

In-frame seven amino-acid duplication in *AIP* arose over the last 3000 years, disrupts protein interaction and stability and is associated with gigantism

Roberto Salvatori^{1,*}, Serban Radian^{2,3,*}, Yoan Diekmann^{4,*}, Donato Iacovazzo^{2,*}, Alessia David⁵, Plamena Gabrovska², Giorgia Grassi², Anna-Marie Bussell⁶, Karen Stals⁶, Astrid Weber⁷, Richard Quinton⁸, Elizabeth C Crowne⁹, Valentina Corazzini¹, Lou Metherell², Tara Kearney¹⁰, Daniel Du Plessis¹⁰, Ajay Kumar Sinha¹¹, Atik Baborie¹¹, Anne-Lise Lecoq^{12,13}, Philippe Chanson^{12,13}, Olaf Ansorge¹⁴, Sian Ellard^{6,15}, Peter J Trainer¹⁶, David Balding^{4,17}, Mark G Thomas⁴ and Márta Korbonits²

¹Johns Hopkins University School of Medicine, Baltimore, Maryland, USA, ²William Harvey Research Institute, Barts and the London School of Medicine, Queen Mary University of London, London, UK, ³Department of Endocrinology, C.I. Parhon National Institute of Endocrinology, 'C. Davila' University of Medicine and Pharmacy, Bucharest, Romania, ⁴Research Department of Genetics, Evolution and Environment, University College London, London, UK, ⁵Centre of Bioinformatics and System Biology, Department of Life Sciences, Imperial College London, London, UK, ⁶Department of Molecular Genetics, Royal Devon and Exeter Foundation Trust, Exeter, UK, ⁷Department of Clinical Genetics, Liverpool Women's Hospital, Liverpool, UK, ⁸Department of Endocrinology, Newcastle-upon-Tyne Hospitals & Institute of Genetic Medicine, Newcastle University, Newcastle, UK, ⁹Bristol Royal Hospital for Children, University Hospitals Bristol Foundation Trust, Bristol, UK, ¹⁰Endocrinology and Neuropathology Unit, Salford Royal Hospital, Manchester, UK, ¹¹The Walton Centre for Neurology and Neurosurgery, Liverpool, UK, ¹²Assistance Publique-Hôpitaux de Paris, Hôpital de Bicêtre, Service d'Endocrinologie et des Maladies de la Reproduction and Centre de Référence des Maladies Endocriniennes Rares de la Croissance, Le Kremlin-Bicêtre, France, ¹³Inserm 1185, Fac Med Paris Sud, Univ Paris-Sud, Université Paris-Saclay, Le Kremlin-Bicêtre, France, ¹⁴Neuropathology, University of Oxford, Oxford, UK, ¹⁵Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, UK, ¹⁶Department of Endocrinology, Christie Hospital, Manchester, UK, and ¹⁷Centre for Systems Genomics, Schools of Biosciences and of Mathematics & Statistics, University of Melbourne, Melbourne, Australia

*(R Salvatori, S Radian, Y Diekmann and D Iacovazzo contributed equally to this work)

Correspondence should be addressed to M Korbonits
Email
m.korbonits@qmul.ac.uk

Abstract

Objective: Mutations in the aryl hydrocarbon receptor-interacting protein (*AIP*) gene are associated with pituitary adenoma, acromegaly and gigantism. Identical alleles in unrelated pedigrees could be inherited from a common ancestor or result from recurrent mutation events.

Design and methods: Observational, inferential and experimental study, including: *AIP* mutation testing; reconstruction of 14 *AIP*-region (8.3 Mbp) haplotypes; coalescent-based approximate Bayesian estimation of the time to most recent common ancestor (tMRCA) of the derived allele; forward population simulations to estimate current number of allele carriers; proposal of mutation mechanism; protein structure predictions; co-immunoprecipitation and cycloheximide chase experiments.

Results: Nine European-origin, unrelated c.805_825dup-positive pedigrees (four familial, five sporadic from the UK, USA and France) included 16 affected (nine gigantism/four acromegaly/two non-functioning pituitary adenoma patients and one prospectively diagnosed acromegaly patient) and nine unaffected carriers. All pedigrees shared a 2.79 Mbp haploblock around *AIP* with additional haploblocks privately shared between subsets of the pedigrees, indicating the existence of an evolutionarily recent common ancestor, the 'English founder', with an estimated median tMRCA of 47 generations (corresponding to 1175 years) with a confidence interval (9–113 generations, equivalent to 225–2825 years). The mutation occurred in a small tandem repeat region predisposed to slipped

strand mispairing. The resulting seven amino-acid duplication disrupts interaction with HSP90 and leads to a marked reduction in protein stability.

Conclusions: The c.805_825dup allele, originating from a common ancestor, associates with a severe clinical phenotype and a high frequency of gigantism. The mutation is likely to be the result of slipped strand mispairing and affects protein–protein interactions and AIP protein stability.

*European Journal of
Endocrinology*
(2017) **177**, 257–266

Introduction

Although the majority of pituitary adenomas are sporadic, they co-occur in families in about 2–5% of the cases (1). Rarely, they form a part of multi-glandular tumor syndromes, such as MEN1 or Carney complex. More frequently, these families present with isolated pituitary adenomas (familial isolated pituitary adenoma (FIPA) syndrome). A significant proportion (~20%) of these pedigrees carry one of a variety of heterozygous alleles in the aryl hydrocarbon receptor-interacting protein gene (*AIP*, MIM: 60555) (2, 3). *AIP* is a 330 amino-acid protein comprised of an N-terminal immunophilin-like domain (4) and a C-terminal domain with three tetratricopeptide (TPR) repeats (5) that are thought to be responsible for protein–protein interactions (1, 6). Disease-associated *AIP* alleles reported to date are mainly the result of nonsense, missense, deletion/insertion or splice site mutations, but a few large deletions and a promoter mutation have also been described (1, 7). *AIP* acts as a tumor suppressor gene, requiring a second somatic hit affecting the wild-type (WT) allele, with loss of heterozygosity identified in a number of cases (6, 8, 9).

Mutations leading to a truncated protein are dispersed over the entire gene, while missense mutations are more common in the C-terminal end of the molecule (6). Several alleles have been reported in different geographic regions and populations, raising the question whether they arose through recurrent mutation events (occurring at mutational hotspots in the gene) or are inherited from a recent common ancestor. The c.910C>T, p.R304* allele, for example, has been identified in several populations independently (Irish, Romanian, English, Italian, Indian, Mexican) (10, 11), consistent with this CpG locus being a mutational hotspot, but in two instances, it has also been shown to give rise locally to numerous patients originating from the same ancestor (Ireland and Italy) (10, 12).

