SEC31A May be Associated with Pituitary Hormone Deficiency and Gonadal Dysgenesis

Running Head: MPHD and gonadal dysgenesis with *SEC31A* variant.

Edward S Tobias[1,2], Angela K Lucas-Herald[3,] Danielle Sagar[4], Augusto C Montezano[5,6], Francisco J Rios[5], Livia De Lucca Camargo[5,6], Graham Hamilton[7], Gabriella Gazdagh[1,2], Louise A Diver[1], Nicola Williams[1], Pawel Herzyk[7,8], Rhian M Touyz[5,6], Andy Greenfield[4,9], Ruth McGowan[1,3], S Faisal Ahmed[3]

1 West of Scotland Centre for Genomic Medicine, Laboratory Medicine Building, Queen Elizabeth University Hospital, Govan Road, Glasgow, G51 4TF, UK
2 Academic Unit of Medical Genetics and Clinical Pathology, University of Glasgow, Queen Elizabeth University Hospital, Glasgow, G51 4TF, UK
3 Developmental Endocrinology Research Group, School of Medicine, Dentistry and Nursing, University of Glasgow, Royal Hospital for Children, 1345 Govan Road, Glasgow, G51 4TF, UK
4 MRC Mammalian Genetics Unit, Harwell Institute, Harwell Campus, Oxfordshire, OX11 0RD, UK.
5 Institute of Cardiovascular and Medical Sciences, British Heart Foundation Centre for Research Excellence, University of Glasgow, 126 University Avenue, Glasgow, G12 8TA, UK.
6 Research Institute of McGill University Health Centre, McGill University, Montreal, Canada.
7 Glasgow Polyomics, College of Medical Veterinary and Life Sciences, Garscube Estate, Switchback Rd, Glasgow, G61 1BD, UK.
8 Institute of Molecular Cell and Systems Biology, College of Medical Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK.
9 Nuffield Department of Women's & Reproductive Health, Institute of Reproductive Sciences, University of Oxford, Oxford, UK

Address for Correspondence

Professor Edward Tobias BSc MBChB PhD FRCP FHEA
Laboratory Medicine Building
Queen Elizabeth University Hospital
1345 Govan Road, Glasgow, G51 4TF, UK
Tel: +44 1413549201
Email: edward.tobias@glasgow.ac.uk

Key words: endocrine, pituitary, DSD, exome, genome, *SEC31A*, COPII
Word count: 1554/1600
Funding: The next generation sequencing was facilitated by grants from the Wellcome Trust ISSF and Scottish Genomes Partnership. This Scottish Genomes Partnership is funded by the Chief Scientist Office of the Scottish Government Health Directorates [SGP/1] and The Medical Research Council Whole Genome Sequencing for Health and Wealth Initiative (MC/PC/15080). AKLH is funded by an NES/CSO Clinical Lectureship.
Acknowledgements: We are grateful to the staff of the sequencing facilities of Glasgow Polyomics, Edinburgh Genomics and the West of Scotland Centre for Genomic Medicine. In addition, we thank Sara Wells, Lydia Teboul, Joffrey Mianné, Gemma Codner and their team for CRISPR-Cas9 mutagenesis. We also thank the husbandry team in Ward 5 of the Mary Lyon Centre (MLC) at Harwell, particularly Jackie Harrison, Gemma Atkins, Anju

Paudyal and Michelle Sandell. Additionally, we thank Martin Fray and the FESA and Gene Delivery teams of the MLC transgenic facility.

ABSTRACT

Purpose: Disorders/differences of sex development (DSD) result from variants in many different human genes but, frequently, have no detectable molecular cause.

Methods: Detailed clinical and genetic phenotyping was conducted on a family with three children. A *Sec31a* animal model and functional studies were used to investigate the significance of the findings.

Results: By trio whole-exome DNA sequencing we detected a heterozygous *de novo* nonsense *SEC31A* variant, in three children of healthy non-consanguineous parents. The children had different combinations of disorders that included complete gonadal dysgenesis and multiple pituitary hormone deficiency. *SEC31A* encodes a component of the COPII coat protein complex, necessary for intracellular anterograde vesicle-mediated transport between the endoplasmic reticulum (ER) and Golgi. CRISPR-Cas9 targeted knockout of the orthologous *Sec31a* gene region resulted in early embryonic lethality in homozygous mice. mRNA expression of ER-stress genes *ATF4* and *CHOP* was increased in the children, suggesting defective protein transport. The pLI score of the gene, from gnomAD data, is 0.02.

Conclusions: *SEC31A* might underlie a previously unrecognised clinical syndrome comprising gonadal dysgenesis, multiple pituitary hormone deficiencies, dysmorphic features and developmental delay. However, a variant that remains undetected, in a different gene, may alternatively be causal in this family.

INTRODUCTION

In differences/disorders of sex development (DSD), chromosomal, gonadal and/or phenotypic sex is atypical [1]. These conditions can result from a pathogenic variant in one of several human genes [2-5]. Many of these genes encode components of the known pro-testis and pro-ovary developmental pathways leading to gonadal differentiation and development. Other DSD genes encode proteins involved in the synthesis of androgens or the molecules, including receptors, that are necessary for androgen action [2].

