An isolated case of lissencephaly caused by the insertion of a mitochondrial genome-derived DNA sequence into the 5′ untranslated region of the *PAFAH1B1* (LIS1) gene

David S. Millar,[1] Carolyn Tysoe,[2] Lazarus P. Lazarou,[3] Daniela T. Pilz,[3] Shehla Mohammed,[4] Katharine Anderson,[5] Nadia Chuzhanova,[6] David N. Cooper[1] and Rachel Butler[3]*

[1]Institute of Medical Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, UK
[2]Royal Devon and Exeter NHS Foundation Trust, Barrack Road, Exeter EX2 5AD, UK
[3]Molecular Genetics, Institute of Medical Genetics, University Hospital of Wales, Heath Park, Cardiff CF14 4XW, UK
[4]Clinical Genetics Service, Guy's Hospital, London SE1 9RT, UK
[5]Department of Community Paediatrics, Seaside View Child Development Centre, Brighton General Hospital, Elm Grove, Brighton BN2 3EW, UK
[6]School of Science and Technology, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK
Correspondence to: Tel: +44 (0)2920 744062; Fax: +44 (0)2920 746551; E-mail: CooperDN@cardiff.ac.uk

Date received (in revised form): 27th May 2010

Abstract

A 130 base pair (bp) insertion (g.-8delCins130) into the 5′ untranslated region of the *PAFAH1B1* (LIS1) gene, seven nucleotides upstream of the translational initiation site, was detected in an isolated case of lissencephaly. The inserted DNA sequence exhibited perfect homology to two non-contiguous regions of the mitochondrial genome (8479 to 8545 and 8775 to 8835, containing portions of two genes, *ATP8* and *ATP6*), as well as near-perfect homology (1 bp mismatch) to a nuclear mitochondrial pseudogene (NUMT) sequence located on chromosome 1p36. This lesion was not evident on polymerase chain reaction (PCR) sequence analysis of either parent, indicating that the mutation had occurred *de novo* in the patient. Experiments designed to distinguish between a mitochondrial and a nuclear genomic origin for the inserted DNA sequence were, however, inconclusive. Mitochondrial genome sequences from both the patient and his parents were sequenced and found to be identical to the sequence inserted into the *PAFAH1B1* gene. Analysis of parental PCR products from the chromosome 1-specific NUMT were also consistent with the interpretation that the inserted sequence had originated directly from the mitochondrial genome. The chromosome 1-specific NUMT in the patient proved to be refractory to PCR analysis, however, suggesting that this region of chromosome 1 could have been deleted or rearranged. Although it remains by far the most likely scenario, in the absence of DNA sequence information from the patient's own chromosome 1-specific NUMT, we cannot unequivocally confirm that the 130 bp insertion originated from mitochondrial genome rather than from the NUMT.

Keywords: *lissencephaly, PAFAH1B1 gene, mitochondrial genome, insertion*

Introduction

Classical (type 1) lissencephaly is a neuronal migration disorder characterised by agyria (absent cerebral convolutions) and pachygyria (reduced, broad cerebral convolutions), a thickened cortex (grey matter), mental retardation and epileptic seizures. It can occur either as part of the contiguous deletion disorder, Miller–Dieker syndrome, or as an isolated condition termed isolated lissencephaly sequence (ILS).[1] Most cases of ILS are caused by defects in either the *PAFAH1B1* (LIS1) or the *DCX* genes. Patients with a *PAFAH1B1* gene alteration have a more severe cerebral phenotype posteriorly, whereas those with a *DCX* gene defect have a more severe cerebral phenotype anteriorly. The heterozygous deletion of a further gene, *YWHAE* (located 1 megabase [Mb] from *PAFAH1B1* on 17p13.3 and encoding the 14-3-3ε protein), together with *PAFAH1B1* in patients with Miller–Dieker syndrome, is known to increase the severity, usually with generalised agyria.[2]

The *PAFAH1B1* gene (MIM# 601545) was the first gene to be implicated in the pathogenesis of lissencephaly and encodes the non-catalytic α-subunit of the intracellular 1b isoform of platelet-activating factor acetylhydrolase.[3] It spans ~92 kilobases (kb) of genomic DNA and contains 11 exons, the first two of which contribute to the 5′ untranslated region (5′ UTR). Although the deletion of the entire *PAFAH1B1* gene is the most common mutation encountered in lissencephaly patients, a considerable number of intragenic lesions have now been reported.[4–8] In general, however, it appears that neither the type nor the position of known intragenic mutations in the *PAFAH1B1* gene are indicative of the likely clinical severity of the condition.[9] Here, we report a highly unusual lissencephaly-causing mutation in the *PAFAH1B1* gene, which involves the insertion of mitochondrial genome-derived DNA sequence into the 5′ UTR of the gene.

