Research Article

Increased Population Risk of AIP-related Acromegaly and Gigantism in Ireland

Serban Radian1,2, Yoan Diekmann3, Plamena Gabrovska1, Brendan Holland1, Lisa Bradley4, Helen Wallace5, Karen Stals6, Anna-Marie Bussell5, Karen McGurren7, Martin Cuesta7, Anthony W. Ryan8, Maria Herincs1, Laura C. Hernández-Ramírez1, Aidan Holland1, Jade Samuels1, Elena Daniela Aflorei5, Sayka Barry1, Judit Dénes1, Ida Pernicova1, Craig E. Stiles1, Giampaolo Trivellin1, Ronan McCloskey1, Michal Azjensztejn9, Noina Abid9, Scott A. Akker1, Moises Mercado10, Mark Cohen11, Rajesh V. Thakker M.D., Sc.D.12, Stephanie Baldeweg13, Ariel Barkan14, Madalina Musat2, Miles Levy15, Steve Orme16, Martina Unterländer17, Joachim Burger17, Ajith V. Kumar18, Sian Ellard6, Joseph McPartlin19, Ross McManus8, Gerard J. Linden20, Brew Atkinson5, David Balding3,21, Amar Agha7, Chris J. Thompson7, Steven J. Hunter5, Mark G. Thomas3, Patrick J Morrison4,22, Mártá Korbonits1

1Centre of Endocrinology, William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom
2Department of Endocrinology, Carol Davila University of Medicine and Pharmacy, C.I. Parhon National Institute of Endocrinology, Bucharest 011863, Romania
3Research Department of Genetics, Evolution and Environment, University College London, London, United Kingdom
4Department of Medical Genetics, Belfast HSC Trust
5Regional Centre for Endocrinology and Diabetes, Royal Victoria Hospital, Belfast, United Kingdom

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/humu.23121.

This article is protected by copyright. All rights reserved.
6Department of Molecular Genetics, Royal Devon and Exeter NHS Foundation Trust/Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, United Kingdom

7Department of Endocrinology and Diabetes, Beaumont Hospital/RCSI Medical School, Dublin, Ireland

8Department of Clinical Medicine and Institute of Molecular Medicine, Trinity College Dublin, Trinity Centre for Health Sciences, St James's Hospital, Dublin, Ireland

9Royal Belfast Hospital for Sick Children, Belfast, United Kingdom

10Endocrinology Service/Experimental Endocrinology Unit, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, IMSS, Mexico City, Mexico

11Department of Endocrinology and Diabetes, Barnet General Hospital, London, United Kingdom

12Academic Endocrine Unit, OCDEM, University of Oxford, Oxford, United Kingdom

13Department of Endocrinology and Diabetes, University College London Hospitals, London, United Kingdom

14Department of Neurosurgery, University of Michigan, Ann Arbor, Michigan, United States of America

15Department of Endocrinology, University Hospitals of Leicester, Leicester, United Kingdom

16Department of Endocrinology, St James's University Hospital, Leeds, United Kingdom

17Institute of Anthropology, Johannes Gutenberg University, Mainz, Germany

18North East Thames Regional Genetics Service, Great Ormond Street Hospital, London, United Kingdom

19Trinity Biobank, Institute of Molecular Medicine, Trinity College Dublin, Trinity Centre for Health Sciences, St James's Hospital, Dublin, Ireland

20Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, United Kingdom

21Schools of Biosciences and of Mathematics & Statistics, University of Melbourne, Parkville, Victoria, Australia
Abstract

The aryl hydrocarbon receptor interacting protein (AIP) founder mutation R304* (or p.R304*; NM_003977.3:c.910C>T, p.Arg304Ter) identified in Northern Ireland (NI) predisposes to acromegaly/gigantism; its population health impact remains unexplored. We measured R304* carrier frequency in 936 Mid Ulster, 1000 Greater Belfast (both in NI) and 2094 Republic of Ireland (ROI) volunteers and in 116 acromegaly/gigantism patients. Carrier frequencies were 0.0064 in Mid Ulster (95%CI=0.0027-0.013; p=0.0005 vs. ROI), 0.001 in Greater Belfast (0.00011-0.0047) and zero in ROI (0-0.0014). R304* prevalence was elevated in acromegaly/gigantism patients in NI (11/87,
12.6%, p<0.05), but not in ROI (2/29, 6.8%) vs. non-Irish patients (0-2.41%). Haploblock conservation supported a common ancestor for all the 18 identified Irish pedigrees (81 carriers, 30 affected). Time to most recent common ancestor (tMRCA) was 2550 (1275-5000) years. tMRCA-based simulations predicted 432 (90-5175) current carriers, including 86 affected (18-1035) for 20% penetrance. In conclusion, R304* is frequent in Mid Ulster, resulting in numerous acromegaly/gigantism cases. tMRCA is consistent with historical/folklore accounts of Irish giants. Forward simulations predict many undetected carriers; geographically-targeted population screening improves asymptomatic carrier identification, complementing clinical testing of patients/relatives. We generated disease awareness locally, necessary for early diagnosis and improved outcomes of AIP-related disease.

