Blueprint Genetics

Whole Exome Family Plus

THIS IS AN UPDATED REPORT. PLEASE SEE THE FOLLOW-UP REPORT TEXT BELOW. THIS REPORT SUPERSEDES THE REPORT ISSUED ON OCT 08, 2018

FOLLOW-UP REPORT

Follow-up Oct 12, 2018

KMT5B c.1840_1844del, p.(Ser614Thrfs*36) has been confirmed using bi-directional Sanger sequencing.

Milja Kaare, PhD, Geneticist

REFERRING HEALTHCARE PROFESSIONAL

NAME	HOSPITAL
<referring professional name>	<hospital name>

PATIENT

NAME	DOB	AGE	GENDER	ORDER ID
<patient name>		6	Female	

PRIMARY SAMPLE TYPE	SAMPLE COLLECTION DATE	CUSTOMER SAMPLE ID
DNA		

SUMMARY OF RESULTS

TEST RESULTS

Analysis of whole exome variants in previously established disease genes

Patient is heterozygous for *KMT5B* c.1840_1844del, p.(Ser614Thrfs*36), which is classified as pathogenic. The variant has occurred *de novo*.

Del/Dup (CNV) analysis

Del/Dup (CNV) analysis did not detect any known disease-causing copy number variation or novel or rare deletion/duplication that was considered deleterious.

VARIANT TABLE: GENETIC ALTERATIONS IN ESTABLISHED DISEASE GENES

GENE	POS	TRANSCRIPT	NOMENCLATURE	CONSEQUENCE	GENOTYPE	CLASSIFICATION
KMT5B	11:67925968	NM_017635.3	c.1840_1844del, p.(Ser614Thrfs*36)	frameshift_variant	HET	**Pathogenic**

ID	gnomAD AC/AN	POLYPHEN	SIFT	MUTTASTER
	0/0	N/A	N/A	N/A

OMIM	PHENOTYPE	INHERITANCE	COMMENT
	Mental retardation, Autism spectrum disorder overgrowth syndrome with intellectual disability	AD	-

Please see APPENDIX 3: Secondary Findings

SEQUENCING PERFORMANCE METRICS

SAMPLE	MEDIAN COVERAGE	PERCENT >= 20X
Index	248	99.48
Mother	208	99.24
Father	285	99.4

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 (± 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 a 6-year-old girl with global developmental delay, intellectual disability, dysmorphic features, macrocephaly, and hypotonia.

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 frameshift variant c.1840_1844del, p.(Ser614Thrfs*36) in *KMT5B*. The variant has occurred *de novo* and is not present in the patient's parents.

This variant has not been observed in the large reference population cohorts of the Genome Aggregation Database (gnomAD, n>120,000 exomes and >15,000 genomes). Database curators have made every effort to exclude individuals with severe pediatric diseases from these cohorts. The five base pair deletion generates a frameshift leading to a premature stop codon in the new reading frame (at codon 649 out of 885 aa). It is predicted to cause loss of normal protein function either through protein truncation or nonsense-mediated mRNA decay. To our knowledge, the variant has not been published in the relevant medical literature or reported in the disease-related variation databases such as ClinVar or HGMD.

KMT5B

Post-translational methylation and demethylation of lysine residues on histone tails is a key dynamic chromatin modification that is mediated by specific methyltransferases (KMTs) and demethylases (KDMs) and underpins gene regulation and several cellular processes. Heterozygous variants in seven KMT and four KDM genes have so far been associated with autosomal and X-linked dominant inherited human developmental disorders in OMIM (PMID: 29276005), including *KMT5B* (also known as *SUV420H1*; MIM *610881). In a large cohort of over 11,730 patients with autism spectrum disorder, intellectual disability, and/or developmental delay a *KMT5B* variant was identified in 7 unrelated patients with autosomal dominant mental retardation. Four of the variants were truncating and three were missense variants. Five of the variants were demonstrated to have occurred *de novo*. One missense variant was maternally inherited without clinical information on the mother, and parental DNA was not available from another patient to determine segregation. (PMID: 25363768, 28191889) Faundes et al. identified two *de novo* heterozygous loss-of-function variants and two *de novo* heterozygous gross deletions in the *KMT5B* gene in four unrelated patients with developmental delay. The patients presented with mild to severe global developmental delay and dysmorphic features including triangular face shape and broad forehead. (PMID: 29276005) To date, HGMD professional mutation database (2018.3) list 13 *KMT5B* variants in association to intellectual disability disorders. The majority of the identified variants are loss-of-function variants.

