraga, M., 2024. The Genomic and Cultural Diversity of the Inka Qhapaq Hucha Ceremony in Chile and Argentina. *Genome Biology and Evolution* 16, evae196. <https://doi.org/10.1093/gbe/evae196>
- De La Fuente Castro, C., Figueiro, G., 2024. Human ancient DNA analysis in Latin America: Current state and challenges. *Archaeometry arcm*.12979. <https://doi.org/10.1111/arcm.12979>
- De Marino, A., Mahmoud, A.A., Bose, M., Bircan, K.O., Terpolovsky, A., Bamunusinghe, V., Bohn, S., Khan, U., Novković, B., Yazdi, P.G., 2022. A comparative analysis of current phasing and imputation software. *PLoS ONE* 17, e0260177. <https://doi.org/10.1371/journal.pone.0260177>
- Der Sarkissian, C., Allentoft, M.E., Ávila-Arcos, M.C., Barnett, R., Campos, P.F., Cappellini, E., Ermini, L., Fernández, R., Da Fonseca, R., Ginolhac, A., Hansen, A.J., Jónsson, H., Korneliussen, T., Margaryan, A., Martin, M.D., Moreno-Mayar, J.V., Raghavan, M., Rasmussen, M., Velasco, M.S., Schroeder, H., Schubert, M., Seguin-Orlando, A., Wales, N., Gilbert, M.T.P., Willerslev, E., Orlando, L., 2015. Ancient genomics. *Phil. Trans. R. Soc. B* 370, 20130387. <https://doi.org/10.1098/rstb.2013.0387>
- Diamond, J., Bellwood, P., 2003. Farmers and Their Languages: The First Expansions. *Science* 300, 597–603. <https://doi.org/10.1126/science.1078208>
- Dillehay, T.D., Ocampo, C., Saavedra, J., Sawakuchi, A.O., Vega, R.M., Pino, M., Collins, M.B., Scott Cummings, L., Arregui, I., Villagran, X.S., Hartmann, G.A., Mella, M., González, A., Dix, G., 2015. New Archaeological Evidence for an Early Human Presence at Monte Verde, Chile. *PLoS ONE* 10, e0141923. <https://doi.org/10.1371/journal.pone.0141923>
- Diversity of thought and data enrich archaeology, 2024. . *Nat Hum Behav* 8, 183–184. <https://doi.org/10.1038/s41562-024-01848-1>
- Do, H., Dobrovic, A., 2012. Dramatic reduction of sequence artefacts from DNA isolated from formalin-fixed cancer biopsies by treatment with uracil-DNA glycosylase. *Oncotarget* 3, 546–558. <https://doi.org/10.18632/oncotarget.503>
- Engineering in the Andes Mountains: History and Design of Inca Suspension Bridges, 2005. , Hispanic Division. Washington, D.C.
- Fash, W., 2002. Religion and Human Agency in Ancient Maya History: Tales from the Hieroglyphic Stairway. *CAJ* 12, 5–19. <https://doi.org/10.1017/S095977430200001X>
- Fash, W.L., 1991. Scribes, warriors, and kings: the city of Copán and the ancient Maya, New aspects of antiquity. Thames and Hudson, New York, N.Y.
- Fash, W.L., Long, K.Z., 1983. Mapa arqueológico del Valle de Copan. In *Introducción a la arqueología de Copán, Honduras, Proyecto Arqueológico Copán, Secretaría de Estado en el Despacho de Cultura y Turismo*. ed.
- Fehren-Schmitz, L., Haak, W., Mächtle, B., Masch, F., Llamas, B., Tomasto Cagigao, E., Sossna, V., Schitteck, K., Isla Cuadrado, J., Eitel, B., Reindel, M., 2014. Climate change underlies global demographic, genetic, and cultural transitions in pre-Columbian southern Peru. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9443–9448. <https://doi.org/10.1073/pnas.1403466111>
- Felsenstein, J., 1971. INBREEDING AND VARIANCE EFFECTIVE NUMBERS IN

- POPULATIONS WITH OVERLAPPING GENERATIONS. *Genetics* 68, 581–597. <https://doi.org/10.1093/genetics/68.4.581>
- Ferguson, K., 2007. The Relationship of the Maya and Teotihuacan: A Mesoamerican Mystery. *Inquiry Journal*, Durham, NH: Hamel Center for Undergraduate Research, University of New Hampshire UNH Undergraduate Research Journal.
- Fernandes, D.M., Sirak, K.A., Ringbauer, H., Sedig, J., Rohland, N., Cheronet, O., Mah, M., Mallick, S., Olalde, I., Culleton, B.J., Adamski, N., Bernardos, R., Bravo, G., Broomandkhoshbacht, N., Callan, K., Candilio, F., Demetz, L., Carlson, K.S.D., Eccles, L., Freilich, S., George, R.J., Lawson, A.M., Mandl, K., Marzaioli, F., McCool, W.C., Oppenheimer, J., Özdoğan, K.T., Schattke, C., Schmidt, R., Stewardson, K., Terrasi, F., Zalzal, F., Antúnez, C.A., Canosa, E.V., Colten, R., Cucina, A., Genchi, F., Kraan, C., La Pastina, F., Lucci, M., Maggiolo, M.V., Marcheco-Teruel, B., Maria, C.T., Martínez, C., París, I., Pateman, M., Simms, T.M., Sivoli, C.G., Vilar, M., Kennett, D.J., Keegan, W.F., Coppa, A., Lipson, M., Pinhasi, R., Reich, D., 2021. A genetic history of the pre-contact Caribbean. *Nature* 590, 103–110. <https://doi.org/10.1038/s41586-020-03053-2>
- Ferraz, T., Suarez Villagran, X., Nägele, K., Radzevičiūtė, R., Barbosa Lemes, R., Salazar-García, D.C., Wesolowski, V., Lopes Alves, M., Bastos, M., Rapp Py-Daniel, A., Pinto Lima, H., Mendes Cardoso, J., Estevam, R., Liryo, A., Guimarães, G.M., Figuti, L., Eggers, S., Plens, C.R., Azevedo Erler, D.M., Valadares Costa, H.A., Da Silva Erler, I., Koole, E., Henriques, G., Solari, A., Martin, G., Serafim Monteiro Da Silva, S.F., Kipnis, R., Müller, L.M., Ferreira, M., Carvalho Resende, J., Chim, E., Da Silva, C.A., Borella, A.C., Tomé, T., Müller Plumm Gomes, L., Barros Fonseca, D., Santos Da Rosa, C., De Moura Saldanha, J.D., Costa Leite, L., Cunha, C.M.S., Viana, S.A., Ozorio Almeida, F., Klokler, D., Fernandes, H.L.A., Talamo, S., DeBlasis, P., Mendonça De Souza, S., De Paula Moraes, C., Elias Oliveira, R., Hünemeier, T., Strauss, A., Posth, C., 2023. Genomic history of coastal societies from eastern South America. *Nat Ecol Evol* 7, 1315–1330. <https://doi.org/10.1038/s41559-023-02114-9>
- Folan, W.J., Gunn, J., Eaton, J.D., Patch, R.W., 1983. Paleoclimatological Patterning in Southern Mesoamerica. *Journal of Field Archaeology* 10, 453–468. <https://doi.org/10.1179/009346983791504219>
- Frankham, R., 1995. Effective population size/adult population size ratios in wildlife: a review. *Genet. Res.* 66, 95–107. <https://doi.org/10.1017/S0016672300034455>
- Furtwängler, A., Reiter, E., Neumann, G.U., Siebke, I., Steuri, N., Hafner, A., Lösch, S., Anthes, N., Schuenemann, V.J., Krause, J., 2018. Ratio of mitochondrial to nuclear DNA affects contamination estimates in ancient DNA analysis. *Sci Rep* 8, 14075. <https://doi.org/10.1038/s41598-018-32083-0>
- Gamba, C., Jones, E.R., Teasdale, M.D., McLaughlin, R.L., Gonzalez-Fortes, G., Mattiangeli, V., Domboróczki, L., Kővári, I., Pap, I., Anders, A., Whittle, A., Dani, J., Raczký, P., Higham, T.F.G., Hofreiter, M., Bradley, D.G., Pinhasi, R., 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nat Commun* 5, 5257. <https://doi.org/10.1038/ncomms6257>
- Garlinghouse, T., 2022. The Aztec civilization: Mexico's last great Indigenous empire. *Live Science*, Live Science.
- Garrido Marques, A., Rubinacci, S., Malaspinas, A.-S., Delaneau, O., Sousa Da Mota, B., 2024. Assessing the impact of post-mortem damage and contamination on imputation performance in ancient DNA. *Sci Rep* 14, 6227. <https://doi.org/10.1038/s41598-024-56584-3>
- Gnecchi-Ruscone, G.A., Sarno, S., De Fanti, S., Gianvincenzo, L., Giuliani, C., Boattini, A., Bortolini, E., Di Corcia, T., Sanchez Mellado, C., Dávila Francia,

- T.J., Gentilini, D., Di Blasio, A.M., Di Cosimo, P., Cilli, E., Gonzalez-Martin, A., Franceschi, C., Franceschi, Z.A., Rickards, O., Sazzini, M., Luiselli, D., Pettener, D., 2019. Dissecting the Pre-Columbian Genomic Ancestry of Native Americans along the Andes–Amazonia Divide. *Molecular Biology and Evolution* 36, 1254–1269. <https://doi.org/10.1093/molbev/msz066>
- González-Martín, A., Gorostiza, A., Regalado-Liu, L., Arroyo-Peña, S., Tirado, S., Nuño-Arana, I., Rubi-Castellanos, R., Sandoval, K., Coble, M.D., Rangel-Villalobos, H., 2015. Demographic History of Indigenous Populations in Mesoamerica Based on mtDNA Sequence Data. *PLoS ONE* 10, e0131791. <https://doi.org/10.1371/journal.pone.0131791>
- González-Oliver, A., Márquez-Morfín, L., Jiménez, J.C., Torre-Blanco, A., 2001. Founding Amerindian mitochondrial DNA lineages in ancient Maya from Xcaret, Quintana Roo: mtDNA Lineages in Ancient Maya. *Am. J. Phys. Anthropol.* 116, 230–235. <https://doi.org/10.1002/ajpa.1118>
- González-Oliver, Pineda-Vázquez, Garfias-Morales, De La Cruz-Laina, Medrano-González, Márquez-Morfín, Ortega-Muñoz, 2018. Genetic Overview of the Maya Populations: Mitochondrial DNA Haplogroups. *Human Biology* 90, 281. <https://doi.org/10.13110/humanbiology.90.4.03>
- Gorostiza, A., Acunha-Alonzo, V., Regalado-Liu, L., Tirado, S., Granados, J., Sámano, D., Rangel-Villalobos, H., González-Martín, A., 2012. Reconstructing the History of Mesoamerican Populations through the Study of the Mitochondrial DNA Control Region. *PLoS ONE* 7, e44666. <https://doi.org/10.1371/journal.pone.0044666>
- Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., Hansen, N.F., Durand, E.Y., Malaspinas, A.-S., Jensen, J.D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H.A., Good, J.M., Schultz, R., Aximu-Petri, A., Butthof, A., Höber, B., Höffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E.S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Ž., Gušić, I., Doronichev, V.B., Golovanova, L.V., Lalueza-Fox, C., De La Rasilla, M., Fortea, J., Rosas, A., Schmitz, R.W., Johnson, P.L.F., Eichler, E.E., Falush, D., Birney, E., Mullikin, J.C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., Pääbo, S., 2010. A Draft Sequence of the Neandertal Genome. *Science* 328, 710–722. <https://doi.org/10.1126/science.1188021>
- Gretzinger, J., Sayer, D., Justeau, P., Altena, E., Pala, M., Dulias, K., Edwards, C.J., Jodoin, S., Lacher, L., Sabin, S., Vâgene, A.J., Haak, W., Ebenesersdóttir, S.S., Moore, K.H.S., Radzeviciute, R., Schmidt, K., Brace, S., Bager, M.A., Patterson, N., Papac, L., Broomandkhoshbacht, N., Callan, K., Harney, É., Iliev, L., Lawson, A.M., Michel, M., Stewardson, K., Zalzal, F., Rohland, N., Kappelhoff-Beckmann, S., Both, F., Winger, D., Neumann, D., Saalow, L., Krabath, S., Beckett, S., Van Twest, M., Faulkner, N., Read, C., Barton, T., Caruth, J., Hines, J., Krause-Kyora, B., Warnke, U., Schuenemann, V.J., Barnes, I., Dahlström, H., Clausen, J.J., Richardson, A., Popescu, E., Dodwell, N., Ladd, S., Phillips, T., Mortimer, R., Sayer, F., Swales, D., Stewart, A., Powlesland, D., Kenyon, R., Ladle, L., Peek, C., Grefen-Peters, S., Ponce, P., Daniels, R., Spall, C., Woolcock, J., Jones, A.M., Roberts, A.V., Symmons, R., Rawden, A.C., Cooper, A., Bos, K.I., Booth, T., Schroeder, H., Thomas, M.G., Helgason, A., Richards, M.B., Reich, D., Krause, J., Schiffels, S., 2022. The Anglo-Saxon migration and the formation of the early English gene pool. *Nature* 610, 112–119. <https://doi.org/10.1038/s41586-022-05247-2>
- Grugni, V., Raveane, A., Ongaro, L., Battaglia, V., Trombetta, B., Colombo, G., Capodiferro, M.R., Olivieri, A., Achilli, A., Perego, U.A., Motta, J., Tribaldos, M., Woodward, S.R., Ferretti, L., Cruciani, F., Torroni, A., Semino, O., 2019.