**Keywords:** AIP; acromegaly and gigantism; evolutionary genetics; founder mutation; population screening

**Introduction**

Germline aryl hydrocarbon receptor interacting-protein (AIP; MIM# 605555) mutations cause autosomal dominant familial isolated pituitary adenomas (FIPA) most commonly manifesting as acromegaly or gigantism. Due to incomplete penetrance, the disease can also manifest as apparently sporadic pituitary adenoma (PA) (Daly, et al., 2007; Leontiou, et al., 2008; Vierimaa, et al., 2006). These mostly GH- (somatotrophinomas) and/or prolactin-secreting tumours develop in 17-20% of AIP mutation (AIPmut) carriers, typically before 30 years of age (Beckers, et al., 2013; Hernández-Ramírez, et al., 2015). A diagnostic delay of several years from onset of symptoms is typical in somatotrophinomas, allowing tumour expansion and prolonged GH excess, leading to significant complications (Reddy, et al., 2010). Despite major therapeutical advances, management of somatotrophinomas remains challenging, especially in young patients with large/invasive tumours,
two frequent characteristics of AIPmut-positive patients (Beckers, et al., 2013). Identification of AIPmut carriers enables earlier clinical diagnosis and treatment, improving outcomes (Hernández-Ramírez, et al., 2015). To this end, predictive genetic testing of relatives of carriers is recommended (Korbonits, et al., 2012). General population screening is costly but targeted screening could be cost-effective when a disease-associated allele is geographically highly localised (Schiavi, et al., 2012), as might be expected for evolutionarily recent mutations. To date, population screening for genetic pituitary disease has not been performed.

We have previously inferred a recent origin for the NM_003977.3:c.910C>T (p.Arg304Ter) allele (p.R304* or R304*) in six FIPA pedigrees of Irish ancestry, including individuals born in the 18th and 20th centuries, in a small region within Mid Ulster, Northern Ireland (NI) (Chahal, et al., 2011; Stals, et al., 2011). Using the age estimate of R304*, we predicted through forward simulations a large number of current carriers (Chahal, et al., 2011). Therefore, we hypothesised that R304* would be frequent in the Irish general population, in particular in Mid Ulster, and should lead to a high proportion of R304*-positive patients with somatotrophinomas. In this study, we conducted a population screening in Mid Ulster, in comparison with two large population samples from Greater Belfast, NI and from the Republic of Ireland (ROI). In addition, we identified R304* carriers among patients with acromegaly/gigantism in two main Irish endocrine referral centres (Belfast and Dublin) and studied R304*-positive patients in our International FIPA Consortium database (Hernández-Ramírez, et al., 2015). We developed a new coalescent-based simulation approach to more precisely estimate the time to most recent common ancestor (tMRCA) of the Irish R304* allele based on haplotype data, and applied forward simulation to predict the number of carriers currently alive.
**Materials and Methods**

**Subjects**

We performed a cross-sectional genetic analysis study of general population adult individuals from three regions in Ireland: screening of volunteers from Mid Ulster (Local Government Districts of Magherafelt, Cookstown and Dungannon of NI) and genotyping of two large control groups from the Greater Belfast region – PRIME study participants (Linden, et al., 2012) – and the ROI Trinity Biobank (Table 1). Sample size calculations and screening procedures are described in the Supp. Methods.

A prospective cohort of unselected patients with acromegaly or gigantism were recruited at the University Hospital in Belfast, the only tertiary pituitary referral centre in NI (n=87, representing 60% of the NI Acromegaly Registry patients), and at the Beaumont Hospital, Dublin, the largest pituitary referral centre in ROI and closest to Mid Ulster (n=29). Additional R304*-positive pedigrees (5 Irish and 5 non-Irish) and 13 AIPmut-negative Irish FIPA pedigrees (Hernández-Ramírez, et al., 2015) were also included. Ethnicity – as assessed through family history – was unambiguous for all pedigrees. R304*-positive patients were previously reported (Chahal, et al., 2011; Hernández-Ramírez, et al., 2015; Igreja, et al., 2010; Leontiou, et al., 2008; Stals, et al., 2011; Williams, et al., 2014), while the other patients and subjects have not been previously described. Patients were tested for AIPmuts, as previously described (Igreja, et al., 2010). The following pedigree codes were used: ‘Sp’ (sporadic), ‘FIPA’ (familial) and ‘Screening’ (screening-detected). Tooth-extracted DNA from a pituitary giant’s skeleton at Trinity College Dublin (Cunningham, 1892) was sequenced for AIPmut, as previously described (Chahal, et al., 2011). The study protocol was approved by the local Ethics Committees; all study participants gave written informed consent.
Evolutionary Genetics and Statistical Analysis