Here, we report a family with 3 children (from unaffected, non-consanguineous parents) who are clinically affected by different combinations of pituitary hormone deficiencies, DSD and cranio-facial abnormalities. High-throughput DNA sequencing revealed, in all three siblings, a truncating variant in the *SEC31A* gene which encodes a component of the COPII-complex that coats the vesicles mediating ER-to-Golgi vesicular transport. Skeletal

dysplasias can occur as a result of disrupted ER-to-Golgi transport (in both anterograde and retrograde directions). We now report and discuss the possible association of significant DSD in individuals with a nonsense variant in a highly conserved and ubiquitously expressed gene encoding a key protein required for anterograde ER-Golgi vesicular transport.

MATERIALS and METHODS

Ethics

West of Scotland Ethics approval (14/WS/0036) and R&D permission (GN14KH079) were obtained from the respective committees for whole exome and whole genome sequencing and for the data storage and analysis.

METHODS

Detailed methods can be found in the Supplemental Information.

Microarray analysis and DNA sequencing

Chromosomal microarray analysis was undertaken on all members of the family, by array comparative genomic hybridisation (aCGH). Full exome and whole genome sequencing was performed.

Animal model

All mouse experimentation was approved by the Animal Welfare and Ethical Review Body (AWERB) at MRC Harwell. Mice used were bred with licensed approval from the UK Home Office (PPL 70/8898). The *Sec31a* deletion allele (SEC31A-DEL1055-EM1-B6) was generated by pronuclear injection of Cas9 RNA and two pairs of guide RNAs (gRNAs) into 1-cell C57BL/6J mouse embryos. Blastocysts were immunostained and quantitative reverse transcription PCR (qRT-PCR) for *Sec31a* analysis was undertaken.

mRNA expression of ER stress gene in circulating cells

mRNA expression of ER stress genes was measured in triplicate in peripheral blood mononuclear cells (PBMCs) using quantitative real-time polymerase chain reaction. Analysis was via Mann Whitney U.

RESULTS

Clinical phenotype

The three siblings investigated were born to the same, non-consanguineous, healthy parents. Their presentations are listed in Table 1 and a family pedigree is demonstrated in Fig. S1. Child 3 was first to present, secondary to his DSD phenotype. During clinic visits for Child 3 aged 1 year, parents discussed the clinical presentation of Child 1 and Child 2.

Genetic analysis

Conventional karyotyping revealed that the oldest child (child 1), phenotypically female, was 46, XY whilst the younger children (child 2 and child 3), phenotypically female and male, were 46, XX and 46, XY, respectively.

Chromosomal microarray analysis was undertaken on all family members. A heterozygous deletion of a genomic region of 493 kb at chromosome 12q13.2 was detected in all three children. It was considered likely benign as it did not contain any genes known to be associated with a similar phenotype and it was also detected in their father, who was phenotypically healthy with a normal growth hormone level.

Testing by DNA fragment length analysis for the polyalanine tract within the *SOX3* gene on the X chromosome, showed the same repeat length in all family members. Therefore, no evidence was detected of either the 33 bp duplication previously reported to be associated with X-linked intellectual disability and growth hormone deficiency or the 21 bp duplication associated with congenital hypopituitarism.

No pathogenic variant was detected on analysis of the DNA of the oldest child by the DSD clinical diagnostic gene panel used at the time in the clinical molecular diagnostic laboratory [6]. Subsequently, however, analysis of all five family members' whole-exome and whole-genome sequencing variant call format (VCF) files was performed. No pathogenic variants were identified using an autosomal recessive model of inheritance, or following close examination of the genomic region at chromosome 12q13.2 for a *de novo* or maternally-inherited variant.

Using a *de novo* dominant model, however, a single heterozygous nonsense variant in *SEC31A* (on chromosome 4) was identified in the DNA of all three siblings,

NC_000004.12:g.82842323G>A (GRCh38). This variant, NM_001077207.4:c.2785C>T was absent in the DNA extracted from the peripheral blood samples of both parents, in all overlapping reads identified. At the protein level, the variant, NP_001070675.1:p.(Gln929Ter), is predicted to cause termination of protein translation in exon 22 (of a total of 27 exons). Study of the relevant BAM data files and also Sanger sequencing were undertaken and confirmed the presence of the variant, heterozygously, in the peripheral blood DNA of all three of the children (in 19/54, 26/55 and 30/55 Illumina sequence reads) and also its absence in both of the parents (in DNA extracted from cells obtained from the parents' peripheral blood, skin and saliva). Approximately 50% of each child's filtered heterozygous peripheral blood DNA variants were inherited from each parent, confirming parental relationships.

Animal model

Previous studies have demonstrated that *Sec31a* is expressed strongly in the mouse in all fetal gonad lineages tested and, notably, is present at 11.5 dpc, the sex-determining stage, in XY and XX supporting cell precursors [7]. Following CRISPR-Cas9 deletion of exon 21 of *Sec31a* (Supplementary Information), a significant reduction in the transcript was detected by qRT-PCR in heterozygous mutant embryos, suggestive of nonsense-mediated decay of the mutated transcript. Embryonic lethality was observed for homozygous mice (Supplementary Fig. S2). *Sec31a*-heterozygous-mice were grossly normal but detailed phenotypes are unavailable.

