RESEARCH NOTE

Molecular genetic analysis of consanguineous families with primary microcephaly identified pathogenic variants in the *ASPM* gene.

MUZAMMIL AHMAD KHAN[1, 2, *], CHRISTIAN WINDPASSINGER[3, *], MUHAMMAD ZEESHAN ALI[2], MUHAMMAD ZUBAIR[2, 4], HADIA GUL[5], SAFDAR ABBAS[2], SAADULLAH KHAN[1, 6], MUHAMMAD BADAR[2], RAMZI M MOHAMMAD[1], ZAFAR NAWAZ [1, 7]

1 *Translational Research Institute, Academic Health System, Hamad Medical Corporation, Doha, 3050, Qatar*

2 *Gomal Centre of Biochemistry and Biotechnology, Gomal University Dera Ismail Khan, Khyber-Pakhtoonkhwa, 29050, Pakistan*

3 *Institute of Human Genetics, Medical University of Graz, Graz 8010, Austria*

4 *Department of Cell and Developmental Biology, School of Life Sciences, University of Science and Technology China*

5 *Department of Biological Sciences, Faculty of Sciences, Gomal University Dera Ismail Khan, Khyber-Pakhtoonkhwa, 29050, Pakistan*

6 *Department of Biotechnology & Genetic Engineering, Kohat University of Science & Technology, Kohat, Khyber Pakhtunkhwa, Pakistan*

7 *Diagnostic Genomic Division,* Department of Laboratory Medicine & Pathology, *Hamad Medical Corporation, Doha, 3050, Qatar*

***** These authors have equally contributed in this work

Running Title:

Molecular analysis of MCPH genes

Keywords: Primary Microcephaly, Pakistani Families, Exome Sequencing, *ASPM*, Armadillo Type fold

Introduction

Autosomal recessive primary microcephaly is a rare genetic disorder that is characterized by reduced head circumference and a varying degree of intellectual disability. Genetic studies on consanguineous families with primary microcephaly have identified 15 (MCPH) causative genes that include *MCPH1, WDR62, CDK5RAP2, CASC5, ASPM, CENPJ, STIL, CEP135, CEP152, ZNF335, PHC1, CDK6, CENPE, SASS6* and *MFSD2A* (Alakbarzade *et al.* 2015, Khan *et al.* 2014, Morris-Rosendahl and Kaindl 2015). Physiologically, most of these MCPH proteins are involved in cell cycle and its regulation. Herein the present clinical genetic study, we present two consanguineous Pakistani families segregating primary microcephaly and intellectual disability. These families were ascertained from the Saraiki ethnic part of Khyber-Pukhtunkhwa province in Pakistan. Whole exome sequencing in one family revealed a novel one bp deletion NM_018136.4: c.10013delA (p.Asp3338Valfs*2), while the other family showed a previously reported nonsense mutation NM_018136.4: c.9730C>T {rs199422195 (p.Arg3244*)} in *ASPM* gene. The novel frame-shift mutation (p.Asp3338Valfs*2) in *ASPM* presumably truncates the protein synthesis that results in loss of armadillo type fold domain.

In the broad sense, most of the identified MCPH genes have either structural or physiologic role in cell cycle and its regulation, neurogenesis and ciliogenesis (Barbelanne & Tsang, 2014). Morphologically, microcephalic patients have normal brain architecture and the reduce brain volume is due to small cerebral cortex (Woods *et al.* 2005). The genetic studies on cohorts and consanguineous families have shown that *ASPM* (MCPH5 locus) and *WDR62* (MCPH2) are the most frequent genes reported in primary microcephaly (Mahmood *et al.* 2011). At cellular level, ASPM is required for normal functioning of mitotic

spindle in embryonic neuroblasts. The *Aspm* knockout studies in mice have shown that reduction in brain size is due to mis-orientation of mitotic spindle fibers, which also affects the ratio of symmetric to asymmetric cell division and thus decrease the number of neuronal cells (do Carmo Avides & Glover, 1999).