- Analysis of the human Y-chromosome haplogroup Q characterizes ancient population movements in Eurasia and the Americas. *BMC Biol* 17, 3. <https://doi.org/10.1186/s12915-018-0622-4>
- Guan, Y., 2014. Detecting Structure of Haplotypes and Local Ancestry. *Genetics* 196, 625–642. <https://doi.org/10.1534/genetics.113.160697>
- Günther, T., Jakobsson, M., 2016. Genes mirror migrations and cultures in prehistoric Europe — a population genomic perspective. *Current Opinion in Genetics & Development* 41, 115–123. <https://doi.org/10.1016/j.gde.2016.09.004>
- Harney, É., Patterson, N., Reich, D., Wakeley, J., 2021. Assessing the performance of qpAdm: a statistical tool for studying population admixture. *Genetics* 217, iyaa045. <https://doi.org/10.1093/genetics/iyaa045>
- Haug, G.H., Günther, D., Peterson, L.C., Sigman, D.M., Hughen, K.A., Aeschlimann, B., 2003. Climate and the Collapse of Maya Civilization. *Science* 299, 1731–1735. <https://doi.org/10.1126/science.1080444>
- Hedenstierna-Jonson, C., Kjellström, A., Zachrisson, T., Krzewińska, M., Sobrado, V., Price, N., Günther, T., Jakobsson, M., Götherström, A., Storå, J., 2017. A female Viking warrior confirmed by genomics. *American J Phys Anthropol* 164, 853–860. <https://doi.org/10.1002/ajpa.23308>
- Heintzman, P.D., Froese, D., Ives, J.W., Soares, A.E.R., Zazula, G.D., Letts, B., Andrews, T.D., Driver, J.C., Hall, E., Hare, P.G., Jass, C.N., MacKay, G., Southon, J.R., Stiller, M., Woywitka, R., Suchard, M.A., Shapiro, B., 2016. Bison phylogeography constrains dispersal and viability of the Ice Free Corridor in western Canada. *Proc. Natl. Acad. Sci. U.S.A.* 113, 8057–8063. <https://doi.org/10.1073/pnas.1601077113>
- Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A., Wilson, A.C., 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature* 312, 282–284. <https://doi.org/10.1038/312282a0>
- Hodell, D.A., Curtis, J.H., Brenner, M., 1995. Possible role of climate in the collapse of Classic Maya civilization. *Nature* 375, 391–394. <https://doi.org/10.1038/375391a0>
- Hodge, M.G., Neff, H., Blackman, M.J., Minc, L.D., 1993. Black-on-Orange Ceramic Production in the Aztec Empire's Heartland. *Latin Am. antiq.* 4, 130–157. <https://doi.org/10.2307/971799>
- Hui, R., D'Atanasio, E., Cassidy, L.M., Scheib, C.L., Kivisild, T., 2020. Evaluating genotype imputation pipeline for ultra-low coverage ancient genomes. *Sci Rep* 10, 18542. <https://doi.org/10.1038/s41598-020-75387-w>
- International Human Genome Sequencing Consortium, Whitehead Institute for Biomedical Research, Center for Genome Research:, Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, Christina, Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., The Sanger Centre:, Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Washington University Genome Sequencing Center, Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissole, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx,

P.J., Clifton, S.W., US DOE Joint Genome Institute:, Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.-F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Baylor College of Medicine Human Genome Sequencing Center:, Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., RIKEN Genomic Sciences Center:, Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Genoscope and CNRS UMR-8030:, Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Department of Genome Analysis, Institute of Molecular Biotechnology:, Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., GTC Sequencing Center:, Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Beijing Genomics Institute/Human Genome Center:, Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Multimegabase Sequencing Center, The Institute for Systems Biology:, Hood, L., Rowen, L., Madan, A., Qin, S., Stanford Genome Technology Center:, Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., University of Oklahoma's Advanced Center for Genome Technology:, Roe, B.A., Chen, F., Pan, H., Max Planck Institute for Molecular Genetics:, Ramser, J., Lehrach, H., Reinhardt, R., Cold Spring Harbor Laboratory, Lita Annenberg Hazen Genome Center:, McCombie, W.R., De La Bastide, M., Dedhia, N., GBF—German Research Centre for Biotechnology:, Blöcker, H., Hornischer, K., Nordsiek, G., *Genome Analysis Group (listed in alphabetical order, also includes individuals listed under other headings):, Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.-C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G.R., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kasprzyk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F.A., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.-P., Yeh, R.-F., Scientific management: National Human Genome Research Institute, US National Institutes of Health:, Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Stanford Human Genome Center:, Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., University of Washington Genome Center:, Olson, M.V., Kaul, R., Raymond, Christopher, Department of Molecular Biology, Keio University School of Medicine:, Shimizu, N., Kawasaki, K., Minoshima, S., University of Texas Southwestern Medical Center at Dallas:, Evans, G.A., Athanasiou, M., Schultz, R., Office of Science, US Department of Energy:, Patrinos, A., The Wellcome Trust:, Morgan, M.J., 2001. Initial sequencing and analysis of the human genome. *Nature* 409, 860–921. <https://doi.org/10.1038/35057062>

- Jenkins, T.L., 2024. *MAPMIXTURE*: An R package and web app for spatial visualisation of admixture and population structure. *Molecular Ecology Resources* 24, e13943. <https://doi.org/10.1111/1755-0998.13943>
- Jones, E.R., Gonzalez-Fortes, G., Connell, S., Siska, V., Eriksson, A., Martiniano, R., McLaughlin, R.L., Gallego Llorente, M., Cassidy, L.M., Gamba, C., Meshveliani, T., Bar-Yosef, O., Müller, W., Belfer-Cohen, A., Matskevich, Z., Jakeli, N., Higham, T.F.G., Currat, M., Lordkipanidze, D., Hofreiter, M., Manica, A., Pinhasi, R., Bradley, D.G., 2015. Upper Palaeolithic genomes reveal deep roots of

- modern Eurasians. *Nat Commun* 6, 8912. <https://doi.org/10.1038/ncomms9912>
- Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., Orlando, L., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29, 1682–1684. <https://doi.org/10.1093/bioinformatics/btt193>
- Kaplan, L., n.d. Alaska Native Language Center. Inuit or Eskimo: Which name to use? URL https://www.uaf.edu/anlc/research-and-resources/resources/archives/inuit_or_eskimo.php#:~:text=Alaska%20Natives%20increasingly%20prefer%20to,other%20organizations%20use%20%22Eskimo%22
- Kazarina, A., Gerhards, G., Petersone-Gordina, E., Kimsis, J., Pole, I., Zole, E., Leonova, V., Ranka, R., 2019. Analysis of the bacterial communities in ancient human bones and burial soil samples: Tracing the impact of environmental bacteria. *Journal of Archaeological Science* 109, 104989. <https://doi.org/10.1016/j.jas.2019.104989>
- Kennett, D.J., Lipson, M., Prufer, K.M., Mora-Marín, D., George, R.J., Rohland, N., Robinson, M., Trask, W.R., Edgar, H.H.J., Hill, E.C., Ray, E.E., Lynch, P., Moes, E., O'Donnell, L., Harper, T.K., Kate, E.J., Ramos, J., Morris, J., Gutierrez, S.M., Ryan, T.M., Culleton, B.J., Awe, J.J., Reich, D., 2022. South-to-north migration preceded the advent of intensive farming in the Maya region. *Nat Commun* 13, 1530. <https://doi.org/10.1038/s41467-022-29158-y>
- Kistler, L., Ware, R., Smith, O., Collins, M., Allaby, R.G., 2017. A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Research* 45, 6310–6320. <https://doi.org/10.1093/nar/gkx361>
- Kjær, K.H., Winther Pedersen, M., De Sanctis, B., De Cahsan, B., Korneliussen, T.S., Michelsen, C.S., Sand, K.K., Jelavić, S., Ruter, A.H., Schmidt, A.M.A., Kjeldsen, K.K., Tesakov, A.S., Snowball, I., Gosse, J.C., Alsos, Inger G., Wang, Y., Dockter, C., Rasmussen, M., Jørgensen, M.E., Skadhauge, B., Prohaska, A., Kristensen, J.Å., Bjerager, M., Allentoft, M.E., Coissac, E., PhyloNorway Consortium, Alsos, Inger Greve, Coissac, E., Rouillard, A., Simakova, A., Fernandez-Guerra, A., Bowler, C., Macias-Fauria, M., Vinner, L., Welch, J.J., Hidy, A.J., Sikora, M., Collins, M.J., Durbin, R., Larsen, N.K., Willerslev, E., 2022. A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. *Nature* 612, 283–291. <https://doi.org/10.1038/s41586-022-05453-y>
- Korneliussen, T.S., Albrechtsen, A., Nielsen, R., 2014. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* 15, 356. <https://doi.org/10.1186/s12859-014-0356-4>
- Kousathanas, A., Leuenberger, C., Link, V., Sell, C., Burger, J., Wegmann, D., 2017. Inferring Heterozygosity from Ancient and Low Coverage Genomes. *Genetics* 205, 317–332. <https://doi.org/10.1534/genetics.116.189985>
- Kravanja, P., Golob, A., Concato, M., Leskovar, T., Zupanič Pajnič, I., 2025. Effects of different environmental factors on preservation of DNA in petrous bones: A comparative study of two Slovenian archaeological sites. *Forensic Science International* 371, 112495. <https://doi.org/10.1016/j.forsciint.2025.112495>
- Krejci, E., Culbert, T.P., 1995. The emergence of Lowland Maya civilization: the transition from the preclassic to the early classic; a conference at Hildesheim, Germany, November 1992, *Acta mesoamericana*. Saurwein, Möckmühl.
- Kumar, S., Bellis, C., Zlojutro, M., Melton, P.E., Blangero, J., Curran, J.E., 2011. Large scale mitochondrial sequencing in Mexican Americans suggests a reappraisal of Native American origins. *BMC Evol Biol* 11, 293. <https://doi.org/10.1186/1471-2148-11-293>
- Lazaridis, I., Patterson, N., Mitnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., Berger, B., Economou, C.,

