

Whole Exome Family Plus

REFERRING HEALTHCARE PROFESSIONAL

NAME **HOSPITAL**

PATIENT

NAME **DOB** **AGE** **GENDER** **ORDER ID**
 Female
PRIMARY SAMPLE TYPE **SAMPLE COLLECTION DATE** **CUSTOMER SAMPLE ID**

SUMMARY OF RESULTS

PRIMARY FINDINGS

Analysis of whole exome sequence variants in previously established disease genes

The patient is heterozygous for *SMC1A* c.1246G>A, p.(Glu416Lys), which is likely pathogenic. The variant has occurred *de novo*. The patient is heterozygous for *MYH10* c.1470T>G, p.(Asn490Lys), which is a variant of uncertain significance (VUS). The variant has occurred *de novo*.

Del/Dup (CNV) analysis

Negative

PRIMARY FINDINGS: SEQUENCE ALTERATIONS IN ESTABLISHED DISEASE GENES

GENE	TRANSCRIPT	NOMENCLATURE	GENOTYPE	CONSEQUENCE	INHERITANCE	CLASSIFICATION
SMC1A	NM_006306.3	c.1246G>A, p.(Glu416Lys)	HET	missense_variant	X-linked	Likely pathogenic
	ID	ASSEMBLY	POS	REF/ALT		
	gnomAD AC/AN	POLYPHEN	SIFT	MUTTASTER	PHENOTYPE	
	0/0	possibly damaging	deleterious	disease causing	Cornelia de Lange syndrome	
MYH10	NM_001256012.1	c.1470T>G, p.(Asn490Lys)	HET	missense_variant	Other	Variant of uncertain significance
	ID	ASSEMBLY	POS	REF/ALT		
	gnomAD AC/AN	POLYPHEN	SIFT	MUTTASTER	PHENOTYPE	
	0/0	possibly damaging	deleterious	disease causing		

SEQUENCING PERFORMANCE METRICS

SAMPLE	MEDIAN COVERAGE	PERCENT \geq 20X
Index	134	99.24
Mother	147	99.25
Father	167	99.46

TEST INFORMATION

Blueprint Genetics Whole Exome Family Plus Test (version 2, Feb 9, 2018) consists of sequence analysis of all protein coding genes in the genome for the proband and affected/unaffected family members, coupled with Whole Exome Deletion/Duplication (CNV) Analysis. The test targets all protein coding exons, exon-intron boundaries (\pm 20 bps) and selected non-coding, deep intronic variants (listed in Appendix). This test should be used to detect single nucleotide variants and small insertions and deletions (INDELs) up to 220 bps and copy number variations defined as single exon or larger deletions and duplications. This test should not be used for the detection of repeat expansion disorders, or diseases caused by mitochondrial DNA (mtDNA) mutations. The test does not recognize balanced translocations or complex inversions, and it may not detect low-level mosaicism.

Analysis of Whole Exome Family Plus Test is primarily focused on established disease genes that have been previously associated with genetic disorders. The genes with known clinical association include those curated by Blueprint Genetics (BpG) and included in BpG diagnostic panels (>2400 genes). These genes are supplemented with genes included in The [Clinical Genomics Database](#) (>3350 genes), and the Developmental Disorders Genotype-Phenotype Database ([DD2GP](#)) (>1640 genes). Total number of genes that are considered as clinically associated in the Whole Exome Family Plus analysis is >3750 (and the number is constantly updated).

If analysis of exome variants in previously established disease genes is inconclusive, exome variant data is also analyzed for variants that are not located within known clinically associated genes, but have properties that make them candidates for potentially disease-causing variants (please see Appendix: Summary of the Test). If over time other patients with similar phenotype and variants in the same gene are identified, the variant may be reclassified as a likely cause of the disorder.

STATEMENT

CLINICAL HISTORY

Patient is an 18-year-old female with short stature and microcephaly, dysmorphic features, myopia, hearing loss, severe global developmental delay and seizure disorder, possibly in keeping with a diagnosis of Cornelia de Lange syndrome. Additionally, a feature of particular interest in the patient is an upper maxilla central incisor, generally seen as a minor sign of holoprosencephaly, suggesting an underlying sonic hedgehog (SHH) signalling defect. Patient is not known to have brain anomalies within the HPE spectrum, but rather mild vermian hypoplasia, a midline defect and mega cisterna magna, a nonspecific finding.