Human ER stress response analysis

To assess the functional effect upon the endoplasmic reticulum (ER) in the children possessing the detected *SEC31A* variant, expression analysis of genes encoding ER stress-related proteins was undertaken (Supplementary Fig. S3). mRNA expression was compared to 2 control samples (from the children's parents who are confirmed not to have the same *SEC31A* mutation). Comparisons were made using Mann Whitney U via Graphpad Prism v 8.0. Children 2 and 3 showed evidence of increased *ATF4* (3.3 fold, p=0.007) and *CHOP* expression (4.5 fold, p=0.01), consistent with ER stress, relative to controls. Child 1 was not included as she was receiving antihypertensive treatment, which may alter ER stress response.

DISCUSSION

Dysregulation of intracellular transport between the ER and the Golgi apparatus is becoming increasingly recognised as a process that can underlie human developmental disorders. For instance, recently, in 14 patients with Saul-Wilson syndrome (microcephalic dwarfism) a *de novo* pathogenic variant was identified in *COG4*, a gene encoding a protein required for retrograde Golgi to ER transport [8]. In fact, it has now been established that several other genetic conditions result from the disruption of Golgi-ER vesicular trafficking. These include cranio-lenticulo-sutural dysplasia (CLSD) and also a recently described developmental syndrome caused by *de novo* heterozygous *ARCN1* coat protein encoding gene variants disrupting retrograde Golgi to ER transport, which involves significant genital abnormalities, most commonly hypospadias and microphallus [9].

In the family described here, the oldest child has XY gonadal dysgenesis with sex reversal and developmental delay, the second child has isolated growth hormone deficiency and the youngest child has a combination of XY DSD characterised by proximal hypospadias, bilateral undescended testes and micropenis in combination with multiple pituitary hormone deficiencies and developmental delay. Whole exome sequencing revealed in all three children a *de novo* heterozygous truncating variant in the gene encoding the COPII-complex protein, SEC31A, required for anterograde ER-to-Golgi transport.

Whilst it is not possible to describe this variant as pathogenic according to the latest ACMG guidelines for variant interpretation there are several features that are nevertheless of interest in this regard. The variant is *de novo*, with the parents of the children being unrelated and healthy. In addition, several of the clinical features evident in members of this family were previously described in individuals (with CLSD) possessing pathogenic variants in genes encoding other COPII proteins, such as *SEC23A* [10 11]. These features include asymmetric, low-set and posteriorly rotated ears, maxillary hypoplasia and intellectual disability. Moreover, siblings homozygous for an exon 22 *SEC31A* nonsense variant were reported to have a very severe congenital neurodevelopmental and skeletal syndrome (Halperin-Birk syndrome) with intrauterine growth retardation, fatal by four years of age [12]. The endocrine and genital characteristics were, however, not described in that reported family. Furthermore, the increased ER stress response gene expression detected here in two children is consistent with that previously reported in *SEC31A* knockout cells (11).

The heterozygous mutant mice appeared grossly phenotypically normal, but it is possible that undetected abnormalities were present. Alternatively, a dosage sensitivity may be present in humans that is absent in mice or there may be a difference between the species in the level of retained functionality of the prematurely truncated protein, as the human gene has far greater transcript complexity.

SEC31A has a loss intolerance probability (pLI) score of 0.02 in the gnomAD database, currently containing 50 observed loss-of-function (LoF) heterozygous variants (with 117 expected) in >50 individuals, assumed to be healthy. This would suggest that an undetected variant in a different gene (including one at 12q13.2) may still be causal in this family.

In conclusion, we report three clinically affected siblings of non-consanguineous parents and suggest the possibility that there is a previously unrecognised association between the human *SEC31A* gene and a clinical syndrome comprising endocrine abnormalities, DSD, dysmorphic features and developmental delay. This would further support the increasingly recognised importance of genetic abnormalities in ER-to-Golgi intracellular transport in human developmental disorders. However, at this stage, particularly in the absence of additional reported similar families in association with a heterozygous *SEC31A* variant, we recognise that a gene other than *SEC31A* may be responsible.

STATEMENTS AND DECLARATIONS

FUNDING

The next generation sequencing was facilitated by grants from the Wellcome Trust ISSF and Scottish Genomes Partnership. This Scottish Genomes Partnership is funded by the Chief Scientist Office of the Scottish Government Health Directorates [SGP/1] and The Medical Research Council Whole Genome Sequencing for Health and Wealth Initiative (MC/PC/15080). AKLH is funded by an NES/CSO Clinical Lectureship (PCL/21/05).

COMPETING INTERESTS

There are no conflicts of interest to declare.

AUTHOR CONTRIBUTIONS

EST, RM and SFA planned the project, obtained ethics approval and led the writing and analysis. AKLH prepared drafts of the report, acquired blood samples and undertook analysis and interpretation of the data. DS and AG undertook the mouse model work. ACM, FJR, LDLC and RMT undertook and supervised the ER stress work. GH, SGP, LAW, NW, PH and GG were responsible for genetic and proteomic analysis and interpretation. All authors reviewed the final manuscript.

