

SHORT REPORT

Variant in the X-chromosome spliceosomal gene *GP*KOW causes male-lethal microcephaly with intrauterine growth restriction

Renée Carroll^{1,2,7}, Raman Kumar^{1,2,7}, Marie Shaw^{1,2}, Jennie Slee³, Vera M Kalscheuer⁴, Mark A Corbett^{1,2} and Jozef Gecz^{*,1,2,5,6}

Congenital microcephaly, with or without additional developmental defects, is a heterogeneous disorder resulting from impaired brain development during early fetal life. The majority of causative genetic variants identified thus far are inherited in an autosomal recessive manner and impact key cellular pathways such as mitosis, DNA damage response and repair, apoptosis and splicing. Here, we report a novel donor splice site variant in the G-patch domain and KOW motifs (*GP*KOW) gene (NG_021310.2:g.6126G>A, NM_015698.4:c.331+5G>A) that segregates with affected and carrier status in a multigenerational family with an X-linked perinatal lethal condition characterized by severe microcephaly and intrauterine growth restriction (IUGR). *GP*KOW is a core member of the spliceosome that has been shown in numerous model organisms and in human cells to be essential for survival. By investigating *GP*KOW transcripts in lymphoblastoid cell lines (LCLs) of three carrier females, we show that the *GP*KOW c.331+5G>A variant disrupts normal splicing of its pre-mRNAs. In a clonal culture expressing only the c.331+5G>A allele isolated from one carrier female LCL, we observed an 80% reduction in wild type *GP*KOW mRNA, 70% reduction in the full length *GP*KOW protein and the presence of a truncated *GP*KOW protein with possible dominant negative effect. Based on our and published data we propose that the *GP*KOW gene is essential for fetal development and when disrupted, leads to a severe, male-lethal phenotype characterised by microcephaly and IUGR.

European Journal of Human Genetics (2017) 25, 1078–1082; doi:10.1038/ejhg.2017.97; published online 14 June 2017

INTRODUCTION

Congenital microcephaly, a condition where head circumference at birth is significantly below the mean for age and sex, is the result of impaired brain development during early fetal life.¹ While the genetic causes are diverse, causative variants predominantly show an autosomal recessive mode of inheritance and affect fundamental cellular pathways such as the assembly and function of the mitotic apparatus; DNA replication, damage and repair; apoptosis signaling and mRNA splicing.^{1,2} Why the developing fetal brain is particularly sensitive to even subtle alterations in pathways core to all living cells is still being investigated.^{2,3}

We studied a previously reported multigenerational family with a male lethal condition primarily characterized by severe microcephaly with IUGR and variable presentation of additional developmental defects,⁴ with carrier females possibly presenting with a mild phenotype. We identified a rare variant in a canonical donor splice site of the X-linked splicing factor gene *GP*KOW as the most likely cause of this condition.

MATERIALS AND METHODS

Details of materials and methods can be found in the Supplementary Information.

RESULTS

X-exome sequencing of obligate carrier III-2 (Figure 1a) and subsequent variant filtering identified five novel heterozygous variants (Supplementary Table 1), of which the *GP*KOW (NG_021310.2:g.6126G>A; NM_015698.4:c.331+5G>A) and *IRS4* (NG_013239.1:g.7189C>T; NM_003604.2:c.2156C>T) variants were considered candidates. However, only the *GP*KOW c.331+5G>A splice-site variant segregated in nine additional family members, including one affected male fetus (Figure 1a and Supplementary Table 3). The *GP*KOW c.331+5G>A did not have a damaging CADD score⁵ (Supplementary Table 1), however different *in silico* prediction tools calculated a decrease in splicing efficiency as compared to wild type sequence (Supplementary Table 2). *GP*KOW has a probability of loss-of-function intolerance Z score of 0.98 in ExAC,⁶ suggesting that disruption of this gene is likely to be damaging (Supplementary Table 1).

