

# Genomic analysis of Korean patients with microcephaly

Jiwon Lee<sup>1</sup>, Jong Eun Park<sup>2</sup>, Chung Lee<sup>3</sup>, Ah Reum Kim<sup>1</sup>, Byung Joon Kim<sup>1</sup>, Woong Yang Park<sup>3</sup>, Chang-Seok Ki<sup>4</sup>, and Jeehun Lee<sup>1</sup>

<sup>1</sup>Samsung Medical Center

<sup>2</sup>Hanyang University Guri Hospital

<sup>3</sup>Samsung Genome Institute

<sup>4</sup>GC Genome

May 5, 2020

## Abstract

Microcephaly is a prevalent phenotype of the patients with neurodevelopmental problems, often with genetic causes. We comprehensively investigated the clinical phenotypes and genetic background of microcephaly in Korean patients. We enrolled 40 patients with microcephaly. We analyzed clinical phenotypes and radiologic images and conducted whole exome sequencing (WES) and an analysis of copy number variation (CNV). Infantile hypotonia and developmental delay were present in all patients. Thirty-four patients (85%) showed primary microcephaly. The diagnostic yield from the WES and CNV analyses was 47.5%. With WES, we detected pathogenic or likely pathogenic variants that were already known to be related to microcephaly in 12 patients (30%); nine of these were de novo variants with autosomal dominant inheritance. Two unrelated patients had mutations in the *KMT2A* gene. In 10 other patients, we found mutations in the *GNB1*, *GNAO1*, *TCF4*, *ASXL1*, *SMC1A*, *VPS13B*, *ACTG1*, *EP300*, and *KMT2D* genes. Seven patients (17.5%) were diagnosed as having pathogenic CNVs. Korean patients with microcephaly show a different genetic spectrum from that of patients with microcephaly of other ethnicities. WES with a CNV analysis represents an effective approach for diagnosing the underlying causes of microcephaly.

## Keywords

microcephaly, whole exome sequencing, chromosomal microarray, low-depth whole genome sequencing, Korea

## 1 | Introduction

Microcephaly is a prevalent phenotype of patients with neurodevelopmental problems (Abuelo, 2007; Vargas, Allred, Leviton, & Holmes, 2001). It is defined as an occipitofrontal circumference (OFC) lower than the third percentile or more than two standard deviations below the mean for sex, age, and ethnicity. Patients with microcephaly show various neurologic manifestations, including psychomotor problems, delayed developments, and epilepsy, which are frequently accompany by facial dysmorphism, skeletal anomalies, and congenital structural anomalies in the major organs (Vargas, Allred, Leviton, & Holmes, 2001). Recognition of microcephaly can prompt clinicians to investigate its causes.

The fundamental size of the brain is determined by neuronal progenitor cells formed at conception and that undergo cell divisions. During this process, microcephaly may be caused by deficiency of neuropils, increased apoptosis of progenitor cells, or improper mitosis of these cells, which is defined as primary microcephaly (Gilmore, & Walsh, 2013). Most primary microcephaly occurs due to the failure of neurogenesis or destructive prenatal events associated with environmental or maternal conditions. On the other hand, children with secondary microcephaly are born with normal head circumference (HC)s and then show a progressive decrease HC with age. Secondary microcephaly represents abnormal neuronal development after birth or a perinatal brain insults (Woods, & Parker, 2013). Children with microcephaly have OFCs at the outer limits of

the normal distribution and low for their age group. Both categories of microcephaly have various causes, including congenital infection, perinatal problems such as hypoxia or a maternal medical condition, or genetic causes (Abuelo, 2007), and the distinction between them helps clinicians isolate the preferential etiologies. In addition, accompanying anomalies or facial dysmorphism can provide clues to determine the exact cause of microcephaly (von der Hagen et al., 2014).