Previous *NIPBL* sequencing in 2009 identified three polymorphisms. Previous CMA and karyotype were normal (46,XX)

There is no parental consanguinity and parents are healthy.

Consent has been received to report secondary findings for all family members.

CLINICAL REPORT

Whole-exome sequence analysis of variants in previously established disease genes

Given that there is no reported family history of the same disease, the exome data of the patient and parents were analysed for *de novo* variants and variants following recessive inheritance pattern. To account for incomplete penetrance of pathogenic variants, also rare inherited heterozygous variants were analysed.

Sequence analysis using the Blueprint Genetics (BpG) Whole Exome Family Plus identified a heterozygous missense variant *SMC1A* c.1246G>A, p.(Glu416Lys) and a heterozygous missense variant *MYH10* c.1470T>G, p.(Asn490Lys).

Both variants have occurred *de novo*; they were not detected in the patient's parents. The sequencing coverage at this genomic position in *SMC1A* is 134 reads in the mother and 60 reads in the father. The sequencing coverage at this genomic position in *MYH10* is 152 reads in the mother and 151 reads in the father.

***SMC1A* c.1246G>A, p.(Glu416Lys)**

This variant is absent in [gnomAD](#), a large reference population database (n>120,000 exomes and >15,000 genomes) which aims to exclude individuals with severe pediatric disease. The variant affects a highly conserved amino acid in the coiled-coil domain of the protein (UniProtKB - [Q14683](#)). There is small physicochemical difference between glutamic acid and lysine (Grantham score 56, [0-215]), and all *in silico* tools utilized predict the alteration to be deleterious. To the best of our knowledge, this variant has not been described in the medical literature, or reported in the disease-related variation database [HGMD](#). The variant has been detected by other laboratories in the context of clinical testing and submitted to ClinVar (variation ID [287649](#)).

SMC1A

The *SMC1A* protein encoded by *SMC1A*-gene (OMIM* [300040](#)) is the human homolog of the yeast Smc1 gene, a core component of the cohesin complex forming a heterodimer with Smc3. The cohesin complex plays a critical role in sister chromatid cohesion as well as a role in regulating gene expression by long-range enhancer-promoter interactions (PMID [15458660](#)). *SMC1A*, although residing on chromosome Xp11.22, incompletely escapes X-inactivation (PMID [7757075](#)). *SMC1A* pathogenic variants are known to cause a phenotype resembling Cornelia de Lange syndrome (CdLS, OMIM# [300590](#)).

CdLS is a multisystem disorder characterized by intrauterine growth retardation, short stature, typical face, hirsutism, congenital anomalies of especially the distal upper limbs, and intellectual and developmental disabilities. Congenital heart disease (CHD), including ventricular septal defects, atrial septal defects, pulmonic stenosis, tetralogy of Fallot (TOF), hypoplastic left heart syndrome, and bicuspid aortic valve have been diagnosed in CdLS patients. Behavioral characteristics include autism spectrum disorder, and a predisposition to engage with challenging behavior, especially self-injurious behavior. Two forms of CdLS are

known, a classic and a mild form with similar medical issues; however, greater cognitive impairment is detected in individuals with classic CdLS. CdLS is associated with variants in a series of genes; variants in *NIPBL* (~70–75%) and *SMC1A* (~5%) are the most prevalent (PMID [7757075](#), GeneReview [NBK1104](#)).