ETHICS APPROVAL STATEMENT

West of Scotland Ethics approval (14/WS/0036) and R&D permission (GN14KH079) were obtained from the respective committees for whole exome and whole genome sequencing and for the data storage and analysis. Written informed consent was obtained from the parents of the patients for the analysis and for the publication of this case report.

CONSENT TO PUBLISH

Written informed consent has been obtained for publication by the family.

Child	Karyotype	Birthweight (kg)	Birth gestation (weeks)	Facial features	Neurodevelopmental delay	Pituitary hormone deficiencies	Genital phenotype	Other features
1 – phenotypic female	46,XY	2.1	Term	Nil of note.	Mainstream school with additional educational support.	Low AMH for male. Normal gonadal, thyroid and adrenal function. No abnormality on MRI.	Clitoromegaly. Normal vagina and uterus on imaging and EUA. Bilateral testicular tissue with substantial fibrotic atrophic change and occasional placental alkaline phosphatase (PLAP) positive cells, suggestive of germ cell tumours on laparoscopic bilateral gonadectomy aged 5 years.	Obesity and hypertension requiring amlodipine therapy
2 – phenotypic female	46,XX	3.8	Term	Cup-shaped ears	Mainstream school – no support required.	GH deficiency. Normal gonadotrophin levels, adrenal and thyroid function. Ectopic posterior pituitary on MRI.	Typical female	Nil of note.
3 – phenotypic male	46,XY	1.8	32	Plagiocephaly, protruberant posteriorly-rotated, cup-shaped ears, sparse eyebrows, sub-mucous cleft palate.	Mainstream school with additional educational support.	Hypogonadotrophic hypogonadism, (T response to prolonged hCG stimulation; low normal AMH; no response of LH or FSH on LHRH stimulation), GH deficiency, hypothyroidism, adrenal insufficiency. Ectopic posterior pituitary on MRI.	Perineal hypospadias, small phallus, bilateral undescended testes. Bilateral orchidopexy and 2-stage repair for proximal hypospadias by age 3 years.	BP on the 98th centile for age and height. Chiari 1 malformation of spine with associated spinal syrinx and epilepsy (well controlled with carbamazepine therapy).

Table 1. Clinical phenotypes of the children. Abbreviations: AMH: anti-Müllerian hormone; BP: blood pressure; EUA: examination under anaesthetic; FSH: Follicle Stimulating Hormone; GH: Growth Hormone; hCG: human chorionic gonadotrophin; LH: Luteinising Hormone; LHRH: Luteinising Hormone Releasing Hormone; MRI: magnetic resonance imaging; T: testosterone.

REFERENCES

1. Lee PA, et al. Consensus statement on management of intersex disorders. Pediatr, 118, e488-e500. (2006).
2. Alimussina M, et al. Genetic testing of XY newborns with a suspected disorder of sex development. Curr Opin Pediatr, 30, 548-57. (2018).
3. Audi L, et al. Genetics in Endocrinology: Approaches to molecular genetic diagnosis in the management of differences/disorders of sex development (DSD): position paper of EU COST Action BM 1303 "DSDnet". Eur J Endocrinol, 179, R197-R206. doi: 10.1530/eje-18-0256 (2018).
4. Buonocore F, et al. Next-generation sequencing reveals novel genetic variants (SRY, DMRT1, NR5A1, DHH, DHX37) in adults with 46, XY DSD. J Endocr Soc, 3, 2341-60. (2019).
5. Baxter RM, et al. Exome sequencing for the diagnosis of 46, XY disorders of sex development. J Clin Endocrinol Metab, 100, E333-E44. (2015).
6. Lucas-Herald AK, et al. Serum Anti-Müllerian Hormone in the Prediction of Response to hCG Stimulation in Children With DSD. J Clin Endocrinol Metab, 105, 1608-16. doi: 10.1210/clinem/dgaa052 (2020).
7. Jameson SA, et al. Temporal transcriptional profiling of somatic and germ cells reveals biased lineage priming of sexual fate in the fetal mouse gonad. PLoS Genet, 8, e1002575. (2012).
8. Ferreira CR, et al. A recurrent de novo heterozygous COG4 substitution leads to saul-wilson syndrome, disrupted vesicular trafficking, and altered proteoglycan glycosylation. Am J Hum Genet, 103, 553-67. (2018).
9. Ritter AL, et al. Expanding the phenotypic spectrum of ARCN1-related syndrome. Genet Med, doi: 10.1016/j.gim.2022.02.005 (2022).
10. Boyadjiev SA, et al. Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. Nat Genet, 38, 1192-97. (2006).
11. Fromme JC, et al. The genetic basis of a craniofacial disease provides insight into COPII coat assembly. Dev Cell, 13, 623-34. (2007).
12. Halperin D, et al. SEC31A mutation affects ER homeostasis, causing a neurological syndrome. J Med Genet, 56, 139-48. (2019).

SUPPLEMENTARY INFORMATION

SUPPLEMENTARY MATERIALS and METHODS

Sequencing analysis

Whole exome sequencing of the children and their unaffected parents was undertaken through Glasgow Polyomics. Exome enrichment was performed using a set of overlapping hybridization capture probes (Agilent). Full exome sequencing of DNA was performed on an Illumina NextSeq500 machine and FastQ files generated. Following sequence read trimming by Cutadapt, sequence alignment and mapping to reference was performed by the Burrows-Wheeler Aligner (BWA). Duplicate sequence reads were located and marked by Picardtools and local re-alignment around indels performed. Base quality scores were adjusted using genomic analysis toolkit (GATK) BaseRecalibrator. Variant identification was achieved using GATK HaplotypeCaller and VariantFiltration tools.