Methods

PCR and sequencing

Exons 2 to 11 of the human *PAFAH1B1* gene were polymerase chain reaction (PCR) amplified (primer details in Table S1) and the resulting products were gel-purified (Gel Extraction Kit, Qiagen, Crawley, UK). Purified PCR products were then sequenced using BigDye v3.1 (Applied Biosystems, Warrington, UK) and analysed on an ABI 3100 genetic analyser (Applied Biosystems, Warrington, UK). PCR primers specific for either the chromosome 1 nuclear mitochondrial pseudogene (NUMT) sequence (accession number NT_004350.19) or the mitochondrial genome sequence (accession number NC_001807.4) were as follows: chromosome 1-specific primers — CHR15 (5′ TGTGGAGCAAACCAGTTTTATG 3′, 47366 to 47387) and CHR13 (5′ CTGAGTAGCTACAG TACATATC 3′, 49106 to 49127); mitochondrial genome-specific primers — MITO5 (5′ GTGGAGC AAACCACAGTTTC 3′, 8185 to 8204) and MITO3 (5′ GGATGAAGCAGATAGTGAGG 3′, 9849 to 9869). PCR was performed using the Expand™ high fidelity system (Roche, East Sussex, UK) using a hot start of 98°C for three minutes followed by 95°C for two minutes, 35 cycles at 95°C for 45 seconds, 60°C for 30 seconds and 72°C for one minute. For the last 25 cycles, the elongation step at 72°C was increased by five seconds per cycle, followed by a further incubation at 72°C for ten minutes. PCR products were then sequenced with primers common to both the chromosome 1 sequence and the mitochondrial genome as follows: CHRMITF (5′ AACCAACACCTCTTTACAGTG 3′; Chr1 47524 to 47544 and mitochondrial genome 8344 to 8364) and CHRMITR (5′ TGAGTAGGCTGATGGTTTCG 3′; Chr1 48139 to 48158 and mitochondrial genome 8959 to 8978).

Chromosome 1-specific regions around the identified NUMT were PCR amplified with different combinations of the primers listed in Table S2. All PCRs were performed using the Expand™ high-fidelity system under the following conditions: 98°C for three minutes, followed by 95°C for two minutes, 35 cycles of 95°C for 45 seconds, 60°C for 30 seconds and 68°C for four minutes. For the last 25 cycles, the elongation step at 68°C was increased by five seconds per cycle. This was followed by a further incubation at 68°C for ten minutes. To ensure that the PCR products were

derived from chromosome 1, partial sequencing of each product obtained was performed with the appropriate primers using BigDye v3.1 and analysed on an ABI 3100 genetic analyser.

Bioinformatics analysis

To elucidate the mechanism of the *PAFAH1B1* insertion, a BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome) of the inserted fragment was performed against the Human Genomic Sequences and Transcripts Database (http://www.ncbi.nlm.nih.gov/Tools). The MegaBLAST option, designed to identify highly homologous sequences, was used. Two sequences were found to yield a significant alignment – namely, a sequence from the mitochondrial genome (100 per cent homology; ref NC_001807.4) and a NUMT sequence located on chromosome 1 (98 per cent homology; ref NT_004350.19).

Sequences flanking the inserted fragments in the mitochondrial genome, as well as the sequences flanking the insertion site in the *PAFAH1B1* gene, were screened for the presence of direct repeats, inverted repeats and symmetric elements (both within each sequence and between them) by means of complexity analysis.[10]

RNA secondary structure was analysed using RNAfold from the Vienna RNA package (http://rna.tbi.univie.ac.at)[11] using default parameters. Identification of putative splice sites was performed using the Berkeley *Drosophila* Genome Project package (http://www.fruitfly.org/seq_tools/splice.html)[12] to analyse the wild-type and mutant (insertion-containing) exon 2 sequences of the *PAFAH1B1* gene flanked by 50 base pairs (bp) of intronic DNA sequence. The minimum score for both acceptor and donor splice sites was set at 0.8.

Results

Patient details

The patient was born, after an uneventful pregnancy and delivery, weighing 3.4 kg. He was the younger of two siblings born to non-consanguineous Caucasian parents. His very early motor milestones were acquired age appropriately, and his overall development was normal until eight months of age, at which time he had a febrile fit and lost some of his acquired skills. When referred at the age of 15 months, his development was at an approximate age level of six to eight months. Seizures developed at four years of age, but these were well controlled with sodium valproate and lamotrigine. They were a combination of tonic/clonic, 'star' seizures (tonic extension) and brief atonic attacks. By this time, global developmental delay was evident. He was first seen for a genetics review at the age of ten years. His gait was unsteady, with a tendency to fall frequently or bump into objects. He was able to walk short distances but longer distances required the use of a wheelchair. Although he was able to walk upstairs holding onto a rail, he descended stairs on his bottom. He spoke a few solitary words and could obey simple commands. He was not dysmorphic, but had a degree of brachycephaly. His head circumference was on the 75th centile, weight on the 98th centile and height on the 25th centile. He had relatively small genitalia, but no neurocutaneous stigmata.

The patient had a normal 46XY karyotype, with no evidence of a microdeletion at the Miller–Dieker locus on 17p13.3. His EEG at the age of six years was abnormal, with frequent showers of sharp waves or spikes throughout the recording. This was more obvious in the anterior brain regions, with some emphasis on the left side. A magnetic resonance imaging (MRI) brain scan at eight years of age revealed frontal, parietal, posterior temporal and occipital pachygyria, with maximal cortical thickening posteriorly. The degree of pachygyria was also milder anteriorly. Appearances were consistent with classical lissencephaly at the milder end of the spectrum, and a posterior to anterior severity consistent with the presence of a *PAFAH1B1* (LIS1) gene mutation. In view of the highly unusual *PAFAH1B1* gene insertion found in this patient, genomic DNA from this individual was also screened for *DCX* gene mutations, but none were found.

Microsatellite marker analysis was used to confirm that familial relationships were as stated at referral (data not shown).