Extreme skewing (> 90%) of X chromosome inactivation (XCI) was observed in blood-derived genomic DNA from three females; II-8, III-2 and III-4 (Figure 1a), however skewing of XCI ratios did not co-segregate with obligate carrier status or presence of the *GP*KOW variant. Namely, I-2 and II-2 are obligate carriers and heterozygous for *GP*KOW c.331+5G>A variant yet have random XCI, whereas III-4, who has not conceived affected children and is homozygous wild type

¹The Robinson Research Institute, The University of Adelaide, Adelaide, Australia; ²School of Medicine, The University of Adelaide, Adelaide, Australia; ³Genetic Services of Western Australia, King Edward Memorial Hospital, Subiaco, WA, Australia; ⁴Research Group Development and Disease, Max Planck Institute for Molecular Genetics, Berlin, Germany; ⁵School of Biological Sciences, The University of Adelaide, Adelaide, SA, Australia; ⁶South Australian Health and Medical Research Institute, Adelaide, SA, Australia
*Correspondence: Professor J Gecz, School of Medicine and The Robinson Research Institute, The University of Adelaide, WCH, 9th Floor Rieger Building, 72 King William Road, Adelaide 5000, SA, Australia. Tel: +61 883133245; Fax: +61 881617342; E-mail: jozef.gecz@adelaide.edu.au

⁷These authors contributed equally to this work.

