Whole Exome Family

REFERRING HEALTHCARE PROFESSIONAL

NAME	HOSPITAL

PATIENT

NAME	DOB	AGE	GENDER	ORDER ID
		1	Female	

PRIMARY SAMPLE TYPE	SAMPLE COLLECTION DATE	CUSTOMER SAMPLE ID
DNA		

SUMMARY OF RESULTS

TEST RESULTS

Analysis of variants in previously established disease genes

Patient is heterozygous for *ZBTB18* c.1390C>T, p.(Arg464Cys), which is classified as likely pathogenic. The variant has occurred *de novo*.

Patient is homozygous for *PAH* c.1139C>T, p.(Thr380Met), which is classified as pathogenic for non-phenylketonuria (mild) hyperphenylalaninemia. Parents are carriers.

GENETIC VARIANTS

VARIANT TABLE: Genetic alterations

GENE	POS	ID	INHERITANCE	CONSEQUENCE	TRANSCRIPT	DNA	PROTEIN	GENOTYPE	EXAC AF	EXAC AC/AN	CLASSIFICATION
ZBTB18	1:244218466	-	AD	missense_variant	NM_205768.2	c.1390C>T	p.(Arg464Cys)	HET (de novo)	-	-/-	Likely pathogenic
PAH	12:103237484	rs62642937	AR	missense_variant	NM_000277.1	c.1139C>T	p.(Thr380Met)	HOM	0.00032984	40/121272	Pathogenic

SEQUENCING PERFORMANCE METRICS

MEDIAN COVERAGE	MEAN COVERAGE	PERCENT >= 10X	PERCENT >= 15X	PERCENT >= 20X
Fa:162,Mo:139,Ch:131	Fa:180.74,Mo:157.76,Ch:146.52	Fa:99.7,Mo:99.5,Ch:99.4	Fa:99.5,Mo:99.3,Ch:99.2	Fa:99.3,Mo:99.1,Ch:98.9

TEST INFORMATION

Blueprint Genetics Whole Exome Family test (version 1, Sep 16, 2016) consists of sequence analysis of all protein coding genes in the genome for the proband and affected/unaffected family members. The test is targeting all protein coding exons and exon-intron boundaries. This diagnostic tool should be used to detect mutations such as single nucleotide substitutions and small insertions and deletions (INDELs). The test should not be used for detection of large copy number variations (CNVs), for analysis of sequence repeats or for diagnosis of disorders caused by mutations in the mitochondrial DNA.

Sequence analysis of the Whole Exome Family test is primarily focused on genes that have been previously associated with genetic disorders. The genes with known clinical association include those curated by Blueprint Genetics and included in Blueprint Genetics diagnostic panels (2051 genes). These genes are supplemented with genes included in the Clinical Genomics Database (3358 genes) and the Developmental Disorders Genotype-Phenotype Database (DDG2P) (1642 genes). Total number of genes that are considered as clinically associated in the Whole Exome analysis is 3593.

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

STATEMENT

CLINICAL HISTORY

The patient is a 21-month-old girl presenting with global developmental delay, speech delay, hypotonia, microcephaly and brachycephaly. Assessment for amino acids showed mildly increased phenylalanine levels (162.0) and lowered cysteine levels (29.0). Parents of the patient are consanguineous and there is no family history of the disease.

CLINICAL REPORT

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

Given that there is no family history of the disorder, the exome data of the patient and parents were primarily analysed for *de novo* variants or variants following recessive inheritance pattern.

Sequence analysis using the Blueprint Genetics (BpG) Whole Exome Family test identified a *de novo* heterozygous variant in *ZBTB18*: c.1390C>T, p.(Arg464Cys). The variant is not present in parents, with no clear indication of mosaicism: father has 331 sequence reads and mother 266 at the locus and none of the reads shows the variant. In the patient, the variant is present in 119 of total 252 reads (47%).

The c.1390C>T, p.(Arg464Cys) variant in *ZBTB18* has not been reported in the The Greater Middle East (GME) Variome Project database (http://igm.ucsd.edu/gme/), Exome Aggregation Consortium (ExAC) database or Genome Aggregation Database (gnomAD). *In silico* tools SIFT, PolyPhen and MutationTaster all predict it to be deleterious. The variant affects a highly conserved arginine residue in the zinc finger domain (C2H2-type 3) of the protein. The c.1390C>T, p.(Arg464Cys) variant has been reported as a *de novo* variant in one individual with severe intellectual disability in a recent article by Cohen and colleagues (PMID: 27598823). Also, in the Decipher database there is one patient with the same *de novo* variant, however, detailed phenotypic information is not available (https://decipher.sanger.ac.uk/patient/323435#genotype/snv/13666/browser). The article by Cohen and colleagues (PMID: 27598823) also refers to a patient sequenced as part of the DDD project with another missense variant affecting the same arginine residue, p.(Arg464His).