Mutation nomenclature is based on GenBank accession NM_017635.3 (*KMT5B*) 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

KMT5B c.1840_1844del, p.(Ser614Thrfs*36) is classified as pathogenic based on the current evidence of the variant (recently established association between the gene and the patient's phenotype, rarity in control populations, variant type (frameshift), variant has occurred *de novo*). Disease caused by *KMT5B* variants is inherited in an autosomal dominant manner. The variant has occurred *de novo*, as it was not observed in the parents. For disorders caused by *de novo* mutations, there is a low recurrence risk for the possible siblings because of the possibility of germline mosaicism. Genetic counseling is recommended.

CONFIRMATION

KMT5B c.1840_1844del, p.(Ser614Thrfs*36) has been confirmed using bi-directional Sanger sequencing.

STEP	DATE
Order date	Sep 20, 2018
Sample received	Sep 20, 2018
Reported	Oct 08, 2018
Last reviewed	Oct 12, 2018

On Oct 08, 2018 the statement has been prepared by our geneticists and physicians, who have together evaluated the sequencing results:



Milja Kaare, Ph.D.

Senior Geneticist

Juha Koskenvuo, MD, Ph.D.

Lab Director, Chief Medical Officer

Helena Kääriäinen, MD, Ph.D.

Clinical Consultant

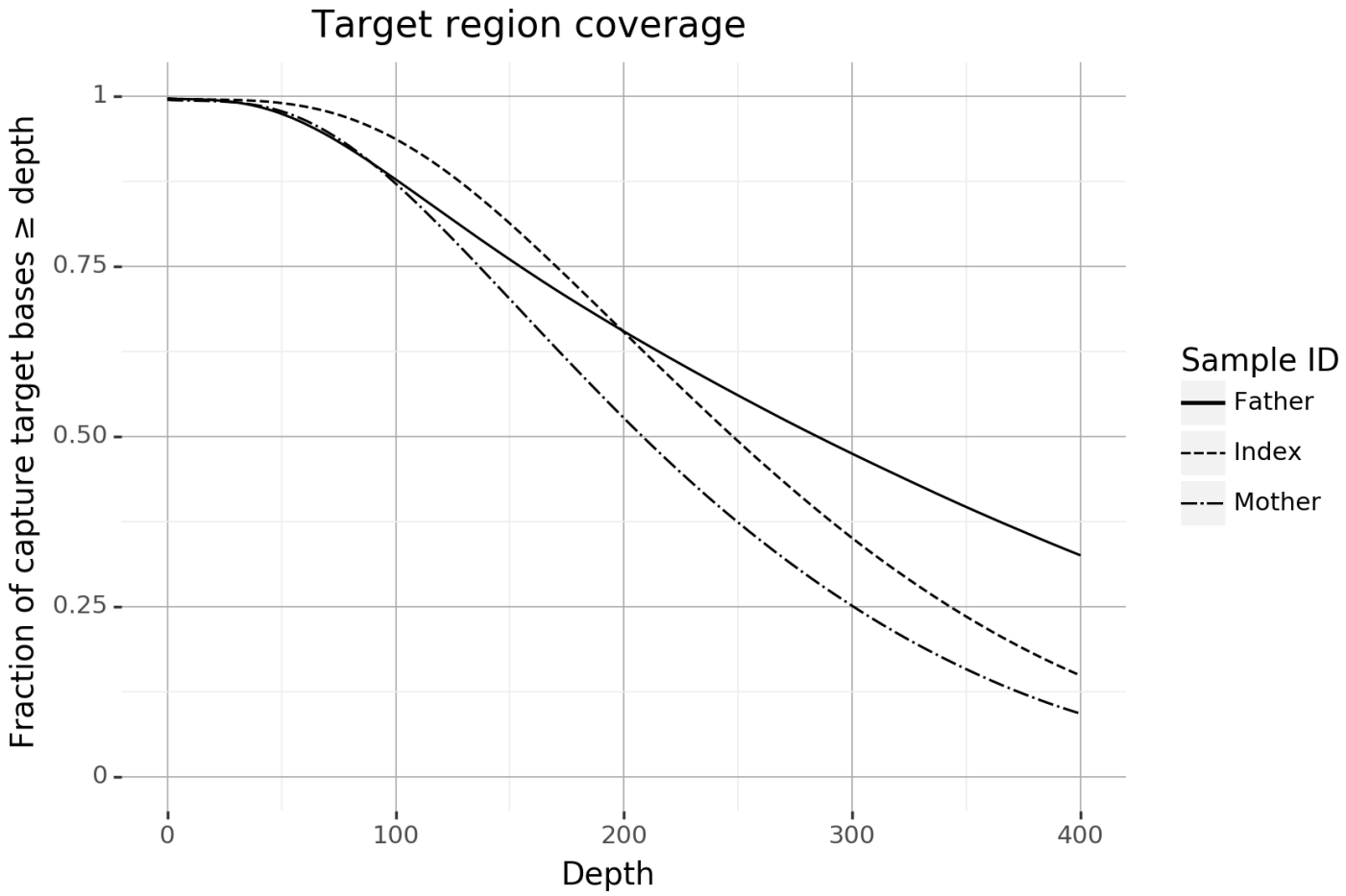
APPENDIX 3: SECONDARY FINDINGS

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).

For any other family members who have opted in for analysis of secondary findings, separate secondary findings statements are attached in Downloads section of Appendix.

NOTES REGARDING SECONDARY FINDINGS

The analysis was negative for secondary findings in the index patient,.


Target region coverage
Fraction of capture target bases ≥ depth
1
0.75
0.50
0.25
0
0
100
200
300
400
Depth
Sample ID
Father
Index
Mother

APPENDIX 5: SUMMARY OF THE TEST

WHOLE EXOME

Laboratory process: Total genomic DNA was extracted from the biological sample using a spin column 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 and purified with SPRI beads. DNA fragments were then end-repaired and sequencing adapters were ligated to both ends of the resulting fragments. Prepared DNA-Adapter libraries were size-selected with SPRI beads to ensure optimal template size and then amplified by ligation-mediated PCR (LM-PCR). The amplified sequencing library was purified using SPRI beads and a hybridization-capture method was applied for enrichment of whole exome and select non-coding regions (xGen Exome Research Panel with custom-designed capture probes, IDT). The enriched sequencing library was amplified by LM-PCR and