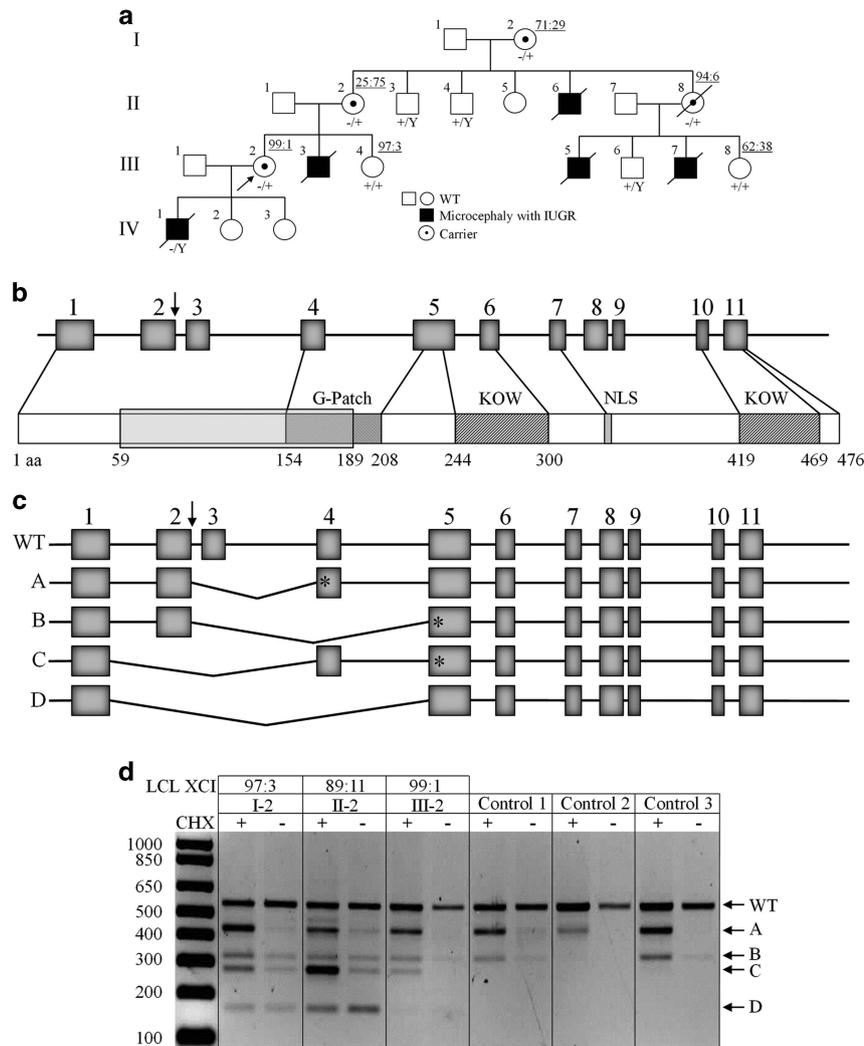


Figure 1 (a) Family pedigree. Genotyped individuals are indicated as WT +/Y or +/+; carrier -/+; affected -/Y. XCI ratios (peripheral blood DNA) for the tested females are shown (Supplementary Table 3). Foetus II-6 was a missed miscarriage at 25/40, III-3 died in the neonatal period at 36/40, III-5 was stillborn at 34/40, and pregnancies III-7 and IV-1 were terminated at 21/40 and 18.5/40, respectively, following diagnosis by ultrasound. IUGR, intrauterine growth restriction. (b) Schematic structure of *GPKOW* gene and protein. Arrow indicates location of +5G>A variant in the donor splice site of exon 2. One G-Patch domain, two KOW motifs and a predicted nuclear localisation sequence (NLS, aa 331–334) are shown. Exons are numbered as in NCBI Refseq NG_021310.2 and the amino acid locations of different domains are from the NCBI Refseq NP_056513.2. The shaded box indicates the aa 59–189 missing from the truncated protein, based on transcript D cDNA sequence. (c) Schematic representation of the four different transcripts identified in the cycloheximide (CHX) assay. Asterisks represent premature termination codons (PTC) introduced due to shifts in reading frame. (d) c.331+5G>A carrier female LCLs show additional aberrant processing of *GPKOW* pre-mRNA. Agarose gel showing *GPKOW* transcripts amplified from carrier and control female LCLs cultured in the presence (+) or absence (-) of CHX. Full-length (WT) and aberrantly-spliced (A–D) transcripts are indicated by arrows. XCI ratios assayed from LCL gDNA from three carrier females are also shown.

(WT) at this locus, has skewed XCI to the ratio of 97:3. Furthermore, skewed XCI, as assessed on venous blood DNA, does not appear to be fully protective against the effect of the *GPKOW* variant in carrier females. For example, II-8 has highly skewed XCI yet her height and occipitofrontal circumference (OFC) measurements are in the <1st and 2nd centiles, respectively. Incidentally, we observed a missense variant in exon 1 of the *AR* gene (NG_009014.2:g.6285T>A, NM_000044.3:c.170T>A, p.(Leu57Gln), rs78686797) that is in linkage disequilibrium with the *GPKOW* c.331+5G>A variant (Supplementary Figure 1c). This was supported by determining the *AR* allele inheritance for affected and unaffected male family members (Supplementary Table 3). Subsequently, we used the methylation status of the *AR* sequence tagged site (NCBI Probe database Pr012386469) that flanks rs78686797 as a proxy to indicate WT or

variant *GPKOW* allele expression. LCLs from carrier females I-2, II-2 and III-2 were all highly skewed toward expression of the WT *GPKOW* allele (Supplementary Figure 3). XCI measured on DNA from LCLs differed to XCI measured from venous blood DNA for obligate carriers I-2 and II-2 (see Supplementary Figure 3A).

Given the severity of the disorder we did not have access to a viable cell source from an affected male. To investigate whether the *GPKOW* c.331+5G>A donor splice site variant alters *GPKOW* pre-mRNA splicing and creates alternative mRNA species harboring premature termination codons (PTC) that are subject to Nonsense Mediated mRNA Decay (NMD), we inhibited NMD in low passage LCLs from three carrier (I-2, II-2 and III-2) and three control females with cycloheximide (CHX). Increased levels of bona fide NMD target *GADD45B* mRNA confirmed NMD inhibition (Supplementary Figure 2).

RT-PCR amplification of *GPKOW* mRNA using primers located within exons 1 and 5 revealed four aberrant *GPKOW* transcripts in addition to the full-length transcript (Figures 1b-d). Transcripts A and B were present in all CHX-treated LCLs suggesting they result from normal splicing events unrelated to the c.331+5G>A variant, and as they were largely undetectable in the untreated LCLs, are naturally degraded by NMD. Sanger sequencing revealed transcript A lacks exon 3 and transcript B lacks exons 3 and 4 (Figures 1c and d and Supplementary Figure 1b), in both cases causing frameshifts and the introduction of a PTC. Transcripts C and D were only amplified from carrier female samples, suggesting they result from incorrect processing of the allele with the variant c.331+5G>A donor splice site. Transcript C was predominantly detected in the CHX-treated samples indicating it is subject to NMD. Sanger sequencing revealed transcript C lacks exons 2 and 3, causing frameshift and the introduction of a PTC. Presence of comparable levels of transcript D in CHX-treated and untreated carrier female LCLs suggested that it is refractory to NMD. Sanger sequencing showed that transcript D lacks exons 2, 3 and 4, maintaining the reading frame (Figures 1c and d and Supplementary Figure 1b). The differences in the abundance of transcripts C and D observed between the carrier female LCLs (Figure 1d) corresponded to the percentage of the cells in each culture expressing the *GPKOW* c.331+5G>A allele, as determined by methylation status of the *AR* c.170T>A, p.(Leu57Gln) allele.