Whole genome sequencing was undertaken through the Scottish Genome Partnership (SGP). Sequencing libraries were prepared for whole genome sequencing using Illumina TruSeq DNA PCR-Free High Throughput library preparation kits. WGS was performed on an Illumina HiSeqX sequencing machine and BWA-MEM was used for read alignment. Mean coverage per sample was 36.7. Variant calling was performed using GATK Haplotype Caller and the data were subsequently filtered bioinformatically to retain exonic and adjacent intronic sequences.

West of Scotland Ethics approval (14/WS/0036) and R&D permission (GN14KH079) were obtained from the respective committees for whole exome and whole genome sequencing and for the data storage and analysis.

Analysis of VCF files using VarSeq

GoldenHelix variant analysis software (VarSeq® v.2.2.1) was used to set up the bioinformatics analysis pipeline for variant filtering and annotation of the VCF files. Custom filters were created by customising prebuilt filter chains available for common inheritance patterns, including *de novo* candidate, dominant heterozygous, compound heterozygous, recessive homozygous and X-linked inheritance. Modifications included the addition of filters utilising public databases such as variant frequencies. The initial trio analysis workflow included the following filter cards: *Read depth*: ≥20; *Genotype quality*: ≥30; *All MAF (minor allele frequency)* (NHLBI ESP6500SI-V2-SSA137 Exomes Variant Frequencies 0.0.30,GHI): <0.01 or missing; *Non-cancer alt allele frequency* (gnomAD Exomes Variant Frequencies 2.1.1, BROAD): <0.01 or missing; *Inheritance-specific filter cards*: 'Mendel error: de novo'; 'Compound heterozygous'; 'Mendel error: transmitted, Zygosity: homozygous, Mother: Heterozygous, Father: Heterozygous' (recessive homozygous pipeline); and 'Zygosity: Heterozygous' and 'Segment after X:1' (X-linked pipeline).

Review of variant in BAM file

Visual inspection of read alignments in BAM files was undertaken using Integrative Genomics Viewer (IGV) software, to confirm the variant's presence and eliminate false positive findings.

Validation of variant by Sanger sequencing
Sanger sequence analysis of the region of the *SEC31A* gene (NM_001077207.4) containing the variant was performed. Primers were designed using the Primer3Plus tool (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and SNPCheck V3 (http://www.genetools.org/SNPCheck/snpcheck.htm). Data were analysed using Mutation Surveyor software. Variant nomenclature and reporting followed current HGVS (www.hgvs.org) and ACMG (www.acmg.net) guidelines, respectively.

Animal model

All mouse experimentation was approved by the Animal Welfare and Ethical Review Body (AWERB) at MRC Harwell. Mice used were bred with licensed approval from the UK Home Office (PPL 70/8898).

Genome editing
The *Sec31a* deletion allele (SEC31A-DEL1055-EM1-B6) was generated by pronuclear injection of Cas9 RNA and two pairs of guide RNAs (gRNAs) into 1-cell C57BL/6J mouse embryos. Each pair of gRNAs flank exon ENSMUSE00001292836 (exon 21 of transcript *Sec31a-201*). Details of microinjection and methodologies were reported in Mianné et al., 2017 [1].

Embryo collection
Noon on the day of the copulatory plug was counted as 0.5 dpc. Adult mice were humanely sacrificed by dislocation of the neck, confirmed by cessation of circulation, and embryos were decapitated in ice-cold, phosphate-buffered saline (PBS) solution.

Blastocyst immunostaining
Uterine horns were harvested from dams and flushed with PBS to obtain blastocysts. Blastocysts were then fixed for 30 mins at room temperature in 4% PFA in PBS, before being used immediately. Blastocysts were washed in PBT (0.1% Tween in PBS) before being permeabilised in 0.5% Tween in PBS. Blastocysts were then blocked in 10% donkey serum in PBS for 1 hour at room temperature before being incubated with primary antibodies against SEC31A (*Sigma-Aldrich*, HPA005457), and the embryonic trophectoderm marker, GATA3 (*R&D*, AF2605) at 1:50 overnight at 4ºC. Blastocysts were then washed in PBT and incubated with the appropriate secondary antibodies at a dilution of 1:200 for 1 hour at room temperature. Blastocysts were washed again with PBT before being incubated in a Vectashield DAPI solution (diluted 1:30 in PBS) overnight at 4ºC prior to imaging. Blastocysts were imaged in a glass bottomed petri dish using an inverted confocal microscope.