ZBTB18 (OMIM *608433; previously known as *ZNF238*) encodes a transcriptional repressor that is involved in various developmental processes such as myogenesis and brain development. *De novo* mutations in the gene have been recently shown to cause intellectual disability with variable features. The article by Cohen and others provides a recent review of a total of 15 patients with pathogenic variants in *ZBTB18* (PMID: 27598823). Pathogenic variants in *ZBTB18* include both truncating and missense variants, indicating that the phenotype is caused by haploinsufficiency. Notably, the gene is located within the 1q43-q44 deletion syndrome region, and individuals with *ZBTB18* mutations have similar symptoms to those with the deletion syndrome. Symptoms present in all patients with *ZBTB18* mutations include intellectual disability and speech delay. Motor delay and hypotonia have been reported in multiple individuals. MRI findings can be normal but corpus callosum hypoplasia has been noted in multiple patients. Absolute microcephaly has been reported in 6 of 15 patients. Craniofacial dysmorphism is present in approximately half of the patients, and behavioral features have been reported for a minority of the patients.

Sequence analysis using the Blueprint Genetics (BpG) Whole Exome Family test also identified a homozygous c.1139C>T, p.(Thr380Met) missense variant in *PAH*. The variant is present as heterozygous in 40 individuals in the ExAC population variation database (allele frequency 0.0003) and the allele frequency is 0.0025 in the GME Variome Project database. *In silico* tools SIFT, PolyPhen and MutationTaster all predict it to be deleterious. Several publications have reported the variant pathogenic for non-phenylketonuria (mild) hyperphenylalaninemia (e.g. PMIDs 7981714, 9399896, 21307867, 27620137).

The *PAH* gene encodes phenylalanine hydroxylase protein that catalyzes the hydroxylation of phenylalanine (OMIM *612349). If the hydroxylation is defected, the concentration of this amino acid can increase in toxic levels. Especially, nerve cells are sensitive to increased concentrations of phenylalanine and without strict dietary restrictions of phenylalanine most children with defected PAH develop irreversible intellectual disability. Phenylalanine hydroxylase deficiencies include phenylketonuria (PKU), non-PKU hyperphenylalaninemia (non-PKU HPA) and a variant PKU. Classical PKU (OMIM #261600, GeneReviews) is caused by a (nearly) complete PAH deficiency and accounts for more than 50% of all HPA cases, while mild variant PKU accounts for some 27% and mild non-PKU HPAs some 18% of all hyperphenylalaninemias (PMID: 26919687). Mild non-PKU HPA can be defined as the group of patients with abnormal phenylalanine concentrations not needing treatment (PMID: 21347590). However, there is no consensus whether this group should include only patients with untreated phenylalanine concentrations between 120 and 360 µmol/L, or also those with untreated phenylalanine concentrations between 360 and 600 µmol/L (PMID: 21347590, https://www.ncbi.nlm.nih.gov/books/NBK1504/). The various forms of PKU are inherited in an autosomal recessive manner.

Mutation nomenclature is based on GenBank accession NM_205768.2 (*ZBTB18*) and NM_000277.1 (*PAH*) 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

We classify the *ZBTB18* c.1390C>T, p.(Arg464Cys) variant as likely pathogenic, because it has occurred *de novo*, it is absent from population variant databases and it is predicted deleterious by multiple *in silico* tools. Moreover, literature and public disease variant databases include two patients with the same variant, and a third patient has a different missense variant p.(Arg464His) affecting the same amino acid residue. The symptoms of the patient show substantial overlap with those described for previously published *ZBTB18* patients.

The patient is homozygous for c.1139C>T, p.(Thr380Met) missense variant in *PAH*. The variant has been established as a pathogenic variant for non-phenylketonuria hyperphenylalaninemia, and therefore it likely explains the mildly increased phenylalanine levels observed in the patient.

Genetic counseling is recommended.