To further investigate the c.331+5G>A variant in isolation from sequence-confirmed WT allele expression, we isolated clonal cultures from the LCL of carrier female II-2. From 13 successfully isolated clones, 12 expressed the WT *GPKOW* allele and one the c.331+5G>A variant allele (Supplementary Figure 3b). Real time quantitative PCR (RT-qPCR) revealed that full length *GPKOW* levels were reduced by 80% in the c.331+5G>A expressing clone as compared to three independent WT expressing clones and three unrelated controls (Figure 2a). Western blot analysis showed a corresponding 70% reduction in full-length *GPKOW* protein, with the presence of an additional band at approximately 40 kDa, corresponding to the size predicted for transcript D translation product (Figures 2b and c). Indirect immunofluorescent detection of *GPKOW* protein in clonal LCL cultures demonstrated significant reduction in *GPKOW* levels but no discernible effect on its localization in the c.331+5G>A variant cells compared to WT cells and unrelated controls (Figure 3).

DISCUSSION

That RNA metabolism is fundamental to normal brain development and neuronal function is demonstrated by the complexity of alternative precursor messenger RNA (pre-mRNA) splicing in the human brain.⁷ In excess of 350 disorders, including many neurodevelopmental disorders, are consequences of genetic variants affecting pre-mRNA splicing.⁸ Furthermore, variants in splicing factors themselves are reported to cause a wide variety of human diseases.⁸⁻¹⁰

In this work we implicate *GPKOW* in a male-lethal developmental disorder. *GPKOW* is a ubiquitously expressed nuclear RNA-binding protein that has been identified as a core member of the spliceosome.¹¹ *GPKOW* associates with catalytically active complexes B^{act} and C^{11,12} and acts as an essential cofactor for the two DEAH-box helicases, DHX16 and DHX38, which sequentially catalyze the key transesterification reactions. *GPKOW* has been classified as essential for human cell survival,¹³ and there are no known paralogs in the human genome. Moreover, there are no homozygous or hemizygous loss of function variants thus far listed in ExAC.⁶ However, DECIPHER lists a patient (#286121) with a 1.4 Mbp deletion including the *GPKOW* gene and a 33 Mbp pathogenic deletion of chromosome 21, presenting microcephaly as one of the clinical symptoms. *GPKOW* is highly conserved and has been shown to be essential for life in numerous model organisms. For example, in *Caenorhabditis elegans*, RNAi manipulation of its *GPKOW* ortholog R11A8.2 results in embryonic lethality in more than 80% of cases and the minority that survived had morphological growth defects.¹⁴ RNAi silencing of the *GPKOW* ortholog CG10324 in *Drosophila melanogaster* neuroblasts is early lethal and causes under proliferation.¹⁵ To our knowledge, no vertebrate model of a *GPKOW* knockout has been generated.

The *GPKOW* c.331+5G>A variant we identified was predicted *in silico* to adversely impact spliceosomal processing of exon 2 donor splice site (Supplementary Table 2). Although adenine at the +5 position occurs in approximately 9% of human donor splice sites genome-wide,¹⁶ there are numerous reported cases of +5G>A disease-associated variants causing aberrant pre-mRNA processing.¹⁷ Blocking NMD in LCLs from three carrier females revealed the expression of the c.331+5G>A allele is correlated with the production of aberrant transcripts C and D. Current evidence shows that *GPKOW* is subject to XCI¹⁸ and as such in a female the gene is expressed only from the active X-chromosome. Using clonal WT and c.331+5G>A

