Genetic causes of MCPH in consanguineous Pakistani families

Kraemer, N. and Picker-Minh, S. and Abbasi, A.A. and Froehler, S. and Ninnemann, O. and Khan, M.N. and Ali, G. and Chen, W. and Kaindl, A.M.

Clinical Genetics
2016 JUN; 89(6): 744-745
2015 NOV 08 (first published online)
doi: 10.1111/cge.12685

Publisher: Wiley-Blackwell

Genetic causes of MCPH in consanguineous Pakistani families

Nadine Kraemer[1-3#], Sylvie Picker-Minh[1,2#], Ansar A. Abbasi[4,5#], Sebastian Fröhler[6#], Olaf Ninnemann[1], Muhammad N. Khan[7], Ghazanfar Ali[7], Wei Chen[6], Angela M. Kaindl[1-3*]

[1]Institute of Cell Biology and Neurobiology, Charité – University Medicine Berlin, Germany
[2]Department of Pediatric Neurology, Charité - University Medicine Berlin, Germany
[3]Sozialpädiatrisches Zentrum (SPZ), Charité – University Medicine Berlin, Germany
[4]Department of Zoology, University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan
[5]Department of Zoology, Mirpur University of Science and Technology (MUST), Mirpur, Pakistan
[6]Genomics Core, Berlin Institute of Health, Berlin, Germany
[7]Department of Biotechnology, University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan

[#]equal contribution

*Corresponding author:
Dr. Angela M. Kaindl, Institute of Cell Biology and Neurobiology, Charité – Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany. Tel / Fax: +49 30 450 528008 / 7566301. Email: angela.kaindl@charite.de.

Running title: Pakistani MCPH mutations

Conflict of interest statement

The authors declare no conflict of interest in the preparation or publication of the data in this manuscript.

Acknowledgements

The authors thank all families for participation in this study. This work was supported by the German Research Foundation (SFB665), the Berlin Institute of Health (BIH), the Charité, and the Higher Education Commission Pakistan.