Fourteen microsatellites (STR) covering 8.3 Mbp around the AIP gene were genotyped in a minimum of one R304* carrier per pedigree, as previously described (Chahal, et al., 2011). Haplotypes were computed using PHASE (Stephens, et al., 2001), incorporating phasing information deduced from genotypes of closely-related carriers (three pedigrees) (Supp. Table S1). To estimate the tMRCA of R304*-containing haplotypes, we developed a novel inference methodology combining an analytical result from coalescence theory and simulations in an Approximate Bayesian Computation framework (Beaumont, et al., 2002). First, we computed the tMRCA based on a smaller haplotype block around R304* that was fully conserved in 18 Irish pedigrees (Supp. Table S2, row E). The full conservation allowed us to apply an analytical formula (Donnelly, et al., 1996). The tMRCA distribution resulting from this first step became the prior distribution for the second step which used a simulation-based approach to incorporate additional conservation data of haploblocks shared only between subsets of pedigrees. We simulated genealogies with recombination and mutation using the program ms (Hudson, 2002) (Supp. Figure S1) and generated distributions of tMRCA from those simulations that were closest to the recombination and mutation pattern inferred from our haplotype data (Supp. Figure S2). Results were based on sex-averaged genetic distances for AIP and microsatellite loci, from HapMap v2 (The HapMap International Consortium, 2007) and Rutgers v3 (Matise, et al., 2007) genetic maps (Supp. Table S1). Next, starting with tMRCA values randomly chosen from the simulated tMRCA distribution (Supp. Figure S3A, B), we took a forward simulation approach to generate possible trajectories for the allele frequency in the population, through binomial sampling. This resulted in distributions of the expected present-day number of carriers, which we conditioned to have at least the minimum number of R304* carriers/generation observed in our dataset (n=27) (Supp. Figure S3C, D). Detailed procedures, including statistical analysis are described as Supp. Methods.
Results

Population Screening Reveals a Significantly High Carrier Frequency of the R304* Allele in Northern Ireland

Several families of patients suffering from acromegaly or gigantism and carrying the Irish R304* allele came from a small geographical area within Mid Ulster in NI (Chahal, et al., 2011; Stals, et al., 2011). In order to estimate the local carrier frequency of R304*, we screened a population sample of 936 adults (Table 1), 90% of which lived in Mid Ulster.

We identified six R304* heterozygous individuals (Supp. Table S3), corresponding to a carrier frequency estimate of 0.0064 (95% probability interval [CI]: 0.0027-0.013). Two related carriers, negative for personal or family history of PA, represented a novel R304* pedigree (Screening 1) (Supp. Tables S3 and S4). The remaining carriers detected through screening were two patients previously diagnosed with somatotrophinomas (pedigrees Sp 4 and FIPA 4) and two unaffected carriers from previously identified R304*-positive pedigrees Sp 4 and FIPA 1 (Supp. Tables S3 and S4). The R304* carrier status was significantly associated with a personal diagnosis of PA (p<0.001, Fisher’s exact test) and with a family history of PA (FIPA or sporadic, p<0.001, Fisher’s exact test).

Mid Ulster screening revealed many other PA patients, either directly (R304*-negative PA patients participating to screening) or indirectly (non-participating PA patients mentioned as family history of R304*-negative screening subjects) (Supp. Table S3).

The Greater Belfast control population sample (n=1000, Table 1) revealed one R304*-positive individual; corresponding to a carrier frequency of 0.001 (95% CI: 0.00011-0.0047). The carrier was unaffected (clinical examination, hormonal measurements and pituitary imaging), did not have a family history of PA, and cascade testing of relatives revealed two carriers, both unaffected (Supp. Table S4, Screening 2 pedigree).
In the ROI control population sample (n=2094, Table 1), no R304*-positive individuals were identified, corresponding to zero carrier frequency (95% CI estimate 0-0.0014). This is in accord with the absence/extremely low R304* carrier frequency observed in diverse population samples included in the Exome Aggregation Consortium (http://exac.broadinstitute.org/): two heterozygotes in 34,856 Europeans (0.00057) and none in 23,450 non-Europeans. The Mid Ulster carrier frequency was significantly higher than the ROI one (p=0.0005, Fisher’s exact test), while the Greater Belfast one did not differ from either ROI or Mid Ulster.