- Bollongino, R., Fu, Q., Bos, K.I., Nordenfelt, S., Li, H., De Filippo, C., Prüfer, K., Sawyer, S., Posth, C., Haak, W., Hallgren, F., Fornander, E., Rohland, N., Delsate, D., Francken, M., Guinet, J.-M., Wahl, J., Ayodo, G., Babiker, H.A., Bailliet, G., Balanovska, E., Balanovsky, O., Barrantes, R., Bedoya, G., Ben-Ami, H., Bene, J., Berrada, F., Bravi, C.M., Brisighelli, F., Busby, G.B.J., Cali, F., Churnosov, M., Cole, D.E.C., Corach, D., Damba, L., Van Driem, G., Dryomov, S., Dugoujon, J.-M., Fedorova, S.A., Gallego Romero, I., Gubina, M., Hammer, M., Henn, B.M., Hervig, T., Hodoglugil, U., Jha, A.R., Karachanak-Yankova, S., Khusainova, R., Khusnutdinova, E., Kittles, R., Kivisild, T., Klitz, W., Kučinskas, V., Kushniarevich, A., Laredj, L., Litvinov, S., Loukidis, T., Mahley, R.W., Melegh, B., Metspalu, E., Molina, J., Mountain, J., Näkkäläjärvi, K., Nesheva, D., Nyambo, T., Osipova, L., Parik, J., Platonov, F., Posukh, O., Romano, V., Rothhammer, F., Rudan, I., Ruizbakiev, R., Sahakyan, H., Sajantila, A., Salas, A., Starikovskaya, E.B., Tarekegn, A., Toncheva, D., Turdikulova, S., Uktveryte, I., Utevska, O., Vasquez, R., Villena, M., Voevoda, M., Winkler, C.A., Yepiskoposyan, L., Zalloua, P., Zemunik, T., Cooper, A., Capelli, C., Thomas, M.G., Ruiz-Linares, A., Tishkoff, S.A., Singh, L., Thangaraj, K., Villems, R., Comas, D., Sukernik, R., Metspalu, M., Meyer, M., Eichler, E.E., Burger, J., Slatkin, M., Pääbo, S., Kelso, J., Reich, D., Krause, J., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513, 409–413. <https://doi.org/10.1038/nature13673>
- Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
- Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Lindo, J., Haas, R., Hofman, C., Apata, M., Moraga, M., Verdugo, R.A., Watson, J.T., Viviano Llave, C., Witonsky, D., Beall, C., Warinner, C., Novembre, J., Aldenderfer, M., Di Rienzo, A., 2018. The genetic prehistory of the Andean highlands 7000 years BP through European contact. *Sci. Adv.* 4, eaau4921. <https://doi.org/10.1126/sciadv.aau4921>
- Liu, Y., Mao, X., Krause, J., Fu, Q., 2021. Insights into human history from the first decade of ancient human genomics. *Science* 373, 1479–1484. <https://doi.org/10.1126/science.abi8202>
- Livi-Bacci, M., 2006. The Depopulation of Hispanic America after the Conquest. *Population & Development Review* 32, 199–232. <https://doi.org/10.1111/j.1728-4457.2006.00116.x>
- Llamas, B., Fehren-Schmitz, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., Nordenfelt, S., Valdiosera, C., Richards, S.M., Rohrlach, A., Romero, M.I.B., Espinoza, I.F., Cagigao, E.T., Jiménez, L.W., Makowski, K., Reyna, I.S.L., Lory, J.M., Torrez, J.A.B., Rivera, M.A., Burger, R.L., Ceruti, M.C., Reinhard, J., Wells, R.S., Politis, G., Santoro, C.M., Standen, V.G., Smith, C., Reich, D., Ho, S.Y.W., Cooper, A., Haak, W., 2016. Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci. Adv.* 2, e1501385. <https://doi.org/10.1126/sciadv.1501385>
- Llamas, B., Valverde, G., Fehren-Schmitz, L., Weyrich, L.S., Cooper, A., Haak, W., 2017. From the field to the laboratory: Controlling DNA contamination in human ancient DNA research in the high-throughput sequencing era. *STAR: Science & Technology of Archaeological Research* 3, 1–14. <https://doi.org/10.1080/20548923.2016.1258824>
- Llorente, M.G., Jones, E.R., Eriksson, A., Siska, V., Arthur, K.W., Arthur, J.W., Curtis,

- M.C., Stock, J.T., Coltorti, M., Pieruccini, P., Stretton, S., Brock, F., Higham, T., Park, Y., Hofreiter, M., Bradley, D.G., Bhak, J., Pinhasi, R., Manica, A., 2015. Ancient Ethiopian genome reveals extensive Eurasian admixture in Eastern Africa. *Science* 350, 820–822. <https://doi.org/10.1126/science.aad2879>
- Maddox, B., 2003. The double helix and the “wronged heroine.” *Nature* 421, 407–408. <https://doi.org/10.1038/nature01399>
- Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., Skoglund, P., Lazaridis, I., Sankararaman, S., Fu, Q., Rohland, N., Renaud, G., Erlich, Y., Willems, T., Gallo, C., Spence, J.P., Song, Y.S., Poletti, G., Balloux, F., Van Driem, G., De Knijff, P., Romero, I.G., Jha, A.R., Behar, D.M., Bravi, C.M., Capelli, C., Hervig, T., Moreno-Estrada, A., Posukh, O.L., Balanovska, E., Balanovsky, O., Karachanak-Yankova, S., Sahakyan, H., Toncheva, D., Yepiskoposyan, L., Tyler-Smith, C., Xue, Y., Abdullah, M.S., Ruiz-Linares, A., Beall, C.M., Di Rienzo, A., Jeong, C., Starikovskaya, E.B., Metspalu, E., Parik, J., Villems, R., Henn, B.M., Hodoglugil, U., Mahley, R., Sajantila, A., Stamatoyannopoulos, G., Wee, J.T.S., Khusainova, R., Khusnutdinova, E., Litvinov, S., Ayodo, G., Comas, D., Hammer, M.F., Kivisild, T., Klitz, W., Winkler, C.A., Labuda, D., Bamshad, M., Jorde, L.B., Tishkoff, S.A., Watkins, W.S., Metspalu, M., Dryomov, S., Sukernik, R., Singh, L., Thangaraj, K., Pääbo, S., Kelso, J., Patterson, N., Reich, D., 2016. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* 538, 201–206. <https://doi.org/10.1038/nature18964>
- Mann, A.E., Sabin, S., Ziesemer, K., Vågene, Å.J., Schroeder, H., Ozga, A.T., Sankaranarayanan, K., Hofman, C.A., Fellows Yates, J.A., Salazar-García, D.C., Frohlich, B., Aldenderfer, M., Hoogland, M., Read, C., Milner, G.R., Stone, A.C., Lewis, C.M., Krause, J., Hofman, C., Bos, K.I., Warinner, C., 2018. Differential preservation of endogenous human and microbial DNA in dental calculus and dentin. *Sci Rep* 8, 9822. <https://doi.org/10.1038/s41598-018-28091-9>
- Maples, B.K., Gravel, S., Kenny, E.E., Bustamante, C.D., 2013. RFMix: A Discriminative Modeling Approach for Rapid and Robust Local-Ancestry Inference. *The American Journal of Human Genetics* 93, 278–288. <https://doi.org/10.1016/j.ajhg.2013.06.020>
- Margaryan, A., Lawson, D.J., Sikora, M., Martin, Racimo, F., Rasmussen, S., Moltke, I., Cassidy, L.M., Jørsboe, E., Ingason, A., Pedersen, M.W., Korneliussen, T., Wilhelmson, H., Buś, M.M., De Barros Damgaard, P., Martiniano, R., Renaud, G., Bhérier, C., Moreno-Mayar, J.V., Fotakis, A.K., Allen, M., Allmæ, R., Molak, M., Cappellini, E., Scorrano, G., McColl, H., Buzhilova, A., Fox, A., Albrechtsen, A., Schütz, B., Skar, B., Arcini, C., Falys, C., Jonson, C.H., Błaszczyk, D., Pezhemsky, D., Turner-Walker, G., Gestsdóttir, H., Lundstrøm, I., Gustin, I., Mainland, I., Potekhina, I., Muntoni, I.M., Cheng, J., Stenderup, J., Ma, J., Gibson, J., Peets, J., Gustafsson, J., Iversen, K.H., Simpson, L., Strand, L., Loe, L., Sikora, Maeve, Florek, M., Vretemark, M., Redknap, M., Bajka, M., Pushkina, T., Søvstø, M., Grigoreva, N., Christensen, T., Kastholm, O., Uldum, O., Favia, P., Holck, P., Sten, S., Arge, S.V., Ellingvåg, S., Moiseyev, V., Bogdanowicz, W., Magnusson, Y., Orlando, L., Pentz, P., Jessen, M.D., Pedersen, A., Collard, M., Bradley, D.G., Jørkov, M.L., Arneborg, J., Lynnerup, N., Price, N., Gilbert, M.T.P., Allentoft, M.E., Bill, J., Sindbæk, S.M., Hedeager, L., Kristiansen, K., Nielsen, R., Werge, T., Willerslev, E., 2020. Population genomics of the Viking world. *Nature* 585, 390–396. <https://doi.org/10.1038/s41586-020-2688-8>
- Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput

- sequencing reads. *EMBnet j.* 17, 10. <https://doi.org/10.14806/ej.17.1.200>
- Martínez de Velasco Cortina, A., Vega Villalobos, M.E. (Eds.), 2015. *The Maya: voices in stone*, Second, revised and enhanced edition. ed. Turner, Madrid.
- Martiniano, R., Cassidy, L.M., Ó'Maoldúin, R., McLaughlin, R., Silva, N.M., Manco, L., Fidalgo, D., Pereira, T., Coelho, M.J., Serra, M., Burger, J., Parreira, R., Moran, E., Valera, A.C., Porfírio, E., Boaventura, R., Silva, A.M., Bradley, D.G., 2017. The population genomics of archaeological transition in west Iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genet* 13, e1006852. <https://doi.org/10.1371/journal.pgen.1006852>
- Masson, M.A., 2012. Maya collapse cycles. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18237–18238. <https://doi.org/10.1073/pnas.1213638109>
- Mata-Míguez, J., Overholtzer, L., Rodríguez-Alegría, E., Kemp, B.M., Bolnick, D.A., 2012. The genetic impact of aztec imperialism: Ancient mitochondrial DNA evidence from Xaltocan, Mexico. *American J Phys Anthropol* 149, 504–516. <https://doi.org/10.1002/ajpa.22152>
- McColl, H., Racimo, F., Vinner, L., Demeter, F., Gakuhari, T., Moreno-Mayar, J.V., Van Driem, G., Gram Wilken, U., Seguin-Orlando, A., De La Fuente Castro, C., Wasef, S., Shoocongdej, R., Souksavatdy, V., Sayavongkhamdy, T., Saidin, M.M., Allentoft, M.E., Sato, T., Malaspinas, A.-S., Aghakhanian, F.A., Korneliussen, T., Prohaska, A., Margaryan, A., De Barros Damgaard, P., Kaewsutthi, S., Lertrit, P., Nguyen, T.M.H., Hung, H., Minh Tran, T., Nghia Truong, H., Nguyen, G.H., Shahidan, S., Wiradnyana, K., Matsumae, H., Shigehara, N., Yoneda, M., Ishida, H., Masuyama, T., Yamada, Y., Tajima, A., Shibata, H., Toyoda, A., Hanihara, T., Nakagome, S., Deviese, T., Bacon, A.-M., Düringer, P., Ponche, J.-L., Shackelford, L., Patole-Edoumba, E., Nguyen, A.T., Bellina-Pryce, B., Galipaud, J.-C., Kinaston, R., Buckley, H., Pottier, C., Rasmussen, S., Higham, T., Foley, R.A., Lahr, M.M., Orlando, L., Sikora, M., Phipps, M.E., Oota, H., Higham, C., Lambert, D.M., Willerslev, E., 2018. The prehistoric peopling of Southeast Asia. *Science* 361, 88–92. <https://doi.org/10.1126/science.aat3628>
- McCombie, W.R., McPherson, J.D., Mardis, E.R., 2019. Next-Generation Sequencing Technologies. *Cold Spring Harb Perspect Med* 9, a036798. <https://doi.org/10.1101/cshperspect.a036798>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Merrill, W.L., Hard, R.J., Mabry, J.B., Fritz, G.J., Adams, K.R., Roney, J.R., MacWilliams, A.C., 2009. The diffusion of maize to the southwestern United States and its impact. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21019–21026. <https://doi.org/10.1073/pnas.0906075106>
- Merriwether, D.A., Reed, D.M., Ferrell, R.E., 1997. Ancient and contemporary mitochondrial DNA variation in the Maya. *Bones of the Maya: studies of ancient skeletons* 208–217.
- Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., De Filippo, C., Sudmant, P.H., Alkan, C., Fu, Q., Do, R., Rohland, N., Tandon, A., Siebauer, M., Green, R.E., Bryc, K., Briggs, A.W., Stenzel, U., Dabney, J., Shendure, J., Kitzman, J., Hammer, M.F., Shunkov, M.V., Derevianko, A.P., Patterson, N., Andrés, A.M., Eichler, E.E., Slatkin, M., Reich, D., Kelso, J., Pääbo, S., 2012. A High-Coverage Genome Sequence from an Archaic Denisovan Individual. *Science* 338, 222–226.

