Exome report

Identical by descent *L1CAM* mutation in two apparently unrelated families with intellectual disability without L1 syndrome

Marie Shaw [a,b,1], Tzu Ying Yap [c,1], Lyndal Henden [d,e], Melanie Bahlo [d,e,f], Alison Gardner [a,b], Vera M. Kalscheuer [g], Eric Haan [b,h], Louise Christie [i], Anna Hackett [i], Jozef Gecz [a,b,*]

[a] Robinson Research Institute, The University of Adelaide, Adelaide, SA 5000, Australia
[b] School of Paediatrics and Reproductive Health, The University of Adelaide, Adelaide, SA 5000, Australia
[c] Discipline of Genetics, School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, SA 5000, Australia
[d] Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
[e] Department of Medical Biology, University of Melbourne, Melbourne, Australia
[f] Department of Mathematics and Statistics, University of Melbourne, Melbourne, Australia
[g] Max Planck Institute for Molecular Genetics, Department Human Molecular Genetics, Ihnestrasse 73, Berlin, D-14195, Germany
[h] SA Pathology, Women's and Children's Hospital, Adelaide, SA 5006, Australia
[i] GOLD NSW, Hunter Genetics, Newcastle, Australia

ARTICLE INFO

ABSTRACT

Mutations in the *L1 Cell Adhesion Molecule* (*L1CAM*) gene (MIM#308840) cause a variety of X-linked recessive neurological disorders collectively called L1 syndrome. Using massively parallel sequencing (MPS) of the X-chromosome exome, we identified a novel missense variant in *L1CAM* in two Caucasian families with mild-moderate intellectual disability without obvious L1 syndrome features. These families were not known to be related. SNP data extracted from MPS identified a 5.6 cM tract of identity by descent (IBD), encompassing the *L1CAM* gene, between the DNA of the two probands. This cannot be explained by chance alone and strongly implies that the two families are related. It also suggests that the *L1CAM* (NM_000425.3, c.604G > A, p.D202N) variant is pathogenic. This report also demonstrates the usefulness of additional information, which can be extracted from exome sequencing data.

© 2015 Elsevier Masson SAS. All rights reserved.

1. Introduction

L1CAM is located in Xq28 and comprises 28 coding exons with a translated protein of 1257 amino acids (NC_000023.10). Mutations in *L1CAM* cause a variety of X-linked recessive neurological disorders, collectively called L1 syndrome, comprising four main phenotypes; X-linked hydrocephalus (MIM#307000), corpus callosum agenesis (MIM#304100), spastic paraplegia type 1 (MIM#303350), and MASA syndrome (MIM#303350). To date, more than 213 disease-causing *L1CAM* mutations have been identified (*L1CAM* mutation database: http://www.l1cammutationdatabase.info/default.aspx) [Vos and Hofstra, 2010].

L1CAM, encoding the L1 cell-to-cell adhesion molecule, acts in a number of ways, with roles in neurite outgrowth during development, neuronal cell migration, axon bundling, synaptogenesis,

myelination, neuronal cell survival, and long-term potentiation [Kenwrick et al., 2000]. *L1CAM* contains six immunoglobulin-like domains, five fibronectin type 3 repeats, as well as trans-membrane and cytoplasmic domains [Moos et al., 1988]. Mutations in key residues/domains across *L1CAM* influence the L1 phenotype, with missense mutations in the extracellular domain causing a less severe phenotype and mutations occurring in the cytoplasmic domain having the least severe effects. In addition, truncating mutations are well known to produce a more severe phenotype. [Bateman et al., 1996; Weller and Gärtner, 2001].

In this study we identified a novel *L1CAM* variant in two apparently unrelated families with mild-moderate intellectual disability without obvious L1 syndrome features. The variant affects a conserved residue in the second immunoglobulin domain and is predicted to be damaging to the L1CAM protein. Using X-chromosome exome data we were able to haplotype the probands from both families, and using a mathematical model, inferred a region of identity by descent (IBD) encompassing *L1CAM* (LOD = 0.77). We discuss the usefulness of extracting additional information from MPS data originally generated for disease variant identification.

* Corresponding author.
  E-mail address: jozef.gecz@adelaide.edu.au (J. Gecz).
[1] Equal first author.