We have previously reported the c.805_825dup, p.F269_H275dup *AIP* exon 6 in-frame duplication

(rs267606578/EF643650) in three affected members of a family from the United Kingdom (3) and in an apparently sporadic giant from France (13). Here, we report seven additional pedigrees (four from the United Kingdom and three from the USA) with the same allele and provide evidence that they are all derived from a common ancestor. Using a coalescent-based approximate Bayesian computation approach, we estimate the time to most recent common ancestor (tMRCA), and by forward simulation, we estimate the current number of carriers. We hypothesize that slipped strand mispairing led to this unusual duplication and the inserted amino-acids render the protein extremely unstable and, at the same time, disrupt the binding site for crucial partner proteins.

Patients and methods

The study protocol was approved by the local ethics committees. All participants provided signed informed consent before the study. *AIP* genotyping (sequencing and multiplex ligation-dependent probe amplification) was performed as previously published (9, 14). We report nine c.805_825dup-positive pedigrees (Table 1, Fig. 1): two previously described, diagnosed in the United Kingdom (3) and France (13), and seven new ones, four diagnosed in the United Kingdom and three in the USA. All individuals have primarily European ancestry, with a known family link to the United Kingdom, but were not known to be related; they were living in France, different United Kingdom counties and two USA states (Maryland and Tennessee) (Fig. 1). Four of the nine pedigrees were FIPA families (Fig. 1), while five patients presented as simplex (apparently sporadic) cases.

We genotyped 14 short tandem repeat loci (microsatellites) around *AIP* (Fig. 2) in the index case of each pedigree, and in a second individual in each of the United Kingdom FIPA 1 and 2 pedigrees, as previously

Table 1 Clinical features of the patients carrying the *AIP* c.805_825dup, p.F269_H275dup mutation.

No	Pedigree ID	Pituitary adenoma phenotype	No. of affected individuals	Affected individuals details (sex, age at onset/diagnosis – years, tumor size, diagnosis)	Treatment for pituitary disease	No. of unaffected carriers	Unaffected carriers details (gender, age at evaluation – years)
1	UK FIPA 1 (3, 27)	GH	4	M, 11/15, Hyperplasia, Gigantism M, 15/26, Macro, Gigantism M, 27.5/29, Macro, Acromegaly M, na/47, Micro, Acromegaly*, GIST	TSS TSS, SSA, DA TCS TSS	5	F, 6; F, 6; F, 34; M, 34; F, 58
2	UK FIPA 2	GH	2	M, 22/23, Macro, Gigantism F, 24/28, nk, Acromegaly	TSS, XRT TSS	2	F, 32; M, 71
3	US FIPA 1	GH/NFPA	2	F, 22/23, Macro, NFPA (silent GH pos) M, 25/26, Macro, Acromegaly (apoplexy)	TSS, XRT TSS 2x, XRT, SSA, pegvisomant	1	F, na, 55 (obligate carrier)
4	US FIPA 2	GH/NFPA	3	M, 17/17, Macro, Gigantism F, 12/13, Macro, Gigantism F, 34/40, Macro, NFPA (immunostaining not available)	TSS TCS+XRT TSS	1	F, na, 60 (obligate carrier)
5	FranceSimplex (13)	GH	1	M, 9/14, Macro, Gigantism, psychosis	TSS+XRT+SSA+DA	–	–
6	UK Simplex 1	GH	1	M, 20/21, Macro, Gigantism	TSS 2x, SSA, pegvisomant	–	–
7	UK Simplex 2	GH	1	M, 13/18, Macro, Gigantism	TSS, XRT	–	–
8	US Simplex	GH	1	M, 17/17, Macro, Gigantism	TSS, SSA, XRT, pegvisomant	–	–
9	UK Simplex 3	GH	1	M, 18/23, Macro, Acromegaly	TCS, XRT	1	F, 66

Definition of gigantism: abnormally high growth speed in children or teenagers with abnormal IGF-1 and GH during OGTT and/or height >3 s.d. above the mean height for age or >2 s.d. over the calculated midparental height.

*Asymptomatic patient prospectively diagnosed with acromegaly following genetic testing, TSS after 3 years of medical therapy. Also has GIST tumor under surveillance.

DA, dopaminergic therapy; GIST, gastrointestinal stromal tumor; na, not applicable; nk, not known; SSA, somatostatin analogue therapy, TCS, transcranial surgery; TSS, trans-sphenoidal surgery; XRT, radiation therapy.

described (10). Haplotypes were inferred using PHASE (15), incorporating phasing information from UK FIPA 1 and 2 pedigrees. To obtain tMRCA estimates, we used a combined approach of coalescent theory calculations and simulation in an approximate Bayesian computation framework (16). We used *ms* (17) for the simulations, combining recombination and mutation rates in order to assess the probability of neither event in the region of interest, as previously described (11). Based on the tMRCA estimates, we employed a forward simulation approach (10, 11) to estimate the current number of carriers per generation. We reject simulations that generated <9 carriers, because we observed in our dataset at least 9 unrelated carriers.

A structural model for the mutated C-terminal domain of the AIP protein was generated by template-based modeling, using the Phyre2 prediction server (18) (top templates PDB: 4aif and 4gcg; Phyre2 confidence score 100.0% and 99.6% respectively). Phyre2 uses the alignment of Hidden Markov models via HHsearch (19)

to significantly improve the accuracy of alignment and detection rate. A match with a confidence score >90% indicates that the core of the protein is modeled at high accuracy (2–4 Å root mean square deviation from the native, true structure). However, surface loops can deviate from the native. Regions not covered by templates are modeled *ab initio* using the Poing algorithm embedded in Phyre2 (18).