Quantitative reverse transcription PCR (qRT-PCR) for Sec31a analysis
Whole embryos were dissected at 11.5 dpc and snap frozen on dry ice. Tissue was then stored at -80ºC until needed. RNA was extracted using the RNeasy Midi kit (*Qiagen*) following the manufacturer's instructions. RNA was eluted in 500µl RNase-free water, in 2 x 250µl batches. cDNA was synthesised in a thermocycler at 37ºC for 2 hours, using a High Capacity Transcription kit [Applied biosystems] according to the manufacturer's instructions. The resulting solution was then diluted 1:6 in RNase-free water. Mastermix solution (1x SYBR©Green Master Mix (*ThermoFisher*), 0.2µM forward and reverse primer) was prepared

and added to a Microamp fast optical reaction plate (*Applied Biosystems*) at 15µl/well. cDNA was then added to this plate at 5µl/well, with each sample plated in duplicate. The plate was loaded into the 7500 Fast real time PCR system (*Applied Biosystems*) and run according to instructions.

Sec31a primers: 5'- CAGTGAGAGCCTTGGATGTGA-3' and 5'- GTTTTAGCTCCCGGCGTCAT-3'

Human ER stress response analysis

Peripheral blood mononuclear cell (PBMC) harvest

Total blood was carefully pipetted over a gradient solution of Ficol:Hipaque Plus (density 1.077 g/ml; Amersham Biosciences) and centrifuged at 400 g for 35 min at 20 ºC with slowest acceleration and deceleration. PBMCs, identified as a cloudy layer were removed with a Pasteur pipette and transferred to a 50 mL tube with ice cold PBS and centrifuged at 400 g for 10 min (full acceleration and deceleration). This step was repeated twice. Finally, the cell pellet was suspended in QIAzol lyses reagent (Qiagen) for RNA isolation.

mRNA analysis

mRNA expression was measured on PBMCs using quantitative real-time polymerase chain reaction. Total RNA was extracted using Qiazol as per manufacturer's instructions (Qiagen, UK). cDNA was prepared by reverse transcription as per manufacturer's instructions (Thermo Fisher, UK) and synthesised on a Multi-Block PCR Thermal Cycler (Satellite 0.2 Thermo Cooler, Thermo Fisher, UK) under the following conditions: 10 minutes at 25°C (primer annealing); 120 minutes at 37°C (cDNA synthesis); 5 minutes at 85°C (reverse transcriptase inactivation) and maintenance at 4°C until thermal cycler termination. Target gene expression was identified using Qiagen QuantiTech primer assays (Qiagen, UK) and SYBR® Green (UK). Primers used are shown in Table S1. Transcript gene expression was normalised using the housekeeping gene *GAPDH* (Qiagen, UK). The $2^{-\Delta\Delta CT}$ method was then used to calculate relative gene expression [2]. Results are expressed as fold increase. mRNA expression was compared to 2 control samples (from the children's parents who are confirmed not to have the same *SEC31A* variant).

Gene	Forward primer	Reverse primer	Manufacturer
GAPDH	GAGTCAACGGATTTGGTCGT	TTGATTTTGGAGGGATCTCG	Qiagen, UK
SEC31A	CAGTGAGAGCCTTGGATGTGA	GTTTTAGCTCCCGGCGTCAT	Qiagen, UK
ATF4	ATGACCGAAATGAGCTTCCTG	GCTGGAGAACCCATGAGGT	Eurofins, UK
ATF6	TCCTCGGTCAGTGGACTCTTA	CTTGGGCTGAATTGAAGGTTTTG	Eurofins, UK
BIP	CACGGTCTTTGACGCCAAG	CCAAATAAGCCTCAGCGGTTT	Eurofins, UK
CHOP	GGAAACAGAGTGGTCATTCCC	CTGCTTGAGCCGTTCATTCTC	Eurofins, UK
XBP1	CCTTGTAGTTGAGAACCAGG	GGGGCTTGGTATATATGTGG	Eurofins, UK
NRF2	TCCAGTCAGAAACCAGTGGAT	GAATGTGTGCGCCAAAAGCTG	Eurofins, UK

Table S1. Details of primers used in mRNA analysis.