Extrapolation of our NI R304* carrier frequency estimates (1/1000 and 6/936) and 95% CIs to the population of Greater Belfast (579,276) and Mid Ulster (139,011), comprising 40% of NI, according to the 2011 census (http://www.ninis2.nisra.gov.uk), yielded an estimated number of carriers of 1470 (95% CI: 439-4530).

**R304* is Common in Irish Patients with Acromegaly/Gigantism**

In order to characterise the extent of R304*-related pituitary disease in Ireland, we tested for AIP alleles in 116 patients with acromegaly/gigantism from the two tertiary endocrine centres close to Mid Ulster. We identified 11 (12.6%) R304*-positive patients with somatotrophinomas in the NI Acromegaly Registry, Belfast, and two (6.8%) in the Dublin centre patients, while five NI (5.7%) and two Dublin patients (6.8%) presented other AIP variants (Supp. Table S5).

The R304*-positive patients belonged to pedigrees FIPA 1-5 and Sp 1-6 (Supp. Table S4). The R304* prevalence among NI patients with somatotrophinomas was higher than that of large non-Irish somatotrophinoma series (Cazabat, et al., 2012; Preda, et al., 2014; Schöfl, et al., 2014; Tichomirowa, et al., 2011) (0-2.41%; p<0.05, Fisher’s exact test). Geographically, R304*-positive patients with somatotrophinomas predominated in Mid Ulster, where they, remarkably, outnumbered AIPmut-negative patients (Figure 1).
Irish R304*-positive patients with somatotrophinomas (excluding those diagnosed prospectively, following predictive genetic testing) displayed the known features of AIPmut-positive PAs: younger age at disease onset/diagnosis, larger pituitary tumours, increased frequency of gigantism and familial disease, compared to AIPmut-negative patients (Table 2).

Our investigations revealed, however, a wider range of phenotypes. While two patients, diagnosed as part of cascade testing in families, harboured macroadenomas and underwent surgery, two prospectively diagnosed R304*-positive patients had pituitary microadenomas and four patients had disease onset after the age of 30 years (Hernández-Ramírez, et al., 2015). Interestingly, none of the patients carrying other AIP variants presented a PA family history.

**All Irish R304* Pedigrees Share a Common Ancestor That Lived Approximately 2500 Years Ago**

Having identified 13 R304*-positive pedigrees through population screening and testing of somatotrophinoma patients in Ireland, we searched the International FIPA Consortium database for additional R304*-positive pedigrees and identified 10 more, resulting in a total of 23 R304*-positive pedigrees (Figure 2). These 10 additional pedigrees were of diverse ethnic backgrounds: six from the UK: one of English and five of Irish ancestry (FIPA 6-8, Sp 7 and 18\textsuperscript{th} century patient pedigrees) and four from other countries (Figure 2 and Supp. Table S4). In this database, we also identified 13 AIPmut-negative Irish FIPA pedigrees (Supp. Table S6).

By reconstructing AIP gene region STR haplotypes in these 23 pedigrees, we determined that 18 of them shared a 0.2-1.95 Mbp fully conserved haploblock around R304*, as well as extended haploblocks covering up to the entire 8.3 Mbp genotyped region (Figure 2 and Supp. Table S1). These data support the existence of a recent common ancestral carrier haplotype, the Irish R304* founder. Fourteen of these 18 pedigrees were living in Ireland and the other four had known Irish ancestry. The five non-Irish pedigrees did not share significant haploblocks with the Irish ones or among each other (Figure 2), supporting the existence of recurrent independent R304* mutation
events at this mutation-prone CpG dinucleotide (Cazabat, et al., 2012; Chahal, et al., 2011; Cuny, et al., 2013; Daly, et al., 2007; de Lima, et al., 2012; Leontiou, et al., 2008; Niyazoglu, et al., 2014; Occhi, et al., 2010; Tichomirowa, et al., 2011; Vierimaa, et al., 2006).

The total number of Irish allele carriers was 90 (37 affected, 53 unaffected), 81 of which are currently alive (30 affected, 51 unaffected). Seven of the 51 unaffected living carriers (13.7%) were detected exclusively by population screening. The highest geographical concentration of R304* carriers in Ireland was observed in Mid Ulster, both for affected (somatotrophinoma) and unaffected individuals (Figure 1), corresponding to eight different pedigrees. A third of the Irish allele carriers (32%, 26/81) lived outside of Ireland: in England (21), Scotland (2), Australia (2) and Canada (1).