The interaction between mutant AIP and HSP90, a well-established AIP-interacting partner (5), was investigated by co-immunoprecipitation. The coding sequence of AIP with the c.805_825dup, p.F269_H275dup mutation (20) was sub-cloned into the pcDNA3.0 vector C-terminally to the Myc tag epitope (3). Myc-tagged p.F269_H275dup AIP and HA-tagged HSP90β were co-transfected in HEK293 cells (10 × 10⁶, 20 μg of total plasmid DNA), using Lipofectamine 2000 (Life Technologies) following the manufacturer's instructions. Cells were harvested 48 h later by trypsinization and resuspended in 1.5 mL of lysis buffer (150 mM NaCl, 10 mM Tris-HCl pH 7.5, 10% v/v glycerol,

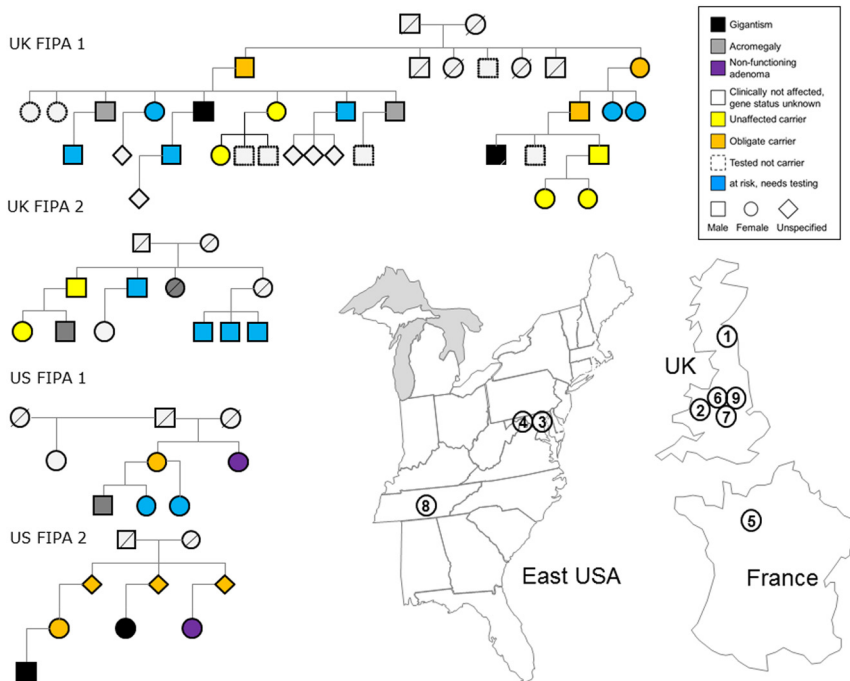


Figure 1

Pedigrees of the FIPA families harboring the c.805_825dup, p.F269_H275dup *AIP* mutation and map of geographical locations of the *AIP* c.805_825dup pedigrees. Pedigree numbers correspond to the first column in Table 1. A full colour version of this figure is available at <http://dx.doi.org/10.1530/EJE-17-0293>.

European Journal of Endocrinology

1% v/v IGEPAL CA-630 (Sigma-Aldrich) supplemented with Complete Protease Inhibitor Cocktail (Roche)). After being cleared by centrifugation, lysates were cleaned up by incubation with 50 µL of Protein G Sepharose 4 Fast Flow (GE Healthcare) and divided in thirds for incubation with 5 µg of anti-Myc (Santa Cruz Biotechnology; clone 9E10) or anti-HA (Sigma-Aldrich H3663) mouse monoclonal antibodies, or mouse IgG (Sigma-Aldrich I5381), as appropriate. The co-immunoprecipitation was performed

as previously described (21), and the eluates were resolved by denaturing polyacrylamide gel electrophoresis followed by Western blot using anti-Myc (Santa Cruz Biotechnology; clone 9E10) and anti-HA (Sigma-Aldrich H6908) antibodies. The experiment was repeated twice for confirmation.

Protein stability was assessed by cycloheximide chase experiments, as previously described (21). Briefly, Myc-tagged WT, p.F269_H275dup and p.R304* AIP

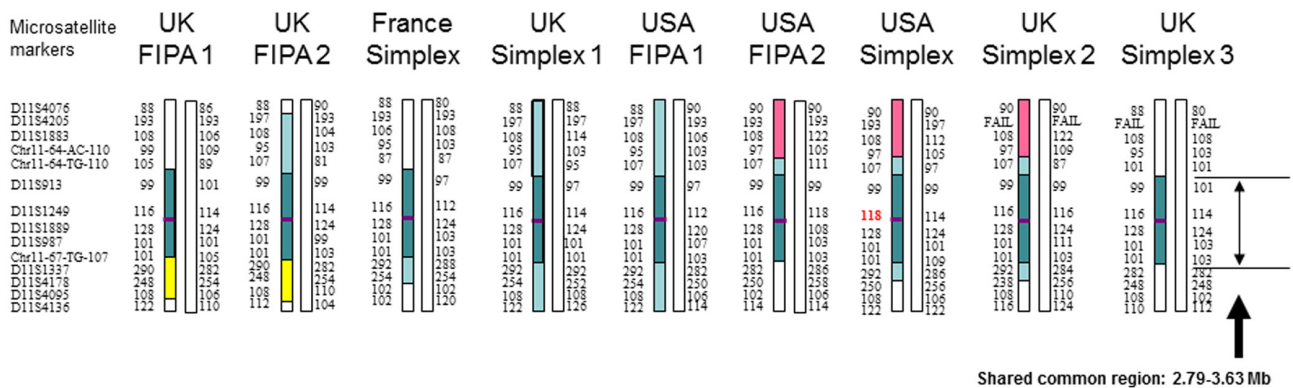


Figure 2

Haplotype analysis of the nine *AIP* c.805_825dup pedigrees. Colored chromosomes carry the *AIP* mutation (purple stripe): dark shading indicates the conserved haploblock shared between all mutated chromosomes; light shading shows portions of the haplotype extending to the entire genotyped region in UK Simplex 1/USA FIPA 1 pedigrees; yellow/pink shading shows additional 'private' sharing between closely related chromosomes. An allele mutation occurred at D11S1249 in patient US Simplex 2 (red type font). The order of individuals is arbitrary to highlight the extended haplotype sharing between pedigrees. Pedigree codes correspond to Table 1. A full colour version of this figure is available at <http://dx.doi.org/10.1530/EJE-17-0293>.

were overexpressed in HEK293 cells, and the cells were subsequently treated with cycloheximide 20 µg/mL (Abcam) for the time indicated. Cell lysates were prepared for polyacrylamide gel electrophoresis and Western blot with anti-Myc (Santa Cruz Biotechnology; clone 9E10) and anti-GAPDH (Santa Cruz Biotechnology FL-335) antibodies. The normalized protein levels were expressed as percentage of those observed at time 0. The results of at least two independent experiments, in triplicates, were pooled together and analyzed using a one-phase decay equation. The degradation constant (K) was compared between the mutants and the WT protein using the extra sum-of-squares *F* test. Significance was set for *P* values <0.05.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Results

The majority of the c.805_825dup-positive affected individuals presented with somatotropinoma (14 of 16 patients), but two clinically non-functioning pituitary adenomas (NFPA) were also observed (Table 1). Two FIPA families were homogenous, with all affected members harboring somatotropinomas and the other two were heterogeneous with both somatotropinomas and clinically NFPA (Fig. 1). From a total of 16 affected allele carriers (4F/12M), 13 patients (10M) presented with signs/symptoms of pituitary macroadenoma: eight gigantism (1F/7M), three acromegaly (3M), 2 clinically non-functioning pituitary adenomas (2F, one with 5–10% GH-positive cells while for the other immunostaining is not available) (Table 1). One asymptomatic male patient (UK FIPA 1), identified prospectively after genetic screening, had a GH-secreting pituitary microadenoma (IGF-1=1.5×upper limit of normal) and a 1 cm gastrointestinal stromal tumor; after three years of medical therapy, his pituitary adenoma was successfully operated on, while he remains under surveillance for the gastrointestinal stromal tumor. One patient, originally operated at the age of 15 years when histology showed somatotroph hyperplasia, was suffering from diabetes and hypertension and died following leg surgery at the

age of 35 years (UK FIPA 1). One subject (France Simplex) affected by gigantism had dilated cardiomyopathy, aortic dilatation and history of infantile psychosis.