In previous studies, genetic causes of microcephaly showed an autosomal recessive inheritance pattern (Darvish et al., 2010; Sajid Hussain et al., 2013). With the development of next-generation sequencing, genetic causes of neurodevelopmental disorders, including microcephaly, have been identified (Hamdan et al., 2014; Najmabadi et al., 2011; Thevenon et al., 2016). To date, only three studies have used whole exome sequencing (WES) to investigate genetic causes of microcephaly in microcephaly cohorts (Boonsawat et al., 2019; Rump et al., 2016; Shaheen et al., 2019). One study determined disease-causing genes in 11 out of 38 patients who had highly heterogeneous clinical and radiologic phenotypes. It revealed that autosomal recessive disorders are highly prevalent among the patients with microcephaly (Rump et al., 2016). Shaheen et al. (2019) identified potentially causal genetic variants in 104 out of 137 families. They found variants in novel and previously reported genes that were the causes of microcephaly with established disease phenotypes. In addition, the authors showed overlap in the genetic causes of microcephaly with genes underlying microcephalic primordial dwarfism. A recent comprehensive genetic analysis based on chromosomal microarray (CMA) and WES detected causative variants in 48.4% of patients (30/62) and analyzed the differences in clinical severity and genetic variants between primary and secondary microcephaly (Boonsawat et al., 2019). Each of these studies covered different characteristics of race, consanguinity, and brain imaging findings. However, there has been no study of an Asian population based exclusively on a microcephaly cohort.

We aimed to comprehensively analyze the genomic and phenotypic features of Korean patients with microcephaly. We tried to delineate clinical features associated with causative genes to contribute to genetic investigations of patients with microcephaly.

## 2 | Materials and Methods

### 2.1 | Patients

We recruited pediatric patients with microcephaly, applying a definition of microcephaly of OFC  $>2$  standard deviations below the mean for same sex, age-matched Korean references, between January 2014 and December 2018 at Samsung Medical Center in Seoul, Korea. All affected individuals were evaluated for clinical features and diagnosed with microcephaly by two pediatric neurologists. We excluded patients with microcephaly that resulted from cerebral insults related to defined perinatal problems such as congenital infection, teratogens, physical environment, or maternal medical condition through laboratory tests, brain imaging, and detailed history taking. We performed WES for all patients and CMA or low-depth whole genome sequencing (LD-WGS) to analyze copy number variation (CNV). We measured weights, heights, and HCs in the same examination and evaluated concordance between the body gauges. Clinical phenotypes included the followings; infantile hypotonia, facial dysmorphism, skeletal and organ anomalies, hearing loss, developmental delay/cognitive impairment, seizure, and affected epilepsy syndrome. Development and cognitive function was assessed using Bayley Scales of Infant Development (BSID-III) or Wechsler Scale of Intelligence for children (WISC-V), depending on the age of the patients. If these scales were not available, we estimated developmental stage or cognitive function based on the clinical status assessed by two pediatric neurologists using the same standards.

### 2.2 | WES

Genomic DNA was extracted from peripheral blood leukocytes using a Wizard Genomic DNA Purification Kit following the manufacturer's instructions (Promega, Madison, WI, USA). SureSelect Human All Exon V5 (Agilent Technologies, Santa Clara, CA, USA) was used for library preparation; and sequencing was performed on an Illumina HiSeq2500 platform (Illumina Inc., San Diego, CA, USA) to generate  $2 \times 100$ -bp paired-end reads. Sequence reads were aligned against the Human Reference Genome build GRCh37 using BWA 0.7.12; duplicated reads were marked with Picard Tools 1.130; local alignment, base quality

recalibration and variant calling were conducted using the Genome Analysis Tool Kit v3.4.0; and annotation and variant effect prediction were performed with SnpEff v4.1g.