To estimate the time to the most recent common ancestor (tMRCA) of the R304* carrier haplotypes among the Irish pedigrees, we developed a novel coalescent-based approach, combining analytical calculations with haplotype simulation in an Approximate Bayesian Computation (Beaumont, et al., 2002) framework (Supp. Methods). The tMRCA estimates of the Irish R304* haplotypes were 102 (51-200) generations - median (95% CI) - based on HapMap, and 101 (50-179) generations, based on Rutgers map genetic distances (Supp. Figure S3A, B). These tMRCA distributions, e.g. 51-200 generations (HapMap-based) overlap substantially with those previously estimated: 17-150 generations (Chahal, et al., 2011) (Pr(X>Y) = 0.79). Assuming a generation time of 25 years (Chahal, et al., 2011), current estimates translate to 2550 (1275-5000) years (HapMap) and 2525 (1250-4475) years (Rutgers). To estimate the number of living subjects carrying the Irish R304* allele, forward simulations conditioned on the number of observed carriers predicted 144 (30-1725) and 141 (30-1430) carriers per generation - median (95% CI) - based on HapMap and Rutgers distances, respectively (Supp. Figure S3C, D). For three overlapping generations alive at present, we estimate the current number of carriers as 3 × 144 = 432 (90-5175; HapMap-based), corresponding to 86 (18-1035) pituitary adenoma patients, assuming 20% penetrance (Beckers, et al., 2013).
**Historical Irish Giants**

The medical literature, as well as Irish folklore, holds numerous descriptions of Irish giants. For some of these presumed pituitary gigantism individuals there is historical and medical evidence for the diagnosis (Supp. Table S7). We have tested DNA samples from two historical giants for AIP derived alleles: Charles Byrne (1761-1783), born in Mid Ulster, who carried the Irish R304* allele (Chahal, et al., 2011) and Cornelius Magrath (1736-1760), born in the Southern coast of Ireland and whose skeleton is conserved at Trinity College, Dublin (Cunningham, 1892), who was R304*-negative. Five additional giants related to known R304* carriers were presumed R304*-positive: Subjects 6, 7, 11, 13 and 19 (Supp. Table S7). Two unrelated Mid Ulster screening participants, R304*-negative themselves, provided photographic evidence of their extremely tall relatives, whom we included as historical Irish giants (Figure 3 and Supp. Table S7).

**Discussion**

In this study we demonstrated that the R304* AIP allele is present in the general population of Ireland and its frequency varies geographically, from a high of 6/936 in Mid Ulster to 1/1000 in the Greater Belfast region (both in NI), while absent in a ROI population sample. These differences were paralleled by the geographical distribution of R304*-positive somatotrophinoma cases, representing a significantly higher proportion of the somatotrophinoma patient population in NI (12.6%), than in large published somatotrophinoma cohorts (maximum 2.41%) (Cazabat, et al., 2012; Preda, et al., 2014; Schöfl, et al., 2014; Tichomirowa, et al., 2011), while in the ROI (Dublin centre) this proportion was intermediate (6.8%). Furthermore, the number of R304* carriers, both patients with somatotrophinomas and unaffected carriers, was highest in Mid Ulster. These data strongly support our hypothesis that Mid Ulster has an elevated R304* carrier frequency. However, 32% of the Irish
R304* allele carriers were identified outside of Ireland, an observation attributable to emigration. Another consequence of the increased R304* carrier frequency in Ireland is that a large proportion (8 out of 21, 38%) of Irish FIPA families are due to R304* - the only AIP mut identified in Irish FIPA so far.

We demonstrate that all 18 R304*-positive pedigrees of Irish origin identified in the study inherited this allele from a common ancestor, the R304* founder. In addition, we have shown that R304* in pedigrees of English, Indian, Romanian, US Italian, Irish and Mexican origin derives from independent recurrent mutational events (Chahal, et al., 2011; Ramirez-Renteria, et al., 2016). Founder AIP mut alleles have also been identified through haplotype analysis in Finnish (Vierimaa, et al., 2006), Italian (Occhi, et al., 2010), Comoros (Cuny, et al., 2013) and English (Salvatori, et al., 2014) populations, involving the p.Q14*, p.R304*, p.G117Afs*39, and p.F269_H275dup alleles, respectively, but the present population frequency of all these founder alleles remains unknown.

The refined coalescent simulation-based Approximate Bayesian Computation approach we describe in this paper provides a practical framework applicable to the estimation of tMRCA of identity-by-descent alleles. Previous estimates using five pedigrees and a simpler approach indicated a more recent common ancestor (66 generations, 95% CI=17-150) (Chahal, et al., 2011), although the difference between new and previous estimates was not statistically significant (Pr(X>Y) = 0.79).

Although the founder probably lived around 2500 years ago, many of today’s carriers are concentrated in a small region, suggesting a limited local migratory activity, despite significant emigration in recent centuries. A population bottleneck followed by re-expansion may explain the Mid Ulster R304* cluster, although genetic drift in a constant sized or growing population is consistent with observed frequencies and estimated age of this allele. Persistence of the R304* allele in this population may suggest the absence of, or only weak, purifying selection acting on this locus. While R304*-associated disease manifestations include hypopituitarism, infertility and shortened life-span, low disease penetrance will lead to a reduced selective disadvantage.
Additionally, it is possible that there were inclusive fitness effects in the unaffected carrier relatives of giants in the past – perhaps as a result of elevated status – that may have counterbalanced selective disadvantages in affected carriers.