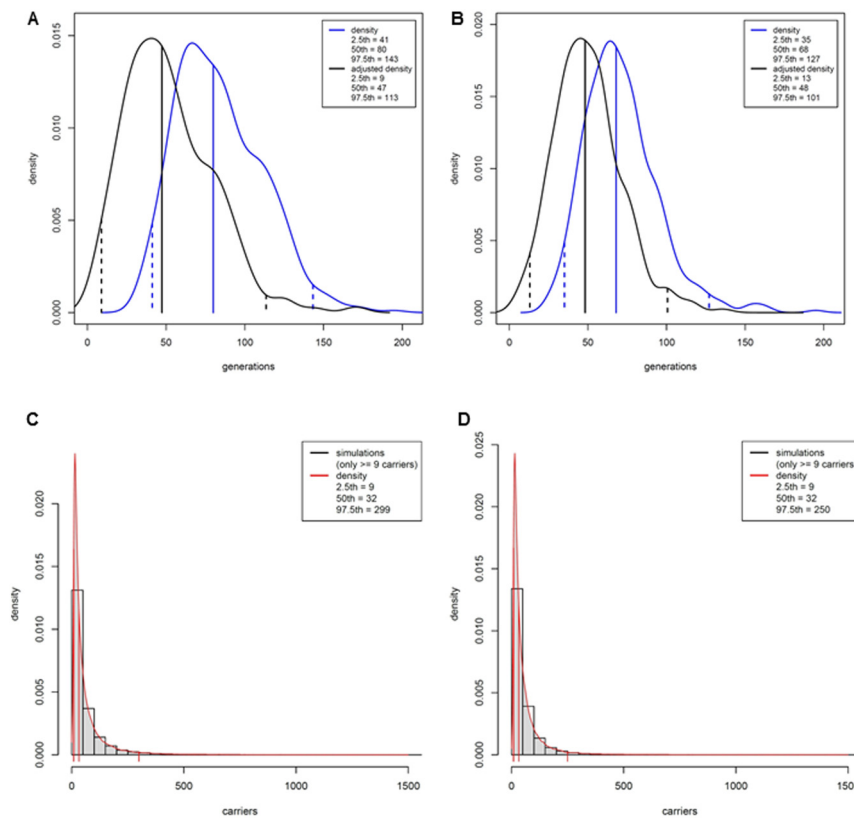
The overall average age of onset of symptoms was 19.1 years (median 18.5, range 9–34).

Haplotype analysis showed a 2.79–3.63 Mbp (1.83–2.59 cM) haploblock around *AIP* shared among c.805_825dup allele-carrying chromosomes in all nine pedigrees, as well as additional haploblocks shared among subsets of chromosomes, strongly supporting the existence of a recent common ancestor; the ‘English founder’ (Fig. 2). We observed a different allele of D11S1249 in one subject (US Simplex), on the c.805_825dup allele-carrying haplotype and conditioned our coalescent simulation results on the occurrence of a mutation at this locus.

We obtained tMRCA estimates of 47 (9–113) generations (median, 95% confidence interval (CI)), based on HapMap genetic distances (22) and 48 (13–101) generations, based on Rutgers map distances (Fig. 3A and B). Assuming a generation time of 25 years (23, 24), these estimates translate to 1175 (225–2825) years and 1200 (325–2525) years respectively.

The estimated current number of carriers per generation – obtained through forward simulations – was 32 (9–299) (median, 95% CI), based on HapMap genetic distances, and 32 (9–250), based on Rutgers map distances (Fig. 3C and D). These results integrate the uncertainty of the tMRCA estimates. Assuming that three 25-year generations co-exist at present, we estimate that 96 (27–897) (median, 95% CI, HapMap-based) c.805_825dup carriers may exist today, of which only 25 have been identified so far.

We studied the structure of the DNA sequence around the 21 base-pair tandem duplication allele. We have identified a region in the normal sequence with two short repeated sequences separated by an intervening sequence (Fig. 4A and B), which is typical for a site prone to ‘replication slippage’. Tandem duplications associated with short direct repeats can be explained by the model of ‘replication slippage’ or ‘slipped strand mispairing’ (25, 26). The key feature of this model is that, during replication, the primer strand transiently dissociates from the template, slips backward or forward, and then re-associates at the short direct repeat resulting in a misaligned configuration (Fig. 4A). If the primer strand containing the newly synthesized second (i.e. 3′) repeat dissociates from the template strand and then misaligns at the first (i.e. 5′) direct repeat, continued DNA synthesis will lead to the insertion of one of two direct repeats plus the intervening sequence. In our case, the repeat

**Figure 3**

Smoothed distribution of approximate Bayesian computation-simulated time to most recent common ancestor (tMRCA) estimates of the *AIP* c.805_825dup allele, unadjusted (blue curves) and adjusted (black curves) (11), based on HapMap (A) and Rutgers (B) genetic distances. Median, 2.5th and 97.5th percentiles are given in the insert. Distribution of numbers of carriers obtained through forward simulations, calculated based on HapMap-based (C) or Rutgers-based (D) time to most recent common ancestor (tMRCA) estimates randomly sampled from the adjusted distributions shown in A&B. Black bars show density histograms, red curves represent the smoothed distributions. Median, 2.5th and 97.5th percentiles are shown in the insert; the lower bound is conditioned to be nine carriers, the minimum observed number of carriers per generation in our cohort. A full colour version of this figure is available at <http://dx.doi.org/10.1530/EJE-17-0293>.