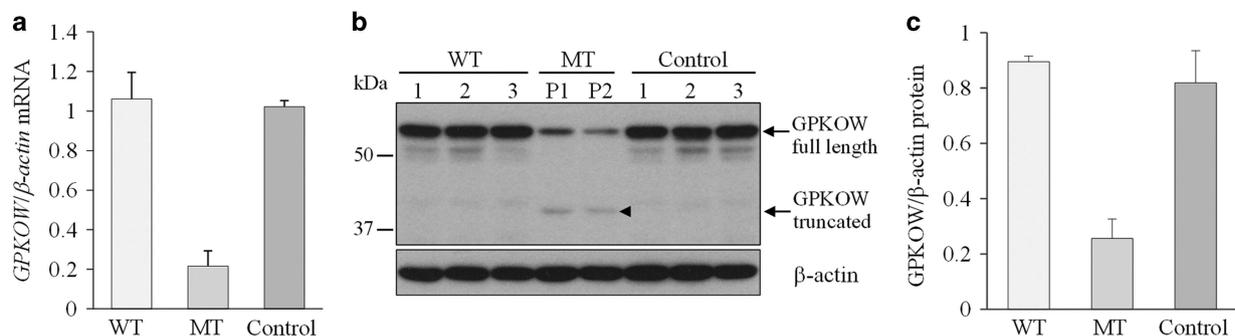


Figure 2 *GPKOW* full length WT mRNA and protein are significantly reduced in the c.331+5G>A LCL clone (MT). (a) Graph showing *GPKOW* mRNA levels relative to *ACTB*. RT-qPCR data from independent WT clones ($n=3$), c.331+5G>A clones ($n=2$; different passages of the same culture) from the carrier female II-2, and normal controls ($n=3$) was averaged and error bars show SD. (b) Western blot showing full length *GPKOW* from independent WT clones ($n=3$), the c.331+5G>A clone ($n=2$; different passages of the same culture; P1, P2), and normal controls ($n=3$). An additional band at around 40kDa in the c.331+5G>A (MT) clones P1 and P2 corresponds to the predicted size of a truncated protein translated from transcript D. (c) Graph showing the mean densitometry values of full length *GPKOW* protein relative to β -actin. Western blot signals in (b) were measured by ImageJ software and normalised to housekeeping β -actin protein. Error bars show SD.