Sequence analysis of the *PAFAH1B1* gene in the patient

PCR amplification and sequencing of the coding region of the *PAFAH1B1* gene (exons 2 to 11) of the proband's DNA revealed a heterozygous 130 bp insertion within exon 2. This insertion was located within the 5′ UTR, seven bp upstream of the translational initiation site (ATG) (between bases 560 and 562 of the reference sequence [accession number NM_000430]; Figure 1). The 130 bp insertion contained two non-templated bases (TT) at its 5′ end and was accompanied by the deletion of a cytosine at position 561 of the *PAFAH1B1* gene (accession number NM_000430). The mutation may therefore be described as g.-8delCins130. No other sequence alterations in any other exon or splice site of the *PAFAH1B1* gene were identified. Sequence analysis of exon 2 of the *PAFAH1B1* gene in both parents indicated only the presence of the wild-type sequence, consistent with the *de novo* occurrence of the mutation in the patient.

Origin of the inserted DNA sequence with homology to the mitochondrial genome

A BLAST search was performed to determine the origin of the inserted sequence. Perfect homology to the mitochondrial genome sequence (8479 to 8545 and 8775 to 8835; accession number NC_001807.4) was noted, as well as near-perfect homology to a NUMT sequence[13] at chromosome 1p36 (47659 to 47727 and 47955 to 48015; accession number NT_004350.19). The homology between the 130 bp inserted sequence and the mitochondrial genome/NUMT was not, however, contiguous; rather, two regions of mitochondrial DNA sequence homology (of length 67 bp and 61 bp, respectively) were noted, which were located 229 bp distant from each other in the mitochondrial genome/NUMT. The 67 bp and 61 bp sequences were both identical to the mitochondrial genome reference sequence, whereas, in the case of the sequence of the telomerically located chromosome 1-specific NUMT, the 67 bp fragment contained one mismatch (Figure 2).

Inspection of the sequence flanking the junction between the 67 bp and 61 bp fragments identified two short imperfect direct repeats, GAAGC and GGAGG, in the mitochondrial genome (8546 to 8550 and 8776 to 8780, respectively; accession number NC_001807.4) which could have mediated the loss of the 229 bp fragment through slipped mispairing. It should be noted that the equivalent sequences in the chromosome 1 NUMT are GAAGT and GGAGG (47728 to 47732 and 47956 to 47960, respectively; accession number NT_004350.19).

The inserted sequence contains portions of two mitochondrial genes, *ATP8* (8367 to 8573; accession number NC_001807.4) and *ATP6* (8528 to

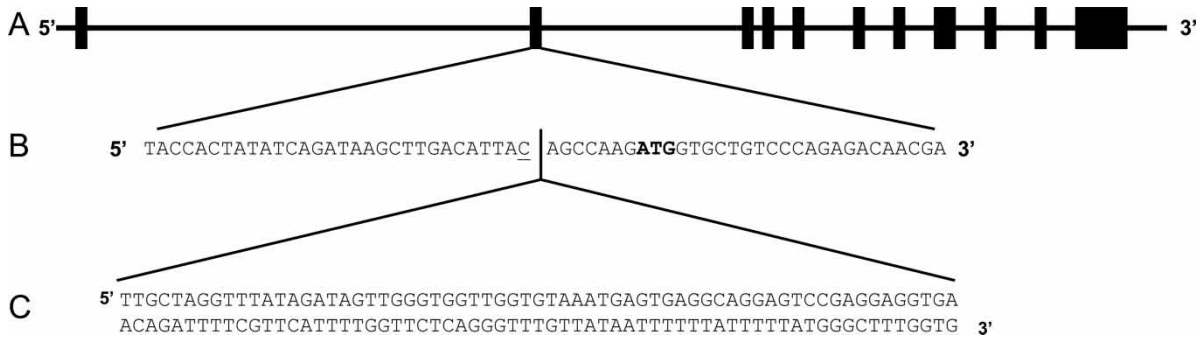

Figure 1. A. Diagram of the *PAFAH1B1* gene on chromosome 17p13.3. B. Wild-type DNA sequence of exon 2 of the *PAFAH1B1* gene flanking the insertion site. The translational initiation site (ATG) is shown in bold. The cytosine residue that was deleted during the insertion is underlined. C. The 130 bp inserted sequence homologous to the mitochondrial genome.