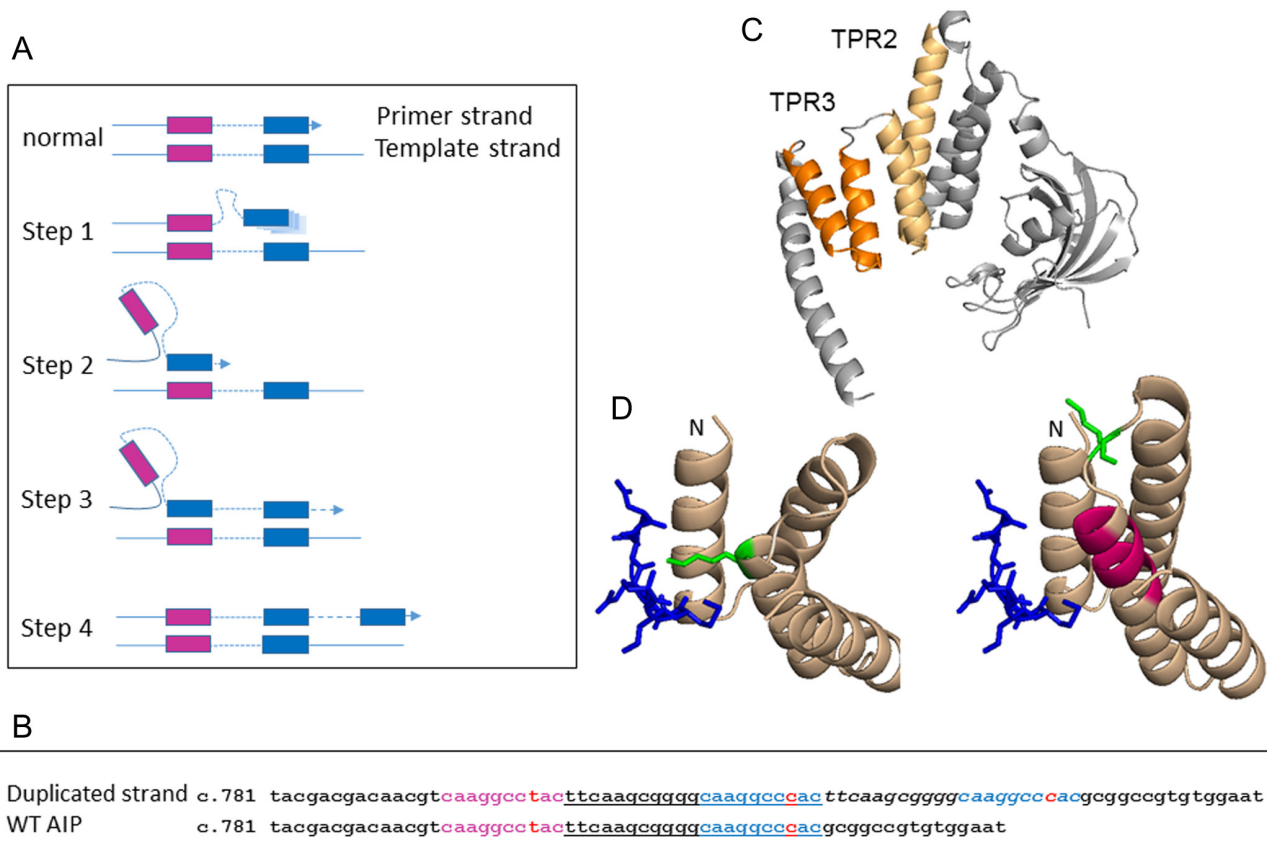
is imperfect (caagcctac and caaggccac) (Fig. 4B), the intervening sequence is ttcaagcgggg and the duplicated region is ttcaagcggggcaaggccac.

The mutation results in the duplication of seven amino-acids in the third TPR domain of AIP (TPR3). These can adopt a helical conformation, thus adding a 1.9 extra helical turn to the TPR3 alpha helix and increasing its length by 10.5 Å. The TPR3 domain is important for the interaction of AIP with other proteins, in particular, HSP90 and TOMM7 peptides (5). Superposition between the WT C-terminal of the AIP molecule (PDB: 4AIF) and the mutant C-terminal AIP model shows that the seven amino acid duplication may displace K266, which is crucial for HSP90 and possibly other partners binding (Fig. 4C, D and E), thus potentially leading to the disruption of AIP function. The results of a co-immunoprecipitation experiment with mutant AIP and HSP90 were consistent with lack of interaction between the two proteins (Fig. 5A), supporting the results of our *in silico* prediction. As compared with WT AIP, the steady state levels of the p.F269_H275dup protein were markedly reduced, and this led us to hypothesize an effect of the mutation on AIP protein stability as well, consistent with what has already been shown for

several other *AIP* mutations (21). Cycloheximide chase experiments showed that the p.F269_H275dup mutation leads to a markedly unstable protein with a degradation constant *K* of 1.169, significantly higher than that of WT AIP (0.0593; $P < 0.001$) and even that of the truncating mutant p.R304* (0.3147; $P < 0.05$) (Fig. 5B), suggesting shorter protein half-life.

Discussion

AIP mutations are reported in about 20% of FIPA pedigrees (1, 27), but a higher percentage has been observed in homogenous acromegaly families. Over 90 *AIP* mutations have been described to date. The R304 locus seems to be the most prevalent mutational 'hot spot', with mutations affecting the c.C910 and c.G911 base-pairs of this CpG site, resulting in R304* and R304Q alleles reported in various countries and unrelated families (28). A few other mutations have also been described as recurring in diverse geographical regions, such as the p.R271W and the p.R81* (3, 29, 30, 31). In this paper, we report nine apparently unrelated pedigrees all carrying the same unique c.805_825dup, p.Phe269_His275dup *AIP* allele.