Individuals with *SMC1A* variants can resemble CdLS, but manifestations are less marked compared to individuals with *NIPBL* variants: growth is less disturbed, facial signs are less marked (except for periocular signs and thin upper vermillion), there are no major limb anomalies, and they have a higher level of cognitive and adaptive functioning. Compared to individuals with a heterozygous *NIPBL* pathogenic variant, the facial features in those with *SMC1A* pathogenic variants include slightly flatter and broader eyebrows, and a broader and longer nasal bridge. Cardiac malformation is less common than in other forms of CdLS. *SMC1A*-CdLS is usually caused by a *de novo* pathogenic variant. Unlike a typical X-linked gene, *SMC1A* is not fully inactivated in the process of X-chromosome inactivation, and females are likely to show signs of the disease. In the setting of maternal inheritance, a heterozygous mother is likely to display some features of CdLS that are milder than those of her affected son. However, to date, too few families with *SMC1A*-CdLS have been identified to fully evaluate this model (PMID: [20301283](#)). In a study by Deardoff et al. (2007), it was found that 10 of 14 total *SMC1A*-mutation-positive individuals were female. Furthermore, authors described similarly affected male and female probands, implying an X-linked dominant mode of expression. Interestingly, several males were rather mildly affected and no more severely affected than many of the *SMC1A*-mutation-positive females. Since *SMC1A* escapes X inactivation, authors hypothesized that the mechanism in affected females could be due to a dominant negative effect of the altered protein and it would be less likely due to decreased protein levels or skewed X inactivation (PMID: [17273969](#)).

Recently, heterozygous loss-of-function variants in *SMC1A* gene have been identified in females with epileptic encephalopathy resembling Rett syndrome, but not having features suggestive of CdLS (PMID [7757075](#), [28166369](#)).

SMC1A-related phenotypes are inherited in X-linked manner. Missense and small in-frame deletion variants have been identified in both genders, but the reported loss-of-function variants have been observed in females suggesting that these types of mutations are not tolerated in males, likely leading to early miscarriages (PMID [7757075](#)).

To date, over 100 variants in *SMC1A* are annotated as disease-causing (DM) in HGMD Professional (version 2020.4) with approximately half being missense variants (54%) and the rest are different types of truncating variants (46%).

MYH10 c.1470T>G, p.(Asn490Lys)

This variant is absent in [gnomAD](#), a large reference population database (n>120,000 exomes and >15,000 genomes) which aims to exclude individuals with severe pediatric disease. The variant affects a highly conserved amino acid in the myosin motor domain of the protein (UniProtKB - [P35580](#)), there is moderate physicochemical difference between asparagine and lysine (Grantham score 94, [0-215]), and all *in silico* tools utilized predict the alteration to be deleterious. To the best of our knowledge, this variant has not been described in the medical literature or reported in disease-related variation databases such as [ClinVar](#) or [HGMD](#).

MYH10

MYH10 gene (MIM *[160776](#)) encodes myosin-10 protein. This protein has a role in cytokinesis, cell shape, and specialized functions such as secretion and capping (UniProtKB - [P35580](#)). Not much is known on the disease association of *MYH10* variants. Tuzovic et al. have reported a novel heterozygous *de novo* nonsense variant in *MYH10* in an 8-year-old patient with intrauterine growth restriction, microcephaly, developmental delay, feeding difficulties, failure to thrive, congenital bilateral hip dysplasia, cerebral and cerebellar atrophy, hydrocephalus, and congenital diaphragmatic hernia (CDH) (PMID: [25003005](#)). Brain MRI showed evidence of diffuse volume loss with decreased thickness of the white matter and increased ventricular size. Jin et al. reported a homozygous p.(Glu1738Lys) variant in a patient with conotruncal heart defects with extracardiac anomalies (PMID: [28991257](#)). Petrovski et al. reported *de novo* nonsense variant in a fetus with bilateral ventriculomegaly and aqueductal stenosis (PMID: [30712878](#)). In addition, multiple publications have reported variants in *MYH10* in individuals with autism spectrum disorder (HGMD professional 2020.4). Hamdan et al. have reported a *de novo* variant of *MYH10* in a more severely affected patient (PMID: [25356899](#)). This patient had severe global developmental delay, she was non-verbal, did not walk, was hypotonic and had cerebral atrophy. Mouse studies with loss-of-function mutations in the same gene supported the disease-association of this gene. Homozygous null mice had 70% fewer, but larger myocytes compared to wild-type mice and also showed an increase in binculeation (PMID: [12893741](#)). Ma and Adelstein reported mice that were homozygous for a hypomorphic

allele *Myh10* p.(Arg709Cys) (PMID: [24825879](#)). Mice homozygous for the hypomorphic allele died due to cardiac failure at embryonic day 14.5, and exhibited a failure in midline fusion resulting in cleft palate, ectopia cordis, diaphragmatic herniation, delayed cerebellar development, and large omphalocele. Heterozygous *Myh10* p.(Arg709Cys) mice had no cardiac or brain lesions, however 50% had omphalocele, and all had diaphragmatic herniation (PMID: [24825879](#)).