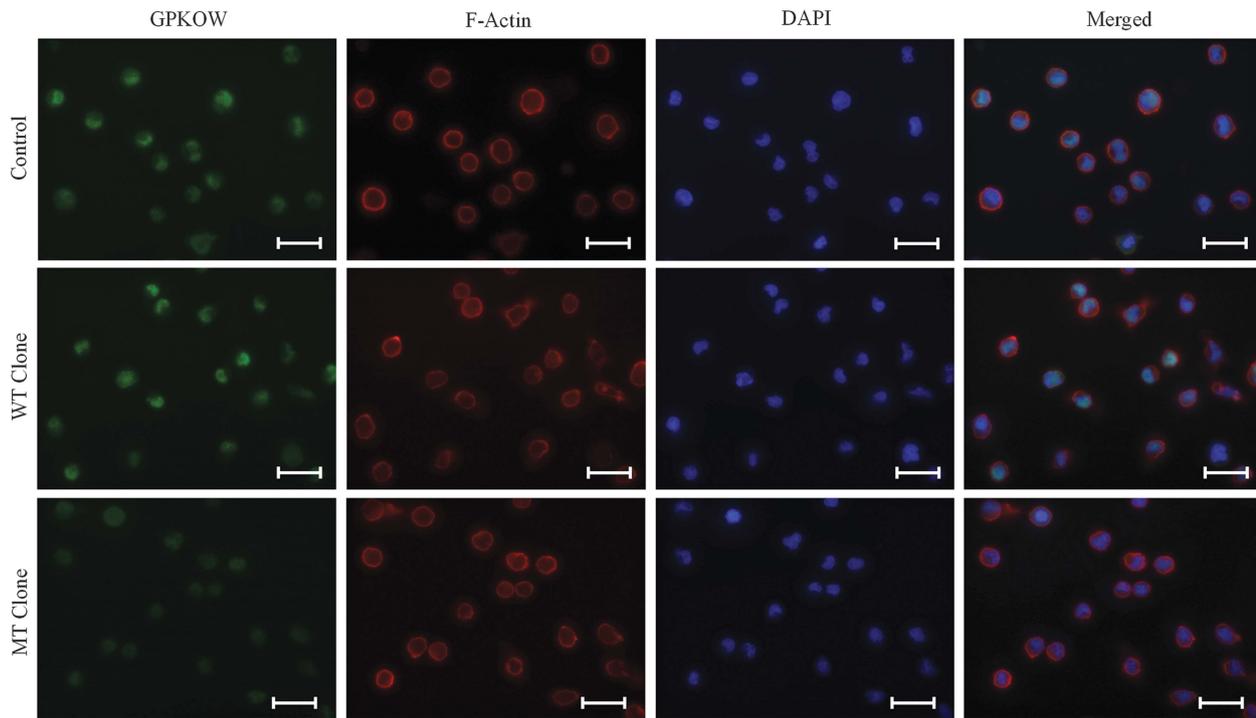


Figure 3 Immunofluorescence showing a dramatic GPLOW reduction in c.331+5G>A clonal LCLs. Representative images of LCL immunofluorescence for GPLOW from normal controls from different individuals ($n=3$), independent WT clones ($n=3$) and c.331+5G>A clone ($n=1$, done in duplicate). Scale bars represent 20 μm .

variant expressing LCL cultures from one carrier female we show the likely effect of the c.331+5G>A variant on GPLOW mRNA splicing. In the c.331+5G>A allele-expressing LCL clone, we observed full-length transcript levels were reduced by 80% as compared to expression levels in WT-expressing cultures (Figure 2a). Importantly, we also observed a corresponding 70% reduction in full-length GPLOW protein, and the presence of a truncated GPLOW protein, corresponding to the transcript D translation product (Figures 2b and c). The truncated protein is predicted to lack amino acids 59–189 (Figures 1b and c), spanning the first two-thirds of the G-patch domain, which has been shown to be important for GPLOW binding to DHX16.¹² Consequently, we predict that the truncated protein, even if folded and localized correctly, would be unlikely to function properly as a cofactor in spliceosomal reactions. We also speculate that the truncated GPLOW protein may have a dominant negative effect by competing with residual WT GPLOW for binding sites and/or binding to and sequestering other spliceosomal factors.

Testing of X-chromosome inactivation provided conflicting data. Three females (II-8, III-2 and III-4) of whom two are obligate carriers (II-8 and III-2), had highly skewed XCI, while two other obligate carriers (I-2 and II-2) showed random XCI. However, when we tested XCI on DNA isolated from LCLs of three obligate carriers, two of whom have variable XCI in blood (I-2 and II-2), we found all three to have highly skewed XCI favoring expression of the wild type GPLOW allele (Supplementary Figure 3a). Skewing of XCI also doesn't appear to be protective of the possible mild phenotype observed in the obligate carrier females (Supplementary Table 3). On aggregate, the combined data on XCI in blood and LCL DNA and on multiple obligate carrier and wild type females in this family show a complex, yet to be resolved mechanism of XCI skewing in different tissues. It also appears that a factor unrelated to the GPLOW variant and

involved in high XCI skewing, might also be segregating in this family (e.g. wild type female III-4 with high skewing of XCI; Figure 1a).

How significantly reduced levels of WT GPLOW along with any function of the truncated GPLOW lead to the observed male-lethal phenotype remains to be determined. However, there are a number of severe congenital disorders with the shared features of microcephaly and growth retardation that are attributed to genetic variants in splicing factors. For example, mandibulofacial dysostosis, Guion-Almeida type (MFDGA, OMIM #610536) is caused by heterozygous variants in *EFTUD2*; Verheij syndrome (OMIM #615583) has recently been attributed to heterozygous variants in *PUF60*,¹⁹ and microcephalic osteodysplastic primordial dwarfism type I (MOPDI, OMIM #210710) caused by recessive variants in *RNU4ATAC*. Furthermore, the particular sensitivity of males to mutations in X-linked splicing factors is demonstrated with the example TARP syndrome (OMIM #311900), a male-lethal developmental disorder caused by mutations in the X-linked gene RNA Binding Motif Protein 10 (*RBM10*).²⁰

Our findings show that a variant in the X-chromosome essential splicing factor GPLOW is the likely cause of a male-lethal disorder featuring microcephaly and IUGR, and add to the growing number spliceosomal gene variants implicated in severe developmental disorders, highlighting the crucial role of pre-mRNA splicing in human development.⁸ Further investigations to elucidate the role of GPLOW in normal development are warranted.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We wish to thank the family involved in this study for their cooperation. The authors would like to thank Evelyn Douglas for assistance. This research was supported by National Health and Medical Research Council of Australia

grants 1041920 and 1091593 to J.G., MS McLeod Research Fellowship to M.C, and by the EU FP7 project GENCODYS, grant number 241995 to VMK. This study makes use of data generated by the DECIPHER Consortium. A full list of centres who contributed to the generation of the data is available from <https://decipher.sanger.ac.uk/> and via email from decipher@sanger.ac.uk. Funding for the project was provided by the Wellcome Trust.