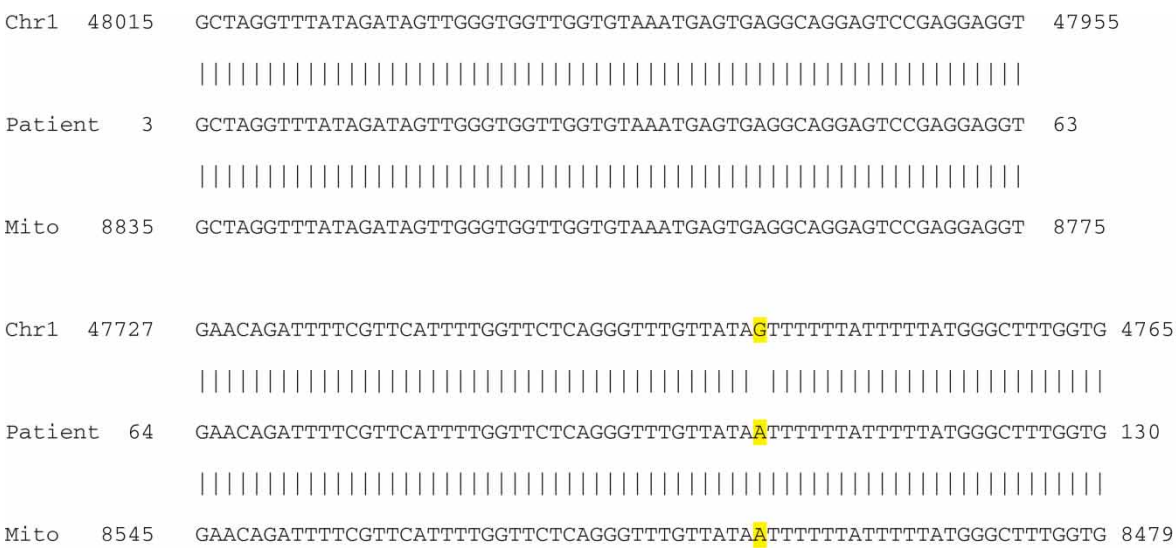

Figure 2. Sequence homology between the 130 bp sequence inserted into exon 2 of the patient's *PAFAH1B1* gene, the mitochondrial genome (8479 to 8835) and the mitochondrial pseudogene (NUMT) sequence located at chromosome 1p36 (48015 to 47659). The mismatch between the inserted sequence and the chromosome 1 NUMT is highlighted in yellow.

9208; accession number NC_001807.4), but whether the inserted sequence was derived from the mitochondrial genome itself or from a reverse transcript of mRNA encoding the *ATP6* and *ATP8* mitochondrial genes cannot be ascertained from the DNA sequence involved.

Since the mitochondrial genome and chromosome 1-specific NUMT sequences differed from each other by only 1 bp over the 130 bp length of the insert, it was considered important to establish whether the patient's own mitochondrial genome and chromosome 1-specific NUMT sequences were identical to their respective published reference sequences. Oligonucleotide primers were therefore designed specifically to PCR amplify DNA fragments corresponding to the inserted sequence from either the mitochondrial genome or chromosome 1. PCR products from the mitochondrial genome of both the patient and his parents were sequenced and found to be identical to (but not, of course, contiguous with) the sequence inserted into the *PAFAH1B1* gene. Similarly, PCR/direct sequencing of the chromosome 1-specific NUMT from both the patient's parents confirmed sequence identity with the standard chromosome 1 reference sequence (ie 1 bp mismatch with respect to the *PAFAH1B1* gene insertion). Attempts to PCR amplify the chromosome 1-specific NUMT sequence from the patient repeatedly failed to yield any PCR product, however. To confirm that the nuclear DNA from the patient was of good quality, PCR amplification of a 3.2 kb fragment containing the *GH1* gene[14] was performed. Successful PCR amplification of this fragment from patient DNA (data not shown) indicated that the lack of PCR amplification of the chromosome 1-specific NUMT was not due to poor DNA quality. The reason why the patient's chromosome 1-specific NUMT was refractory to analysis remains unclear but is potentially interesting, given the possible involvement of this sequence in the *PAFAH1B1* gene insertion. Although it remains the most likely scenario, in the absence of DNA sequence information from the patient's own chromosome 1-specific NUMT, we cannot unequivocally confirm that the 130 bp insertion originated from the mitochondrial genome sequence rather than from the NUMT.

In order to ascertain whether the patient possessed a deletion of chromosome 1p36 encompassing the chromosome 1-specific NUMT, an attempt was made to PCR amplify across the NUMT

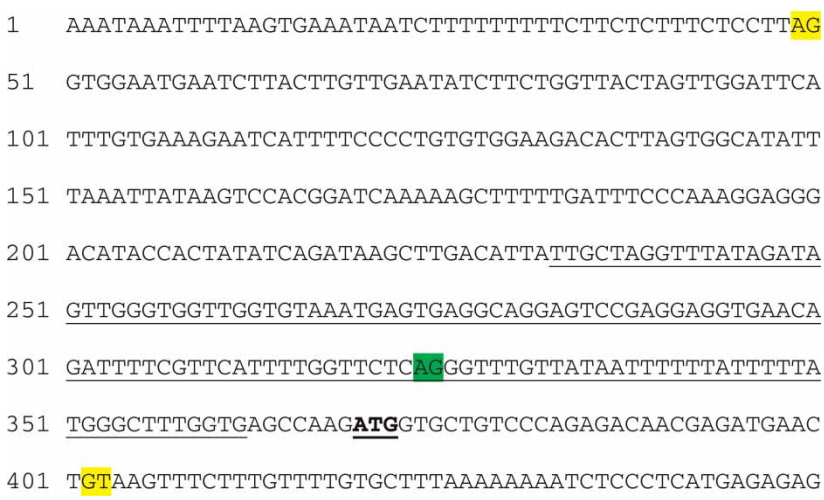

Figure 3. Exon 2 of the *PAFAH1B1* gene in the patient indicating the 130 bp insertion (underlined), the wild-type translational initiation site (ATG; in bold and underlined), the positions of the wild-type exon 2 acceptor and donor splice sites (highlighted in yellow) and the putative novel acceptor splice site (highlighted in green).

sequence. Combinations of different chromosome 1-specific primers (Table S2) were employed to amplify different DNA sequences between 43,267 and 53,538 (accession Number NT_004350.19) in both the patient and his parents. PCR products of the appropriate sizes were amplified from both parents and, on sequencing, were confirmed to match the chromosome 1-specific NUMT sequence. No PCR product from the chromosome 1-specific NUMT was obtained from the patient for any combination of primers used, however, suggesting that this region of chromosome 1 may have been homozygously deleted, rearranged or both.