- <https://doi.org/10.1126/science.1224344>
- Miller, W., Drautz, D.I., Ratan, A., Pusey, B., Qi, J., Lesk, A.M., Tomsho, L.P., Packard, M.D., Zhao, F., Sher, A., Tikhonov, A., Raney, B., Patterson, N., Lindblad-Toh, K., Lander, E.S., Knight, J.R., Irzyk, G.P., Fredrikson, K.M., Harkins, T.T., Sheridan, S., Pringle, T., Schuster, S.C., 2008. Sequencing the nuclear genome of the extinct woolly mammoth. *Nature* 456, 387–390.
<https://doi.org/10.1038/nature07446>
- Minority Rights Group, 2018. Maya Ch'ortí in Honduras.
- Misra, V.N., 2001. Prehistoric human colonization of India. *J Biosci* 26, 491–531.
<https://doi.org/10.1007/BF02704749>
- Mizoguchi, Y., Nakamura, S., 2006. Maya skeletal remains from the Copan and El Puente sites in Honduras. *AS* 114, 75–88. <https://doi.org/10.1537/ase.040331>
- Monroy Kuhn, J.M., Jakobsson, M., Günther, T., 2018. Estimating genetic kin relationships in prehistoric populations. *PLoS ONE* 13, e0195491.
<https://doi.org/10.1371/journal.pone.0195491>
- Moreno-Mayar, J.V., Potter, B.A., Vinner, L., Steinrücken, M., Rasmussen, S., Terhorst, J., Kamm, J.A., Albrechtsen, A., Malaspinas, A.-S., Sikora, M., Reuther, J.D., Irish, J.D., Malhi, R.S., Orlando, L., Song, Y.S., Nielsen, R., Meltzer, D.J., Willerslev, E., 2018a. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* 553, 203–207.
<https://doi.org/10.1038/nature25173>
- Moreno-Mayar, J.V., Sousa Da Mota, B., Higham, T., Klemm, S., Gorman Edmunds, M., Stenderup, J., Iraeta-Orbegozo, M., Laborde, V., Heyer, E., Torres Hochstetter, F., Friess, M., Allentoft, M.E., Schroeder, H., Delaneau, O., Malaspinas, A.-S., 2024. Ancient Rapanui genomes reveal resilience and pre-European contact with the Americas. *Nature* 633, 389–397.
<https://doi.org/10.1038/s41586-024-07881-4>
- Moreno-Mayar, J.V., Vinner, L., de Barros Damgaard, P., de la Fuente, C., Chan, J., Spence, J.P., Allentoft, M.E., Vimala, T., Racimo, F., Pinotti, T., Rasmussen, S., Margaryan, A., Iraeta Orbegozo, M., Mylopotamitaki, D., Wooller, M., Bataille, C., Becerra-Valdivia, L., Chivall, D., Comeskey, D., Devière, T., Grayson, D.K., George, L., Harry, H., Alexandersen, V., Primeau, C., Erlandson, J., Rodrigues-Carvalho, C., Reis, S., Bastos, M.Q.R., Cybulski, J., Vullo, C., Morello, F., Vilar, M., Wells, S., Gregersen, K., Hansen, K.L., Lynnerup, N., Mirazón Lahr, M., Kjær, K., Strauss, A., Alfonso-Durruty, M., Salas, A., Schroeder, H., Higham, T., Malhi, R.S., Rasic, J.T., Souza, L., Santos, F.R., Malaspinas, A.-S., Sikora, M., Nielsen, R., Song, Y.S., Meltzer, D.J., Willerslev, E., 2018b. Early human dispersals within the Americas. *Science* 362, eaav2621. <https://doi.org/10.1126/science.aav2621>
- Mullis, K.B., Faloona, F.A., 1987. [21] Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction, in: *Methods in Enzymology*. Elsevier, pp. 335–350. [https://doi.org/10.1016/0076-6879\(87\)55023-6](https://doi.org/10.1016/0076-6879(87)55023-6)
- Murray, M., Nakamura, S., Fuentes, M., Canizales, R., Talavera, O., Varela, S., Madrid, Z., Carbajal, C., Ogawa, M., Yoneda, M., Cassidy, L.M., Gakuhari, T., Nakagome, S., 2025. Ancient genomes reveal demographic trajectories during the Classic Maya period. *Current Biology* S0960982225005779.
<https://doi.org/10.1016/j.cub.2025.05.002>
- Nägele, K., Posth, C., Iraeta Orbegozo, M., Chinique de Armas, Y., Hernández Godoy, S.T., González Herrera, U.M., Nieves-Colón, M.A., Sandoval-Velasco, M., Mylopotamitaki, D., Radzeviciute, R., Laffoon, J., Pestle, W.J., Ramos-Madrigal, J., Lamnidis, T.C., Schaffer, W.C., Carr, R.S., Day, J.S., Arredondo Antúnez, C., Rangel Rivero, A., Martínez-Fuentes, A.J., Crespo-Torres, E., Roksandic, I., Stone, A.C., Lalueza-Fox, C., Hoogland, M., Roksandic, M., Hofman, C.L.,

- Krause, J., Schroeder, H., 2020. Genomic insights into the early peopling of the Caribbean. *Science* 369, 456–460. <https://doi.org/10.1126/science.aba8697>
- Nägele, K., Rivollat, M., Yu, H., Wang, K., 2022. Ancient genomic research - From broad strokes to nuanced reconstructions of the past. *Journal of Anthropological Sciences* 193–230. <https://doi.org/10.4436/JASS.10017>
- Nakatsuka, N., Holguin, B., Sedig, J., Langenwaller, P.E., Carpenter, J., Culleton, B.J., García-Moreno, C., Harper, T.K., Martin, D., Martínez-Ramírez, J., Porcayo-Michelini, A., Tiesler, V., Villapando-Canchola, M.E., Valdes Herrera, A., Callan, K., Curtis, E., Kearns, A., Iliev, L., Lawson, A.M., Mah, M., Mallick, S., Micco, A., Michel, M., Workman, J.N., Oppenheimer, J., Qiu, L., Zalzal, F., Rohland, N., Punzo Diaz, J.L., Johnson, J.R., Reich, D., 2023. Genetic continuity and change among the Indigenous peoples of California. *Nature* 624, 122–129. <https://doi.org/10.1038/s41586-023-06771-5>
- Nakatsuka, N., Lazaridis, I., Barbieri, C., Skoglund, P., Rohland, N., Mallick, S., Posth, C., Harkins-Kinkaid, K., Ferry, M., Harney, É., Michel, M., Stewardson, K., Novak-Forst, J., Capriles, J.M., Durruty, M.A., Álvarez, K.A., Beresford-Jones, D., Burger, R., Cadwallader, L., Fujita, R., Isla, J., Lau, G., Aguirre, C.L., LeBlanc, S., Maldonado, S.C., Meddens, F., Messineo, P.G., Culleton, B.J., Harper, T.K., Quilter, J., Politis, G., Rademaker, K., Reindel, M., Rivera, M., Salazar, L., Sandoval, J.R., Santoro, C.M., Scheifler, N., Standen, V., Barreto, M.I., Espinoza, I.F., Tomasto-Cagigao, E., Valverde, G., Kennett, D.J., Cooper, A., Krause, J., Haak, W., Llamas, B., Reich, D., Fehren-Schmitz, L., 2020a. A Paleogenomic Reconstruction of the Deep Population History of the Andes. *Cell* 181, 1131–1145.e21. <https://doi.org/10.1016/j.cell.2020.04.015>
- Nakatsuka, N., Luisi, P., Motti, J.M.B., Salemme, M., Santiago, F., D'Angelo del Campo, M.D., Vecchi, R.J., Espinosa-Parrilla, Y., Prieto, A., Adamski, N., Lawson, A.M., Harper, T.K., Culleton, B.J., Kennett, D.J., Lalueza-Fox, C., Mallick, S., Rohland, N., Guichón, R.A., Cabana, G.S., Nores, R., Reich, D., 2020b. Ancient genomes in South Patagonia reveal population movements associated with technological shifts and geography. *Nat Commun* 11, 3868. <https://doi.org/10.1038/s41467-020-17656-w>
- Ochoa-Lugo, de Lourdes Muñoz, Pérez-Ramírez, Beaty, López-Armenta, Cervini-Silva, Moreno-Galeana, Meza, Ramos, Crawford, Romano-Pacheco, 2016. Genetic Affiliation of Pre-Hispanic and Contemporary Mayas through Maternal Linage. *Human Biology* 88, 136. <https://doi.org/10.13110/humanbiology.88.2.0136>
- Okonechnikov, K., Conesa, A., García-Alcalde, F., 2016. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* 32, 292–294. <https://doi.org/10.1093/bioinformatics/btv566>
- Olga Hernández Espinoza, P., Márquez Morfín, L., 2015. Maya paleodemographics: What do we know? *American J Hum Biol* 27, 747–757. <https://doi.org/10.1002/ajhb.22751>
- Oliveira, S., Nägele, K., Carlhoff, S., Pugach, I., Koesbardiati, T., Hübner, A., Meyer, M., Oktaviana, A.A., Takenaka, M., Katagiri, C., Murti, D.B., Putri, R.S., Mahirta, Petchey, F., Higham, T., Higham, C.F.W., O'Connor, S., Hawkins, S., Kinaston, R., Bellwood, P., Ono, R., Powell, A., Krause, J., Posth, C., Stoneking, M., 2022. Ancient genomes from the last three millennia support multiple human dispersals into Wallacea. *Nat Ecol Evol* 6, 1024–1034. <https://doi.org/10.1038/s41559-022-01775-2>
- Orlando, L., Gilbert, M.T.P., Willerslev, E., 2015. Reconstructing ancient genomes and epigenomes. *Nat Rev Genet* 16, 395–408. <https://doi.org/10.1038/nrg3935>
- Owsley, D.W., Jantz, R.L., 2001. Archaeological Politics and Public Interest in Paleoamerican Studies: Lessons from Gordon Creek Woman and Kennewick Man. *Am. antiq.* 66, 565–575. <https://doi.org/10.2307/2694173>

- Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., Reich, D., 2012. Ancient Admixture in Human History. *Genetics* 192, 1065–1093. <https://doi.org/10.1534/genetics.112.145037>
- Patterson, N., Price, A.L., Reich, D., 2006. Population Structure and Eigenanalysis. *PLoS Genet* 2, e190. <https://doi.org/10.1371/journal.pgen.0020190>
- Pedersen, M.W., Ruter, A., Schweger, C., Friebe, H., Staff, R.A., Kjeldsen, K.K., Mendoza, M.L.Z., Beaudoin, A.B., Zutter, C., Larsen, N.K., Potter, B.A., Nielsen, R., Rainville, R.A., Orlando, L., Meltzer, D.J., Kjær, K.H., Willerslev, E., 2016. Postglacial viability and colonization in North America's ice-free corridor. *Nature* 537, 45–49. <https://doi.org/10.1038/nature19085>
- Petkova, D., Novembre, J., Stephens, M., 2016. Visualizing spatial population structure with estimated effective migration surfaces. *Nat Genet* 48, 94–100. <https://doi.org/10.1038/ng.3464>
- PhyloTree, 2016.
- Pickrell, J.K., Reich, D., 2014. Toward a new history and geography of human genes informed by ancient DNA. *Trends in Genetics* 30, 377–389. <https://doi.org/10.1016/j.tig.2014.07.007>
- Popović, D., Molak, M., Ziolkowski, M., Vranich, A., Sobczyk, M., Vidaurre, D.U., Agresti, G., Skrzypczak, M., Ginalski, K., Lamnidis, T.C., Nakatsuka, N., Mallick, S., Baca, M., 2021. Ancient genomes reveal long-range influence of the pre-Columbian culture and site of Tiwanaku. *Sci. Adv.* 7, eabg7261. <https://doi.org/10.1126/sciadv.abg7261>
- Posth, C., Nakatsuka, N., Lazaridis, I., Skoglund, P., Mallick, S., Lamnidis, T.C., Rohland, N., Nägele, K., Adamski, N., Bertolini, E., Broomandkhoshbacht, N., Cooper, A., Culleton, B.J., Ferraz, T., Ferry, M., Furtwängler, A., Haak, W., Harkins, K., Harper, T.K., Hünemeier, T., Lawson, A.M., Llamas, B., Michel, M., Nelson, E., Oppenheimer, J., Patterson, N., Schiffels, S., Sedig, J., Stewardson, K., Talamo, S., Wang, C.-C., Hublin, J.-J., Hubbe, M., Harvati, K., Nuevo Delaunay, A., Beier, J., Francken, M., Kaulicke, P., Reyes-Centeno, H., Rademaker, K., Trask, W.R., Robinson, M., Gutierrez, S.M., Prufer, K.M., Salazar-García, D.C., Chim, E.N., Müller Plumm Gomes, L., Alves, M.L., Liryo, A., Inglez, M., Oliveira, R.E., Bernardo, D.V., Barioni, A., Wesolowski, V., Scheifler, N.A., Rivera, M.A., Plens, C.R., Messineo, P.G., Figuti, L., Corach, D., Scabuzzo, C., Eggers, S., DeBlasis, P., Reindel, M., Méndez, C., Politis, G., Tomasto-Cagigao, E., Kennett, D.J., Strauss, A., Fehren-Schmitz, L., Krause, J., Reich, D., 2018. Reconstructing the Deep Population History of Central and South America. *Cell* 175, 1185–1197.e22. <https://doi.org/10.1016/j.cell.2018.10.027>
- Price, A.L., Tandon, A., Patterson, N., Barnes, K.C., Rafaels, N., Ruczinski, I., Beaty, T.H., Mathias, R., Reich, D., Myers, S., 2009. Sensitive Detection of Chromosomal Segments of Distinct Ancestry in Admixed Populations. *PLoS Genet* 5, e1000519. <https://doi.org/10.1371/journal.pgen.1000519>
- Price, T.D., Burton, J.H., Sharer, R.J., Buikstra, J.E., Wright, L.E., Traxler, L.P., Miller, K.A., 2010. Kings and commoners at Copan: Isotopic evidence for origins and movement in the Classic Maya period. *Journal of Anthropological Archaeology* 29, 15–32. <https://doi.org/10.1016/j.jaa.2009.10.001>
- Price, T.D., Nakamura, S., Suzuki, S., Burton, J.H., Tiesler, V., 2014. New isotope data on Maya mobility and enclaves at Classic Copan, Honduras. *Journal of Anthropological Archaeology* 36, 32–47. <https://doi.org/10.1016/j.jaa.2014.02.003>
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., De Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The*