2. Clinical description

2.1. Family 1 (Supplementary Figure 1)

The uncle (II-2) was first seen in the genetics clinic aged 38 years and again at 39 years. The pregnancy and delivery was uneventful at 41 weeks gestation. He was hypotonic in infancy and had difficulties with breast feeding. His milestones were delayed with standing at 21 months, walking at 23 months and speech at 18 months. As a child he preferred his own company and had some socialization difficulties with anxiety. He attended a special class from year 5 in primary school but found mainstream high school difficult. He then attended life skill programs and is now on a disability pension and lives with his mother. As an adult he is a loner but is not shy and is conversant in subjects in which he has an interest. He has an excellent memory and can recall statistics of AFL football and politics. He becomes anxious with changes to routines and in noisy, crowded environments. Recently (since the age of 45 years) his anxiety has escalated and he has had panic episodes in response to challenging life events. He has hypertension and haemochromatosis. On examination his head circumference was 58 cm (>98th percentile), height was 171 cm (25th percentile) and weight was 110 kg (>97th percentile, BMI 37). He has a mild-moderate intellectual disability and is non-dysmorphic with normal hands and thumbs (Fig. 1a, Supplementary Figure 1). His neurological examination was normal. There is no history of seizures and EEGs have been normal. A CT head scan performed at age 10 years, showed mild lateral ventricle dilation and mild generalised cerebral atrophy. Repeat CT head scans at 39 years and again at 45 years showed only dysgenesis of the corpus callosum affecting the body and splenium. X-inactivation studies in his mother were normal.

The nephew (III-1) was initially seen aged 13 years, and then again aged 17 years. The pregnancy was complicated by gestational diabetes and pre-eclampsia and was induced at 38 weeks gestation. His birth weight was 2860 gm (25th percentile), head circumference was 33 cm (10th percentile) and length was 48 cm (50th percentile). Apgars were 9 at 1 min and 9 at 5 min. He had no hypotonia and breast fed well. His milestones were only slightly delayed with sitting at 4 months, crawling at 8.5 months and walking at 18.5 months. At 6 months the baby health clinic initiated a referral to a paediatrician due to concerns about increasing head circumference. An MRI showed partial agenesis of the corpus callosum. At age 4 years he had a nephrectomy due to a small
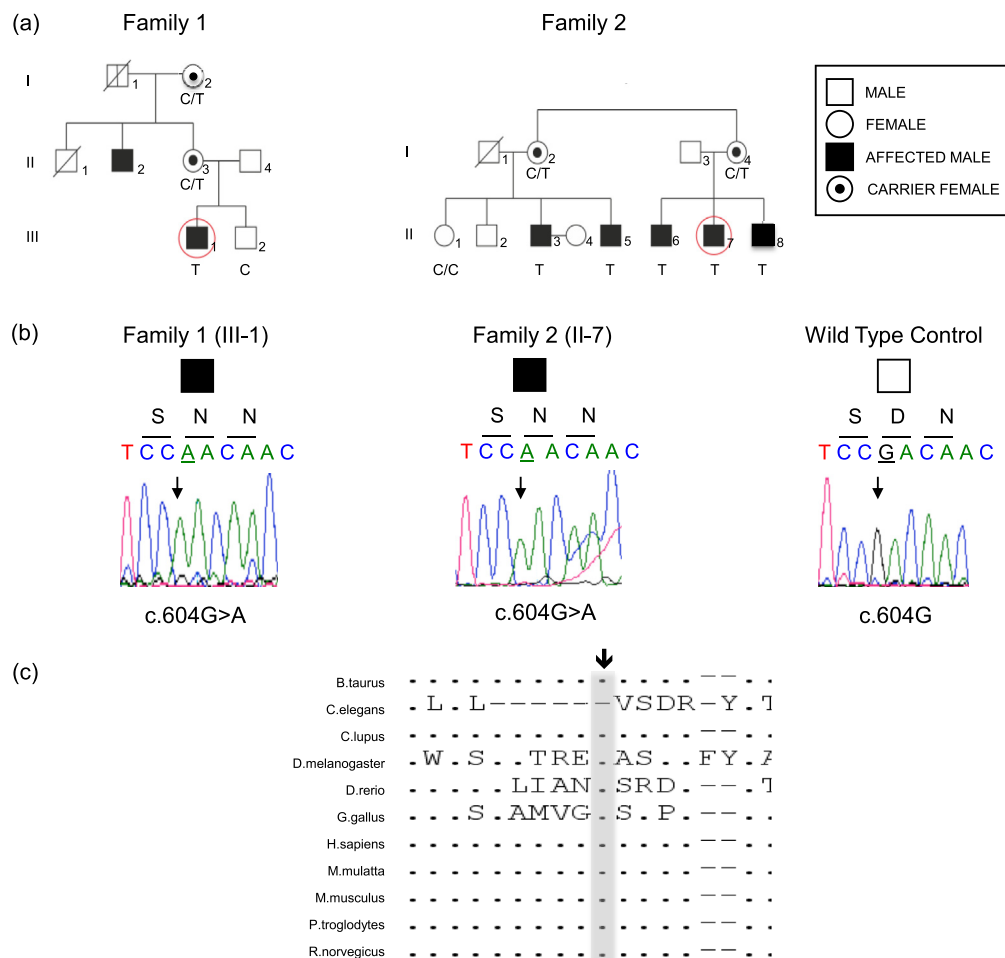