There are currently 3 variants in *MYH10* annotated as disease-causing (DM) in the HGMD Professional variant database (version 2020.4): 2 nonsense and one missense. These are mentioned above in the works of Petrovski *et al.*, Tuzovic *et al.* and Jin *et al.* Additionally, there are 6 variants annotated as possibly disease-causing (DM?), associated with autism spectrum disorder and/or intellectual disability. Currently, there is only 1 *MYH10* nonsense variant classified as pathogenic in clinical testing in the ClinVar database seen in clinical testing (January 2021).

Mutation nomenclature is based on GenBank accession NM_006306.3 (*SMC1A*) and NM_001256012.1 (*MYH10*) with nucleotide one being the first nucleotide of the translation initiation codon ATG.

Upon request, filtered variant files and raw data files from the whole exome analysis can also be provided.

CONCLUSION

SMC1A c.1246G>A, p.(Glu416Lys) is classified as likely pathogenic, based on the established association between the gene and the patient's phenotype, the variant's rarity in control populations, *in silico* predicted pathogenicity, and *de novo* occurrence. Disease caused by *SMC1A* variants is inherited in an X-linked dominant manner. The clinical presentation of heterozygous females may be influenced by skewed X-inactivation. In this case, the variant has occurred *de novo*; it was not detected in the patient's parents. Maternity and paternity were confirmed based on Whole Exome Family sequence data. Therefore, the recurrence risk to any future siblings of the patient is low, although there remains a small risk of recurrence due to the possibility of parental germline mosaicism. Genetic counseling is recommended.

MYH10 c.1470T>G, p.(Asn490Lys) is classified as a variant of uncertain significance (VUS), as there is insufficient evidence to evaluate its clinical relevance. It is currently unclear whether this variant is contributing to the phenotype of the patient. In this case, the variant has occurred *de novo*; it was not detected in the patient's parents. Maternity and paternity were confirmed based on Whole Exome Family sequence data. This variant should not be used for clinical decision-making or risk evaluation in family members. Management of the patient and family should be based on clinical evaluation and judgment. Genetic counseling is recommended.

The identified *MYH10* c.1470T>G, p.(Asn490Lys) variant is not eligible for the [VUS Clarification Service](#) at this time as family member testing is not sufficient to result in reclassification to likely pathogenic (please refer to our [variant classification schemes](#) on our website for additional information). The BpG VUS Clarification Service is offered when testing additional family members is likely to result in reclassification of the variant to likely pathogenic. Testing of the VUS in family members is available as part of our [Familial Variant Testing](#) service.

STEP	DATE
Order date	
Sample received	
Sample in analysis	
Reported	

On Feb 1st, 2021 the statement has been prepared by our geneticists and physicians, who have together evaluated the sequencing results:



Maria Calvo, Ph.D.

Geneticist



Juha Koskenvuo, MD, Ph.D.