- 1 Alcántara D, O'Driscoll M: Congenital microcephaly. *Am J Med Genet C Semin Med Genet* 2014; **166C**: 124–139.
- 2 Gilmore EC, Walsh CA: Genetic causes of microcephaly and lessons for neuronal development. *Wiley Interdiscip Rev Dev Biol* 2013; **2**: 461–478.
- 3 Zhang X, Chen MH, Wu X *et al*: Cell-Type-Specific Alternative Splicing Governs Cell Fate in the Developing Cerebral Cortex. *Cell* 2016; **166**: 1147–1162, e1115.
- 4 Hockey A, Crowhurst J, Cullity G: Microcephaly, holoprosencephaly, hypokinesia—second report of a new syndrome. *Prenat Diagn* 1988; **8**: 683–686.
- 5 Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J: A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014; **46**: 310–315.
- 6 Lek M, Karczewski KJ, Minikel EV *et al*: Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016; **536**: 285–291.
- 7 Li Q, Lee JA, Black DL: Neuronal regulation of alternative pre-mRNA splicing. *Nat Rev Neurosci* 2007; **8**: 819–831.
- 8 Chabot B, Shkreta L: Defective control of pre-messenger RNA splicing in human disease. *J Cell Biol* 2016; **212**: 13–27.
- 9 Daguene E, Dujardin G, Valcarcel J: The pathogenicity of splicing defects: mechanistic insights into pre-mRNA processing inform novel therapeutic approaches. *EMBO reports* 2015; **16**: 1640–1655.
- 10 Lehalle D, Wiczorek D, Zechi-Ceide RM *et al*: A review of craniofacial disorders caused by spliceosomal defects. *Clin Genet* 2015; **88**: 405–415.
- 11 Agafonov DE, Deckert J, Wolf E *et al*: Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. *Mol Cell Biol* 2011; **31**: 2667–2682.
- 12 Zang S, Lin TY, Chen X *et al*: GPKOW is essential for pre-mRNA splicing in vitro and suppresses splicing defect caused by dominant-negative DHX16 mutation in vivo. *Biosci Rep* 2014; **34**: e00163.

- 13 Blomen VA, Majek P, Jae LT *et al*: Gene essentiality and synthetic lethality in haploid human cells. *Science* 2015; **350**: 1092–1096.
- 14 Piano F, Schetter AJ, Morton DG *et al*: Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*. *Curr Biol* 2002; **12**: 1959–1964.
- 15 Neumuller RA, Richter C, Fischer A, Novatchkova M, Neumuller KG, Knoblich JA: Genome-wide analysis of self-renewal in *Drosophila* neural stem cells by transgenic RNAi. *Cell Stem Cell* 2011; **8**: 580–593.
- 16 Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R: Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res* 2006; **34**: 3955–3967.
- 17 Caminsky N, Mucaki EJ, Rogan PK: Interpretation of mRNA splicing mutations in genetic disease: review of the literature and guidelines for information-theoretical analysis. *F1000Research* 2014; **3**: 282.
- 18 Cotton AM, Price EM, Jones MJ, Balaton BP, Kobor MS, Brown CJ: Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Hum Mol Genet* 2015; **24**: 1528–1539.
- 19 El Chehadeh S, Kerstjens-Frederikse WS, Thevenon J *et al*: Dominant variants in the splicing factor PUF60 cause a recognizable syndrome with intellectual disability, heart defects and short stature. *Eur J Hum Genet* 2016; **25**: 43–51.
- 20 Johnston JJ, Teer JK, Cherukuri PF *et al*: Massively parallel sequencing of exons on the X chromosome identifies RBM10 as the gene that causes a syndromic form of cleft palate. *Am J Hum Genet* 2010; **86**: 743–748.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)