Fig. 1. (a) Pedigrees of families with mutation in *L1CAM*. Males indicated by boxes and females by circles. X-chromosome exome sequenced proband is indicated by a red circle. *L1CAM* genotype is indicated where DNA analysis was carried out. (b) *L1CAM* sequencing from the proband of family 1 and 2 and wild-type control. Primers flanking *L1CAM* exon 7 were used to amplify and sequence gDNA, F - 5′- GTG TTA GCG GGG GTG TCT T-3′ and R - 5′- ATT GCA AGG CTG AGG GAC T-3′. Arrow indicates nucleotide variant site (c.604G > A, p.D202N). (c) Protein alignment of *L1CAM* orthologs. Arrow indicates L1CAM variant amino acid, from aspartic acid, D to asparagine, N at position 202/1258. *B. taurus* (NP_001179364.1). *C.elegans* (NP_001033394.1). *C. lupus* (XP_549364.2). *D.melanogaster* (NP_511090.1). *D. rerio* (NP_571458.1). *G. gallus* (NP_990484.1). *H. sapiens* (NP_000416.1). *M. mulatta* (XP_001087861.1). *M. musculus* (NP_032504.3). *P. troglodytes* (NP_001180609.1). *R. norvegicus* (NP_059041.1). Aligned using Megalign ClustalW.: (*DNAStar*). Dots represent residues matching the consensus sequence.

dysplastic kidney and recurrent reflux with urinary tract infections. He was prescribed Dexamethazone for challenging impulsive behaviors as a child. His behavior has improved in high school where he likes to engage in sport and social activities and is currently on no medication. There is no history of seizures. He is in a support class in a mainstream high school. On examination, his head circumference was 59 cm (>98th percentile), height was 167 cm (75th percentile) and his weight was average. He has a mild intellectual disability and is non-dysmorphic with normal hands and thumbs (Fig. 1a, Supplementary Figure 1). His neurological examination was unremarkable. A repeat MRI head scan at age 12 years showed mild-moderate thinning of the posterior periventricular matter and a hypoplastic corpus callosum.

2.2. Family 2 (Supplementary Figure 1)

Due to DNA availability at the time, X-chromosome exome sequencing was carried on the affected individual II-7, however subsequent clinical follow up on this individual and his affected brothers II-6 and II-8 was not possible. As a result only the affected individuals II-3 and II-5, from this family, were subsequently clinically re-assessed.

Brother 1 (II-3) was reviewed aged 43 years. He was hypotonic in early childhood, had early mild global developmental delay and was subsequently assessed to have borderline-low normal intellectual abilities. He tended to be a loner as a child but was neither shy nor aggressive. He attended normal school and completed four years of secondary education. He began to have seizures at 10–12 years of age and was treated with antiepileptic medication; his last seizure was in his late teens and medication was ceased at around 20 years of age without recurrence of seizures. After leaving school, he had several jobs in open employment prior to becoming a carer for his wife, who received a disability pension. The couple did not have children. His intellectual abilities appeared to be within the normal range at the time of the review. His head circumference was 59 cm (>98th percentile), height was 179 cm (75th percentile) and weight was 109 kg (>97th percentile, BMI 34). He was not dysmorphic and had normal thumbs (Fig. 1a, Supplementary Figure 1). Neurological examination was unremarkable. Brain imaging was not performed.