- American Journal of Human Genetics 81, 559–575.
<https://doi.org/10.1086/519795>
- Raghavan, M., DeGiorgio, M., Albrechtsen, A., Moltke, I., Skoglund, P., Korneliussen, T.S., Grønnow, B., Appelt, M., Gulløv, H.C., Friesen, T.M., Fitzhugh, W., Malmström, H., Rasmussen, S., Olsen, J., Melchior, L., Fuller, B.T., Fahrni, S.M., Stafford, T., Grimes, V., Renouf, M.A.P., Cybulski, J., Lynnerup, N., Lahr, M.M., Britton, K., Knecht, R., Arneborg, J., Metspalu, M., Cornejo, O.E., Malaspinas, A.-S., Wang, Y., Rasmussen, M., Raghavan, V., Hansen, T.V.O., Khusnutdinova, E., Pierre, T., Dneprovsky, K., Andreasen, C., Lange, H., Hayes, M.G., Coltrain, J., Spitsyn, V.A., Götherström, A., Orlando, L., Kivisild, T., Villems, R., Crawford, M.H., Nielsen, F.C., Dissing, J., Heinemeier, J., Meldgaard, M., Bustamante, C., O'Rourke, D.H., Jakobsson, M., Gilbert, M.T.P., Nielsen, R., Willerslev, E., 2014a. The genetic prehistory of the New World Arctic. *Science* 345, 1255832. <https://doi.org/10.1126/science.1255832>
- Raghavan, M., Skoglund, P., Graf, K.E., Metspalu, M., Albrechtsen, A., Moltke, I., Rasmussen, S., Stafford Jr, T.W., Orlando, L., Metspalu, E., Karmin, M., Tambets, K., Rootsi, S., Mägi, R., Campos, P.F., Balanovska, E., Balanovsky, O., Khusnutdinova, E., Litvinov, S., Osipova, L.P., Fedorova, S.A., Voevoda, M.I., DeGiorgio, M., Sicheritz-Ponten, T., Brunak, S., Demeshchenko, S., Kivisild, T., Villems, R., Nielsen, R., Jakobsson, M., Willerslev, E., 2014b. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* 505, 87–91. <https://doi.org/10.1038/nature12736>
- Raghavan, M., Steinrücken, M., Harris, K., Schiffels, S., Rasmussen, S., DeGiorgio, M., Albrechtsen, A., Valdiosera, C., Ávila-Arcos, M.C., Malaspinas, A.-S., Eriksson, A., Moltke, I., Metspalu, M., Homburger, J.R., Wall, J., Cornejo, O.E., Moreno-Mayar, J.V., Korneliussen, T.S., Pierre, T., Rasmussen, M., Campos, P.F., Damgaard, P. de B., Allentoft, M.E., Lindo, J., Metspalu, E., Rodríguez-Varela, R., Mansilla, J., Henrickson, C., Seguin-Orlando, A., Malmström, H., Stafford, T., Shringarpure, S.S., Moreno-Estrada, A., Karmin, M., Tambets, K., Bergström, A., Xue, Y., Warmuth, V., Friend, A.D., Singarayer, J., Valdes, P., Balloux, F., Lebreiro, I., Vera, J.L., Rangel-Villalobos, H., Pettener, D., Luiselli, D., Davis, L.G., Heyer, E., Zollikofer, C.P.E., Ponce de León, M.S., Smith, C.I., Grimes, V., Pike, K.-A., Deal, M., Fuller, B.T., Arriaza, B., Standen, V., Luz, M.F., Ricaut, F., Guidon, N., Osipova, L., Voevoda, M.I., Posukh, O.L., Balanovsky, O., Lavryashina, M., Bogunov, Y., Khusnutdinova, E., Gubina, M., Balanovska, E., Fedorova, S., Litvinov, S., Malyarchuk, B., Derenko, M., Moshier, M.J., Archer, D., Cybulski, J., Petzelt, B., Mitchell, J., Worl, R., Norman, P.J., Parham, P., Kemp, B.M., Kivisild, T., Tyler-Smith, C., Sandhu, M.S., Crawford, M., Villems, R., Smith, D.G., Waters, M.R., Goebel, T., Johnson, J.R., Malhi, R.S., Jakobsson, M., Meltzer, D.J., Manica, A., Durbin, R., Bustamante, C.D., Song, Y.S., Nielsen, R., Willerslev, E., 2015. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* 349, aab3884. <https://doi.org/10.1126/science.aab3884>
- Ramos-Madrugal, J., Fritz, G.J., Schroeder, B., Smith, B., Sánchez-Barreiro, F., Carøe, C., Runge, A.K.W., Boer, S., McGrath, K., Vieira, F.G., Liu, S., Da Fonseca, R.R., Guo, C., Zhang, G., Petersen, B., Sicheritz-Pontén, T., Gopalakrishnan, S., Gilbert, M.T.P., Wales, N., 2025. The genomic origin of early maize in eastern North America. *Cell* 188, 33-43.e16.
<https://doi.org/10.1016/j.cell.2024.11.003>
- Rasmussen, M., Anzick, S.L., Waters, M.R., Skoglund, P., DeGiorgio, M., Stafford, T.W., Rasmussen, S., Moltke, I., Albrechtsen, A., Doyle, S.M., Poznik, G.D., Gudmundsdottir, V., Yadav, R., Malaspinas, A.-S., V, S.S.W., Allentoft, M.E., Cornejo, O.E., Tambets, K., Eriksson, A., Heintzman, P.D., Karmin, M.,

- Korneliussen, T.S., Meltzer, D.J., Pierre, T.L., Stenderup, J., Saag, L., Warmuth, V.M., Lopes, M.C., Malhi, R.S., Brunak, S., Sicheritz-Ponten, T., Barnes, I., Collins, M., Orlando, L., Balloux, F., Manica, A., Gupta, R., Metspalu, M., Bustamante, C.D., Jakobsson, M., Nielsen, R., Willerslev, E., 2014. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* 506, 225–229. <https://doi.org/10.1038/nature13025>
- Rasmussen, M., Guo, X., Wang, Y., Lohmueller, K.E., Rasmussen, S., Albrechtsen, A., Skotte, L., Lindgreen, S., Metspalu, M., Jombart, T., Kivisild, T., Zhai, W., Eriksson, A., Manica, A., Orlando, L., De La Vega, F.M., Tridico, S., Metspalu, E., Nielsen, K., Ávila-Arcos, M.C., Moreno-Mayar, J.V., Muller, C., Dortch, J., Gilbert, M.T.P., Lund, O., Wesolowska, A., Karmin, M., Weinert, L.A., Wang, B., Li, J., Tai, S., Xiao, F., Hanihara, T., van Driem, G., Jha, A.R., Ricaut, F.-X., de Knijff, P., Migliano, A.B., Gallego Romero, I., Kristiansen, K., Lambert, D.M., Brunak, S., Forster, P., Brinkmann, B., Nehlich, O., Bunce, M., Richards, M., Gupta, R., Bustamante, C.D., Krogh, A., Foley, R.A., Lahr, M.M., Balloux, F., Sicheritz-Pontén, T., Villems, R., Nielsen, R., Wang, J., Willerslev, E., 2011. An Aboriginal Australian Genome Reveals Separate Human Dispersals into Asia. *Science* 334, 94–98. <https://doi.org/10.1126/science.1211177>
- Rasmussen, M., Li, Y., Lindgreen, S., Pedersen, J.S., Albrechtsen, A., Moltke, I., Metspalu, M., Metspalu, E., Kivisild, T., Gupta, R., Bertalan, M., Nielsen, K., Gilbert, M.T.P., Wang, Y., Raghavan, M., Campos, P.F., Kamp, H.M., Wilson, A.S., Gledhill, A., Tridico, S., Bunce, M., Lorenzen, E.D., Binladen, J., Guo, X., Zhao, J., Zhang, X., Zhang, H., Li, Z., Chen, M., Orlando, L., Kristiansen, K., Bak, M., Tommerup, N., Bendixen, C., Pierre, T.L., Grønnow, B., Meldgaard, M., Andreasen, C., Fedorova, S.A., Osipova, L.P., Higham, T.F.G., Ramsey, C.B., Hansen, T.V.O., Nielsen, F.C., Crawford, M.H., Brunak, S., Sicheritz-Pontén, T., Villems, R., Nielsen, R., Krogh, A., Wang, J., Willerslev, E., 2010. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* 463, 757–762. <https://doi.org/10.1038/nature08835>
- Rasmussen, M., Sikora, M., Albrechtsen, A., Korneliussen, T.S., Moreno-Mayar, J.V., Poznik, G.D., Zollikofer, C.P.E., Ponce de León, M.S., Allentoft, M.E., Moltke, I., Jónsson, H., Valdiosera, C., Malhi, R.S., Orlando, L., Bustamante, C.D., Stafford, T.W., Meltzer, D.J., Nielsen, R., Willerslev, E., 2015. The ancestry and affiliations of Kennewick Man. *Nature* 523, 455–458. <https://doi.org/10.1038/nature14625>
- Regueiro, M., Alvarez, J., Rowold, D., Herrera, R.J., 2013. On the origins, rapid expansion and genetic diversity of native Americans from hunting-gatherers to agriculturalists. *American J Phys Anthropol* 150, 333–348. <https://doi.org/10.1002/ajpa.22207>
- Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L.F., Maricic, T., Good, J.M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E.E., Stoneking, M., Richards, M., Talamo, S., Shunkov, M.V., Derevianko, A.P., Hublin, J.-J., Kelso, J., Slatkin, M., Pääbo, S., 2010. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468, 1053–1060. <https://doi.org/10.1038/nature09710>
- Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Parra, M.V., Rojas, W., Duque, C., Mesa, N., García, L.F., Triana, O., Blair, S., Maestre, A., Dib, J.C., Bravi, C.M., Bailliet, G., Corach, D., Hünemeier, T., Bortolini, M.C., Salzano, F.M., Petzl-Erler, M.L., Acuña-Alonzo, V., Aguilar-Salinas, C., Canizales-Quinteros, S., Tusié-Luna, T., Riba, L., Rodríguez-Cruz, M., Lopez-Alarcón, M., Coral-Vazquez, R., Canto-Cetina, T., Silva-Zolezzi, I., Fernandez-Lopez, J.C., Contreras, A.V., Jimenez-Sanchez, G.,