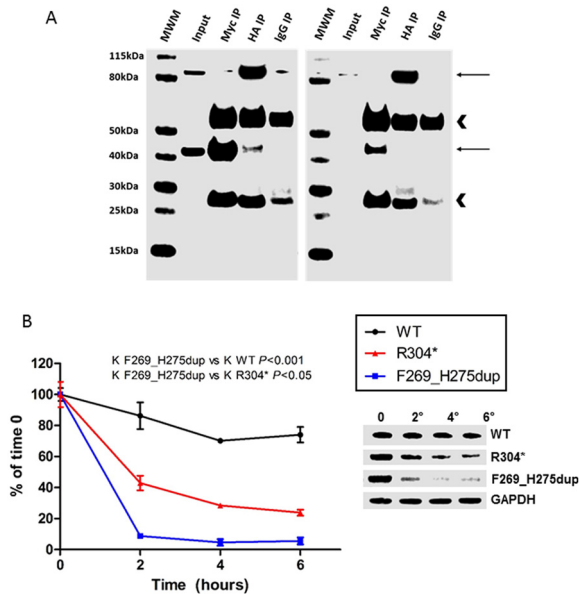
**Figure 4**

(A) Schematic drawing of the suggested replication slippage mechanism of the c.805_825dup, p.F269_H275dup AIP mutation. The purple and blue colorings match the sequences in (B). During DNA replication, the primer strand containing the newly synthesized second repeat transiently dissociates from the template, slips backward and then re-associates at the first short direct repeat of the template strand, resulting in a misaligned configuration. Continued DNA synthesis will lead to the insertion of intervening sequence and the second direct repeat, resulting in the mutated allele. (B) DNA sequences of WT *AIP* (bottom row) and of the mutated allele (top row), starting at the c.781 base-pair (NM_003977.2). Note the two 9 base-pair long repeats, the 5' one marked in purple (caaggcctac) and the 3' one marked in blue (caagcggcgctgtggaat). These are imperfect repeats as the 7th base-pair (in red font) is t in the 5' repeat and c in the 3' repeat. The intervening sequence between the two direct repeats is tc~~aaagcgggg~~caagqccc. The underlined sequence is duplicated in the upper sequence, i.e. the mutated allele, in which the first copy of the duplicated sequence is underlined and the second copy of the duplicated sequence is italic. This duplication leads to three copies of the 9 bp repeat: the purple is the 5' repeat, while the blue is the 3' repeat which has been duplicated. (C, D and E) Cartoon representation of the AIP protein structure based on the crystal structure of the N and the C-terminal domains of the protein (4, 5). The first and second alpha helices of the WT AIP TPR2 and TPR3 motifs (C) are presented in light orange and dark orange respectively. Details of the wild-type (D) and mutant (E) AIP bound to the SRMEEVD peptide, a fragment of the HSP90 partner protein (shown in blue). The seven amino-acid duplication in the first helix of AIP TPR3 is shown in purple. Residue K266, which in the wild-type AIP interacts with the HSP90 peptide (or with TOMM20 peptide) (5) is displayed as a green stick. A full colour version of this figure is available at <http://dx.doi.org/10.1530/EJE-17-0293>.

Seven of these pedigrees were not previously reported, and we provide updated information for the other two (3, 9, 13, 20).

When the same allele is identified in different families, the question arises whether this is the result of

independent recurrent mutational events in a region of the genome that is more prone to mutations ('hot spot', such as CpG sites) (32) or the result of spreading of the disease-associated allele in the population originating from a single founder mutation event. The c.805_825dup,

**Figure 5**

(A) The p.F269_H275dup AIP mutation disrupts the interaction with HSP90. Myc-tagged wild-type (WT) (left panel) or p.F269_H275dup (right panel) AIP were co-transfected with HA-tagged HSP90 β in HEK293 cells. Co-immunoprecipitation was performed using anti-Myc or anti-HA mouse antibodies or mouse IgG. Eluates were resolved by denaturing polyacrylamide gel electrophoresis followed by Western blot using anti-Myc and anti-HA antibodies. Positive interaction was demonstrated between WT AIP and HSP90 (a band is also visible in the negative IgG control for HSP90 but not for AIP). No interaction was seen between p.F269_H275dup AIP and HSP90. The steady state levels of mutant AIP were reduced as compared with WT AIP (a band was visible in the Input only after increasing the contrast settings, figure not shown), suggesting that the mutation affects the stability of AIP as well (Fig. 5B). Top arrow: HA-HSP90, bottom arrow: Myc-AIP. Arrowheads: heavy (top) and light (bottom) chains of mouse immunoglobulins. MWMM: molecular weight marker. IP: immunoprecipitation.

(B) Cycloheximide chase experiment in HEK293 cells overexpressing Myc-tagged WT, p.R304* and p.F269_H275dup AIP. The transfected cells were incubated in the presence of cycloheximide 20 μ g/mL for the time indicated. The left panel shows the degradation curves of each protein obtained after plotting the normalized protein levels as percentage of those observed at time 0. The right panel shows representative Western blot images. The degradation speed (K) of the p.F269_H275dup mutant protein was significantly higher than that of WT AIP and that of the truncating mutant p.R304*. A full colour version of this figure is available at <http://dx.doi.org/10.1530/EJE-17-0293>.

(p.Phe269_His275dup) involves the duplication of 21 base-pairs and probably occurred via the 'slipped strand mispairing' mechanism (25, 26, 33) (Fig. 4A). Certain structural features of the genome can predispose a particular region to rearrangement, such as repeated sequences located relatively close to each other, raising the possibility that similar duplications might occur independently. On the other hand, while the patients described here reside in three different areas of the world, all the individuals have ancestral links to the United Kingdom. Furthermore, the steady flow of immigration from Britain to North America in recent centuries and the close proximity of England and France are consistent with an evolutionarily recent common ancestor. Indeed, the result of our haplotype analysis strongly indicates that all patients share a recent common ancestor, the 'English founder', who lived approximately 1175 years ago. This estimate has, however, a wide 95% confidence interval of 225–2825 years. The cause of the geographic spreading of this allele is unknown, although mobility in the areas where carriers reside was historically high – particularly since the industrial revolution (34). The penetrance of the pituitary phenotype in patients with mutated AIP is known to be low (about 20%) (27). This is confirmed in the largest c.805_825dup-positive pedigree (UK FIPA 1), with an estimated penetrance of 23% (assuming that 50% of the relatives not genotyped would be carriers). Because of this low penetrance, AIP mutations do not prevent reproduction in the majority of carriers. Furthermore, even patients who develop pituitary adenomas may maintain normal fertility until the adenoma is large enough to cause hypogonadism. Whether the mutated allele confers some kind of survival advantage is unknown (11).

It is important to note that four of the patients did not have a family history of pituitary adenoma. Following our haplotype analysis, these are unlikely to be *de novo* mutations. These patients were tested for AIP mutations because of their clinical characteristics (large adenomas developing at young age, with onset of symptoms ranging between 9 and 20 years), confirming previous reports of relatively high prevalence of AIP mutations in apparently sporadic GH-secreting macroadenomas of the youth (27, 35, 36).

The c.805_825dup AIP allele is pathogenic, as indicated by the clinical, *in silico* and experimental data. The mutation is predicted to affect AIP structure and function by altering a crucial interaction site between the third TPR domain of AIP (Fig. 4) and one of its most important partners, HSP90 (5), and this has been

confirmed experimentally (Fig. 5A). Mutations at protein–protein binding sites are a common cause of disease as they can disrupt the interaction between molecules in cellular protein networks, which is required for correct protein function (37). An alternative pathogenic mechanism by which the c.805_825dup (p.F269_H275dup) affects AIP function is by causing the disruption of AIP protein folding, leading to increased proteasomal degradation, as we have previously shown for numerous AIP missense mutations (21). Here, we provided evidence that the p.F269_H275dup mutation markedly affects protein stability, leading to a protein that is significantly less stable compared with the common p.R304* truncating mutant.