Letter to the editor

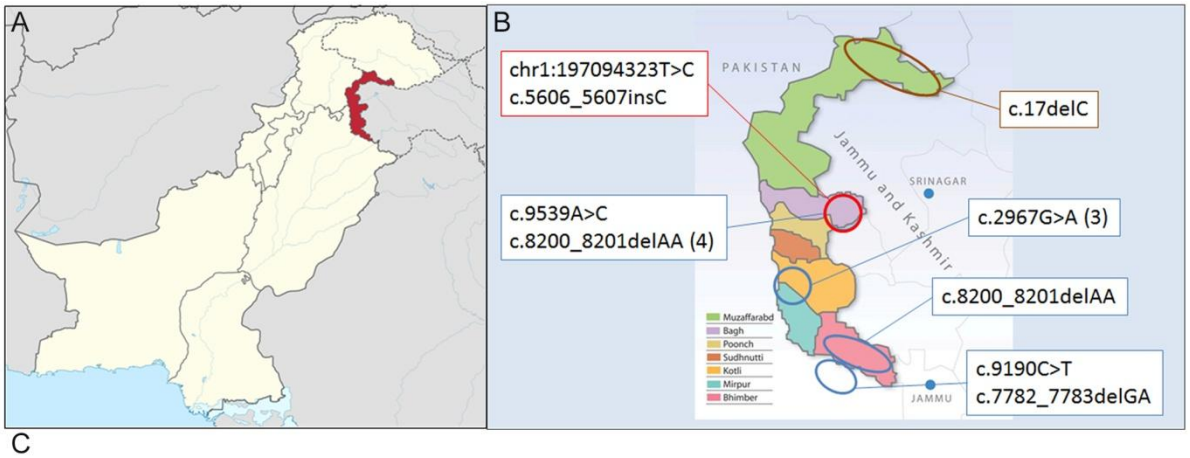
Autosomal recessive primary microcephaly (MCPH) is a rare and genetic heterogeneous neurodevelopmental disorder characterized by a pronounced reduction of brain volume and intellectual disability (1, 2). To date, 13 MCPH genes have been identified (reviewed in (3)), eight of them in the last five years, showing the rapid progress in this field. Still, families with an MCPH phenotype who do not carry mutations in these known MCPH genes exist, indicating further genetic heterogeneity. Here, we report the results of a genetic study on a cohort of 15 consanguineous families (46 affected individuals) with an MCPH phenotype from the Azad Jammu and Kashmir (AJK) territory in Northern Pakistan (Table 1, Supplemental Figure 1). All families were examined clinically on field trips with microcephaly and intellectual disability detected in all affected individuals (Supplemental Table 1), but no other obvious malformations. Applying whole exome sequencing, according to methods described previously (4), followed by Sanger sequencing, we identified mutations in the four known MCPH genes *ASPM* (MIM*605481), *CENPJ* (MIM*609279), *MCPH1* (MIM*607117), and *STIL* (MIM*181590) in these families. We detected one novel homozygous frameshift mutation (c.5606_5607insC, p.H1870Tfs*26) in the *ASPM* gene predicted to result in a truncated protein in one family. Further 12 MCPH families carried previously reported homozygous mutations in the *ASPM* gene: (i) the frameshift mutation c.8200_8201delAA (p.N2734Lfs*16) (5) in five families, (ii) the nonsense mutation c.2967G>A (p.W989*) (6) in three families, and (iii) the missense mutation c.9539A>C (p.Q3180P) (7), the nonsense mutations c.9190C>T (p.R3064*) (6) and c.2938C>T (p.R980*) (http://www.ncbi.nlm.nih.gov/clinvar/variation/21569/), and the frameshift mutation c.7782_7783delGA (p.K2595Sfs*6) (6), in one family each. Family MC6 carried a novel homozygous splice site mutation on chromosome position chr1:197125193T>C (GRCh38.p2 assembly, NC_000001.11) in addition to the missense mutation c.9539A>C (p.Q3180P). In one family the homozygous frameshift mutation c.17delC (p.S7Lfs*4) (8) in the *CENPJ* gene was identified. Family MC4 carried a heterozygous missense mutation in the *MCPH1* (c.2422G>A; p.V808I) as well as in the *STIL* gene (c.1136C>T in transcript variants 1-4; c.995C>T in transcript variants 5 and 6; p.S379F in isoforms 1-3; p.S332F in isoforms 4 and 5). While homozygosity of these mutations is known to cause MCPH1 and MCPH7, respectively, it can only be speculated at this point that a combination of heterozygous mutations in two different MCPH genes can potentially result in a MCPH phenotype. However, none of these two mutations are predicted to be pathogenic with respect to PolyPhen2, SIFT, MutationTaster or CADD. The *MCPH1* and the *STIL* mutation have been

observed in only 3/120,734 alleles and 291/120,486 alleles, respectively, in ExAc. Further functional analysis and analysis of other possible candidate genes will be necessary to identify the MCPH-causing mutation in this family. In summary, we have identified both novel and known mutations in MCPH genes with the majority of families carrying mutations in the *ASPM* gene. This study implies that the high incidence of MCPH in Northern Pakistan can to a large extent be attributed to *ASPM* mutations, which coincides with a high percentage of *ASPM* mutations in MCPH worldwide.