Identification of asymptomatic AIPmut carriers enables prospective early diagnosis of PAs, followed by prompt and effective treatment, with improved chances of curing the disease and preventing complications (Hernández-Ramírez, et al., 2015; Williams, et al., 2014). To this end, cascade testing of AIPmut-positive patient relatives is recommended; however, limited knowledge of family relations, non-disclosure of information between relatives and non-acceptance of testing may hinder these efforts. While carrier identification in affected individuals is now becoming standard, detection of unaffected carriers and their endocrine screening is a challenge for the future. Supported by substantial patient and community involvement, and by local media coverage, our study has created awareness about AIP-related disease in Ireland, a necessary precursor for improved disease recognition and acceptance of genetic testing. Although population screening is an unbiased approach to carrier identification and could detect unaffected carriers without family history of pituitary disease, further studies are needed to justify this approach. Our study provides a proof-of-principle that asymptomatic R304* carriers in Ireland can be identified through population screening, by identifying two novel R304*-positive pedigrees comprising a relatively large proportion (7/52, 13.5%) of the known number of living unaffected carriers.

Given the tMRCA-predicted number of Irish R304* carriers (432, 95% CI: 90-5175, 86 affected expected at 20% penetrance) and that of known living carriers (81, of which 30 affected), we estimate the number of undetected carriers as the difference between predicted and known carriers, resulting in around 351 undetected carriers, including 56 affected patients. Many of these undetected carriers may actually live outside of Ireland, in countries with significant Irish immigration, such as Great Britain, Canada and Australia, where 32% of the known living Irish allele carriers reside, and the USA, where AIP mutation testing has been limited to date. For NI, alternative
calculations by extrapolating population carrier frequency estimates (Mid Ulster and Greater Belfast) to regional population size (approximately 40% of NI), estimated 340 to 5138 undetected carriers, a range comparable to the tMRCA-predicted estimate.

Several key features contribute to our study’s results and impact: 1) patient and general public involvement in its design and conduct, 2) analysis of multiple Irish general population samples and of a large unbiased patient cohort - the NI Acromegaly Registry, 3) analysis of the largest R304* pedigree collection to date, thanks to international collaboration, and 4) the incorporation of historical and ancient DNA data into our analyses. In comparison, studies describing other AIPmut founder alleles, as well as our own previous study of the Irish R304* allele, examined limited numbers of AIPmut pedigrees and relatively small numbers of patients and population controls, without providing estimates of population carrier frequencies of these alleles and their geographical distribution (Chahal, et al., 2011; Cuny, et al., 2013; Occhi, et al., 2010; Salvatori, et al., 2014; Vierimaa, et al., 2006).

Our study also has limitations. For practical reasons, only three large Irish general population samples were screened for R304*. Regarding PA types included in the study, we chose to include only patients with somatotrophinomas, as this is largely the commonest PA type in AIPmut carriers and the only one for which population-level data was available. Patients with prolactinomas can also harbour AIP mutations, although significantly less frequently than somatotrophinomas (Cazabat, et al., 2012; Cuny, et al., 2013; Preda, et al., 2014; Tichomirowa, et al., 2011); the frequency of R304* in Irish patients with prolactinomas remains to be determined. Sampling of patients with somatotrophinomas from ROI was limited to the Dublin centre, based on its proximity to Mid Ulster, and fewer patients were analysed, compared to NI. Future studies will help to better characterise the R304* carrier frequency in the ROI general and PA patient populations. Not all living members of R304*-positive pedigrees were available for genetic and clinical testing; this is a common difficulty in genetic studies and clinical practice. Our evolutionary genetics approach compensated for this, by
estimating the number of undetected allele carriers. Overall, we believe that these limitations have relatively little effect on our conclusions.

Our work suggested several measures to address the challenging task of identifying the large number of undetected Irish R304* allele carriers, in order to improve the outcomes of PA patients. 1) We propose R304* testing of patients of Irish descent with somatotrophinomas, especially ones with younger (< 30 years of age) onset, in Ireland and abroad, particularly in the Great Britain, USA, Canada and Australia, where large Irish immigrant populations exist. 2) Working closely together with patients and their communities has improved knowledge about AIP-related disease, which may lead to earlier recognition of disease, and hopefully increased acceptance of genetic testing. 3) Population screening in Ireland, especially in Mid Ulster, might be effective, but more study is needed regarding the psychological and economic implications, as well as further data on the clinical benefits of R304* screening (Hernández-Ramírez, et al., 2015).

