

CLINICAL NEXT GENERATION SEQUENCING REVEALS AN *H3F3A* GENE AS A NEW POTENTIAL GENE CANDIDATE FOR MICROCEPHALY ASSOCIATED WITH SEVERE DEVELOPMENTAL DELAY, INTELLECTUAL DISABILITY AND GROWTH RETARDATION

Maver A¹, Čuturilo G^{2,3}, Ruml Stojanović J³, Peterlin B^{1,*}

*Corresponding Author: Professor Borut Peterlin, Clinical Institute of Genomic Medicine, University Medical Center Ljubljana, Šljajmerjeva 4, 1000 Ljubljana, Slovenia. Tel: +38615401137. E-mail: borut.peterlin@kclj.si

ABSTRACT

Microcephaly is characterized by significant clinical and genetic heterogeneity, therefore reaching the genetic diagnosis remains challenging in this group of disorders. We describe a case of a girl with secondary microcephaly, associated with severe developmental delay, intellectual disability, growth retardation and dysmorphic features. For purposes of clinical genetic diagnostic testing, we performed trio whole exome sequencing in the proband and unaffected parents. We found a heterozygous *de novo* missense variant in the *H3F3A* gene in the proband (NM_002107.4: c.185T>G), which is absent from the gnomAD and from the Slovenian Genome databases. The identified variant affects a highly conserved leucine residue at position 62 of the histone H3 protein (H3.3) and is predicted to affect the physicochemical properties of the affected protein. Mouse models, which demonstrated involvement of H3.3 protein in the control of neuronal- and glial-specific gene expression patterns that control synaptic connectivity and behavioral plasticity. Additionally, we also identified similar cases reported in the ClinVar database. These arguments support the possible pathogenic role of the reported genetic variant and thus suggest a novel molecular mechanism for this syndromic form of microcephaly.

Keywords: Growth delay; *H3F3A* gene; intellectual disability; Microcephaly; Severe developmental delay

INTRODUCTION

The clinical sign of abnormally small head, usually defined as head circumference below the third percentile, or at least two standard deviations below the mean from an age- and sex-matched reference population, affects approximately 2.0-3.0% of the world population [1,2]. It is classified into primary if present at birth and secondary if developing thereafter [2]. Microcephaly can be an isolated sign or present as one of the associated features defining various genetic syndromes. Over 900 distinct phenotypes and about 800 genes have so far been recorded in the OMIM database [3], which implies significant clinical and genetic heterogeneity.

Consequently, establishing genetic etiology is challenging; genetic pathogenesis was reported in about 29.0% of cases based on metabolic and classical genetic testing [4], while using next generation sequencing (NGS) improved diagnostic yield. A recent study [3], employing molecular karyotyping and exome sequencing, improved diagnostic rates to 58.0% with primary microcephaly, 27.0% with secondary microcephaly, and to 15.0% in patients with unknown onset. Moreover, whole exome sequencing in the clinical setting may identify not only mutations in genes already known to be associated with investigated phenotypes, but also discover new potential genes and mechanisms for human disorders [5,6]. In this study, we describe a new association between microcephaly accompanied with severe developmental delay and intellectual disability and a potentially new gene *H3F3A* involved in maintaining heterochromatin and telomeric integrity.

CLINICAL REPORT

The proband is a girl, the only child born to healthy Caucasian, non consanguineous, healthy parents, who were at the time of birth 26 (mother) and 29 (father) years old.

¹ Clinical Institute of Genomic Medicine, University Medical Center Ljubljana, Ljubljana, Slovenia

² Faculty of Medicine, University of Belgrade, 11000 Belgrade, Serbia

³ Department of Clinical Genetics, University Children's Hospital, 11000 Belgrade, Serbia

The pregnancy was uneventful, with no evidence of infection, drug exposure or radiation. Delivery was at the 41st week of pregnancy by cesarean section due to the breech presentation. At birth, the proband weighed 2550 g, 4th centile, 46 cm, 4th centile and head circumference 33 cm, 5th centile. She was hypotonic and had slight transitory breathing problems with no malformations except small (3 mm), hemodynamically insignificant atrial septum defect. Magnetic resonance imaging (MRI) demonstrated hypoplasia of corpus callosum and cerebellum as well as thin layer of frontal and parietal periventricular gliosis. Visual evoked potentials demonstrated abnormal cortical responses with low amplitudes, suggesting a dysfunction of afferent optical system. At age 5, the proband weighed 16.5 kg, 10th centile, height of 94 cm, $p < 1$ st centile, and head circumference 45.5 cm, $p < 1$ st centile. She has severe global developmental delay, she is able to sit independently, but has not developed walking or speech. Her social interaction is poor. She has a wide, depressed nasal bridge, hypertelorism, hypotonic face, poorly formed, posteriorly set ears (Figure 1) and mild flexion contractures of fingers.