- Gómez-Vázquez, M.J., Molina, J., Carracedo, Á., Salas, A., Gallo, C., Poletti, G., Witonsky, D.B., Alkorta-Aranburu, G., Sukernik, R.I., Osipova, L., Fedorova, S.A., Vasquez, R., Villena, M., Moreau, C., Barrantes, R., Pauls, D., Excoffier, L., Bedoya, G., Rothhammer, F., Dugoujon, J.-M., Larrouy, G., Klitz, W., Labuda, D., Kidd, J., Kidd, K., Di Rienzo, A., Freimer, N.B., Price, A.L., Ruiz-Linares, A., 2012. Reconstructing Native American population history. *Nature* 488, 370–374. <https://doi.org/10.1038/nature11258>
- Roca-Rada, X., Souilmi, Y., Teixeira, J.C., Llamas, B., 2020. Ancient DNA Studies in Pre-Columbian Mesoamerica. *Genes* 11, 1346. <https://doi.org/10.3390/genes11111346>
- Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., Reich, D., 2015. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. B* 370, 20130624. <https://doi.org/10.1098/rstb.2013.0624>
- Rohrlach, A.B., Tuke, J., Popli, D., Haak, W., 2023. BREADR: An R Package for the Bayesian Estimation of Genetic Relatedness from Low-coverage Genotype Data. <https://doi.org/10.1101/2023.04.17.537144>
- Rubinacci, S., Ribeiro, D.M., Hofmeister, R.J., Delaneau, O., 2021. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nat Genet* 53, 120–126. <https://doi.org/10.1038/s41588-020-00756-0>
- Rue, D., Webster, D., Traverse, A., 2002. LATE HOLOCENE FIRE AND AGRICULTURE IN THE COPAN VALLEY, HONDURAS. *Ancient Mesoam* 13, 267–272. <https://doi.org/10.1017/S095653610213210X>
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A., 1988. Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase. *Science* 239, 487–491. <https://doi.org/10.1126/science.2448875>
- Salazar, L., Burger, R., Forst, J., Barquera, R., Nesbitt, J., Calero, J., Washburn, E., Verano, J., Zhu, K., Sop, K., Kassadjikova, K., Ibarra Asencios, B., Davidson, R., Bradley, B., Krause, J., Fehren-Schmitz, L., 2023. Insights into the genetic histories and lifeways of Machu Picchu's occupants. *Sci. Adv.* 9, eadg3377. <https://doi.org/10.1126/sciadv.adg3377>
- Salter-Townshend, M., Myers, S., 2019. Fine-Scale Inference of Ancestry Segments Without Prior Knowledge of Admixing Groups. *Genetics* 212, 869–889. <https://doi.org/10.1534/genetics.119.302139>
- Sánchez-Quinto, F., Schroeder, H., Ramirez, O., Ávila-Arcos, M.C., Pybus, M., Olalde, I., Velazquez, A.M.V., Marcos, M.E.P., Encinas, J.M.V., Bertranpetit, J., Orlando, L., Gilbert, M.T.P., Lalueza-Fox, C., 2012. Genomic Affinities of Two 7,000-Year-Old Iberian Hunter-Gatherers. *Current Biology* 22, 1494–1499. <https://doi.org/10.1016/j.cub.2012.06.005>
- Sandoval, K., Buentello-Malo, L., Peñaloza-Espinosa, R., Avelino, H., Salas, A., Calafell, F., Comas, D., 2009. Linguistic and maternal genetic diversity are not correlated in Native Mexicans. *Hum Genet* 126, 521–531. <https://doi.org/10.1007/s00439-009-0693-y>
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463–5467. <https://doi.org/10.1073/pnas.74.12.5463>
- Sankararaman, S., Sridhar, S., Kimmel, G., Halperin, E., 2008. Estimating Local Ancestry in Admixed Populations. *The American Journal of Human Genetics* 82, 290–303. <https://doi.org/10.1016/j.ajhg.2007.09.022>
- Saupe, T., Montinaro, F., Scaggion, C., Carrara, N., Kivisild, T., D'Atanasio, E., Hui, R., Solnik, A., Lebrasseur, O., Larson, G., Alessandri, L., Arienzo, I., De Angelis, F., Rolfo, M.F., Skeates, R., Silvestri, L., Beckett, J., Talamo, S., Dolfini, A., Miari, M., Metspalu, M., Benazzi, S., Capelli, C., Pagani, L., Scheib, C.L., 2021.

- Ancient genomes reveal structural shifts after the arrival of Steppe-related ancestry in the Italian Peninsula. *Current Biology* 31, 2576–2591.e12.
<https://doi.org/10.1016/j.cub.2021.04.022>
- Scheib, C.L., Li, H., Desai, T., Link, V., Kendall, C., Dewar, G., Griffith, P.W., Mörseburg, A., Johnson, J.R., Potter, A., Kerr, S.L., Endicott, P., Lindo, J., Haber, M., Xue, Y., Tyler-Smith, C., Sandhu, M.S., Lorenz, J.G., Randall, T.D., Faltyskova, Z., Pagani, L., Danecek, P., O'Connell, T.C., Martz, P., Boraas, A.S., Byrd, B.F., Leventhal, A., Cambra, R., Williamson, R., Lesage, L., Holguin, B., Ygnacio-De Soto, E., Rosas, J., Metspalu, M., Stock, J.T., Manica, A., Scally, A., Wegmann, D., Malhi, R.S., Kivisild, T., 2018. Ancient human parallel lineages within North America contributed to a coastal expansion. *Science* 360, 1024–1027.
<https://doi.org/10.1126/science.aar6851>
- Schele, L., 1992. The Founders of Lineages at Copan and other Maya Sites. *Ancient Mesoam* 3, 135–144. <https://doi.org/10.1017/S0956536100002352>
- Schele, L., Mathews, P., 1998. The code of kings: the language of seven sacred Maya temples and tombs. Scribner, New York.
- Scherer, A.K., 2007. Population structure of the classic period Maya. *American J Phys Anthropol* 132, 367–380. <https://doi.org/10.1002/ajpa.20535>
- Schroeder, H., Sikora, M., Gopalakrishnan, S., Cassidy, L.M., Maisano Delser, P., Sandoval Velasco, M., Schraiber, J.G., Rasmussen, S., Homburger, J.R., Ávila-Arcos, M.C., Allentoft, M.E., Moreno-Mayar, J.V., Renaud, G., Gómez-Carballa, A., Laffoon, J.E., Hopkins, R.J.A., Higham, T.F.G., Carr, R.S., Schaffer, W.C., Day, J.S., Hoogland, M., Salas, A., Bustamante, C.D., Nielsen, R., Bradley, D.G., Hofman, C.L., Willerslev, E., 2018. Origins and genetic legacies of the Caribbean Taino. *Proc. Natl. Acad. Sci. U.S.A.* 115, 2341–2346.
<https://doi.org/10.1073/pnas.1716839115>
- Schubert, M., Lindgreen, S., Orlando, L., 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes* 9, 88.
<https://doi.org/10.1186/s13104-016-1900-2>
- Severson, A.L., Byrd, B.F., Mallott, E.K., Owings, A.C., DeGiorgio, M., de Flamingh, A., Nijmeh, C., Arellano, M.V., Leventhal, A., Rosenberg, N.A., Malhi, R.S., 2022. Ancient and modern genomics of the Ohlone Indigenous population of California. *Proc. Natl. Acad. Sci. U.S.A.* 119, e2111533119.
<https://doi.org/10.1073/pnas.2111533119>
- Sharer, R.J., 2009. Daily life in Maya civilization, 2nd ed. ed, The Greenwood Press “Daily life through history” series. Greenwood Press, Westport, Conn.
- Shendure, J., Balasubramanian, S., Church, G.M., Gilbert, W., Rogers, J., Schloss, J.A., Waterston, R.H., 2017. DNA sequencing at 40: past, present and future. *Nature* 550, 345–353. <https://doi.org/10.1038/nature24286>
- Silva, M.A.C.E., Ferraz, T., Hünemeier, T., 2022. A genomic perspective on South American human history. *Genet. Mol. Biol.* 45, e20220078.
<https://doi.org/10.1590/1678-4685-gmb-2022-0078>
- Skoglund, P., Mallick, S., Bortolini, M.C., Chennagiri, N., Hünemeier, T., Petzl-Erler, M.L., Salzano, F.M., Patterson, N., Reich, D., 2015. Genetic evidence for two founding populations of the Americas. *Nature* 525, 104–108.
<https://doi.org/10.1038/nature14895>
- Skoglund, P., Storå, J., Götherström, A., Jakobsson, M., 2013. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *Journal of Archaeological Science* 40, 4477–4482.
<https://doi.org/10.1016/j.jas.2013.07.004>
- Söchtig, J., Álvarez-Iglesias, V., Mosquera-Miguel, A., Gelabert-Besada, M., Gómez-Carballa, A., Salas, A., 2015. Genomic insights on the ethno-history of the Maya and the ‘Ladinos’ from Guatemala. *BMC Genomics* 16, 131.

- <https://doi.org/10.1186/s12864-015-1339-1>
- Somerville, A.D., Fauvelle, M., Froehle, A.W., 2013. Applying new approaches to modeling diet and status: isotopic evidence for commoner resiliency and elite variability in the Classic Maya lowlands. *Journal of Archaeological Science* 40, 1539–1553. <https://doi.org/10.1016/j.jas.2012.10.029>
- Sousa Da Mota, B., Rubinacci, S., Cruz Dávalos, D.I., G. Amorim, C.E., Sikora, M., Johannsen, N.N., Szmyt, M.H., Włodarczak, P., Szczepanek, A., Przybyła, M.M., Schroeder, H., Allentoft, M.E., Willerslev, E., Malaspinas, A.-S., Delaneau, O., 2023. Imputation of ancient human genomes. *Nat Commun* 14, 3660. <https://doi.org/10.1038/s41467-023-39202-0>
- Stoneking, M., Arias, L., Liu, D., Oliveira, S., Pugach, I., Rodriguez, J.J.R.B., 2023. Genomic perspectives on human dispersals during the Holocene. *Proc. Natl. Acad. Sci. U.S.A.* 120, e2209475119. <https://doi.org/10.1073/pnas.2209475119>
- Suzuki, S., Nakamura, S., Price, T.D., 2020. Isotopic proveniencing at Classic Copan and in the southern periphery of the Maya Area: A new perspective on multi-ethnic society. *Journal of Anthropological Archaeology* 60, 101228. <https://doi.org/10.1016/j.jaa.2020.101228>
- The 1000 Genomes Project Consortium, Corresponding authors, Auton, A., Abecasis, G.R., Steering committee, Altshuler, D.M., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., Flicek, P., Gabriel, S.B., Gibbs, R.A., Green, E.D., Hurles, M.E., Knoppers, B.M., Korbel, J.O., Lander, E.S., Lee, C., Lehrach, H., Mardis, E.R., Marth, G.T., McVean, G.A., Nickerson, D.A., Schmidt, J.P., Sherry, S.T., Wang, J., Wilson, R.K., Production group, Baylor College of Medicine, Gibbs, R.A., Boerwinkle, E., Doddapaneni, H., Han, Y., Korchina, V., Kovar, C., Lee, S., Muzny, D., Reid, J.G., Zhu, Y., BGI-Shenzhen, Wang, J., Chang, Y., Feng, Q., Fang, X., Guo, X., Jian, M., Jiang, H., Jin, X., Lan, T., Li, G., Li, J., Li, Yingrui, Liu, S., Liu, Xiao, Lu, Y., Ma, X., Tang, M., Wang, B., Wang, G., Wu, H., Wu, R., Xu, X., Yin, Y., Zhang, D., Zhang, W., Zhao, J., Zhao, M., Zheng, X., Broad Institute of MIT and Harvard, Lander, E.S., Altshuler, D.M., Gabriel, S.B., Gupta, N., Coriell Institute for Medical Research, Gharani, N., Toji, L.H., Gerry, N.P., Resch, A.M., European Molecular Biology Laboratory, European Bioinformatics Institute, Flicek, P., Barker, J., Clarke, L., Gil, L., Hunt, S.E., Kelman, G., Kulesha, E., Leinonen, R., McLaren, W.M., Radhakrishnan, R., Roa, A., Smirnov, D., Smith, R.E., Streeter, I., Thormann, A., Toneva, I., Vaughan, B., Zheng-Bradley, X., Illumina, Bentley, D.R., Grocock, R., Humphray, S., James, T., Kingsbury, Z., Max Planck Institute for Molecular Genetics, Lehrach, H., Sudbrak, R., Albrecht, M.W., Amstislavskiy, V.S., Borodina, T.A., Lienhard, M., Mertes, F., Sultan, M., Timmermann, B., Yaspo, M.-L., McDonnell Genome Institute at Washington University, Mardis, E.R., Wilson, R.K., Fulton, L., Fulton, R., US National Institutes of Health, Sherry, S.T., Ananiev, V., Belaia, Z., Beloslyudtsev, D., Bouk, N., Chen, C., Church, D., Cohen, R., Cook, C., Garner, J., Hefferon, T., Kimelman, M., Liu, C., Lopez, J., Meric, P., O'Sullivan, C., Ostapchuk, Y., Phan, L., Ponomarov, S., Schneider, V., Shekhtman, E., Sirotkin, K., Slotta, D., Zhang, H., University of Oxford, McVean, G.A., Wellcome Trust Sanger Institute, Durbin, R.M., Balasubramaniam, S., Burton, J., Danecek, P., Keane, T.M., Kolb-Kokocinski, A., McCarthy, S., Stalker, J., Quail, M., Analysis group, Affymetrix, Schmidt, J.P., Davies, C.J., Gollub, J., Webster, T., Wong, B., Zhan, Y., Albert Einstein College of Medicine, Auton, A., Campbell, C.L., Kong, Y., Marcketta, A., Baylor College of Medicine, Gibbs, R.A., Yu, F., Antunes, L., Bainbridge, M., Muzny, D., Sabo, A., Huang, Z., BGI-Shenzhen, Wang, J., Coin, L.J.M., Fang, L., Guo, X., Jin, X., Li, G., Li, Q., Li, Yingrui, Li, Z., Lin, H., Liu, B., Luo, R., Shao, H., Xie, Y., Ye, C., Yu, C., Zhang, F., Zheng, H., Zhu, H., Bilkent