In summary, we report the same AIP allele occurring in nine apparently unrelated pedigrees from three different (but linked through gene flow) countries and show that all the affected subjects, the majority suffering from young-onset somatotropinomas, share a common ancestral haplotype. The mutation results in the duplication of seven amino-acids in third TPR domain of AIP, leading to the disruption of protein–protein interactions and markedly reduced protein stability.

Declaration of interest

R S and M K serve on Pfizer Ltd's advisory board. The other authors declare no conflict of interest.

Funding

S R was supported by the EC-REA Marie Skłodowska-Curie IEF postdoctoral fellowship (grant 303006/2011). A D was supported by an MRC postdoctoral fellowship (grant MR/K021613/1). D I is supported by a Diabetes UK George Alberti Clinical Research Training Fellowship. M K's work on familial pituitary adenomas is supported by the Medical Research Council, Barts and the London Charity and Pfizer Ltd.

Acknowledgments

The authors thank Prof. James Lupski and Dr Claudia M B de Carvalho Fonseca, Baylor College of Medicine, Houston, TX, USA, for their advice regarding the slipped strand mispairing mechanism of mutation. They also acknowledge Dr Laura Cristina Hernández-Ramírez (currently at NIH, Bethesda) for her work on the previous optimization of the experimental techniques employed and Prof. Paul Chapple (QMUL, London) for helpful comments on the manuscript.

References

- Beckers A, Aaltonen LA, Daly AF & Karhu A. Familial isolated pituitary adenomas (FIPA) and the pituitary adenoma predisposition due to mutations in the aryl hydrocarbon receptor interacting protein (AIP) gene. *Endocrine Reviews* 2013 **34** 239–277. (doi:10.1210/er.2012-1013)
- Daly AF, Tichomirowa MA, Petrossians P, Heliövaara E, Jaffrain-Rea ML, Barlier A, Naves LA, Ebeling T, Karhu A, Raappana A *et al.* Clinical characteristics and therapeutic responses in patients with germ-line AIP mutations and pituitary adenomas: an international collaborative study. *Journal of Clinical Endocrinology and Metabolism* 2010 **95** E373–E383. (doi:10.1210/jc.2009-2556)
- Leontiou CA, Gueorguiev M, van der Spuy J, Quinton R, Lolli F, Hassan S, Chahal HS, Igreja SC, Jordan S, Rowe J *et al.* The role of the aryl hydrocarbon receptor-interacting protein gene in familial and sporadic pituitary adenomas. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 2390–2401. (doi:10.1210/jc.2007-2611)
- Linnert M, Haupt K, Lin YJ, Kissing S, Paschke AK, Fischer G, Weiward M & Lucke C. NMR assignments of the FKBP-type PPIase domain of the human aryl-hydrocarbon receptor-interacting protein (AIP). *Biomolecular NMR Assignments* 2012 **6** 209–212. (doi:10.1007/s12104-012-9359-0)
- Morgan RM, Hernandez-Ramirez LC, Trivellin G, Zhou L, Roe SM, Korbonits M & Prodromou C. Structure of the TPR domain of AIP: lack of client protein interaction with the C-terminal alpha-7 helix of the TPR domain of AIP is sufficient for pituitary adenoma predisposition. *PLoS ONE* 2012 **7** e53339. (doi:10.1371/journal.pone.0053339)
- Trivellin G & Korbonits M. AIP and its interacting partners. *Journal of Endocrinology* 2011 **210** 137–155. (doi:10.1530/JOE-11-0054)
- Chahal HS, Chapple JP, Frohman LA, Grossman AB & Korbonits M. Clinical genetic and molecular characterization of patients with familial isolated pituitary adenomas (FIPA). *Trends in Endocrinology and Metabolism* 2010 **21** 419–427. (doi:10.1016/j.tem.2010.02.007)
- Villa C, Lagonigro MS, Magri F, Koziak M, Jaffrain-Rea ML, Brauner R, Bouligand J, Junier MP, Di Rocco F, Sainte-Rose C *et al.* Hyperplasia-adenoma sequence in pituitary tumorigenesis related to aryl hydrocarbon receptor interacting protein gene mutation. *Endocrine-Related Cancer* 2011 **18** 347–356. (doi:10.1530/ERC-11-0059)
- Igreja S, Chahal HS, King P, Bolger GB, Srirangalingam U, Guasti L, Chapple JP, Trivellin G, Gueorguiev M, Guegan K *et al.* Characterization of aryl hydrocarbon receptor interacting protein (AIP) mutations in familial isolated pituitary adenoma families. *Human Mutation* 2010 **31** 950–960. (doi:10.1002/humu.21292)
- Chahal HS, Stals K, Unterlander M, Balding DJ, Thomas MG, Kumar AV, Besser GM, Atkinson AB, Morrison PJ, Howlett TA *et al.* AIP mutation in pituitary adenomas in the 18th century and today. *New England Journal of Medicine* 2011 **364** 43–50. (doi:10.1056/NEJMoa1008020)
- Radian S, Diekmann Y, Gabrovská P, Holland B, Bradley L, Wallace H, Stals K, Bussell AM, McGurren K, Cuesta M *et al.* Increased population risk of AIP-related acromegaly and gigantism in Ireland. *Human Mutation* 2017 **38** 78–85. (doi:10.1002/humu.23121)
- Occhi G, Jaffrain-Rea ML, Trivellin G, Albiger N, Ceccato F, De Menis E, Angelini M, Ferasin S, Beckers A, Mantero F *et al.* The R304X mutation of the aryl hydrocarbon receptor interacting protein gene in familial isolated pituitary adenomas: mutational hot-spot or founder effect? *Journal of Endocrinological Investigation* 2010 **33** 800–805. (doi:10.1007/BF03350345)
- Cazabat L, Bouligand J, Salenave S, Bernier M, Gaillard S, Parker F, Young J, Guiochon-Mantel A & Chanson P. Germline AIP mutations in apparently sporadic pituitary adenomas: prevalence in a prospective single-center cohort of 443 patients. *Journal of Clinical Endocrinology and Metabolism* 2012 **97** E663–E670. (doi:10.1210/jc.2011-2291)
- Vierimaa O, Georgitsi M, Lehtonen R, Vahteristo P, Kokko A, Raitila A, Tuppurainen K, Ebeling TM, Salmela PI, Paschke R *et al.* Pituitary adenoma predisposition caused by germline mutations in the AIP gene. *Science* 2006 **312** 1228–1230. (doi:10.1126/science.1126100)
- Stephens M, Smith NJ & Donnelly P. A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 2001 **68** 978–989. (doi:10.1086/319501)
- Beaumont MA, Zhang W & Balding DJ. Approximate Bayesian computation in population genetics. *Genetics* 2002 **162** 2025–2035.