METHODS

Exome Sequencing. Exome sequencing was performed in the affected proband and unaffected parents. Briefly, the library preparation was performed using the TruSeq protocol (Illumina Inc., San Diego, CA, USA), followed by the capture of exonic targets using the xGen Exome Research Panel v1.0 using Integrated DNA Technologies (IDT), probes (Coralville, IA, USA). The exome capture targeted 39 Mb of exonic regions of 19,396 genes in the human genome (hg19). Sequencing was subsequently performed using the pair-end sequencing protocol on NextSeq 550 (Illumina Inc.) in 2×150 cycles. Sequencing data was processed using an in-house analysis pipeline, based on the combination of Burrows-Wheeler (BWA) aligner (v0.7.2) (<http://bio-bwa.sourceforge.net/>) and GATK software (v3.2) (<https://software.broadinstitute.org/gatk/>) for variant calling. Duplicate sequences were removed using Picard MarkDuplicates (<https://broadinstitute.github.io/picard/>), followed by base quality score recalibration, variant calling, variant quality score recalibration and variant filtering using the tools in the GATK framework [7].

Variant Analysis. The resulting variants were collected using the VariantTools software (<https://github.com/vatlab/variant-tools>) and their transcript and protein consequences predicted using the ANNOVAR (<http://annovar.openbioinformatics.org/en/latest/>) and SNPeff software (<http://snpeff.sourceforge.net/>) [8-10]. Variant consequences were predicted based on Refseq gene models (<https://www.ncbi.nlm.nih.gov/refseq/>). Precomputed



Figure 1. Facial features of the patient with a heterozygous *de novo* H3F3A gene variant p.(Leu62Arg) include wide and depressed nasal bridge, hypertelorism, hypotonic face and poorly formed, posteriorly set ears.

pathogenicity predictions for missense variants were obtained from the dbNSFP v2 database (<https://google.com/site/popgen/dbNSFP>) [11] and the MutationTaster, SIFT, Polyphen2, MetaSVM, CADD and REVEL (which are included in the dbNSFP database referred above) estimates were used in prioritization of pathogenic variants. Evolutionary conservation rates of the variant sites was based on GERP++ rejected substitution (RS) scores [12].

Variant frequency information for worldwide populations was based on the data from gnomAD project (gnomad.broadinstitute.com). We used the in-house population variant frequency estimates based on the internal data of 3000 Slovenian exomes as a source of variation frequency in the background population. We used ClinVar as the source of information on variant-disease associations [13].

Variant Filtration Strategy. The annotated variants were filtered using three strategies, based on the assumed pattern of inheritance, autosomal dominant (*de novo* variants), autosomal recessive (homozygous and compound heterozygous variants) and X-linked (*de novo* and maternally inherited variants). We used the frequency threshold of 0.01% in any of the surveyed populations in the dominant scenario and the frequency threshold of 0.1% for the autosomal recessive and X-linked scenarios. We performed manual interpretation of the remaining variants with consideration of clinical overlap, variant rarity in the general population, theoretical pathogenicity predictions

and evolutionary conservation. At this stage, all the variants were also manually inspected at read level to control the quality of the variant calls.

RESULTS

Exome analysis. Initially, we focused the search for causative variants in the proband to panels of genes, possibly associated with the clinical condition, including 1607 genes in an expanded developmental delay panel, 611 genes in an undiagnosed metabolic disorders panel and 155 genes in an intellectual disability panel. This initial search failed to reveal any causative variants.

Subsequently, we performed the search for candidate *de novo* and biallelic variants in novel genes. This revealed the presence of a heterozygous *de novo* missense variant in the *H3F3A* gene (NM_002107.4: c.185T>G) detected in the proband and absent in the parents. This variant is predicted to result in a p.(Leu62Arg) amino acid substitution in the Histone H3 protein.