Our results open the way for future studies of several key aspects of AIP-related disease. The Irish R304* cohort is unique in having a large number of carriers sharing the same pathogenic gene alteration and promoter area within the fully conserved haploblock, yet pedigrees display differences in penetrance and patient phenotypes differ significantly. Currently, it is not known what influences disease penetrance in AIPmut-positive individuals. Different individual AIPmut alleles (Igreja, et al., 2010) or classes (e.g. truncating vs. non-truncating) (Hernández-Ramírez, et al., 2015), changes in gene expression regulation (Cooper, et al., 2013) and other epistatic interactions may each play a role, as will other genetic and environmental factors. Further genetic studies of this cohort could help the identification of pituitary disease-modifying genes. Follow-up of clinically-identified and screening-detected pedigrees will provide new data about the natural history of disease in carriers, optimal follow-up schedules, efficiency and outcome of population screening.
Irish folklore has numerous stories regarding Irish giants and the remains of some of these giants have been studied in the past (Chahal, et al., 2011; Cunningham, 1892; Frankcom and Musgrave, 1976). Our data provides an explanation for the observation made by the pioneering anthropologist James C. Prichard in 1826: *In Ireland men of uncommon stature are often seen, and even a gigantic form and stature occur there much more frequently than in this island [Britain]... We can hardly avoid the conclusion that there must be some peculiarity in Ireland which gives rise to these phenomena* (Prichard, 1826).

**Supporting information**

Supp. Methods, Supp. Figures S1-S3, Supp. Tables S1-S7 and references (Austerlitz, et al., 2003; Bergland, 1965; Burger, et al., 2007; Carleton, 1996; Colombo, 2007; Fawcett, 1909; Frankcom and Musgrave, 1976; Georgitsi, et al., 2007; Lin, et al., 2007; Lynass, 1842; Oriola, et al., 2013; Prezio, et al., 1961; Raitila, et al., 2010; Stephens, et al., 1998; Templeton, 2006; Trivellin, et al., 2014) can be found in the online version of this article under “Supporting Information”.

**Acknowledgements**

We are grateful to Professor Paul Glacken, Anatomy Department, Trinity College and to Mr Gerard Kearns, Professor of Maxillofacial Surgery, Dublin for their help of obtaining tooth samples from the skeleton in Trinity College. We are very grateful to those general population screening subjects who provided data and materials included in Figure 3 and Supp. Table S7. We gratefully acknowledge the assistance of Mr. Rusko N. Slavov, Patricia and Michael Holland, and all our Mid Ulster friends who helped in the screening study.
This work was supported by the Wellcome Trust (Clinical Training fellowship 097970/Z/11/Z to C.E.S, Wellcome Senior Investigator award 100719/Z/12/Z to M.G.T and Wellcome Senior Investigator award 098395/Z/12/A to S.E.), the European Commission Research Executive Agency (Marie Currie Fellowship 303006/2011 to S.R.), Barts and the London Charity (to M.K), Pfizer UK (WS 733753 to M.K), Science Foundation Ireland (09/IN.1/B2640 to R.M.M.), Consejo Nacional de Ciencia y Tecnologia (scholarship to L.C.H.R.), FIPA Patients Charity, Belfast City Hospital Charitable fund, and the generous support of the Holland family.

Conflict of interest:

MK received grant funding from Novartis and Pfizer and is a member of Medical Advisory Board of Pfizer Inc. All the other authors have nothing to declare.

References


Cunningham DJ. 1892. The skull and some of the other bones of the skeleton of Cornelius Magrath, the Irish giant. The Journal of the Anthropological Institute of Great Britain and Ireland 21:40-41.


This article is protected by copyright. All rights reserved.


**Figure legends**

**Figure 1** Geographical distribution of R304* allele carriers in Ireland, compared to AIPmut-negative patients with somatotrophinomas. R304* carriers, unaffected (R304* unaffected, green columns) and patients with somatotrophinomas (R304* GH, red columns) were compared to AIPmut-negative patients with somatotrophinomas (AIP-neg GH, blue columns). Only subjects currently residing in Ireland were included; R304*-positive patients with other PA types (n=5) are not shown. The map was arbitrarily divided into regions separated by dotted borders, red for Mid Ulster. Ninety percent of the Mid Ulster screening subjects resided here, including all six R304* carriers. Each small graph summarizes data from the underlying region; the y-axis scale is identical for all graphs.
Figure 2 Microsatellite haplotypes of chromosome 11q12.2-13.3 of R304*-positive pedigrees. Irish (first two rows) and non-Irish pedigrees (third row) are shown. Marker alleles are displayed as amplicon sizes. Dark shading: haploblock shared between all Irish pedigrees (0.2-1.95 Mbp long); light shading: additional shared haploblocks. Thick horizontal lines represent AIP alleles (black = wild-type, yellow = R304*); intervals between markers are not drawn to scale. Sp = sporadic PA.