University, Alkan, C., Dal, E., Kahveci, F., Boston College, Marth, G.T., Garrison, E.P., Kural, D., Lee, W.-P., Fung Leong, W., Stromberg, M., Ward, A.N., Wu, J., Zhang, M., Broad Institute of MIT and Harvard, Daly, M.J., DePristo, M.A., Handsaker, R.E., Altshuler, D.M., Banks, E., Bhatia, G., Del Angel, G., Gabriel, S.B., Genovese, G., Gupta, N., Li, H., Kashin, S., Lander, E.S., McCarroll, S.A., Nemesh, J.C., Poplin, R.E., Cold Spring Harbor Laboratory, Yoon, S.C., Lihm, J., Makarov, V., Cornell University, Clark, A.G., Gottipati, S., Keinan, A., Rodriguez-Flores, J.L., European Molecular Biology Laboratory, Korbel, J.O., Rausch, T., Fritz, M.H., Stütz, A.M., European Molecular Biology Laboratory, European Bioinformatics Institute, Flicek, P., Beal, K., Clarke, L., Datta, A., Herrero, J., McLaren, W.M., Ritchie, G.R.S., Smith, R.E., Zerbino, D., Zheng-Bradley, X., Harvard University, Sabeti, P.C., Shlyakhter, I., Schaffner, S.F., Vitti, J., Human Gene Mutation Database, Cooper, D.N., Ball, E.V., Stenson, P.D., Illumina, Bentley, D.R., Barnes, B., Bauer, M., Keira Cheetham, R., Cox, A., Eberle, M., Humphray, S., Kahn, S., Murray, L., Peden, J., Shaw, R., Icahn School of Medicine at Mount Sinai, Kenny, E.E., Louisiana State University, Batzer, M.A., Konkel, M.K., Walker, J.A., Massachusetts General Hospital, MacArthur, D.G., Lek, M., Max Planck Institute for Molecular Genetics, Sudbrak, R., Amstislavskiy, V.S., Herwig, R., McDonnell Genome Institute at Washington University, Mardis, E.R., Ding, L., Koboldt, D.C., Larson, D., Ye, Kai, McGill University, Gravel, S., National Eye Institute, NIH, Swaroop, A., Chew, E., New York Genome Center, Lappalainen, T., Erlich, Y., Gymrek, M., Frederick Willems, T., Ontario Institute for Cancer Research, Simpson, J.T., Pennsylvania State University, Shriver, M.D., Rutgers Cancer Institute of New Jersey, Rosenfeld, J.A., Stanford University, Bustamante, C.D., Montgomery, S.B., De La Vega, F.M., Byrnes, J.K., Carroll, A.W., DeGorter, M.K., Lacroute, P., Maples, B.K., Martin, A.R., Moreno-Estrada, A., Shringarpure, S.S., Zakharia, F., Tel-Aviv University, Halperin, E., Baran, Y., The Jackson Laboratory for Genomic Medicine, Lee, C., Cerveira, E., Hwang, J., Malhotra, A., Plewczynski, D., Radew, K., Romanovitch, M., Zhang, C., Thermo Fisher Scientific, Hyland, F.C.L., Translational Genomics Research Institute, Craig, D.W., Christoforides, A., Homer, N., Izatt, T., Kurdoglu, A.A., Sinari, S.A., Squire, K., US National Institutes of Health, Sherry, S.T., Xiao, C., University of California, San Diego, Sebat, J., Antaki, D., Gujral, M., Noor, A., Ye, Kenny, University of California, San Francisco, Burchard, E.G., Hernandez, R.D., Gignoux, C.R., University of California, Santa Cruz, Haussler, D., Katzman, S.J., James Kent, W., University of Chicago, Howie, B., University College London, Ruiz-Linares, A., University of Geneva, Dermitzakis, E.T., University of Maryland School of Medicine, Devine, S.E., University of Michigan, Abecasis, G.R., Min Kang, H., Kidd, J.M., Blackwell, T., Caron, S., Chen, W., Emery, S., Fritsche, L., Fuchsberger, C., Jun, G., Li, B., Lyons, R., Scheller, C., Sidore, C., Song, S., Sliwerska, E., Taliun, D., Tan, A., Welch, R., Kate Wing, M., Zhan, X., University of Montréal, Awadalla, P., Hodgkinson, A., University of North Carolina at Chapel Hill, Li, Yun, University of North Carolina at Charlotte, Shi, X., Quitadamo, A., University of Oxford, Lunter, G., McVean, G.A., Marchini, J.L., Myers, S., Churchhouse, C., Delaneau, O., Gupta-Hinch, A., Kretzschmar, W., Iqbal, Z., Mathieson, I., Menelaou, A., Rimmer, A., Xifara, D.K., University of Puerto Rico, Oleksyk, T.K., University of Texas Health Sciences Center at Houston, Fu, Yunxin, Liu, Xiaoming, Xiong, M., University of Utah, Jorde, L., Witherspoon, D., Xing, J., University of Washington, Eichler, E.E., Browning, B.L., Browning, S.R., Hormozdiari, F., Sudmant, P.H., Weill Cornell Medical College, Khurana, E., Wellcome Trust Sanger Institute, Durbin, R.M., Hurles, M.E., Tyler-Smith, C., Albers, C.A., Ayub, Q., Balasubramaniam,

S., Chen, Y., Colonna, V., Danecek, P., Jostins, L., Keane, T.M., McCarthy, S., Walter, K., Xue, Y., Yale University, Gerstein, M.B., Abyzov, A., Balasubramanian, S., Chen, J., Clarke, D., Fu, Yao, Harmanci, A.O., Jin, M., Lee, D., Liu, J., Jasmine Mu, X., Zhang, J., Zhang, Yan, Structural variation group, BGI-Shenzhen, Li, Yingrui, Luo, R., Zhu, H., Bilkent University, Alkan, C., Dal, E., Kahveci, F., Boston College, Marth, G.T., Garrison, E.P., Kural, D., Lee, W.-P., Ward, A.N., Wu, J., Zhang, M., Broad Institute of MIT and Harvard, McCarroll, S.A., Handsaker, R.E., Altshuler, D.M., Banks, E., Del Angel, G., Genovese, G., Hartl, C., Li, H., Kashin, S., Nemesh, J.C., Shakir, K., Cold Spring Harbor Laboratory, Yoon, S.C., Lihm, J., Makarov, V., Cornell University, Degenhardt, J., European Molecular Biology Laboratory, Korbel, J.O., Fritz, M.H., Meiers, S., Raeder, B., Rausch, T., Stütz, A.M., European Molecular Biology Laboratory, European Bioinformatics Institute, Flicek, P., Paolo Casale, F., Clarke, L., Smith, R.E., Stegle, O., Zheng-Bradley, X., Illumina, Bentley, D.R., Barnes, B., Keira Cheetham, R., Eberle, M., Humphray, S., Kahn, S., Murray, L., Shaw, R., Leiden University Medical Center, Lameijer, E.-W., Louisiana State University, Batzer, M.A., Konkel, M.K., Walker, J.A., McDonnell Genome Institute at Washington University, Ding, L., Hall, I., Ye, Kai, Stanford University, Lacroute, P., The Jackson Laboratory for Genomic Medicine, Lee, C., Cerveira, E., Malhotra, A., Hwang, J., Plewczynski, D., Radew, K., Romanovitch, M., Zhang, C., Translational Genomics Research Institute, Craig, D.W., Homer, N., US National Institutes of Health, Church, D., Xiao, C., University of California, San Diego, Sebat, J., Antaki, D., Bafna, V., Michaelson, J., Ye, Kenny, University of Maryland School of Medicine, Devine, S.E., Gardner, E.J., University of Michigan, Abecasis, G.R., Kidd, J.M., Mills, R.E., Dayama, G., Emery, S., Jun, G., University of North Carolina at Charlotte, Shi, X., Quitadamo, A., University of Oxford, Lunter, G., McVean, G.A., University of Texas MD Anderson Cancer Center, Chen, K., Fan, X., Chong, Z., Chen, T., University of Utah, Witherspoon, D., Xing, J., University of Washington, Eichler, E.E., Chaisson, M.J., Hormozdiari, F., Huddleston, J., Malig, M., Nelson, B.J., Sudmant, P.H., Vanderbilt University School of Medicine, Parrish, N.F., Weill Cornell Medical College, Khurana, E., Wellcome Trust Sanger Institute, Hurles, M.E., Blackburne, B., Lindsay, S.J., Ning, Z., Walter, K., Zhang, Yujun, Yale University, Gerstein, M.B., Abyzov, A., Chen, J., Clarke, D., Lam, H., Jasmine Mu, X., Sisu, C., Zhang, J., Zhang, Yan, Exome group, Baylor College of Medicine, Gibbs, R.A., Yu, F., Bainbridge, M., Challis, D., Evani, U.S., Kovar, C., Lu, J., Muzny, D., Nagaswamy, U., Reid, J.G., Sabo, A., Yu, J., BGI-Shenzhen, Guo, X., Li, W., Li, Yingrui, Wu, R., Boston College, Marth, G.T., Garrison, E.P., Fung Leong, W., Ward, A.N., Broad Institute of MIT and Harvard, Del Angel, G., DePristo, M.A., Gabriel, S.B., Gupta, N., Hartl, C., Poplin, R.E., Cornell University, Clark, A.G., Rodriguez-Flores, J.L., European Molecular Biology Laboratory, European Bioinformatics Institute, Flicek, P., Clarke, L., Smith, R.E., Zheng-Bradley, X., Massachusetts General Hospital, MacArthur, D.G., McDonnell Genome Institute at Washington University, Mardis, E.R., Fulton, R., Koboldt, D.C., McGill University, Gravel, S., Stanford University, Bustamante, C.D., Translational Genomics Research Institute, Craig, D.W., Christoforides, A., Homer, N., Izatt, T., US National Institutes of Health, Sherry, S.T., Xiao, C., University of Geneva, Dermitzakis, E.T., University of Michigan, Abecasis, G.R., Min Kang, H., University of Oxford, McVean, G.A., Yale University, Gerstein, M.B., Balasubramanian, S., Habegger, L., Functional interpretation group, Cornell University, Yu, H., European Molecular Biology Laboratory, European Bioinformatics Institute, Flicek, P., Clarke, L., Cunningham, F., Dunham, I., Zerbino, D., Zheng-Bradley, X., Harvard University, Lage, K., Berg Jespersen,

J., Horn, H., Stanford University, Montgomery, S.B., DeGorter, M.K., Weill Cornell Medical College, Khurana, E., Wellcome Trust Sanger Institute, Tyler-Smith, C., Chen, Y., Colonna, V., Xue, Y., Yale University, Gerstein, M.B., Balasubramanian, S., Fu, Yao, Kim, D., Chromosome Y group, Albert Einstein College of Medicine, Auton, A., Marcketta, A., American Museum of Natural History, Desalle, R., Narechania, A., Arizona State University, Wilson Sayres, M.A., Boston College, Garrison, E.P., Broad Institute of MIT and Harvard, Handsaker, R.E., Kashin, S., McCarroll, S.A., Cornell University, Rodriguez-Flores, J.L., European Molecular Biology Laboratory, European Bioinformatics Institute, Flicek, P., Clarke, L., Zheng-Bradley, X., New York Genome Center, Erlich, Y., Gymrek, M., Frederick Willems, T., Stanford University, Bustamante, C.D., Mendez, F.L., David Poznik, G., Underhill, P.A., The Jackson Laboratory for Genomic Medicine, Lee, C., Cerveira, E., Malhotra, A., Romanovitch, M., Zhang, C., University of Michigan, Abecasis, G.R., University of Queensland, Coin, L., Shao, H., Virginia Bioinformatics Institute, Mittelman, D., Wellcome Trust Sanger Institute, Tyler-Smith, C., Ayub, Q., Banerjee, R., Cerezo, M., Chen, Y., Fitzgerald, T.W., Louzada, S., Massaia, A., McCarthy, S., Ritchie, G.R., Xue, Y., Yang, F., Data coordination center group, Baylor College of Medicine, Gibbs, R.A., Kovar, C., Kalra, D., Hale, W., Muzny, D., Reid, J.G., BGI-Shenzhen, Wang, J., Dan, X., Guo, X., Li, G., Li, Yingrui, Ye, C., Zheng, X., Broad Institute of MIT and Harvard, Altshuler, D.M., European Molecular Biology Laboratory, European Bioinformatics Institute, Flicek, P., Clarke, L., Zheng-Bradley, X., Illumina, Bentley, D.R., Cox, A., Humphray, S., Kahn, S., Max Planck Institute for Molecular Genetics, Sudbrak, R., Albrecht, M.W., Lienhard, M., McDonnell Genome Institute at Washington University, Larson, D., Translational Genomics Research Institute, Craig, D.W., Izatt, T., Kurdoglu, A.A., US National Institutes of Health, Sherry, S.T., Xiao, C., University of California, Santa Cruz, Haussler, D., University of Michigan, Abecasis, G.R., University of Oxford, McVean, G.A., Wellcome Trust Sanger Institute, Durbin, R.M., Balasubramanian, S., Keane, T.M., McCarthy, S., Stalker, J., Samples and ELSI group, Chakravarti, A., Knoppers, B.M., Abecasis, G.R., Barnes, K.C., Beiswanger, C., Burchard, E.G., Bustamante, C.D., Cai, H., Cao, H., Durbin, R.M., Gerry, N.P., Gharani, N., Gibbs, R.A., Gignoux, C.R., Gravel, S., Henn, B., Jones, D., Jorde, L., Kaye, J.S., Keinan, A., Kent, A., Kerasidou, A., Li, Yingrui, Mathias, R., McVean, G.A., Moreno-Estrada, A., Ossorio, P.N., Parker, M., Resch, A.M., Rotimi, C.N., Royal, C.D., Sandoval, K., Su, Y., Sudbrak, R., Tian, Z., Tishkoff, S., Toji, L.H., Tyler-Smith, C., Via, M., Wang, Y., Yang, H., Yang, L., Zhu, J., Sample collection, British from England and Scotland (GBR), Bodmer, W., Colombians in Medellín, Colombia (CLM), Bedoya, G., Ruiz-Linares, A., Han Chinese South (CHS), Cai, Z., Gao, Y., Chu, J., Finnish in Finland (FIN), Peltonen, L., Iberian Populations in Spain (IBS), Garcia-Montero, A., Orfao, A., Puerto Ricans in Puerto Rico (PUR), Dutil, J., Martinez-Cruzado, J.C., Oleksyk, T.K., African Caribbean in Barbados (ACB), Barnes, K.C., Mathias, R.A., Hennis, A., Watson, H., McKenzie, C., Bengali in Bangladesh (BEB), Qadri, F., LaRocque, R., Sabeti, P.C., Chinese Dai in Xishuangbanna, China (CDX), Zhu, J., Deng, X., Esan in Nigeria (ESN), Sabeti, P.C., Asogun, D., Folarin, O., Happi, C., Omoniwa, O., Stremlau, M., Tariyal, R., Gambian in Western Division – Mandinka (GWD), Jallow, M., Sisay Joof, F., Corrah, T., Rockett, K., Kwiatkowski, D., Indian Telugu in the UK (ITU) and Sri Lankan Tamil in the UK (STU), Kooner, J., Kinh in Ho Chi Minh City, Vietnam (KHV), Tinh Hiê`n, T., Dunstan, S.J., Thuy Hang, N., Mende in Sierra Leone (MSL), Fonnie, R., Garry, R., Kanneh, L., Moses, L., Sabeti, P.C., Schieffelin, J., Grant, D.S., Peruvian in