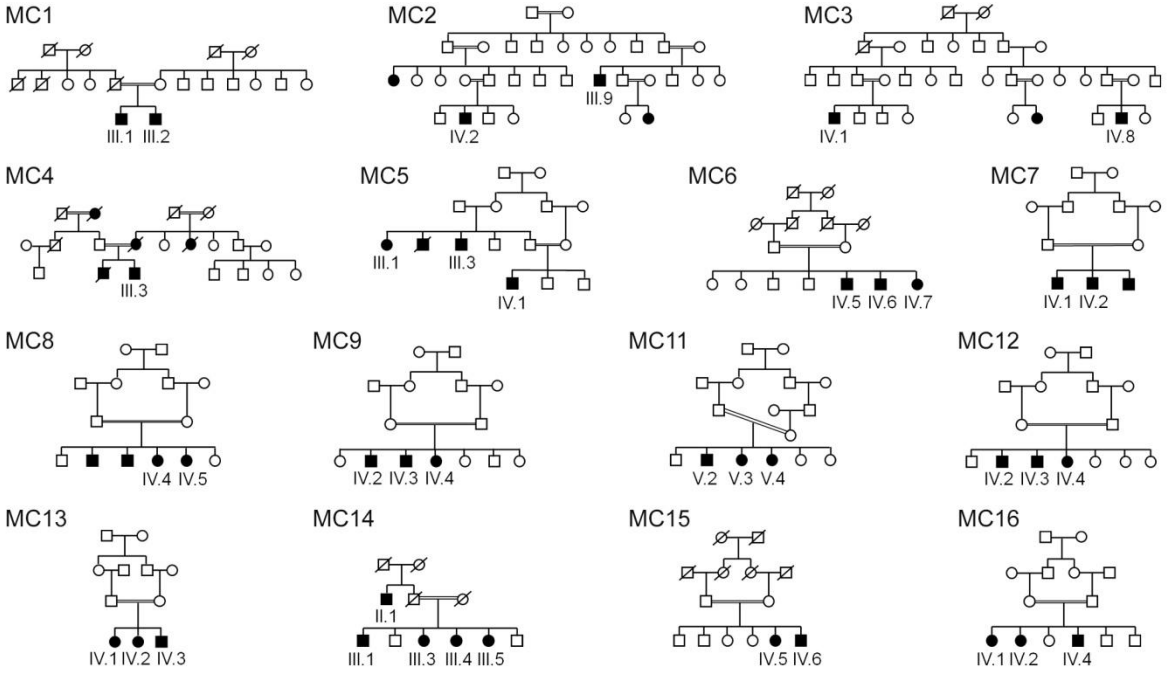
Supplemental Figure 1. Identified MCPH gene mutations in relation to Pakistani cohort residence, pedigrees and pictures of patients.

(A) Map of Pakistan with Azad Jammu and Kashmir (AJK) territory marked in red. (http://commons.wikimedia.org/wiki/File:Azad_Kashmir_in_Pakistan.svg#/media/File:Azad_Kashmir_in_Pakistan.svg). (B) MCPH gene mutations identified in families in the AJK territory. The two novel mutations identified in the *ASPM* gene are marked with a red box. Known mutations in the *ASPM* gene are marked with blue boxes. The mutation identified in the *CENPJ* gene is marked with a brown box. (Modified map from http://www.hrw.org/reports/2006/pakistan0906/1.htm). (C) Pedigrees of consanguineous families and pictures of affected family members.

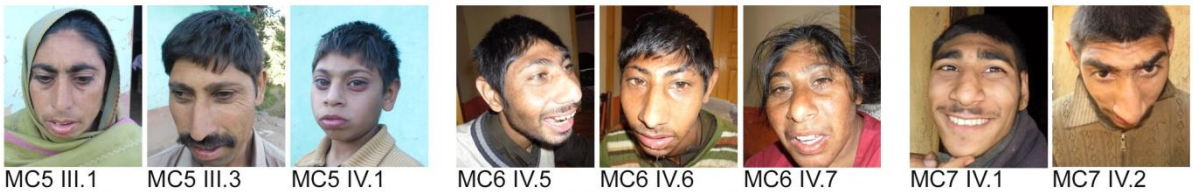
Supplemental Table 1. Head circumferences (OFC) with standard deviations (SD) of patients.