Analysis of the inserted sequence and the site of insertion in the *PAFAH1B1* gene

The 130 bp sequence inserted into the 5′ UTR of the *PAFAH1B1* gene contains two out-of-frame ATGs that could, at least in principle, serve as alternative translational initiation codons. The sequences flanking these ATGs GTGTAA**ATG**A (positions 263 to 272; Figure 3) and ATTTTT**ATG**G (positions 344 to 353; Figure 3) — do not match the Kozak consensus sequence (GCC(A/G)NN**ATG**G),[15] however, unlike the wild-type sequence (GCCAAG**ATG**G) in the *PAFAH1B1* gene. This suggests that neither of these sites is likely to be able to play a role in translational initiation. The insertion occurs at nucleotide position −7 relative to the ATG, immediately adjacent to the 5′ end of the Kozak consensus sequence.

Since the identified insertion lies within the 5′ UTR of the *PAFAH1B1* gene, it has the potential to have an impact on RNA secondary structure. Using RNAfold, the optimal secondary structure minimum free energy was determined to be −249.0 kcal/mol for the wild-type 5′ UTR, whereas for the mutant 5′ UTR (containing the insertion), this value was −299.6 kcal/mol. This suggests that the stability of the 5′ UTR may not have been dramatically altered by the insertion; however, the predicted secondary structure of the mutant 5′ UTR molecule was clearly very different from that of the wild-type 5′ UTR (Figure S1).

Differences in pre-mRNA structure resulting from single bp substitutions have been reported to result in aberrant splicing.[16] Prediction of splice sites in the wild-type *PAFAH1B1* sequence using NNSPLICE software[12] attributed the experimentally validated exon 2 acceptor (position 50; Figure 3) and donor (position 401; Figure 3) splice sites with potential splice site scores of 0.97 and 1.00, respectively. When the analysis was repeated for the mutant *PAFAH1B1* sequence, an additional

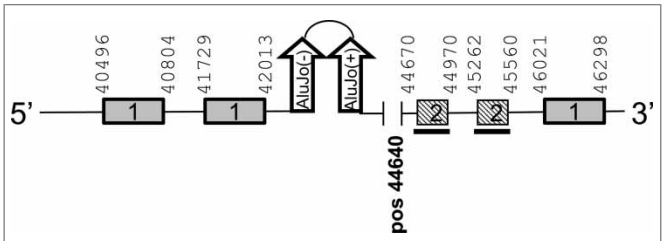

Figure 4. Repetitive sequence elements found in the vicinity of the *PAFAH1B1* insertion breakpoint at position 44640 indicated by vertical bars. Homologous regions are shown by differently shaded rectangles. Inverted repeats comprising two *Alujo* repeats are indicated by arrows.

acceptor splice site was predicted (position 325, score 0.91; Figure 3) in addition to the wild-type splice sites. Without *in vitro* splicing analysis, however, it remains unclear whether this additional acceptor splice site would be functionally significant.

Mechanism of mutagenesis

No sequence homology was found between the site of insertion in the *PAFAH1B1* gene and the mitochondrial genome that could have explained how the mitochondrial DNA fragment became integrated at this position. Two highly homologous regions, marked 1 and 2 in Figure 4, were, however, identified in the vicinity of the *PAFAH1B1* gene. These regions could have led to a double-strand break through non-B slipped structure formation.

Discussion

Sequence analysis of the coding region of the lissencephaly patient's *PAFAH1B1* gene revealed a heterozygous 130 bp insertion within exon 2. Since this lesion could not be detected by sequence analysis of either parent, it would appear that the mutation occurred *de novo* in the patient. The inserted sequence was found to be perfectly homologous to two non-contiguous (separated by 229 bp) sequences in the mitochondrial genome reference sequence, of length 67 bp and 61 bp, respectively. We propose that the intervening 229 bp fragment may have been lost during the process of insertion into the *PAFAH1B1* gene. This postulate is supported by the presence of two short imperfect direct repeats (GGAGG and GAAGC), flanking the junction between the 67 bp and 61 bp fragments, which could have mediated the loss of the 229 bp fragment through slipped mispairing.

The inserted sequence contains portions of two mitochondrial genes, *ATP8* and *ATP6*, but the sequence itself is too short to be able to determine unequivocally whether it was derived from the mitochondrial genome or from a reverse transcript of mRNA derived from the two mitochondrial genes. On the basis that (i) there is no correlation between NUMT frequency and the abundance of their cognate transcripts[17] and (ii) there is no preference for the integration of NUMTs from transcribed regions as opposed to non-transcribed regions,[18,19] we therefore surmise that the composite 130 bp insertion is likely to have been of mitochondrial genome origin.

Near-perfect homology (one bp mismatch) was, however, also found between the 130 bp insertion in the *PAFAH1B1* gene and a NUMT sequence at chromosome 1p36, raising the possibility that the insert could conceivably have been of nuclear origin. Experiments designed to distinguish between a mitochondrial and a nuclear genomic origin were, however, inconclusive. Analysis of parental PCR products from both the mitochondrial genome and the chromosome 1-specific NUMT was consistent with the interpretation that the insertion had originated directly from the mitochondrial genome (although a heterozygous chromosome 1-specific NUMT deletion would not have been detected). This concurs with the finding that most newly arising NUMTs derive from the independent insertion of mitochondrial genome fragments, rather than from the duplicational transposition of pre-existing NUMTs.[20] No PCR product from the chromosome 1-specific NUMT could be obtained from the patient, however, suggesting that this region of chromosome 1 may have been specifically deleted, rearranged or both. If confirmed, the concomitant occurrence of a chromosome 1 rearrangement could be suggestive of the involvement of this locus in the insertion into the *PAFAH1B1* gene in our

patient; however, it might also be sheer coincidence.