The identified variant was absent in all populations of the gnomAD project and was also not detected in our in-house database of 3000 exomes. Furthermore, we observed a depletion of the overall rate of missense variants in the *H3F3A* gene in the gnomAD database, indicating an evolutionary constraint against missense variation in this gene, with only five missense variants in *H3F3A* observed in the gnomAD database populations (Z score for missense variant constraint in gnomAD dataset equaled 3.2).

The identified variant affects a highly conserved leucine residue at position 62 of the Histone H3 protein and the GERP rejected substitution score indicated high conservation rate (GERP RS score 4.97) of the variant nucleotide position across species. The substitution of the aliphatic hydrophobic leucine residue with a positively charged hydrophilic arginine residue affects the polarity at position 62 and thus predicted to affect the physicochemical properties of the affected protein domain in the proximity of the DNA binding site. All the utilized tools for *in silico* prediction of the variant's effect on protein function predict its damaging effect in consensus, including MutationTaster, SIFT, Polyphen2, MetaSVM, CADD and REVEL (Figure 2).

DISCUSSION

We report the *H3F3A* gene as a new candidate gene, associated with microcephaly, severe global developmental delay and intellectual disability. Additionally, prenatal and postnatal growth retardation, structural abnormalities of central nervous system and dysmorphological features were present in the reported patient.

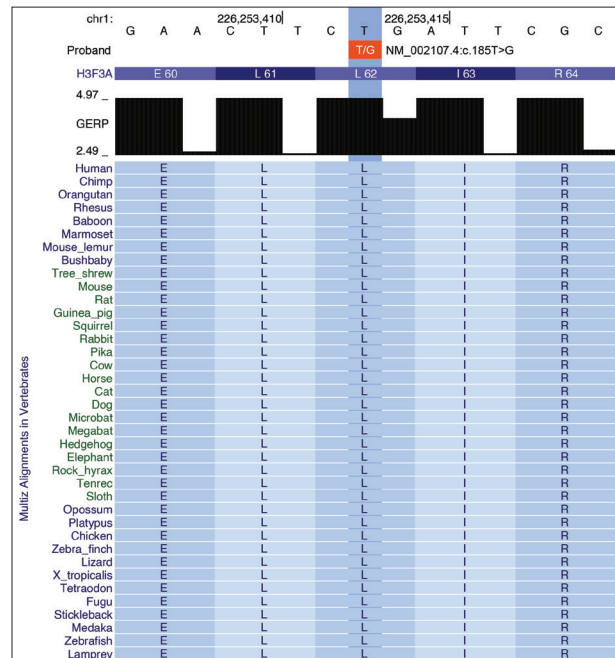


Figure 2. The identified *H3F3A* gene variant affects a highly evolutionary conserved leucine at amino acid position 62 in the Histone H3 protein, which is invariant in all the surveyed species (according to Multiz alignments). Figure was generated using the University of CA Santa Cruz (UCSC) genome browser (<https://genome.ucsc.edu/>).

The *H3F3A* gene, together with the *H3F3B* gene, codes the replacement histone H3.3. The H3.3-containing nucleosomes remain highly dynamic and are involved in the control of neuronal- and glial-specific gene expression patterns throughout life, which are required to support synaptic connectivity and behavioral plasticity [14]. H3.3 knock down resulted in decreased numbers of dendritic spines on pyramidal cells and global reduction of both excitatory and inhibitory synapses in the mouse model [14]. Disruption of the *H3F3A* gene in transgenic mice resulted in stunted growth and neuromuscular deficits [15]. Somatic mutations in the *H3F3* gene were reported in pediatric glioblastoma [16] and in giant cell tumors of the bone [17]. Furthermore, it was recently suggested that the pathogenic mechanism of germline histone mutations is distinct from that of the published cancer-associated somatic histone mutations, and may converge on control of cell proliferation [18].

In our patient, we identified a heterozygous *de novo* missense variant in the *H3F3A* gene (NM_002107.4: c.185 T>G), which is absent from the gnomAD and from the Slovenian Genome databases. The identified variant affects a highly conserved leucine residue at position 62 of the Histone H3 protein and is predicted to affect the physicochemical properties of the affected protein domain.