Lab Director, Chief Medical Officer



Milja Kaare, Ph.D., CLG

Senior Geneticist

APPENDIX 3: SECONDARY FINDINGS

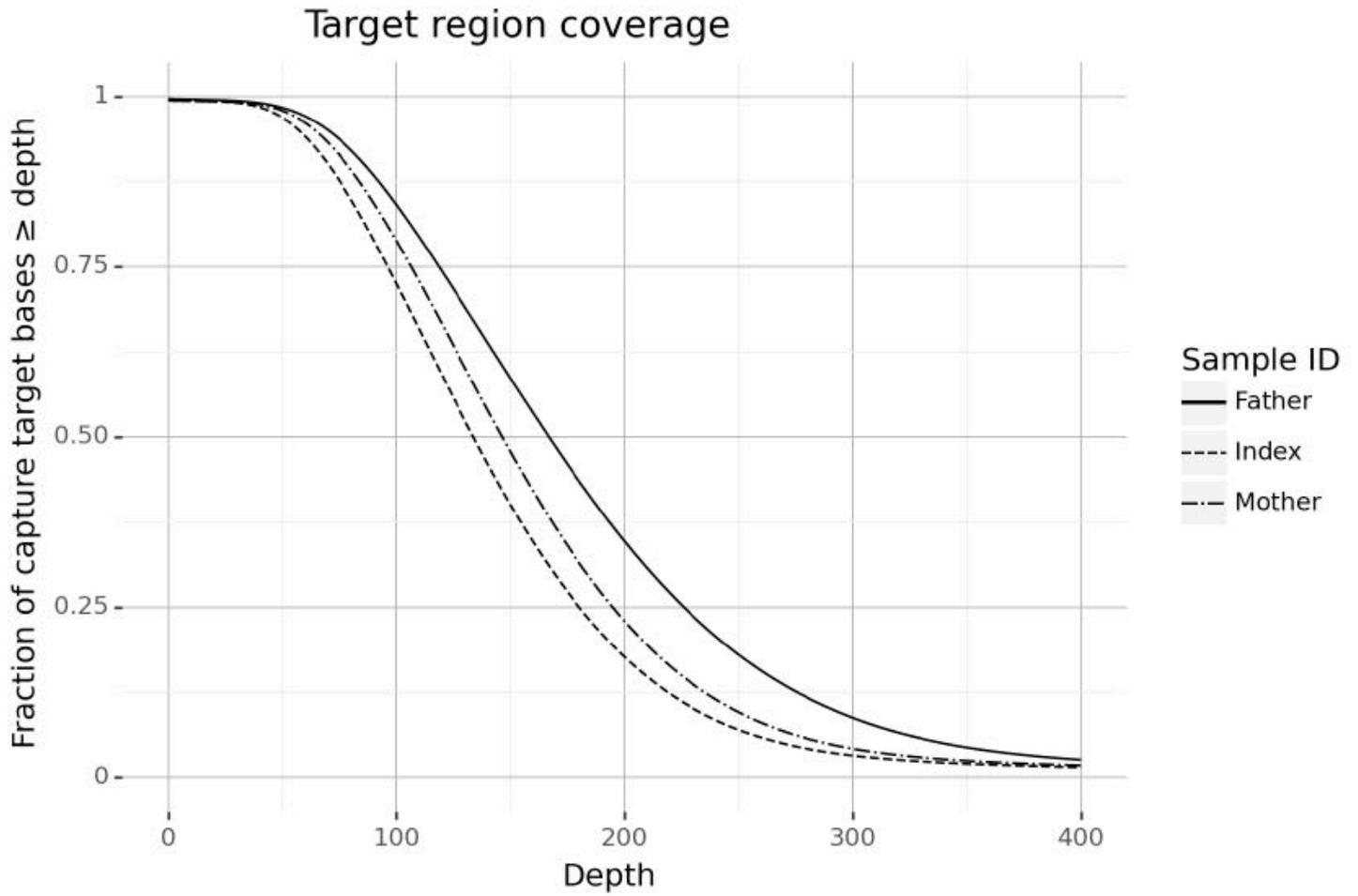
The patient was opted-in for an analysis of secondary findings, which are sequence variants unrelated to the indication for ordering the sequencing, but of medical value for patient care. Whole Exome data of the patient was analyzed for secondary findings in 59 genes according to recommendations of American College of Medical Genetics and Genomics (ACMG; PMID [27854360](#)).

NOTES REGARDING SECONDARY FINDINGS

The analysis was negative for secondary findings.

For any other family members who have opted in for secondary findings analysis, separate statements are available under the order ID in Nucleus.

Readability of the coverage plot may be hindered by faxing. A high quality coverage plot can be found with the full report on nucleus.blueprintgenetics.com.



APPENDIX 5: SUMMARY OF THE TEST

WHOLE EXOME

Laboratory process: When required, the total genomic DNA was extracted from the biological sample using bead-based method. DNA quality and quantity were assessed using electrophoretic methods. After assessment of DNA quality, qualified genomic DNA sample was randomly fragmented using non-contact, isothermal sonochemistry processing. Sequencing library was prepared by ligating sequencing adapters to both ends of DNA fragments. Sequencing libraries were size-selected with bead-based method to ensure optimal template size and amplified by polymerase chain reaction (PCR). Regions of interest (exons and intronic targets) were targeted using hybridization-based target capture method. The quality of the completed sequencing library was controlled by ensuring the correct template size and quantity, and to eliminate the presence of leftover primers and adapter-adapter dimers. Ready sequencing libraries that passed the quality control were sequenced using the Illumina's sequencing-by-synthesis method using paired-end sequencing (150 by 150 bases). Primary data analysis converting images into base calls and associated quality scores was carried out by the sequencing instrument using Illumina's proprietary software, generating CBCL files as the final output.