- 17 Hudson RR. Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* 2002 **18** 337–338. (doi:10.1093/bioinformatics/18.2.337)
- 18 Kelley LA, Mezulis S, Yates CM, Wass MN & Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 2015 **10** 845–858. (doi:10.1038/nprot.2015.053)
- 19 Soding J. Protein homology detection by HMM-HMM comparison. *Bioinformatics* 2005 **21** 951–960. (doi:10.1093/bioinformatics/bti125)
- 20 Lecoq AL, Bouligand J, Hage M, Cazabat L, Salenave S, Linglart A, Young J, Guiochon-Mantel A, Chanson P & Kamenicky P. Very low frequency of germline GPR101 genetic variation and no biallelic defects with AIP in a large cohort of patients with sporadic pituitary adenomas. *European Journal of Endocrinology* 2016 **174** 523–530. (doi:10.1530/EJE-15-1044)
- 21 Hernandez-Ramirez LC, Martucci F, Morgan RM, Trivellin G, Tilley D, Ramos-Guajardo N, Iacovazzo D, D'Acquisto F, Prodromou C & Korbonits M. Rapid proteasomal degradation of mutant proteins is the primary mechanism leading to tumorigenesis in patients with missense AIP mutations. *Journal of Clinical Endocrinology and Metabolism* 2016 **101** 3144–3154. (doi:10.1210/jc.2016-1307)
- 22 Frazer KA, Eskin E, Kang HM, Bogue MA, Hinds DA, Beilharz EJ, Gupta RV, Montgomery J, Morenzoni MM, Nilsen GB *et al.* A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* 2007 **448** 1050–1053. (doi:10.1038/nature06067)
- 23 Thomas MG, Stumpf MP & Harke H. Evidence for an apartheid-like social structure in early Anglo-Saxon England. *Proceedings of Biological Sciences* 2006 **273** 2651–2657. (doi:10.1098/rspb.2006.3627)
- 24 Fenner JN. Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *American Journal of Physical Anthropology* 2005 **128** 415–423. (doi:10.1002/ajpa.20188)
- 25 Levinson G & Gutman GA. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution* 1987 **4** 203–221. (doi:10.1093/oxfordjournals.molbev.a040442)
- 26 Lupski JR & Stankiewicz P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genetics* 2005 **1** e49. (doi:10.1371/journal.pgen.0010049)
- 27 Hernandez-Ramirez LC, Gabrovská P, Denes J, Stals K, Trivellin G, Tilley D, Ferrau F, Evanson J, Ellard S, Grossman AB *et al.* Landscape of familial isolated and young-onset pituitary adenomas: prospective diagnosis in AIP mutation carriers. *Journal of Clinical Endocrinology and Metabolism* 2015 **100** E1242–E1254. (doi:10.1210/jc.2015-1869)
- 28 Stals K, Trivellin G & Korbonits M. AIP mutation in pituitary adenomas. *New England Journal of Medicine* 2011 **364** 1973–1974. (doi:10.1056/NEJMc1101859)
- 29 Daly AF, Vanbellinghen JF, Khoo SK, Jaffrain-Rea ML, Naves LA, Guitelman MA, Murat A, Emy P, Gimenez-Roqueplo AP, Tamburrano G *et al.* Aryl hydrocarbon receptor-interacting protein gene mutations in familial isolated pituitary adenomas: analysis in 73 families. *Journal of Clinical Endocrinology and Metabolism* 2007 **92** 1891–1896. (doi:10.1210/jc.2006-2513)
- 30 Jennings JE, Georgitsi M, Holdaway I, Daly AF, Tichomirowa M, Beckers A, Aaltonen LA, Karhu A & Cameron FJ. Aggressive pituitary adenomas occurring in young patients in a large Polynesian kindred with a germline R271W mutation in the AIP gene. *European Journal of Endocrinology* 2009 **161** 799–804. (doi:10.1530/EJE-09-0406)
- 31 Toledo RA, Lourenco DM Jr, Liberman B, Cunha-Neto MB, Cavalcanti MG, Moyses CB, Toledo SP & Dahia PL. Germline mutation in the aryl hydrocarbon receptor interacting protein gene in familial somatotropinoma. *Journal of Clinical Endocrinology and Metabolism* 2007 **92** 1934–1937. (doi:10.1210/jc.2006-2394)
- 32 Holliday R & Grigg GW. DNA methylation and mutation. *Mutation Research* 1993 **285** 61–67. (doi:10.1016/0027-5107(93)90052-H)
- 33 Thomas MG, Cook CE, Miller KW, Waring MJ & Hagemberg E. Molecular instability in the COII-tRNA(Lys) intergenic region of the human mitochondrial genome: multiple origins of the 9-bp deletion and heteroplasmy for expanded repeats. *Philosophical Transactions of The Royal Society B Biological Sciences* 1998 **353** 955–965. (doi:10.1098/rstb.1998.0260)
- 34 Weale ME, Weiss DA, Jager RF, Bradman N & Thomas MG. Y chromosome evidence for Anglo-Saxon mass migration. *Molecular Biology and Evolution* 2002 **19** 1008–1021. (doi:10.1093/oxfordjournals.molbev.a004160)
- 35 Tichomirowa MA, Barlier A, Daly AF, Jaffrain-Rea ML, Ronchi C, Yaneva M, Urban JD, Petrossians P, Elenkova A, Tabarin A *et al.* High prevalence of AIP gene mutations following focused screening in young patients with sporadic pituitary macroadenomas. *European Journal of Endocrinology* 2011 **165** 509–515. (doi:10.1530/EJE-11-0304)
- 36 Cuny T, Pertuit M, Sahnoun-Fathallah M, Daly AF, Occhi G, Odou MF, Tabarin A, Nunes ML, Delemer B, Rohmer V *et al.* Genetic analysis in young patients with sporadic pituitary macroadenomas: Beside AIP don't forget MEN1 genetic analysis. *European Journal of Endocrinology* 2013 **168** 533–541. (doi:10.1530/EJE-12-0763)
- 37 David A, Razali R, Wass MN & Sternberg MJ. Protein-protein interaction sites are hot spots for disease-associated nonsynonymous SNPs. *Human Mutation* 2012 **33** 359–363. (doi:10.1002/humu.21656)

Received 13 April 2017

Revised version received 29 May 2017

Accepted 19 June 2017