SUPPLEMENTARY RESULTS

Family structure

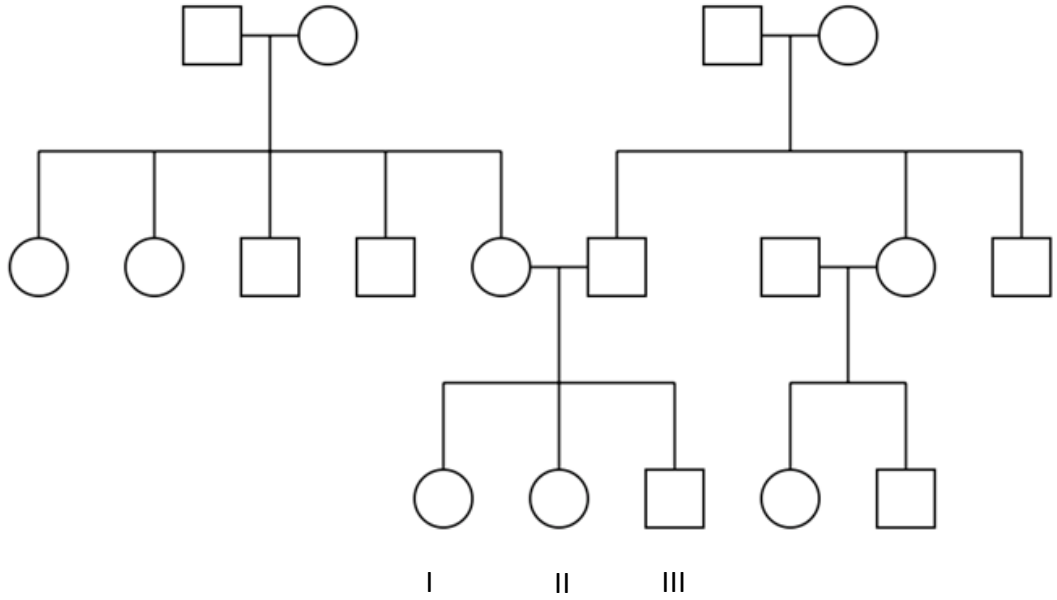
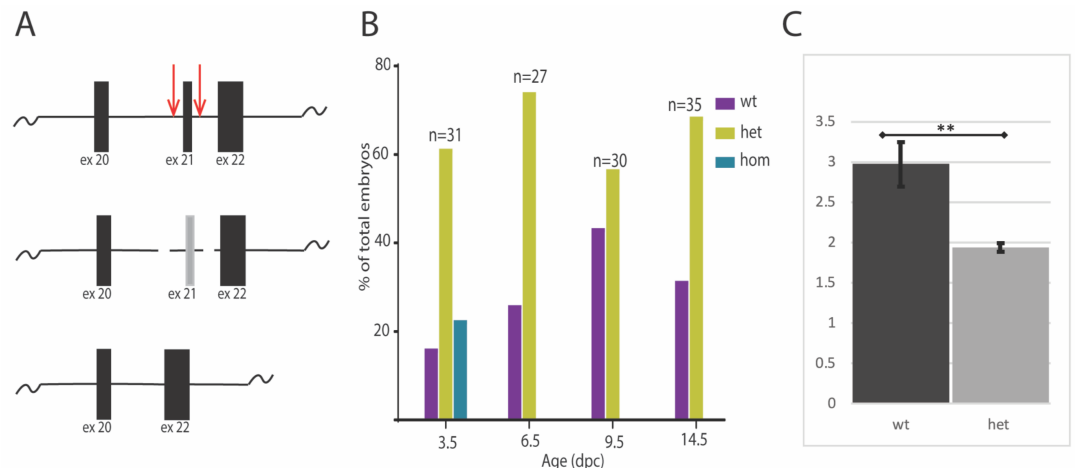


Figure S1. Pedigree showing the family structure with the three affected siblings represented in the third generation. I: Child 1; II: Child 2; III: Child 3.

Animal model: disruption of *Sec31a* in mice causes early embryonic lethality



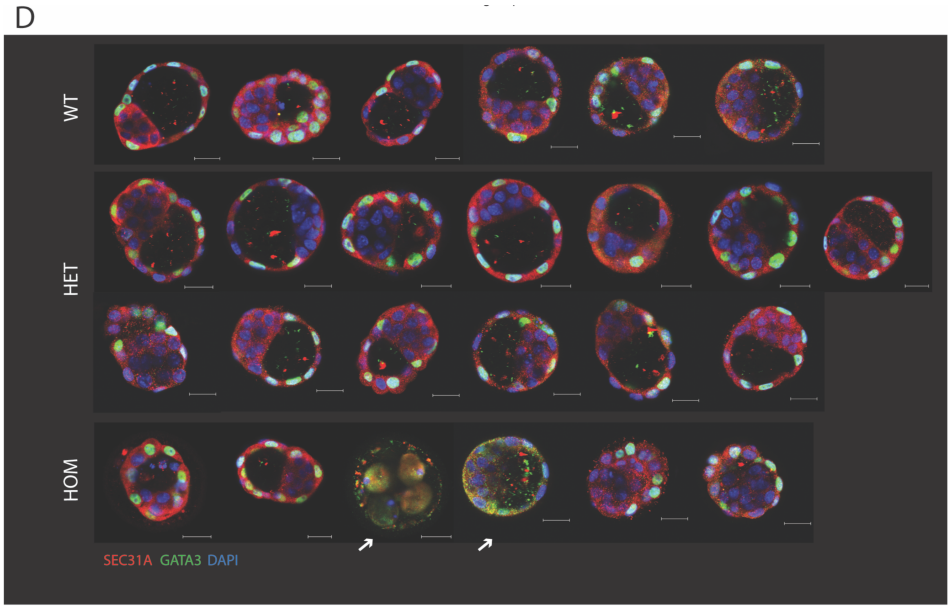


Figure S2. Disruption to *Sec31a* causes early embryonic lethality in mice. (A) CRISPR-Cas9 was used to target exon 21 of *Sec31a* (based on numbering of transcript 201). Two pairs of single-guide RNAs used (red arrows) flank the exon. A 1055 nucleotide deletion removes the entire exon, resulting in a predicted frame-shift and premature stop codon; (B) Inter-cross matings between heterozygotes (het) for the deletion mutation yielded no detectable homozygous (hom) foetuses at 14.5, 9.5 and 6.5 days *post coitum* (dpc), but these were detected at 3.5 dpc. The numbers of embryos/fetuses analysed at each stage is shown above each cohort; (C) Quantification of *Sec31a* transcript by qRT-PCR reveals a significant reduction in levels in heterozygous mutant embryos at 11.5 dpc compared to wild-type (wt) controls at the same stage. **, $p \leq 0.05$ (student's *t*-test); (D) Analysis at 3.5 dpc reveals homozygous mutant blastocysts. Two of these (white arrows) have negligible levels of SEC31A protein, and one is clearly retarded. SEC31A protein in other homozygous blastocysts may be maternal in origin.