CONFIRMATION

ZBTB18 c.1390C>T, p.(Arg464Cys) and *PAH* c.1139C>T, p.(Thr380Met) were confirmed by bidirectional Sanger sequencing.

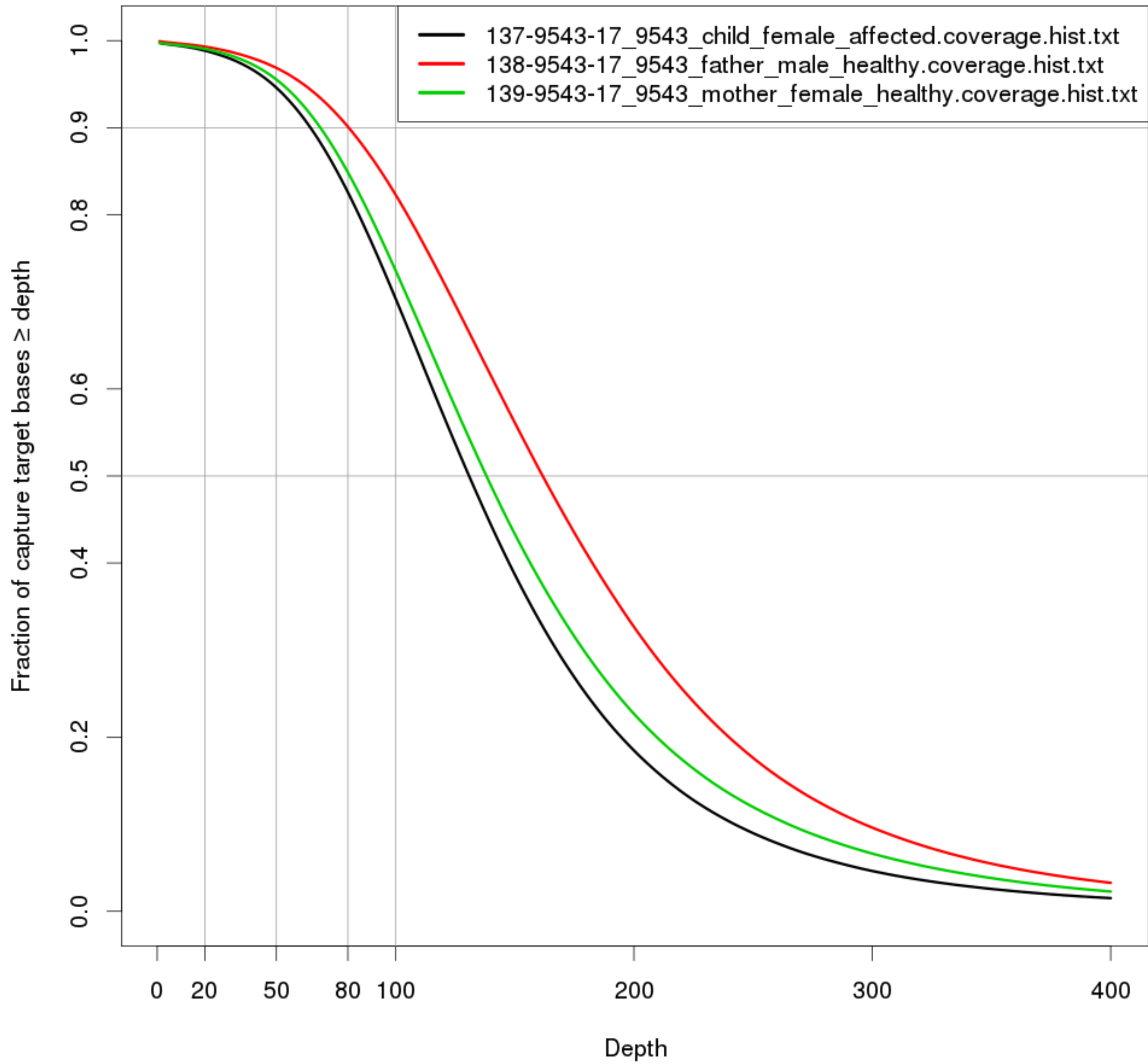
STEP	DATE
Order date	Jan 09, 2017
Sample received	Jan 09, 2017
Reported	Jan 30, 2017

On Mar 01, 2017 the statement has been prepared by our geneticists and physicians, who have together evaluated the sequencing results:

Sari Tuupanen, Ph.D.	Juha Koskenvuo, MD, Ph.D.	Helena Kääriäinen, MD, Ph.D.
Geneticist	Lab Director, Chief Medical Officer	Clinical Consultant

APPENDIX 2: SEQUENCING COVERAGE

COVERAGE PLOT



APPENDIX 3: SUMMARY OF METHODS

WHOLE EXOME

Total genomic DNA was extracted from the biological sample using a spin column method. DNA quality and quantity were assessed through gel electrophoresis and fluorometric analysis, respectively.