Although we can offer no formal confirmation that the 130 bp insertion either alters or abolishes the expression of the *PAFAH1B1* gene, we believe that this is very likely to be the case on account of the site of insertion within the 5′ UTR only seven nucleotides upstream of the translational initiation site. 5′ UTRs contain a variety of *cis*-regulatory elements which influence translation and hence, when disrupted by mutation, can cause inherited disease.[21,22] We have presented some indirect evidence to support the view that this large insertion may have affected the translation of the *PAFAH1B1* mRNA by either destabilising the secondary structure of the 5′ UTR or by altering the splicing phenotype. Gross insertions into 5′ UTRs are extremely unusual as a cause of inherited disease; a comparable example of a 75 bp insertion (of non-mitochondrial origin) has, however, been reported in the 5′ UTR of the *RFXAP* gene of a patient with major histocompatibility complex (MHC) class II deficiency, which served to silence the gene.[23] This example certainly supports the contention that such an insertion within the 5′ UTR can have a profound effect on the expression of the downstream gene. In view of the comparatively mild clinical phenotype exhibited by the lissencephaly patient reported here, however, we speculate that some residual *PAFAH1B1* gene expression may have served to ameliorate the severity of the condition in this particular case.

No sequence homology was found between the mitochondrial DNA insert and the site of insertion in the *PAFAH1B1* gene. Taken together with the identification of two non-templated bases (TT) at the 5′ end of the inserted sequence and the deletion of a cytosine 5′ to the breakpoint in the *PAFAH1B1* gene, this lesion would therefore appear to be compatible with a mutational model that combines features of non-homologous end-joining with 'NUMT-mediated double-strand break repair'.[13,24] We speculate that some of the repetitive sequence elements that we identified in the vicinity of the *PAFAH1B1* gene could have given rise to a double-strand break which was then subsequently repaired by integration of the mitochondrial DNA fragment.

Mitochondrial DNA/NUMT insertions are not infrequent in the human nuclear genome; indeed, Hazkani-Covo *et al.*[13] identified a total of 871 NUMTs in a standard human genome, equivalent to 0.0087 per cent of the entire genome sequence. Intriguingly, NUMTs may have a propensity to integrate within human gene regions, as opposed to intergenic regions,[25] and some of these NUMTs are polymorphic in terms of their presence/absence in the genome.[26] More rarely, the integration of a NUMT into a human gene can give rise to an inherited disease. Probably, the best characterised example of this type of pathological mutation is the 72 bp insertion into exon 14 of the *GLI3* gene in a sporadic case of Pallister–Hall syndrome.[27] Several other examples of intragenic NUMT insertions causing inherited disease have also been reported, however. Thus, the insertion of a 93 bp mitochondrial DNA fragment into the *MCOLN1* gene is responsible for an inherited case of mucolipidosis type IV.[28] Additional examples of mitochondrial DNA insertions have been reported in the *USH1C* (36 bp, Usher syndrome)[29] and *F7* (251 bp, factor VII deficiency)[30] genes. In the context of the mutation reported here, however, the mitochondrial DNA insertion polymorphism into intron I of the human *FOXO1A* gene is perhaps the most intriguing, since this 39 bp insertion is derived from the DNA sequence between nucleotides 8531 and 8569 of the mitochondrial genome (accession number NC_001807.1) containing the *ATP8* and *ATP6* genes.[31] The sequence inserted into the *FOXO1A* gene therefore overlaps with the 130 bp *PAFAH1B1* gene insert reported here by 14 bases (8532 to 8545; accession number NC_001807.4). It remains to be seen if this particular region of the human mitochondrial genome exhibits an increased propensity for insertion into the nuclear genome.

Acknowledgments

The authors would like to thank Dr Elaine Hughes (consultant paediatric neurologist) and Dr Sian Bennett (consultant community paediatrician) for their contributions in the clinical investigation of this patient.