purified using SPRI beads. The quality of the completed sequencing library was controlled by ensuring the correct template size and quantity and to eliminate the presence of leftover primer-dimers. Each captured library passing quality control was sequenced using the Illumina sequencing system with paired-end sequencing (150 by 150 bases). Sequencing-derived raw image files were processed using a base-calling software (Illumina) and the sequence data was transformed into FASTQ format.
Bioinformatics and quality control: The bioinformatics analysis began with quality control of raw sequence reads. Clean 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). The panel content was sliced from high-quality exome sequencing data acquired as presented above. The sequencing depth and coverage for the tested sample were calculated based on the alignments. 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 as well, after which raw sequence reads were processed into variants by a proprietary bioinformatics pipeline. 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, which processes aligned sequence reads. 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>=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.
- 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 (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 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).
Confirmation: Pathogenic and likely pathogenic variants that established a molecular diagnosis were confirmed with bi-directional Sanger sequencing unless all of the following criteria were fulfilled: 1) the variant quality score (QS) was above the internal threshold for a true positive call, 2) an unambiguous manual curation of the variant region using IGV was concordant with the variant call and 3) previous Sanger confirmation of the same variant has been performed at least three times in our laboratory. Reported variants of uncertain significance were confirmed with bi-directional Sanger sequencing only if the QS was below our internally defined score of a true positive call. CNVs (Dels/Dups) were confirmed using a quantitative-PCR assay if they covered less than 10 target exons or were not confirmed at least three times previously at our laboratory.
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). 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.
Regulation and accreditations: This test has not been cleared or approved by the FDA. 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.

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.

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