Bioinformatics and quality control: Base called raw sequencing data was transformed into FASTQ format using Illumina's software (bcl2fastq). Sequence reads of each sample were mapped to the human reference genome (GRCh37/hg19). Burrows-Wheeler Aligner (BWA-MEM) software was used for read alignment. Duplicate read marking, local realignment around indels, base quality score recalibration and variant calling were performed using GATK algorithms (Sentieon) for nDNA. Variant data for was annotated using a collection of tools (VcfAnno and VEP) with a variety of public variant databases, including but not limited to, gnomAD, ClinVar and HGMD. The median sequencing depth and coverage across the target regions for the tested sample were calculated based on MQ0 aligned reads. The sequencing run included in-process reference sample(s) for quality control, which passed our thresholds for sensitivity and specificity. The patient's sample was subjected to thorough quality control measures including assessments for contamination and sample mix-up. Copy number variations (CNVs), defined as single exon or larger deletions or duplications (Del/Dups), were detected from the sequence analysis data using a proprietary bioinformatics pipeline. The difference between observed and expected sequencing depth at the targeted genomic regions was calculated and regions were divided into segments with variable DNA copy number. The expected sequencing depth was obtained by using other samples processed in the same sequence analysis as a guiding reference. The sequence data was adjusted to account for the effects of varying guanine and cytosine content.

Interpretation: Our variant classification follows the [Blueprint Genetics Variant Classification Schemes](#) modified from the [ACMG guideline 2015](#). Minor modifications were made to increase the reproducibility of the variant classification and to improve the clinical validity of the report. Likely benign and benign variants were not reported. The pathogenicity potential of the identified variants were assessed by considering the predicted consequence, the biochemical properties of the codon change, the degree of evolutionary conservation as well as a number of reference population databases and mutation databases such as, but not limited, to the 1000 Genomes Project, [gnomAD](#), [ClinVar](#) and HGMD. For missense variants, *in silico* variant prediction tools such as SIFT, PolyPhen, MutationTaster were used to assist with variant classification. In addition, the clinical relevance of any identified CNVs was evaluated by reviewing the relevant literature and databases such as 1000 Genomes Project, [Database of Genomic Variants](#), ExAC, [DECIPHER](#). The clinical evaluation team assessed the pathogenicity of the identified variants by evaluating the information in the patient referral, reviewing the relevant literature and manually inspecting the sequencing data if needed. Reporting was carried out using HGNC-approved gene nomenclature and mutation nomenclature following the HGVS guidelines.

In addition to analysis of variants in previously established disease genes, variants in genes where disease association has not yet been established were considered as potentially disease-causing using the following scheme:

- For probands who were whole-exome sequenced with parents, all coding region de novo variants were considered as candidate variants.

- Novel (absent in gnomAD) heterozygous, truncating variants (nonsense, frameshift, canonical splice site variants) in genes predicted to be intolerant for loss-of-function variation based on ExAC variant data. Genes were determined as intolerant if probability of loss-of-function intolerance score $pLI \geq 0.9$. The closer pLI is to one, the more LoF intolerant the gene appears to be. Genes with $pLI \geq 0.9$ are defined as an extremely LoF intolerant set of genes.
- Rare ($<1\%$ MAF in gnomAD), truncating homozygous or (predicted) compound heterozygous variants, or a combination of rare truncating and rare missense variant that is predicted deleterious by multiple *in silico* tools.

In addition, only variants in genes whose known expression pattern and function are considered relevant for the phenotype are included (eg, variants in genes exclusively expressed in a muscular tissue are not considered as a candidate for a central nervous system disease). Candidate variants are not validated by Sanger sequencing, but their quality is inspected by visualization of sequence reads and evaluation of quality metrics, and only likely true variants are reported.