Materials and Methods

Family recruitment and Sample collection

Prior approval was obtained from ethical review board of Gomal Centre of Biochemistry and Biotechnology, Gomal University, D.I.Khan, Pakistan, to conduct the present molecular study, and the families were enrolled after taking the informed written consent. The exome data analysis was performed in Hamad Medical Corporation, Doha, Qatar with the approval of data sharing agreement. For molecular investigation, we were able to collect the peripheral blood samples from individuals III-1, IV-1, IV-3 and IV-5 of family A, while individuals IV-3, IV-4, V-4, III-1, III-2 and V-3 were among the volunteer in family B. The DNA was isolated from these samples using standard phenol-chloroform assay.

Whole Exome Sequencing and prioritization of candidate pathogenic variant

Paired end whole exome sequencing was performed on NextSeq500 plate-form (Illumina, USA), using Nextera rapid capture exome kit for library construction, following the manufacturer protocol. The raw data analyses were accomplished through BaseSpace integrated applications to get BAM and annotated VCF files. Subsequently, the variant filtration was carried out on genetalk software to identify the pathogenic variants (Kamphans and Krawitz 2012). In addition to that, computational based prediction of most probable candidate genes/variants and its association with disease phenotype was performed through exomewalker (Smedley *et al.* 2014), Exomiser (Robinson *et al.* 2014) and PhenIX softwares

(Zemojtel *et al.* 2014). These softwares perform phenotype and protein interaction based prioritization of most plausibly implicated genes.

Sanger sequencing and mutation analysis

For subsequent mutation detection and segregation analysis, Sanger DNA sequencing of *ASPM* gene was performed. Primers sequences were designed through Primer3web software, v.4.0.0 (URL: http://bioinfo.ut.ee/primer3/). DNA sequencing was carried out by using BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems) and the products were analyzed on ABI 3130*xl* Genetic Analyzer (ABI, USA).

Results

Herein this study, two consanguineous families were recruited from the rural areas of Pakistan. The genealogical analysis of both families revealed their Saraiki based ethnicity. Pedigree analysis demonstrated autosomal recessive mode of disease segregation. The general clinical description of patients, from both families, presented non-syndromic congenital primary microcephaly. The patients presented reduced occipitofrontal head circumference along with moderate intellectual disability. The patients had severe speech disability, and thus were facing difficulty in conveying their message. Psychologically, the patients were having aggressive attitude due to which they had tendency of self beating behavior. The patients never attended any training institute and were unable to perform arithmetic measures or any conceptual task, but still they had milder category of recognition power, like recognizing route to their home and family relatives. The additional clinical features of all families are described in table 1.

Molecular findings:

Exome sequence analysis of individual IV-3 and the subsequent segregation study through Sanger sequencing (III-1, IV-1, IV-3 and IV-5) identified a novel deletion mutation NM_018136.4:c.10013delA in the 26th exon of *ASPM*. The identified pathogenic variant was also prioritized by the candidate gene identification softwares using exome variant file (VCF file). Protein sequence analysis revealed that this mutation (p.Asp3338Valfs*2) is present in the armadillo-type fold domain. This out-of-frame deletion shifts the reading frame and truncate the protein synthesis (p.Asp3338Valfs*2) due to premature stop codon and produces 3338 amino acid shortened protein with deleted armadillo type fold domain. The identified sequence variant is not present in 100 normal controls, 1000 genome browser (http://www.1000genomes.org/1000-genomes-browsers), exome variant server (http://evs.gs.washington.edu/EVS/) and exome aggregation consortium databases (http://exac.broadinstitute.org/). The mutation analysis of second family (IV-4 and V-4) revealed the previously reported nonsense mutation c.9730C>T {rs199422195 (p.Arg3244*)} in 24th exon (NM_018136.4) of *ASPM* gene (Figure 1).

Discussion

Abnormal spindle-like microcephaly associated protein (ASPM) is encoded by *ASPM* gene that spans 62.567 Mb long region on chromosome 1q31.3 [Dec. 2013 (GRCh38/hg38)]. Its longest transcript (NM_018136) consists of 28 exons and encodes for a protein with 3477 amino acids. ASPM is reported to be crucial for the normal functioning of mitotic spindle fibers during embryonic neuroblasts formation. The ASPM protein has several functional domains: the calponin homology domains, P-loop containing nucleoside triphosphatase hydrolase domains (P-loop NTPase), armadillo-type fold domain and IQ motifs (Interpro accession ID: Q8IZT6). Among these functional domains the calponin homology domain is associated with regulation of contractility and organization of actin cytoskeleton in smooth

muscles (Bramham *et al.* 2002), the P-loop NTPase domains performs hydrolysis of β-γ phosphate bond in nucleoside triphosphate (NTP) (Leipe *et al.* 2004), the IQ motif acts as an EF-hand binding site and some time as a phosphorylation site of protein kinase C (Baudier *et al.* 1991) and the armadillo-type fold domain is suggestively involved in associating cadherin to the cytoskeleton (Huber *et al.* 1997).