Human ER stress response analysis

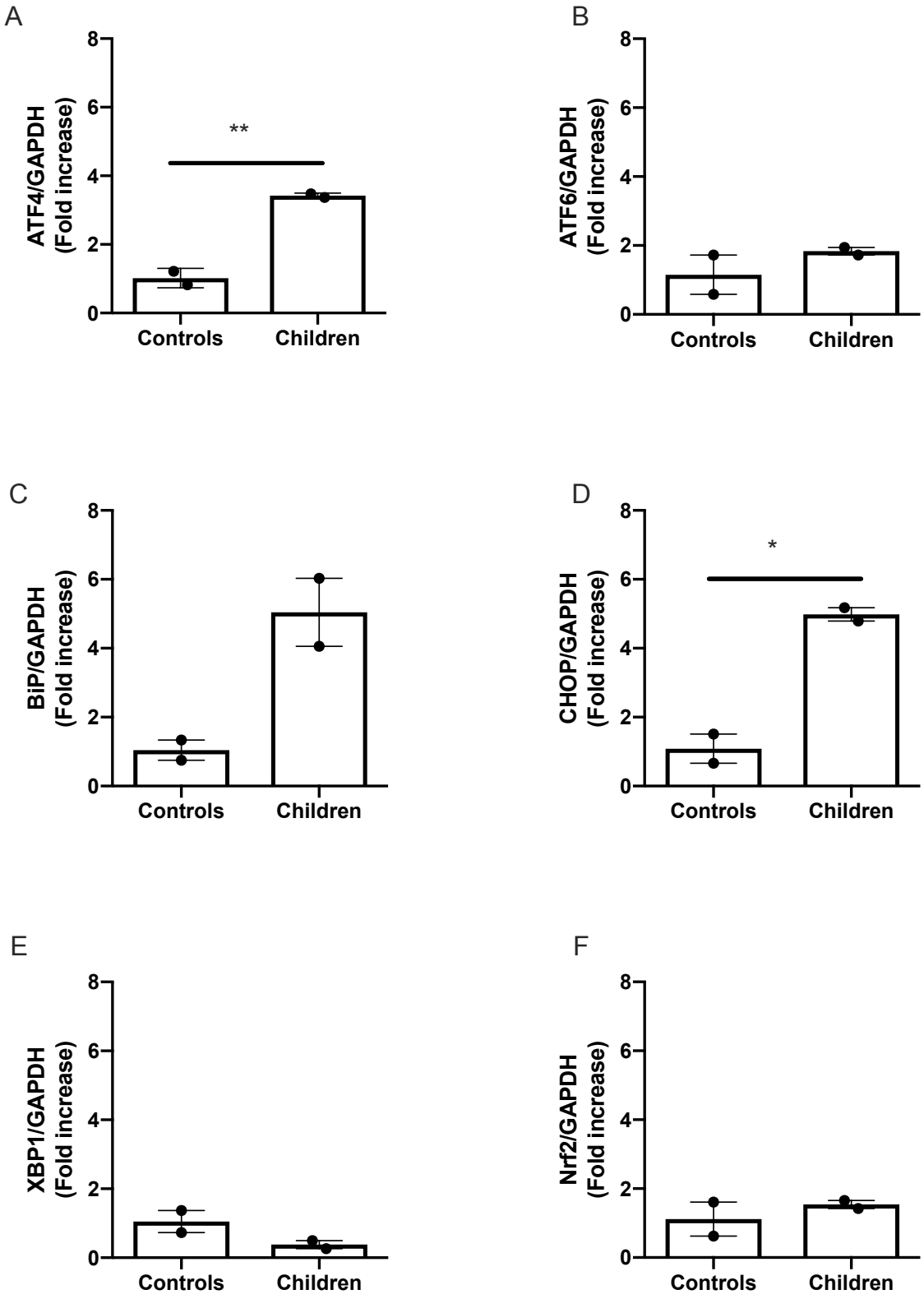


Figure S3. Evidence of ER stress in individuals with the *SEC31A* variant.
The mRNA levels of *ATF4* (A) and *CHOP* (D) were increased in the children compared to 2

parental controls. There were no changes in mRNA levels of *ATF6* (B) or *NRF2* (F). There was a trend to *BIP* (C) being increased and *XBP1* (E) expression being reduced. Total RNA was extracted from PBMC, gene expression was determined by real-time PCR and normalized by GAPDH. Analysis was performed in triplicate, with the average of the results taken. Analysis was via Mann Whitney U. *p<0.05. **p<0.01

SUPPLEMENTARY REFERENCES

1. Mianné J, Codner GF, Caulder A, Fell R, Hutchison M, King R, Stewart ME, Wells S, Teboul L. Analysing the outcome of CRISPR-aided genome editing in embryos: Screening, genotyping and quality control. *Methods* 2017;**121**:68-76
2. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(− ΔΔC(T)) method. *Methods* 2001;**25**(4):402-408