For proband and family members who were opted-in for analysis of secondary findings from the WES data, 59 clinically actionable genes were analyzed and reported for secondary findings according to recommendations by ACMG (PMID 27854360) with minor modifications aiming to increase the clarity of the classifications of the reportable variants (please see our website/clinical interpretation). Secondary findings are not analyzed or reported for deceased individuals or fetal samples.

Confirmation of sequence alterations: Sequence variants classified as pathogenic, likely pathogenic and variants of uncertain significance (VUS) were confirmed using bi-directional Sanger sequencing when they did not meet our stringent NGS quality metrics for a true positive call. In addition, prenatal case with diagnostic findings were confirmed.

Confirmation of copy number variants: CNVs (Deletions/Duplications) were confirmed using a digital PCR assay if they covered less than 10 exons (heterozygous), less than 3 exons (homo/hemizygous) or were not confirmed at least three times previously at our laboratory. Furthermore, CNVs of any size were not confirmed when the breakpoints of the call could be determined.

Analytic validation: This laboratory-developed test has been independently validated by Blueprint Genetics. The sensitivity of this panel is expected to be in the same range as the validated whole exome sequencing laboratory assay used to generate the panel data (sensitivity for SNVs 99.65%, and indels 1-50 bps 99.07%, one-exon deletion 92.3% and two exons CNV 100%, and specificity $>99.9\%$ for most variant types). It does not detect very low level mosaicism as a variant with minor allele fraction of 14.6% can be detected in 90% of the cases.

Test restrictions: A normal result does not rule out the diagnosis of a genetic disorder since some DNA abnormalities may be undetectable by the applied technology. Test results should always be interpreted in the context of clinical findings, family history, and other relevant data. Inaccurate, or incomplete information may lead to misinterpretation of the results.

Technical limitations: This test does not detect the following: complex inversions, gene conversions, balanced translocations, repeat expansion disorders unless specifically mentioned, non-coding variants deeper than ± 20 base pairs from exon-intron boundary unless otherwise indicated (please see the list of non-coding variants covered by the test). Additionally, this test may not reliably detect the following: low level mosaicism, stretches of mononucleotide repeats, indels larger than 50bp, single exon deletions or duplications, and variants within pseudogene regions/duplicated segments. The sensitivity of this test may be reduced if DNA is extracted by a laboratory other than Blueprint Genetics. Laboratory error is also possible. Please see the Analytic validation above.

Regulation and accreditations: This test was developed and its performance characteristics determined by Blueprint Genetics (see Analytic validation). It has not been cleared or approved by the US Food and Drug Administration. This analysis has been performed in a CLIA-certified laboratory (#99D2092375), accredited by the College of American Pathologists (CAP #9257331) and by FINAS Finnish Accreditation Service, (laboratory no. T292), accreditation requirement SFS-EN ISO 15189:2013. All the tests are under the scope of the ISO 15189 accreditation (excluding digital PCR confirmation).

Please refer to Appendix 8 of the report in Nucleus ordering and reporting portal for full list of non-coding variants included in the Whole Exome analysis.

Please refer to Appendix 7 of the report in Nucleus ordering and reporting portal for full list of non-coding variants included in the Whole Exome analysis.

GLOSSARY OF USED ABBREVIATIONS:

AD = autosomal dominant

AR = autosomal recessive

gnomAD = genome Aggregation Database (reference population database; >138,600 individuals)

gnomAD AC/AN = allele count/allele number in the genome Aggregation Database (gnomAD)

HEM = hemizygous

HET = heterozygous

HOM = homozygous

ID = rsID in dbSNP

MutationTaster = *in silico* prediction tools used to evaluate the significance of identified amino acid changes.

Nomenclature = HGVS nomenclature for a variant in the nucleotide and the predicted effect of a variant in the protein level

OMIM = Online Mendelian Inheritance in Man®

PolyPhen = *in silico* prediction tool used to evaluate the significance of amino acid changes.

POS = genomic position of the variant in the format of chromosome:position

SIFT = *in silico* prediction tool used to evaluate the significance of amino acid changes.

Transcript = GenBank accession for reference sequence used for variant nomenclature