Brother 2 (II-5) was seen aged 30 years. He was hypotonic in early childhood and had early global developmental delay with sitting at 7.5 months, crawling at 14 months, walking at 18 months, and started to use single words after two years of age. Developmental assessment at 25 months showed the following developmental levels: gross motor 16 months, fine motor/adaptive 12–15 months, language 12 months and social behaviour 15 months. He was subsequently assessed to have mild intellectual disability. He was initially employed in a sheltered workshop but subsequently in a government supported position in general employment because of significant ability. He had strabismus (bilateral esotropia) as an infant. He was described as a quiet child and a loner but was neither shy nor aggressive. He has not had seizures. His head circumference was 57 cm (75th percentile), his height was 173 cm (25th percentile). He gained weight steadily from 16 years of age, reaching a peak weight of 150 kg (>97th percentile, BMI 50). He was not dysmorphic and had normal thumbs (Fig. 1a, Supplementary Figure 1). Neurological examination was unremarkable. Brain imaging was not performed. He was found dead in bed at 37 years of age − the cause of death was not determined.

3. Methods

The research was approved by the Women's and Children's Health Network Human Research Ethics Committee. DNA extracted from whole-blood (QIAamp DNA blood maxi kit; Qiagen, Limburg, Netherlands), from these two families was part of a large X-chromosome exome sequencing (X-chromosome exome) study [Hu et al., 2015 Feb 3]. Only one affected individual from each family was sequenced (Fig. 1a). Family 1 III-1 (GAIIx; Illumina, San Diego, USA), Family 2 II-7 (HiSeq; Illumina, San Diego, USA). Confirmation of variants, and segregation analysis, by Sanger sequencing was carried out using standard methods. The L1CAM variant (NM_000425.3, c.604G > A, p.D202N) was not present in the Exome Variant Server (http://evs.gs.washington.edu/EVS/), ExAC Browser (http://exac.broadinstitute.org/), and multiple SNP databases as listed by the UCSC genome browser (http://genome.ucsc.edu/index.html) at December 2014.

Processing of the X-chromosome exome data was carried out. The raw reads were aligned using Novoalign (http://www.novocraft.com/main/index.php), PCR duplicates were removed with Picard (http://picard.sourceforge.net/) MarkDuplicates and variant calling was performed for all possible HapMap phase II SNPs with sufficient coverage using SAMtools packages (http://samtools.sourceforge.net/) (See Supplementary Tables 1 and 2 for summary statistics). VCF2LINKDATAGEN [Smith et al., 2011] was then used to produce genotype data for both individuals in the style of SNP chip data using population allele frequencies from HapMap phase II SNPs. LINKDATAGEN [Bahlo and Bromhead, 2009] was used to obtain genotype data with SNPs in approximate linkage equilibrium and high heterozygosity according to the HapMap CEU population data.

A first order continuous time Hidden Markov Model (HMM) was implemented to detect relatedness on the X chromosome based on the original formulation by Purcell et al. [Purcell et al., 2007]. The HMM has been designed to infer IBD sharing on the X chromosome using SNP array genotype data between male pairs only and therefore has only two IBD states; sharing zero or one allele IBD. The probabilities associated with sharing zero or one allele IBD are calculated by a method-of-moments approach and these values represent the initial probabilities in the HMM. The emission probabilities are functions of the IBD state, observed genotypes, population allele frequencies and genetic map positions that are taken from the HapMap phase II data (http://www.hapmap.org) while the transition probabilities require knowledge of the IBD state, initial probabilities, recombination rate, the distance between adjacent SNPs (measured in cM) and the number of meiosis separating the two individuals. The number of meiosis was estimated using a maximum likelihood approach.

The posterior probability of IBD = 1 was calculated using standard hidden Markov Model algorithms. The Viterbi algorithm was used to determine the most likely state sequence. The LOD score was calculated using MERLIN [Abecasis et al., 2002] with a pedigree with the inferred number of meiosis and a fully penetrant X-linked recessive model.