Further evidence for the potential role of the *H3F3A* gene in neurodevelopment is also based on the ClinVar reports of two patients with likely pathogenic heterozygous

missense mutations in the *H3F3A* gene, c.377A>G (p.Gln126Arg), who both demonstrated a neurological phenotype including hypotonia, global developmental delay, intellectual disability, visual impairment and dysmorphic features (ClinVar accession: VCV000520774.1) [13].

The limitation of our case report is lack of functional studies and lack of additional patients with mutations in the *H3F3A* gene and the suggested phenotype. Nevertheless, we provide evidence on a potentially new mechanism involving histones in the regulation of gene expression in the brain as the cause of a syndromic microcephaly.

Declaration of Interest. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Funding. This study was supported by the funding of Slovenian Research agency projects J3-8205 and J3-9280 and the framework programme P3-0326.

REFERENCES

1. Kalmin MM, Gower EW, Stringer EM, Bowman NM, McQuade ET, Westreich D. Misclassification in defining and diagnosing microcephaly. *Paediatr Perinat Ep.* 2019; 33(4): 286-390.
2. Woods GC, Parker A. Investigating microcephaly. *Arch Dis Child.* 2013; 98(9):707-713.
3. Boonsawat P, Joset P, Steindl K, Oneda B, Gogoll L, Azzarello-Burri S, *et al.* Elucidation of the phenotypic spectrum and genetic landscape in primary and secondary microcephaly. *Genet Med.* 2019; 21(9): 2043-2058.
4. Hagen M, Pivarsci M, Liebe J, Bernuth H, Didonato N, Hennermann JB, *et al.* Diagnostic approach to microcephaly in childhood: A two center study and review of the literature. *Dev Med. Child Neurol.* 2014; 56(8): 732-41.
5. Maver A, Čuturilo G, Kovanda A, Miletić A, Peterlin B. Rare missense TUBGCP5 gene variant in a patient with primary microcephaly. *Eur J Med Genet.* 2019; 65(12): 103598.
6. Tumienè B, Maver A, Writzl K, Hodžić A, Čuturilo G, Kuzmanić Šamija R, *et al.* Diagnostic exome sequencing of syndromic epilepsy patients in clinical practice. *Clin Genet.* 2018; 93(5):1057-1062.
7. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011; 43(5): 491-498.
8. Lucas AF, Wang G, Scheet P, Peng B. Integrated annotation and analysis of genetic variants from next-generation sequencing studies with variant tools. *Bioinformatics.* 2012; 28(3): 421-422.
9. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010; 38(16): e164-e164.
10. Cingolani P, Platts A, Wang L, Coon M, Nguyen T, Wang L, *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin).* 2012; 6(2): 80-92.
11. Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: A database of human non synonymous SNVs and their functional predictions and annotations. *Hum Mutat.* 2013; 34(9): E2393-E402.
12. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol.* 2010; 6(12): e1001025.
13. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, *et al.* ClinVar: Public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* 2014; 42(Database issue): D980-D985.
14. Maze I, Wenderski W, Noh K-M, Bagot RC, Tzavaras N, Purushothaman I, *et al.* Critical role of histone turnover in neuronal transcription and plasticity. *Neuron.* 2015; 87(1): 77-94.
15. Couldrey C, Carlton M, Nolan P, Colledge W, Evans M. A retroviral gene trap insertion into the Histone 3.3A gene causes partial neonatal lethality, stunted growth, neuromuscular deficits and male sub-fertility in transgenic mice. *Hum Mol Genet.* 1999; 8(13): 2489-2495.
16. Schwartzenuber J, Korshunov A, Liu X-Y, Jones DT, Pfaff E, Jacob K, *et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature.* 2012; 482(7384): 226-231.
17. Behjati S, Tarpey PS, Presneau N, Scheipl S, Pillay N, Loo P, *et al.* Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat Genet.* 2013; 45(12): 1479-1482.
18. Nair DR, Bhoj E. 3548 De novo germline variants in Histone 3 Family 3A (H3F3A) and Histone 3 Family 3B (H3F3B) cause a severe neurodegenerative disorder and functional effects unique from their somatic mutations. (Abstract). *J Clin Transl Sci.* 2019; 3(Suppl 1): 103. doi: 10.1017/cts.2019.235. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6799666/>).