References

1. Wynshaw-Boris, A. (2007), 'Lissencephaly and LIS1: Insights into the molecular mechanisms of neuronal migration and development', *Clin. Genet*. Vol. 72, pp. 296–304.
2. Cardoso, C., Leventer, R.J., Ward, H.L., Toyo-Oka, K. *et al.* (2003), 'Refinement of a 400-kb critical region allows genotypic differentiation between isolated lissencephaly, Miller-Dieker syndrome, and other phenotypes secondary to deletions of 17p13.3', *Am. J. Hum. Genet*. Vol. 72, pp. 918–930.
3. Lo Nigro, C., Chong, C.S., Smith, A.C., Dobyns, W.B. *et al.* (1997), 'Point mutations and an intragenic deletion in LIS1, the lissencephaly causative gene in isolated lissencephaly sequence and Miller-Dieker syndrome', *Hum. Mol. Genet*. Vol. 6, pp. 157–164.
4. Cardoso, C., Leventer, R.J., Dowling, J.J., Ward, H.L. *et al.* (2002), 'Clinical and molecular basis of classical lissencephaly: Mutations in the *LIS1* gene (*PAFAH1B1*)', *Hum. Mutat*. Vol. 19, pp. 4–15.
5. Saillour, Y., Carion, N., Quelin, C., Leger, P.L. *et al.* (2009), '*LIS1*-related isolated lissencephaly: Spectrum of mutations and relationships with malformation severity', *Arch. Neurol.* Vol. 66, pp. 1007–1015.
6. Mei, D., Lewis, R., Parrini, E., Lazarou, L.P. *et al.* (2008), 'High frequency of genome deletions and a duplication in the *L1S1* gene in lissencephaly: Implications for molecular diagnosis', *J. Med. Genet*. Vol. 45, pp. 355–361.
7. Human Gene Mutation Database. Available at http://www.hgmd.org
8. Stenson, P.D., Mort, M., Ball, E.V., Howells, K. *et al.* (2009), 'The Human Gene Mutation Database: 2008 update', *Genome Med*. Vol. 1, p. 13.
9. Uyanik, G., Morris-Rosendahl, D.J., Stiegler, J., Klapecki, J. *et al.* (2007), 'Location and type of mutation in the LIS1 gene do not predict phenotypic severity', *Neurology* Vol. 69, pp. 442–447.
10. Gusev, V.D., Nemytikova, L.A. and Chuzhanova, N.A. (1999), 'On the complexity measures of genetic sequences', *Bioinformatics* Vol. 15, pp. 994–999.
11. Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R. *et al.* (2008), 'The Vienna RNA Websuite', *Nucleic Acids Res.* Vol. 36, pp. W70–W74.
12. Reese, M.G., Eeckman, F.H., Kulp, D. and Haussler, D. (1997), 'Improved splice site detection in Genie', *J. Comp. Biol*. Vol. 4, pp. 311–323.
13. Hazkani-Covo, E., Zeller, R.M. and Martin, W. (2010), 'Molecular poltergeists: Mitochondrial DNA copies (*numts*) in sequenced nuclear genomes', *PLoS Genet*. Vol. 6, p. e1000834.
14. Millar, D.S., Lewis, M.D., Horan, M., Newsway, V. *et al.* (2003), 'Novel mutations of the growth hormone 1 (*GH1*) gene disclosed by modulations of the clinical selection criteria for individuals with short stature', *Hum. Mutat*. Vol. 21, pp. 424–440.
15. Kozak, M. (2002), 'Pushing the limits of the scanning mechanism for initiation of translation', *Gene* Vol. 299, pp. 1–34.
16. Buratti, E. and Baralle, F.E. (2004), 'Influence of RNA secondary structure on the pre-mRNA splicing process', *Mol. Cell. Biol*. Vol. 24, pp. 10505–10514.
17. Woischnik, M. and Moraes, C.T. (2002), 'Pattern of organization of human mitochondrial pseudogenes in the nuclear genome', *Genome Res.* Vol. 12, pp. 885–893.
18. Henze, K. and Martin, W. (2001), 'How do mitochondrial genes get into the nucleus?', *Trends Genet*. Vol. 17, pp. 383–387.
19. Mourier, T., Hansen, A.J., Willerslev, E. and Arctander, P. (2001), 'The Human Genome Project reveals a continuous transfer of large mitochondrial fragments to the nucleus', *Mol. Biol. Evol.* Vol. 18, pp. 1833–1837.
20. Bensasson, D., Feldman, M.W. and Petrov, D.A. (2003), 'Rates of DNA duplication and mitochondrial DNA insertion in the human genome', *J. Mol. Evol.* Vol. 57, pp. 343–354.
21. Pickering, B.M. and Willis, A.E. (2005), 'The implications of structured 5′ untranslated regions on translation and disease', *Semin. Cell Devel. Biol*. Vol. 16, pp. 39–47.
22. Chatterjee, S. and Pal, J.K. (2009), 'Role of 5′- and 3′-untranslated regions of mRNAs in human diseases', *Biol. Cell* Vol. 101, pp. 251–262.
23. van Eggermond, M.C., Tezcan, I., Heemskerk, M.H. and van den Elsen, P.J. (2008), 'Transcriptional silencing of *RFXAP* in MHC class II-deficiency', *Mol. Immunol*. Vol. 45, pp. 2920–2928.
24. Hazkani-Covo, E. and Covo, S. (2008), 'Numt-mediated double-strand break repair mitigates deletions during primate genome evolution', *PLoS Genet*. Vol. 4, p. e1000237.
25. Ricchetti, M., Tekaiam, F. and Dujon, B. (2004), 'Continued colonization of the human genome by mitochondrial DNA', *PLoS Biol*. Vol. 2, p. E273.
26. Yuan, J.D., Shi, J.X., Meng, G.X., An, L.G. *et al.* (1999), 'Nuclear pseudogenes of mitochondrial DNA as a variable part of the human genome', *Cell Res.* Vol. 9, pp. 281–290.
27. Turner, C., Killoran, C., Thomas, N.S., Rosenberg, M. *et al.* (2003), 'Human genetic disease caused by *de novo* mitochondrial-nuclear DNA transfer', *Hum. Genet*. Vol. 112, pp. 303–309.
28. Goldin, E., Stahl, S., Cooney, A.M., Kaneski, C.R. *et al.* (2004), 'Transfer of a mitochondrial DNA fragment to *MCOLN1* causes an inherited case of mucolipidosis IV', *Hum. Mutat*. Vol. 24, pp. 460–465.
29. Chen, J.M., Chuzhanova, N., Stenson, P.D., Férec, C. *et al.* (2005), 'Meta-analysis of gross insertions causing human genetic disease: novel mutational mechanisms and the role of replication slippage', *Hum. Mutat*. Vol. 25, pp. 207–221.
30. Borensztajn, K., Chafa, O., Alhenc-Gelas, M., Salha, S. *et al.* (2002), 'Characterization of two novel splice site mutations in human factor VII gene causing severe plasma factor VII deficiency and bleeding diathesis', *Br. J. Haematol*. Vol. 117, pp. 168–171.
31. Giampieri, C., Centurelli, M., Bonafè, M., Olivieri, F. *et al.* (2004), 'A novel mitochondrial DNA-like sequence insertion polymorphism in intron I of the *FOXO1A* gene', *Gene* Vol. 327, pp. 215–219.