Lima, Peru (PEL), Gallo, C., Poletti, G., Punjabi in Lahore, Pakistan (PJL), Saleheen, D., Rasheed, A., Scientific management, Brooks, L.D., Felsenfeld, A.L., McEwen, J.E., Vaydylevich, Y., Green, E.D., Duncanson, A., Dunn, M., Schloss, J.A., Wang, J., Yang, H., Writing group, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Min Kang, H., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R., 2015. A global reference for human genetic variation. *Nature* 526, 68–74.
<https://doi.org/10.1038/nature15393>

The International HapMap Consortium, Genotyping centres: Baylor College of Medicine and ParAllele BioScience, Gibbs, R.A., Belmont, J.W., Hardenbol, P., Willis, T.D., Yu, F., Chinese HapMap Consortium, Yang, H., Ch'ang, L.-Y., Huang, W., Liu, B., Shen, Y., Tam, P.K.-H., Tsui, L.-C., Waye, M.M.Y., Wong, J.T.-F., Zeng, C., Zhang, Q., Illumina, Chee, M.S., Galver, L.M., Kruglyak, S., Murray, S.S., Oliphant, A.R., McGill University and Génome Québec Innovation Centre, Montpetit, A., Hudson, T.J., Chagnon, F., Ferretti, V., Leboeuf, M., Phillips, M.S., Verner, A., University of California at San Francisco and Washington University, Kwok, P.-Y., Duan, S., Lind, D.L., Miller, R.D., Rice, J.P., Saccone, N.L., Taillon-Miller, P., Xiao, M., University of Tokyo and RIKEN, Nakamura, Y., Sekine, A., Sorimachi, K., Tanaka, T., Tanaka, Y., Tsunoda, T., Yoshino, E., Wellcome Trust Sanger Institute, Bentley, D.R., Deloukas, P., Hunt, S., Powell, D., Whitehead Institute/MIT Center for Genome Research, Altshuler, D., Gabriel, S.B., Community engagement/public consultation and sample-collection groups: Beijing Normal University and Beijing Genomics Institute, Zhang, H., Zeng, C., Health Sciences University of Hokkaido, Eubios Ethics Institute and Shinshu University, Matsuda, I., Fukushima, Y., Macer, D.R., Suda, E., Howard University and University of Ibadan, Rotimi, C.N., Adebamowo, C.A., Aniagwu, T., Marshall, P.A., Matthew, O., Nkwodimmah, C., Royal, C.D.M., University of Utah, Leppert, M.F., Dixon, M., Analysis Groups: Cold Spring Harbor Laboratory, Stein, L.D., Cunningham, F., Kanani, A., Thorisson, G.A., Johns Hopkins University School of Medicine, Chakravarti, A., Chen, P.E., Cutler, D.J., Kashuk, C.S., University of Oxford, Donnelly, P., Marchini, J., McVean, G.A.T., Myers, S.R., University of Oxford, Wellcome Trust Centre for Human Genetics, Cardon, L.R., Abecasis, G.R., Morris, A., Weir, B.S., US National Institutes of Health, Mullikin, J.C., Sherry, S.T., Feolo, M., Whitehead Institute/MIT Center for Genome Research, Altshuler, D., Daly, M.J., Schaffner, S.F., Ethical, Legal and Social Issues: Chinese Academy of Social Sciences, Qiu, R., Genetic Interest Group, Kent, A., Howard University, Dunston, G.M., Kyoto University, Kato, K., Nagasaki University, Niikawa, N., University of Montréal, Knoppers, B.M., University of Oklahoma, Foster, M.W., Vanderbilt University, Clayton, E.W., Wang, V.O., Wellcome Trust, Watkin, J., SNP Discovery: Baylor College of Medicine, Gibbs, R.A., Belmont, J.W., Sodergren, E., Weinstock, G.M., Washington University, Wilson, R.K., Fulton, L.L., Wellcome Trust Sanger Institute, Rogers, J., Whitehead Institute/MIT Center for Genome Research, Birren, B.W., Scientific Management: Chinese Academy of Sciences, Han, H., Chinese Ministry of Science and Technology, Wang, H., Genome Canada, Godbout, M., Wallenburg, J.C., Génome Québec, L'Archevêque, P., Bellemare, G., Japanese Ministry of Education, Culture, Sports, Science and Technology, Todani, K., Fujita, T., Tanaka, S., The SNP Consortium, Holden, A.L., Lai, E.H., US National Institutes of Health, Collins, F.S., Brooks, L.D., McEwen, J.E., Guyer, M.S., Jordan, E., Peterson, J.L., Spiegel, J., Sung, L.M., Zacharia, L.F., Wellcome Trust, Kennedy, K., Dunn, M.G., Seabrook, R., Shillito, M., Skene, B., Stewart, J.G., Initial Planning Groups: Populations and Ethical, Legal and Social Issues Group, Valle, D.L.,

- Clayton, E.W., Jorde, L.B., Belmont, J.W., Chakravarti, A., Cho, M.K., Duster, T., Foster, M.W., Jasperse, M., Knoppers, B.M., Kwok, P.-Y., Licinio, J., Long, J.C., Marshall, P.A., Ossorio, P.N., Wang, V.O., Rotimi, C.N., Royal, C.D.M., Spallone, P., Terry, S.F., Methods Group, Lander, E.S., Lai, E.H., Nickerson, D.A., Abecasis, G.R., Altshuler, D., Bentley, D.R., Boehnke, M., Cardon, L.R., Daly, M.J., Deloukas, P., Douglas, J.A., Gabriel, S.B., Hudson, R.R., Hudson, T.J., Kruglyak, L., Kwok, P.-Y., Nakamura, Y., Nussbaum, R.L., Royal, C.D.M., Schaffner, S.F., Sherry, S.T., Stein, L.D., Tanaka, T., 2003. The International HapMap Project. *Nature* 426, 789–796. <https://doi.org/10.1038/nature02168>
- Tiesler, V., Sedig, J., Nakatsuka, N., Mallick, S., Lazaridis, I., Bernardos, R., Broomandkhoshbacht, N., Oppenheimer, J., Lawson, A.M., Stewardson, K., Rohland, N., Kennett, D.J., Price, T.D., Reich, D., 2022. Life and death in early colonial Campeche: new insights from ancient DNA. *Antiquity* 96, 937–954. <https://doi.org/10.15184/aqy.2022.79>
- Tiwanaku Part 1: The City, 2021. . Ancient Americas.
- Tsosie, K.S., Begay, R.L., Fox, K., Garrison, N.A., 2020. Generations of genomes: advances in paleogenomics technology and engagement for Indigenous people of the Americas. *Current Opinion in Genetics & Development* 62, 91–96. <https://doi.org/10.1016/j.gde.2020.06.010>
- Unterländer, M., Palstra, F., Lazaridis, I., Pilipenko, A., Hofmanová, Z., Groß, M., Sell, C., Blöcher, J., Kirsanow, K., Rohland, N., Rieger, B., Kaiser, E., Schier, W., Pozdniakov, D., Khokhlov, A., Georges, M., Wilde, S., Powell, A., Heyer, E., Currat, M., Reich, D., Samashev, Z., Parzinger, H., Molodin, V.I., Burger, J., 2017. Ancestry and demography of Iron Age nomads of the Eurasian Steppe. *Nat Commun* 8, 14615. <https://doi.org/10.1038/ncomms14615>
- Van Der Valk, T., Pečnerová, P., Díez-del-Molino, D., Bergström, A., Oppenheimer, J., Hartmann, S., Xenikoudakis, G., Thomas, J.A., Dehasque, M., Sağlıcan, E., Fidan, F.R., Barnes, I., Liu, S., Somel, M., Heintzman, P.D., Nikolskiy, P., Shapiro, B., Skoglund, P., Hofreiter, M., Lister, A.M., Götherström, A., Dalén, L., 2021. Million-year-old DNA sheds light on the genomic history of mammoths. *Nature* 591, 265–269. <https://doi.org/10.1038/s41586-021-03224-9>
- Van Dop Dejesus, J., 2015. Aztec civilization. *National Geographic*, National Geographic.
- van Oven, M., Kayser, M., 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* 30, E386–E394. <https://doi.org/10.1002/humu.20921>
- Verdugo, C., Zhu, K., Kassadjikova, K., Berg, L., Forst, J., Galloway, A., Brady, J.E., Fehren-Schmitz, L., 2020. An investigation of ancient Maya intentional dental modification practices at Midnight Terror Cave using anthroposcopic and paleogenomic methods. *Journal of Archaeological Science* 115, 105096. <https://doi.org/10.1016/j.jas.2020.105096>
- Villa-Islas, V., Izarraras-Gomez, A., Larena, M., Campos, E.M.P., Sandoval-Velasco, M., Rodríguez-Rodríguez, J.E., Bravo-Lopez, M., Moguel, B., Fregel, R., Garfias-Morales, E., Medina Tretmanis, J., Velázquez-Ramírez, D.A., Herrera-Muñoz, A., Sandoval, K., Nieves-Colón, M.A., Zepeda García Moreno, G., Villanea, F.A., Medina, E.F.V., Aguayo-Haro, R., Valdiosera, C., Ioannidis, A.G., Moreno-Estrada, A., Jay, F., Huerta-Sanchez, E., Moreno-Mayar, J.V., Sánchez-Quinto, F., Ávila-Arcos, M.C., 2023. Demographic history and genetic structure in pre-Hispanic Central Mexico. *Science* 380, eadd6142. <https://doi.org/10.1126/science.add6142>
- Watson, J.D., Crick, F.H.C., 1953. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* 171, 737–738. <https://doi.org/10.1038/171737a0>

- Webster, D., Sanders, W.T., Van Rossum, P., 1992. A Simulation of Copan Populatin History and its Implications. *Ancient Mesoam* 3, 185–197. <https://doi.org/10.1017/S095653610000239X>
- Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H.-J., Kronenberg, F., Salas, A., Schönherr, S., 2016. HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res* 44, W58–W63. <https://doi.org/10.1093/nar/gkw233>
- Willerslev, E., Meltzer, D.J., 2021. Peopling of the Americas as inferred from ancient genomics. *Nature* 594, 356–364. <https://doi.org/10.1038/s41586-021-03499-y>