Literature survey has shown that mutations in ARM domain have previously been reported to be implicated in MCPH (Gul *et al.* 2007; Nicholas *et al.* 2009). ARM domain is a 42 amino acid conserved entity that is present in various proteins, e.g. beta-catenin, alpha-importin, plakoglobin, which together make a distinct class of protein family called ARM-repeat family proteins. Tandem repetition of ARM repeats makes a superhelical fold which provide a versatile interacting platform to the protein and therefore ARM domain containing proteins shown to have several independent molecular and biological functions in the cell (Coates 2003). String matching analysis using protein-protein interaction (STRING, protein interaction database; URL: http://string-db.org/) did not show any significant change in the interaction of ASPM with its known interactors, which probably indicate the involvement of some other protein interactors involved in the neurogenesis pathway with whom ASPM interacts via its ARM domain.

Conclusion:

We performed mutation analysis of *ASPM* in two consanguineous Pakistani families presenting primary microcephaly. One family revealed a novel single base deletion of adenine nucleotide (c.10013delA) in 26th exon of *ASPM*, while the other family exhibited a reported mutation (c.9730C>A) in its 24th exon. The novel frameshift mutation presumably leads to a truncated protein with the deletion of an armadillo-type fold domain. The study added a new pathogenic variant in the mutational spectrum of armadillo-type fold domain of

ASPM. This will help in elucidating the novel protein interactors involved in neurogenesis pathway. Moreover; screening of additional Saraiki ethnic families may contribute in devising a molecular diagnostic test that might be helpful in genetic counseling of non-consanguineous families.

Acknowledgements:

We pay special thanks to all family members for their volunteer participation in this study. The current study is partially supported by startup research grant from Higher Education Commission of Pakistan (*M-IPFP/HRD/HEC/2011/346*) for molecular screening of neurological disorders. MAK appreciates the valuable contribution of his beloved nephew "Abdul Ahad Khan Gandapur" in recruiting the family.

Conflict of interest: Authors declared that they have no competing or financial interest

References

1. Alakbarzade V., Hameed A., Quek D. Q. Y., Chioza B. A., Baple E. L., Cazenave-Gassiot A., et al. 2015 A partially inactivating mutation in the sodium-dependent lysophosphatidylcholine transporter MFSD2A causes a non-lethal microcephaly syndrome. *Nature. Genet.* **47**, 814-817.
2. Barbelanne M. and Tsang W. Y. 2014 Molecular and cellular basis of autosomal recessive primary microcephaly. *Biomed. Res. Int.* 2014, 547986.
3. Baudier J., Deloulme J. C., Van Dorsselaer A., Black D. and Matthes H. W. 1991 Purification and characterization of a brain specific protein kinase C substrate, neurogranin (p17). Identification of a consensus amino acid sequence between neurogranin and neuromodulin (GAP43) that corresponds to the protein kinase C phosphorylation site and the calmodulin-binding domain. *J. Biol. Chem.* **266**, 229-237.