A
B
chr1:197094323T>C
c.5606_5607insC
c.9539A>C
c.8200_8201delAA (4)
PAKISTAN
Jammu and Kashmir
SRINAGAR
JAMMU
c.17delC
c.2967G>A (3)
c.8200_8201delAA
c.9190C>T
c.7782_7783delGA
Muzaffarabd
Bagh
Poonch
Sudhnutti
Kotli
Mirpur
Bhimber

C


MC1
III.1 III.2
MC2
III.9
IV.2
MC3
IV.1
IV.8
MC4
III.3
MC5
III.1
III.3
IV.1
MC6
IV.5 IV.6 IV.7
MC7
IV.1 IV.2
MC8
IV.4 IV.5
MC9
IV.2 IV.3 IV.4
MC11
V.2 V.3 V.4
MC12
IV.2 IV.3 IV.4
MC13
IV.1 IV.2 IV.3
MC14
II.1
III.1 III.3 III.4 III.5
MC15
IV.5 IV.6
MC16
IV.1 IV.2 IV.4


MC1 III.1
MC1 III.2
MC2 IV.2
MC2 III.9
MC3 IV.1
MC3 IV.8
MC4 III.3


MC5 III.1
MC5 III.3
MC5 IV.1
MC6 IV.5
MC6 IV.6
MC6 IV.7
MC7 IV.1
MC7 IV.2


MC8 IV.4
MC8 IV.5
MC9 IV.2
MC9 IV.3
MC9 IV.4
MC11 V.2
MC11 V.3
MC11 V.4

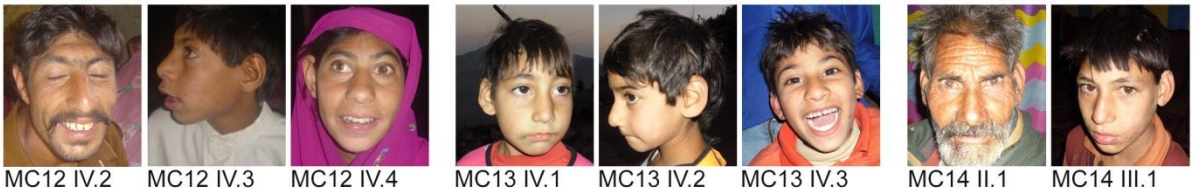

MC12 IV.2
MC12 IV.3
MC12 IV.4
MC13 IV.1
MC13 IV.2
MC13 IV.3
MC14 II.1
MC14 III.1

Table 1: MCPH gene mutations in cohort of Northern Pakistani descent

Gene	Chromosome position[1]	DNA variation[2]	Exon	Protein variation[3]	Family ID	Reference[4]
ASPM	g.197125190G>A	c.2938C>T	11	p.R980*	MC1	ClinVar database[5]
	g.197125161C>T	c.2967G>A	11	p.W989*	MC2, 3, 12	(6)
	g.197101050_197101051delTT	c.8200_8201delAA	18	p.N2734Lfs*16	MC5, 8, 13, 14, 15	(5)
	g.197090947T>G	c.9539A>C (transcript variant 1)	23	p.Q3180P (isoform 1)	MC6	(7)
		c.4784A>C (transcript variant 2)	22	p.Q1595P (isoform 2)		
	g.197093156G>A	c.9190C>T (transcript variant 1)	21	p.R3064* (isoform 1)	MC9	(6)
		c.4435C>T (transcript variant 2)	20	p.R1479* (isoform 2)		
	g.197101468_197101469delTC	c.7782_7783delGA	18	p.K2595Sfs*6	MC11	(6)
	g.197103644_197103645insG	c.5606_5607insC	18	p.H1870Tfs*26	MC16	novel
	g.197125193T>C[6]	-	-	?	MC6	novel
CENPJ	g.24913009delG	c.17delC	2	p.S7Lfs*4	MC7	(8)

[1]Reference sequences: *ASPM*, NC_000001.11; *CENPJ*, NC_000013.11

[2]Reference sequences: *ASPM*, NM_018136.4 (transcript variant 1), NM_001206846.1 (transcript variant 2); *CENPJ*, NM_018451.4

[3]Reference sequences: ASPM, NP_060606.3 (isoform 1), NP_001193775.1 (isoform 2); CENPJ, NP_060921.3