Table S1. Oligonucleotide primers used for the PCR amplification and sequencing of exons 2–11 of the human *PAFAH1B1* gene.

Name	Sequence (5′ to 3′)	Position[a]
LIS1EX2F	TGTGGAAGACACTTAGTGGCA	44547 – 44567
LIS1EX2R	AAGAGACCTCCCAAAGCTGTA	44760 – 44780
LIS1EX3F	AAGAGTATCTTCAGGGTTAATG	71646 – 71667
LIS1EX3R	TTGTGCGTAACTGTTAACTACA	71890 – 71911
LIS1EX4F	TCTTGAGGATCATAGTTAAGCC	72334 – 72355
LIS1EX4R	TGCAGAAGAATGTTATTTTCAG	72516 – 72537
LIS1EX5F	GAAATCTATCTGTACGTAACTAC	73327 – 73349
LIS1EX5R	ATCTCGGCTCACTGCAAACT	73641 – 73660
LIS1EX6F	AAGGAGTGATGGAGTTGGTG	76489 – 76508
LIS1EX6R	GGGACACTGTACACTGTTAG	76771 – 76790
LIS1EX7F	AACCCCATGGTAAAATCCCAT	78939 – 78959
LIS1EX7R	GGCTGGTCTTCAATTCCTGA	79229 – 79248
LIS1EX8F	TTCTGGGAAGTGTCCTGATG	80374 – 80393
LIS1EX8R	CAGATATCAGCAATAAAACCATG	80684 – 80706
LIS1EX9F	GTCCATACCTAACTTTCTTGTG	82824 – 82844
LIS1EX9R	CATAAAGCATTAATCCCAAAAGG	83028 – 83050
LIS1EX10F	GATGCTATTTAAACATTTTGCC	86481 – 86502
LIS1EX10R	TTTGTCTGGCACTCCAAAATC	86726 – 86746
LIS1EX11F	GGTCTCACTATGTTTGTTGTCCA	88035 – 88057
LIS1EX11R	GGTATCATCAGAGTGCATCCAG	88210 – 88232

[a]Position based on accession number NG_009799.

Table S2. Oligonucleotide primers designed to PCR amplify the chromosome 1-specific NUMT.

Name	Sequence (5′ to 3′)	Position[a]
CHR115	TCACCAAGACCCTACTTCTGA	43267 – 43287
CHR125	TCATCGCCCTTACCACACTG	44637 – 44656
CHR135	TCTCTCCCAGTCCTAGCCG	45677 – 45695
CHR145	CTGCTTCCTAGTCCTGTACG	46870 – 46889
CHR113	CGTGAGGGCCAAGATCTGG	50894 – 50912
CHR123	TCAGTCTCAGTGTCTGTTGCA	51602 – 51622
CHR133	TGGATGCTGACAACACTGTAG	52627 – 52647
CHR143	CTATGAGTCAGAAAGCCAGAC	53518 – 53538

[a]Position based on accession number NT_004350.19.

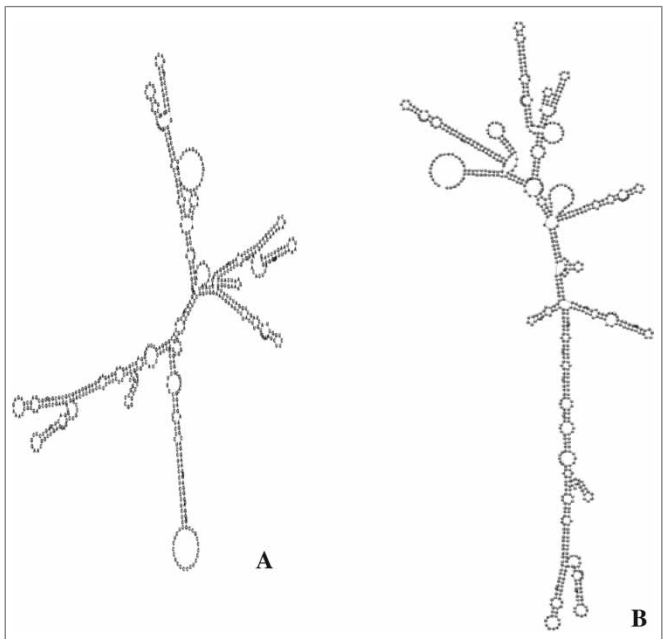

Figure S1. RNA secondary structure prediction using RNAfold. (A) Wild-type 5′ UTR corresponding to exons 1 and 2 of the *PAFAH1B1* gene. (B) 5′ UTR containing the 130 bp mitochondrial DNA insertion.