DNA sample was then shipped to BGI* where whole exome capture and sequencing were performed. After assessment of DNA quality at BGI, qualified genomic DNA sample was randomly fragmented by Covaris technology and the size of the library fragments was mainly distributed between 200bp and 300bp. Then adapters were ligated to both ends of the resulting fragments. Extracted DNA was amplified by ligation-mediated PCR (LM-PCR), purified, and hybridized to the exome array for enrichment (Agilent SureSelect Human All Exon V6). Non-hybridized fragments were then washed out. Captured LM-PCR products were subjected to Agilent 2100 Bioanalyzer and quantitative PCR was used to estimate the magnitude of enrichment. Each qualified captured library was then loaded on Illumina HiSeq platforms, and high-throughput sequencing was performed for each captured library while ensuring that each sample met the desired average sequencing coverage (>100x). Sequencing-derived raw image files were processed by Illumina basecalling software for base-calling with default parameters and the sequence data of each individual was generated as paired-end reads, which was defined as "raw data" and stored in FASTQ format. The bioinformatics analysis began with quality control of raw sequence reads. Clean sequence reads of each sample was mapped to the human reference genome (GRCh37/hg19). Burrows-Wheeler Aligner (BWA) software was used to do the alignment. Local realignment around indels and base quality score recalibration were performed using GATK, with duplicate reads removed by Picard tools. The sequencing depth and coverage for each individual were calculated based on the alignments.

After generation of exome sequencing data at BGI, data was delivered by a secure connection to Blueprint Genetics. Each exome batch was subjected to thorough quality control measures, after which raw sequence reads are transformed into variants by a proprietary bioinformatics pipeline. The pathogenic potential of variants was predicted by taking into account the predicted consequence, biochemical properties of the codon change, the degree of evolutionary conservation as well as allelic frequencies from large population studies, including e.g. data from the 1000 Genomes project, the ExAC consortium and ClinVar archive, as well as an in-house database of previously reported variants. The clinical association of genes was determined using our curated set of clinically associated genes included in Blueprint Genetics diagnostic panels, and in addition information from Clinical Genomics Database (https://research.nhgri.nih.gov/CGD) and The Developmental Disorders Genotype-Phenotype Database (DDG2P) (https://decipher.sanger.ac.uk/info/ddg2p) was used. The exome variant data of the family was filtered based on all possible modes of inheritance of the disorder. 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. Variant(s) classified as pathogenic or likely pathogenic were confirmed using Sanger sequencing. 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.
1. For probands who were whole-exome sequenced with parents, all de novo variants were considered as candidate variants.
2. Novel (not in ExAC, 1000G, EVS) 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 (see http://exac.broadinstitute.org/faq). Genes were determined as intolerant if probability of loss-of-function intolerance score pLI >= 0.9. The closer pLI is to one, the more LoF intolerant the gene appears to be. Genes with pLI >= 0.9 are defined as an extremely LoF intolerant set of genes.
3. Rare (<1% ExAC, 1000 genomes, EVS), 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 methods

In addition, for criteria 2 and 3, only variants in genes whose known expression pattern and function are considered relevant for the phenotype are included (e.g., 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 individuals who were opted-in for analysis of secondary findings from the whole exome data, 56 clinically actionable genes were analysed and reported for secondary findings according to recommendations by American College of Medical Genetics and Genomics (PMID: 23788249).

This laboratory-developed test has been independently validated by Blueprint Genetics. This test has not been cleared or approved by the FDA. A normal result by this test does not rule out the diagnosis of a disorder since some DNA abnormalities may be undetectable by this 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.

*BGI has been selected and audited as the laboratory generating whole-exome sequencing data for Blueprint Genetics Whole Exome products. BGI is one of the leading next-generation sequencing laboratories in the world, has CAP accreditation and a proven track record with providing next-generation sequencing for several international clinical collaborators. Blueprint Genetics ensures high quality and uniform sequence data for each exome by inclusion of stringent quality control steps in the analysis process.