4. Results

The pedigree structure of both families was consistent with X-linked recessive inheritance, however linkage mapping was not performed. Clinical examination did not suggest a specific syndromic form of X linked intellectual disability. Previous candidate gene investigations had not identified the causative gene. Following X-chromosome exome, a total of three variants met our filtration [Weller and Gärtner, 2001] criteria for Family 1, and four for Family 2 (Table 1). The alignment and coverage statistics are in Supplementary Tables 1 and 2 The L1CAM variant (NM_000425.3, c.604G > A, p.D202N) was confirmed by Sanger sequencing (Fig. 1b) and segregated with the phenotype in both families. The aspartic acid (D) residue at position p.202 of L1CAM

Table 1
Summary of variants identified by X chromosome exome sequencing, and key clinical features of affected individuals. (NK=Not Known).

Variants identified:	Family 1		Family 2	
L1CAM (Hg19):	chrX:153136335 C > T		chrX:153136335 C > T	
	L1CAM, NM_000425.4 c.604G > A, p.D202N		**L1CAM, NM_000425.4 c.604G > A, p.D202N**	
Other variants identified:	STARD8, NM_014725.4 c.640G > C, p.R214G		CLCN5, NM_000084.4 ex1+1G > T	
	PCDH11X, NM_032968.4 c.2314G > A, p.V772M		FRMD7, NM_194277.2 c.2138delT, p713(fs)	
			XK, NM_021083.2 c.1108T > G, p.Y370D	

Phenotypic features:	Family 1		Family 2	
	Uncle (II-2)	Nephew (III-1)	Brother 1 (II-3)	Brother 2 (II-5)
Age at examination	39 years	17 years	–	–
intellectual disability	Mild-Moderate	Mild	Borderline – Low normal	Mild
Macrocephaly	Yes – 59 cm (98th percentile)	Yes – 59 cm (98th percentile)	Yes – 59 cm (>98th percentile)	No – 57 cm (90th percentile)
Dysmorphic	No	No	No	No
Seizures	No (EEG normal)	No	Yes (At 10–12 years; EEG mildly abnormal at 4 years, last seizure late teens)	No
Delayed milestones	Yes – Walked 23 months	Yes – Walked 18 months	No – walked 15 months	Yes – Walked 18 months
Speech delay	Yes	Yes	Yes	Yes
Spasticity	No	No	No	No
Adducted thumbs	No	No	No	No
Dysgenesis of corpus callosum	Yes	Yes-partial	NK	NK
Ventriculomegaly	Yes – Lateral ventricles	No	NK	NK
Malformed ventricles	Yes	Yes	NK	NK
Neurological assessment	Normal	Normal	Normal	Normal
Cerebral atrophy	Yes – mild	Yes – mild	NK	NK
Behaviour as a child	Loner/Anxiety	Impulsive	Not aggressive, loner	Not aggressive, loner
Behaviour as an adult	Anxiety	Social	Not shy	Not shy
Weight	Obese (110 kg > 97th percentile, BMI 37)	Normal	Overweight (109 kg, >97th percentile, BMI 34)	Obese (150 kg, >97th percentile, BMI 50)
Height	171 cm (25th percentile)	167 cm (75th percentile)	179 cm (75th percentile)	173 cm (25th percentile)
Other	Haemochromatosis	Dysplastic kidney – nephrectomy age 4 years	NK	Strabismus

is highly conserved (Fig. 1c). The change to asparagine (N) is predicted to be "probably damaging" (PolyPhen2: http://genetics.bwh.harvard.edu/pph2/, SIFT, http://blocks.fhcrc.org/sift/SIFT.html), with a CADD [Kircher et al., 2014] score prediction of "deleterious", at 22.6. This variant affects a conserved residue in the second L1CAM immunoglobulin domain, where pathogenic missense mutations have been observed to cluster [Bateman et al., 1996]. 27 disease-causing mutations have been identified within this Ig2 domain alone (*L1CAM* mutation database: http://www.l1cammutationdatabase.info/default.aspx)[Vos and Hofstra, 2010]. The other variants (Table 1) found were excluded because of multiple reasons: i) none was found in both families; i.e. none mapped within the ancestral haplotype; ii) *STARD8* c.640G > C, p.R214G, *PCDH11X* c.2314G > A, p.V772M and *XK* c.1108T > G, p.Y370D are low frequency variants in EXaC; iii) *CLCN5* ex1+1G > T is unique and affects one isoform of *CLCN5*, a gene involved in kidney-related disorders (MIM#300009); and iv) *FRMD7*, c.2138delT, p713(fs) affects two last amino acids of the FRMD7 protein implicated in congenital nystagmus (MIM#310700).