Haplotypes of 18th century patient, FIPA 1, 2, 3, 6, 7, FIPA UK, Romania, US (Italian) and Sp India pedigrees have been previously published (Chahal, et al., 2011; Stals, et al., 2011).
### Figure 3

New evidence of historical Irish giants. (A) Subject 15 (Supp. Table S7), 211 cm tall, from Maghera, Mid Ulster, uncle of a screening participant, died of a hypertensive crisis one month after photograph was taken - 13th May 1918, Daily Sketch (Canadian newspaper). (B) Group of Irish immigrants from Garvagh (a town close to Mid Ulster) after arrival in Western Australia, 1910. At centre with the pole, Subject 16 (Supp. Table S7), 213 cm tall; standing far left - unrelated subject with pronounced acromegalic features and tall stature (photograph from Garvagh Museum, NI).
Table Captions

**Table 1.** Demographic parameters of Irish population samples.

<table>
<thead>
<tr>
<th>Sample type</th>
<th>Mid Ulster (n=936)</th>
<th>Greater Belfast (n=1000)</th>
<th>ROI (n=2094)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Saliva DNA</td>
<td>Blood DNA</td>
<td>Blood DNA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(n=1095)</td>
<td>(n=1095)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Buccal swab DNA</td>
<td>Buccal swab DNA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(n=999)</td>
<td>(n=999)</td>
</tr>
<tr>
<td>Males - count</td>
<td>435 (46.5%)</td>
<td>1000 (100%)</td>
<td>486 (44.4%)</td>
</tr>
<tr>
<td>(proportion)</td>
<td></td>
<td></td>
<td>559 (55.9%)</td>
</tr>
<tr>
<td>Age - years, median</td>
<td>46 (18-85)</td>
<td>64 (58-72)</td>
<td>37 (19-69)</td>
</tr>
<tr>
<td>(range)</td>
<td></td>
<td></td>
<td>53 (18-76)</td>
</tr>
<tr>
<td>R304* / wt AIP</td>
<td>6</td>
<td>1a</td>
<td>0</td>
</tr>
<tr>
<td>genotype</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>wt/wt AIP genotype</td>
<td>930</td>
<td>999</td>
<td>1095</td>
</tr>
<tr>
<td>R304* carrier</td>
<td>6/936</td>
<td>1/1000</td>
<td>0</td>
</tr>
<tr>
<td>frequency</td>
<td></td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

*71 years old male, normal height, no clinical signs of pituitary disease, normal serum hormone levels (basal IGF-I, prolactin, and GH serum levels during the oral glucose tolerance test (GH-OGTT)), normal pituitary MRI examination - subsequently died of a pulmonary mesothelioma linked to asbestos exposure, aged 72 years; cascade testing of relatives (Screening 2 pedigree) identified two unaffected R304* carriers: 1M, 77 years, clinically unaffected and 1F, 42 years, clinically unaffected, normal pituitary MRI, random GH=4.7 ng/mL, minimal IGF-I increase (1·07 x ULN), normal PRL, GH-OGTT not yet performed. wt, wild-type*
Table 2. Clinical characteristics of Irish R304*-positive somatotrophinoma patients (NI Acromegaly Registry, Dublin patients and FIPA Consortium database) compared to AIPmut-negative Irish patients with somatotrophinomas (NI and Dublin patients).

<table>
<thead>
<tr>
<th>AIP genotype</th>
<th>Heterozygous NM_003977.3:c.910C&gt;T</th>
<th>Homozygous wild-type</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age (range) at onset (years)</td>
<td>17 (6-68) (n=29)</td>
<td>38.5 (14-85) (n=64)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Median age (range) at diagnosis (years)</td>
<td>19.5 (7-68) (n=30)</td>
<td>46 (19-85) (n=95)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Age &lt;30 years at disease onset</td>
<td>27/31 (87.1 %)</td>
<td>21/96 (21.9 %)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mean tumour size (mm)</td>
<td>23.8 ± 10.6 (n=11)</td>
<td>17.2 ± 8.4 (n=43)</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Pituitary macroadenoma</td>
<td>22/24 (91.7 %)</td>
<td>58/70 (82.9 %)</td>
<td>ns</td>
</tr>
<tr>
<td>Gigantism</td>
<td>18/31 (58.1 %)</td>
<td>7/96 (7.3 %)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Family history of PA</td>
<td>24/31 (77.4 %)</td>
<td>3/94 (3.2 %)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

ns = not significant