[4]First publication of the mutation

[5]http://www.ncbi.nlm.nih.gov/clinvar/variation/21569/

[6]Splice site mutation in intron 10

References

1. Kaindl AM, Passemard S, Kumar P et al. Many roads lead to primary autosomal recessive microcephaly. Prog Neurobiol 2010: 90: 363-383.
2. Thornton GK, Woods CG. Primary microcephaly: do all roads lead to Rome? Trends Genet 2009: 25: 501-510.
3. Morris-Rosendahl DJ, Kaindl A. What Next-Generation Sequencing (NGS) technology has enabled us to learn about Primary Autosomal Recessive Microcephaly (MCPH). Mol Cell Probes 2015.
4. von Bernuth H, Ravindran E, Du H et al. Combined immunodeficiency develops with age in Immunodeficiency-centromeric instability-facial anomalies syndrome 2 (ICF2). Orphanet J Rare Dis 2014: 9: 116.
5. Sajid Hussain M, Marriam Bakhtiar S, Farooq M et al. Genetic heterogeneity in Pakistani microcephaly families. Clinical Genet 2013: 83: 446-451.
6. Nicholas AK, Swanson EA, Cox JJ et al. The molecular landscape of ASPM mutations in primary microcephaly. J Med Genet 2009: 46: 249-253.
7. Gul A, Hassan MJ, Mahmood S et al. Genetic studies of autosomal recessive primary microcephaly in 33 Pakistani families: Novel sequence variants in ASPM gene. Neurogenetics 2006: 7: 105-110.
8. Bond J, Roberts E, Springell K et al. A centrosomal mechanism involving CDK5RAP2 and CENPJ controls brain size. Nat Genet 2005: 37: 353-355.

Whole exome sequencing and bioinformatic methods. Afflicted patients from all families were subjected to exome sequencing. Genomic DNA was isolated from blood samples using standard methods. Three micrograms of genomic DNA were enriched using the Agilent Human All Exon V4 kit (Agilent Technologies, Santa Clara, CA, USA) following the manufacturer's protocol. Whole-exome libraries were sequenced on an Illumina HiSeq 2000 system for 2x101 cycles following the manufacturer's instructions (Illumina, San Diego, CA, USA). In total, we obtained 78-119 million paired-end 101 bp reads per sample, of which 97.54-99.8% could be mapped onto the human genome. After removing duplicated reads, which were possibly derived from PCR artifacts, 72-108 million unique reads were mapped to the targeted protein coding regions, resulting in an average of 99-150 fold-coverage within the targeted coding region. Using the GATK 3.2.2, we detected 27,376-28,950 SNPs and INDELs in the exome of each family member, of which 98.3-98.4% were known variants deposited in dbSNP 135. Given the cohort, we searched for recessive inheritance patterns. All raw sequencing reads were mapped onto UCSC hg19 using BWA mem 0.7.10, mappings were converted into BAM file format using samtools 0.1.19. Initial mappings were post-processed using GATK following their 'best practices'. In brief, likely PCR duplicates were detected using Picard 1.92. Reads were then realigned around sites of known INDELs. Finally, raw base quality scores were empirically recalibrated. Single Nucleotide Polymorphisms (SNPs) and INDELs were identified using the HaplotypeCaller from GATK. Variants were classified as novel or known variants according to Single Nucleotide Polymorphism Database dbSNP 135. Functional consequences of each variant were annotated using genetic variant annotation and effect prediction toolbox snpEff 3.0 for UCSC hg19 RefSeq genes and ENSEMBL 75 human gene models. The potential deleterious effect was evaluated using PolyPhen 2 (Polymorphism Phenotyping v2), SIFT (Scale-invariant feature transform), PhyloP, MutationTaster, GERP++ (Genomic Evolutionary Rate Profiling), LRT (Likelihood Ratio Test), and OMIM (Online Mendelian Inheritance in Man), if available. Variants within the cohort were filtered for a recessive inheritance pattern.