To investigate whether this Hg19: chrX:153136335C > T variant was recurrent or whether the two families were distantly related we went on to identify 2075 HapMap II SNPs in the X-chromosome exome data. From these, 271 SNPs were chosen for the HMM analysis. We estimated the genome wide probability of sharing zero alleles IBD to be 0.9 (for unrelated individuals this would be 1) and the number of meiosis that separate the two individuals as 5. A single region of IBD was detected using the posterior probability and the Viterbi algorithm which included the *L1CAM* variant (Fig. 2). The IBD region is approximately 5.6 cM (3.4 Mb) in length and covers 17 markers (out of 271), starting with SNP rs7067140 (Hg19: chrX:150575385) and ending with rs1126762 (Hg19: chrX:154020114). These results provide evidence that the two

families originate from a common ancestor and implicate this *L1CAM* variant to be pathogenic. None of the other 2 and 3 variants respectively fall into this interval.

5. Discussion

Using X-chromosome exome sequencing we have identified a novel missense variant in *L1CAM* in two previously unrelated
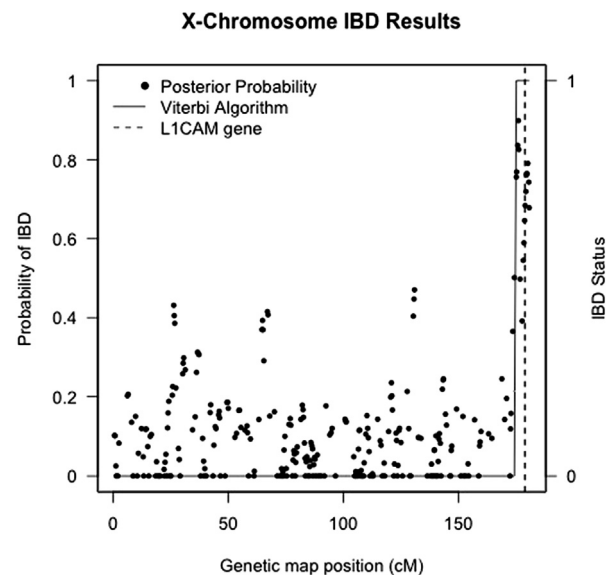

Fig. 2. IBD detection, determined using posterior probability (dots) and the Viterbi algorithm (solid line). The dotted line in the graph shows the location of the *L1CAM* gene.

families, both with mild-moderate intellectual disability and without clinical signs suggestive of a particular syndromic diagnosis. To follow up we asked whether the identified *L1CAM* variant represented two separate events (recurrence) or, was inherited from a common ancestor (IBD). The wealth of SNP data available from our X-chromosome exome sequencing, cross-referenced with the HapMap Phase II project data, enabled us to further explore this question. HMMs have been extensively used in biological sequence analysis, from gene prediction to RNA sequence alignment [Yoon, 2009]. By combing our existing data and applying HMMs, we were able to infer a shared IBD region of approximately 5.6 cM (LOD = 0.77), providing evidence that the two families have a common ancestor, and further implicating *L1CAM* as the likely causal variant.

In silico analyses, segregation testing and the online resources Polyphen2, Sift and CADD (which all predict this variant as deleterious) provide additional evidence of this variant's pathogenicity. The absence of any non-synonymous variation in this domain of L1CAM (e.g. Exome Variant Server: http://evs.gs.washington.edu/EVS) provides more support of its deleterious nature.

A different mutation located in the same nucleotide position (c.604G > T p.D202Y) has been reported previously in a patient with CRASH syndrome [Sztriha et al., 2000], a much more severe phenotype in comparison to our two families. Both variants are scored as "probably damaging" by Polyphen2. Our variant (p.D202N) results in change to an asparagine (N), a hydrophilic uncharged amino acid whereas the Sztriha et al. variant results in change to a tyrosine (Y), which is hydrophobic. The difference in clinical presentations could be the result of differing pathogenicity of the variants based on the amino acid, and subsequent protein structure, as well as other factors of genetic background.