4. Bramham J., Hodgkinson J. L., Smith B. O., Uhrín D., Barlow P. N. and Winder S. J. 2002 Solution structure of the calponin CH domain and fitting to the 3D-helical reconstruction of F-actin:calponin. *Structure.* **10**, 249-258.
5. Coates J. C. 2003 Armadillo repeat proteins: beyond the animal kingdom. *Trends. Cell. Biol.* **13**, 463-471.
6. do Carmo Avides M. and Glover D. M. 1999 Abnormal spindle protein, Asp, and the integrity of mitotic centrosomal microtubule organizing centers. *Science.* **283**, 1733–1735.
7. Gul A., Tariq M., Khan M. N., Hassan M. J., Ali G. and Ahmad W. 2007 Novel protein-truncating mutations in the ASPM gene in families with autosomal recessive primary microcephaly. *J. Neurogenet*. **21**, 153-163.
8. Huber A. H., Nelson W. J. and Weis W. I. 1997 Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell*. **90**, 871-882.
9. Kamphans T. and Krawitz P. M. 2012 GeneTalk: an expert exchange platform for assessing rare sequence variants in personal genomes. *Bioinformatics*. **28**, 2515–2516.
10. Leipe D. D., Koonin E. V. and Aravind L. 2004 STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multip le,complex domain architectures, unusual phyletic patterns,and evolution by horizonta l gene transfer. *J. Mol. Biol.* **343**, 1-28.
11. Mahmood S., Ahmad W. and Hassan M. J. 2011 Autosomal recessive primary microcephaly (MCPH): clinical manifestations, genetic heterogeneity and mutation continuum. *Orphanet. J. Rare. Dis.* **6**, 39.
12. Morris-Rosendahl D. J. and Kaindl A. M. 2015 What next-generation sequencing (NGS) technology has enabled us to learn about primary autosomal recessive microcephaly (MCPH). *Mol. Cell. Probes.* **29,** 271-281.

13. Nicholas A. K., Swanson E. A., Cox J. J., Karbani G., Malik S., Springell K., et al 2009. The molecular landscape of ASPM mutations in primary microcephaly. *J. Med. Genet.* **46**, 249-253.
14. Robinson P. N., Köhler S., Oellrich A., Sanger. Mouse. Genetics. Project., Wang K., Mungall C. J., et al 2014 Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome. Research.* **24**, 340-348.
15. Seelow D. and Schuelke M. 2012 HomozygosityMapper2012-bridging the gap between homozygosity mapping and deep sequencing. *Nucleic. Acids. Res.* **40**(Web Server issue), W516-20.
16. Smedley D., Köhler S., Czeschik J. C., Amberger J., Bocchini C., Hamosh A., et al 2014. Walking the interactome for candidate prioritization in exome sequencing studies of Mendelian diseases. *Bioinformatics.* **30**, 3215-3222.
17. Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B. C., Remm M., et al 2012. Primer3--new capabilities and interfaces. *Nucleic. Acids. Res.* **40**, e115.
18. Woods C. G., Bond J. and Enard W. 2005 Autosomal recessive primary microcephaly (MCPH): a review of clinical, molecular, and evolutionary findings. *Am. J. Hum. Genet.* **76**, 717-728.
19. Zemojtel T., Köhler S., Mackenroth L., Jäger M., Hecht J., Krawitz P., et al 2014. Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome. *Sci. Transl. Med.* **6**, 252ra123.

Received 7 June 2016, in final revised form 1 August 2016; accepted 7 September 2016

Unedited version published online: 12 September 2016

Table 1: Genotype associated phenotypic descriptions of consanguineous microcephalic families

Family ID	**Family A**	**Family B**		
Mutation	c.10013delA (p.Asp3338Valfs*2)	c. 9730C>T (p.Arg3244*) **(rs199422195)**		
Individual ID	IV-3	IV-3	IV-4	V-4
Age (Years)	21 years	19	24	6
Sex	Male	Male	Male	Male
Head Circumference	41 cm	46 cm	46 cm	39 cm
Intellectual disability	Yes	Yes	Yes	Yes
Speech disability	Yes	Yes	Yes	Yes
Epileptic shocks	No	No	No	No
Muscular dystrophy	No	No	No	No
Skeletal abnormality	No	No	No	No
Neurologic defect	No	No	No	No
Behavioral expression	Hyperactive with Jolly mood	Hyperactive with Jolly mood	Hyperactive with Jolly mood	Hyperactive with Jolly mood
Ocular defect	No	No	Strabismus	No
Dermal lesions	No	No	No	No
Any visceral organ defect	No	No	No	No

Figure 1: A: A four generational consanguineous Pakistani family inheriting autosomal recessive primary microcephaly. The sequence chromatograms depict the homozygous deletion of adenine nucleotide in patient IV-2 (c.10013delA), heterozygous status in carrier IV-3 and reference allele in IV-5. **B**: The graphical presentation of family's tree with reported *ASPM* mutation (c.9730C>A).