Family	Pedigree ID	Age of exam [y]	Intellectual disability	OFC [cm]	SD
MC1	III:1	28	-/(+)	43	-9,48
	III:2	21	+	37	-13,93
MC2	III:1	38	++	41	-9,40
	III:9	16	++	38	-14,50
	IV:2	6	-	36	-10,33
MC3	IV:1	7	+	41	-7,33
	IV:6	13	+	48	-3,79
	IV:8	3	-	33	-15,45
MC4	III:3	27	-	49	-5,04
MC5	III:1	35	+	49,5	-3,73
	III:3	30	+	48,26	-5,59
	IV:1	10	+	40,6	-11,26
MC6	IV:5	18	++	43,18	-9,35
	IV:6	14	++	41,91	-9,76
	IV:7	20	++	40,64	-9,64
MC7	IV:1	15	+	43,18	-7,88
	IV:2	20	++	43,18	-9,35
	IV:3	18	++	45,7	-7,48
MC8	IV:2	16	-	49,53	-4,89
	IV:4	10	+	41,91	-6,96
	IV:5	13	++	45,97	-5,24
MC9	IV:2	15	+	40,64	-9,57
	IV:3	16	++	41,91	-11,24
	IV:4	14	++	41,4	-8,93
MC11	V:2	14	+	47,24	-5,66
	V:3	15	+	45,72	-6,28
	V:4	16	+	47,49	-5,47
MC13	IV:1	4	+	40,64	-6,39
	IV:2	3	+	43,18	-4,01
	IV:3	8	++	43,18	-5,88
MC14	II:1	50	+	40,64	-11,23
	III:1	10	-/(+)	45,72	-4,70
	III:3	12	-/(+)	48,26	-3,09
	III:4	25	-/(+)	49,53	-4,64
	III:5	30	-/(+)	45,72	-7,47
MC15	IV:5	16	-/(+)	40,64	-10,74
	IV:6	10	++	45,72	-4,70
MC16	IV:1	7	-/(+)	45,72	-3,59
	IV:2	11	+	42,62	-6,45
	IV:4	2	+	38,1	-7,23

- = no; (+) = very mild; + = mild, ++ = severe

Referenz Nellhaus, G., Pediatrics, 1968:

Male, 18 J, OFC Mean = ca. 55,8 cm; - 2 SD = ca. - 2,7 cm

Male, 16 J., OFC Mean = 55,4 cm; - 2 SD = ca. – 2,4 cm

Male, 15 J., OFC Mean = 55,0 cm; - 2 SD = ca. – 3 cm

Male, 14 J., OFC Mean = 54,6cm; - 2 SD = ca. – 2,6 cm

Male, 10 J., OFC Mean = 53 cm; - 2 SD = ca. – 3,1 cm

Male, 8 J., OFC Mean = 52,3cm; - 2 SD = ca. – 3,1 cm

Male, 7 J., OFC Mean = 52 cm; - 2 SD = ca. - 3 cm

Male, 6 J., OFC Mean = 51,5 cm; - 2 SD = ca. - 3 cm

Male, 3 J., OFC Mean = 50 cm; - 2 SD = ca. – 2,2 cm

Male, 2 J., OFC Mean = 49,3 cm; - 2 SD = ca. – 3,1 cm

Female, 18 J, OFC Mean = ca. 55,1 cm; - 2 SD = ca. – 3 cm

Female, 16 J, OFC Mean = ca. 54,6 cm; - 2 SD = ca. – 2,6 cm

Female, 15 J, OFC Mean = ca. 54,2 cm; - 2 SD = ca. – 2,7 cm

Female, 14 J, OFC Mean = ca. 53,9 cm; - 2 SD = ca. – 2,8 cm

Female, 13 J, OFC Mean = ca. 53,3 cm; - 2 SD = ca. – 2,8 cm

Female, 12 J, OFC Mean = ca. 52,9 cm; - 2 SD = ca. – 3 cm

Female, 11 J, OFC Mean = ca. 52,3 cm; - 2 SD = ca. – 3 cm

Female, 10 J, OFC Mean = ca. 52 cm; - 2 SD = ca. – 2,9 cm

Female, 7 J, OFC Mean = ca. 51,1 cm; - 2 SD = ca. – 3 cm

Female, 4 J, OFC Mean = ca. 49,9 cm; - 2 SD = ca. – 2,9 cm

Female, 3 J, OFC Mean = ca. 49 cm; - 2 SD = ca. – 2,9 cm