Mutations in *L1CAM* are a well-established cause of X-linked intellectual disability, causing neurological abnormalities of varying severity with phenotype being variable, both within, and between families [Finckh et al., 2000]. In our case, the phenotype of these two families (Table 1) was not suggestive of *L1CAM* mutation. Although macrocephaly and dysgenesis of the corpus callosum was present in both affected males from Family 1, neither had hydrocephalus. In family 2, one of the brothers had macrocephaly and no brain imaging was performed. There was no observation of adducted thumbs or spasticity in either family. Both families exhibited speech delay (but not aphasia) and mild intellectual disability. While in a retrospective view this *L1CAM* variant could have been classified as pathogenic at the outset this was not obvious to us at the time. We did not have all necessary clinical, variation and bioinformatic prediction data available and simply by finding the variant twice in two unrelated families it was not prioritised. The 'tipping point' was when we were able to take advantage of the additional information contained within our X-chromosome exome sequencing data to link the two families. Through this we gained confidence in this *L1CAM* variant to be pathogenic, especially as historically identification of two identical variants alone would much rather suggest a low-frequency, non-deleterious recurrent variant. This also enabled a more detailed follow up clinical investigation and eventually diagnosis for both families, which would not have been possible on initial clinical examination alone, given the mild and atypical phenotype.

Our report shows that additional information can be extracted from MPS data initially gathered for variant identification studies.

We suggest that researchers include this analysis pipeline when low-frequency, and therefore likely pathogenic, variants are identified using MPS technology. We also ask the question, how many such X chromosome variants, both low-frequency, identical by decent and therefore likely pathogenic, might be there?

Conflict of interest

Authors do not have any conflict of interest.

Acknowledgements

Appendix A. Supplementary data

Supplementary data related to this article can be found at doi:10.1016/j.ejmg.2015.04.004

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002;30:97—101.

Bahlo M, Bromhead CJ. Generating linkage mapping files from Affymetrix SNP chip data. Bioinformatics 2009;25(15):1961—2.

Bateman A, Jouet M, MacFarlane J, et al. Outline structure of the human L1 cell adhesion molecule and the sites where mutations cause neurological disorders. EMBO J 1996;15(22):6050—9.

Finckh U, Schröder J, Ressler B, et al. Spectrum and detection rate of L1CAM mutations in isolated and familial cases with clinically suspected L1-disease. Am J Med Genet 2000;92(1):40—6.

Hu H, Haas SA, Chelly J, et al. X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes. Mol Psychiatry 2015 Feb 3. http://dx.doi.org/10.1038/mp.2014.193 [Epub ahead of print].

Kenwrick S, Watkins A, De Angelis E. Neural cell recognition molecule L1: relating biological complexity to human disease mutations. Hum Mol Genet 2000;9(6):879—86.

Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46(3):310—5.

Moos M, Tacke R, Scherer H, et al. Neural adhesion molecule L1 as a member of the immunoglobulin superfamily with binding domains similar to fibronectin. Nature 1988;334(6184):701—3.

Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81(3):559—75.

Smith KR, Bromhead CJ, Hildebrand MS, et al. Reducing the exome search space for mendelian diseases using genetic linkage analysis of exome genotypes. Genome Biol 2011;12(9):R85.

Sztriha L, Frossard P, Hofstra RM, et al. Novel missense mutation in the L1 gene in a child with corpus callosum agenesis, retardation, adducted thumbs, spastic paraparesis, and hydrocephalus. J Child Neurol 2000;15(4):239—43.

Vos YJ, Hofstra RM. An updated and upgraded L1CAM mutation database. Hum Mutat 2010;31(1):E1102—9.

Weller S, Gärtner J. Genetic and clinical aspects of X-linked hydrocephalus (L1 disease): mutations in the L1CAM gene. Hum Mutat 2001;18(1):1—12.

Yoon BJ. Hidden Markov models and their applications in biological sequence analysis. Curr Genomics 2009;10(6):402—15.