### 2.3 | Variant filtering steps and data analysis

Variants were filtered and prioritized using a four-step strategy to generate a short candidate variant list for experimental validation (Figure 1). Initially, we removed variants with less than  $10\times$  coverage. Next, variants were limited to those with low population frequency. The minor allele frequency (MAF) threshold was carefully chosen and variants with an MAF  $\geq 1\%$  in the Genome Aggregation Database (gnomAD) (<http://gnomad.broadinstitute.org/>) or the Korean Reference Genome Database (KRGDB) (<http://coda.nih.go.kr/coda/KRGDB/index.jsp>) were removed. The third step was to prioritize variants causing missense, nonsense, frameshifts, and in-frame insertions/deletions variants, or changes affecting consensus splice site sequences. Finally, we performed a gene-specific analysis with an in-silico gene panel composed of 903 genes, filtering for selected phenotype traits Human Phenotype Ontology (HPO)-terms for Microcephaly (HP:0000252) or Online Mendelian Inheritance in Man (OMIM) microcephaly phenotype genes (Supplementary Table 1). To delineate candidate genetic variants, an additional allele analysis was performed under the following conditions: 1) triplicate data with no pathogenic variant (PV)s nor likely pathogenic variant (LPV)s, 2) *de novo*, compound heterozygous, homozygous, or hemizygous variants, 3)  $\geq 2$  alleles in gnomAD or  $\geq 8$  alleles if recessive, 4) a CADD score of  $\geq 15$  or higher and all deleterious predictions in SIFT (<http://sift.jcvi.org>), PolyPhen2 (<http://genetics.bwh.harvard.edu/pph2>), and MutationTaster (<http://mutationtaster.org/>) if missense variants, 5) affected genes with data from animal models and/or functional studies suggesting neurodevelopmental roles.

### 2.4 | CMA and LD-WGS

CMA was performed using an Illumina HumanCytoSNP-12 platform (Illumina, San Diego, CA, USA). The array employs both Copy Number Changes (CNC) and Single Nucleotide Polymorphism (SNP) probes on a whole genome array with a 30 Kb resolution. A higher resolution was used for all regions known to be associated with balanced and unbalanced structural variants. The data were analyzed in Illumina KaryoStudio version 1.4.3.0 (Illumina, San Diego, CA, USA) and reported using NCBI human genome build 37.1 (hg19) following ISCN nomenclature.

The LD-WGS-based CNV analysis was performed by GC Genome (Yongin, Korea). Briefly, genomic DNA was extracted from peripheral-blood leukocytes, sheared to a target size of 250 bp, and sequenced on a NextSeq 500 platform (Illumina, CA, USA) in 75 bp single-read mode. CNV calling for 3.1 million sequence reads was performed using DNACopy software v1.38. Using the target-specific reference mapping ratio values for the target sample and the normal controls, we calculated the copy number status of each targeted sample. The February 2009 Human reference sequence (GRCh37/hg19) was used for genome assembly.

### 2.5 | Interpretation of candidate genetic variants by WES and CNVs by CMA or LD-WGS

Candidate variants were classified according to the standards and guidelines of the American College of Medical Genetics and Genomics (ACMG) and Association for Molecular Pathology (AMP) (Richards et al., 2015). These guidelines recommend classifying variants into five categories: PV, LPV, variant of uncertain significance (VUS), likely benign variant (LBV), and benign variant (BV) based on combined lines of weighted evidences, including population, computational, functional, segregation, *de novo*, allelic, and other data. To assess the frequency of a variant in a control or general population, we used the KRGDB, which consists of publicly available race-matched control data from WGS of 1100 Korean individuals, as well as other public databases such as the 1000 Genomes Project database, ExAC database, and gnomAD. A primary literature review was conducted using various sources cited in the Human Gene Mutation Database (HGMD) professional version, release 2018.02 (<http://www.hgmd.org/>), and ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>), as well as PubMed, to determine the potential pathogenicity of all identified variants. Various *in silico* tools, including SIFT, Polyphen2, and MutationTaster, all of which use missense prediction algorithms, were used to determine the predicted impact of missense change.

To interpret CNVs, we used DGV (<http://dgv.tcag.ca/dgv/app/home>), DECIPHER (<https://decipher.sanger.ac.uk/>), and dbVar (<https://www.ncbi.nlm.nih.gov/dbvar/>) databases as well as OMIM (<https://www.omim.org/>). In addition, the clinical findings were compared with those reported in the literature.

## 2.6 | Confirmation and validation of candidate variants

Candidate variants identified in WES data were confirmed using standard PCR and Sanger sequencing methods. Primer sequences are available upon request. Sequence data were aligned to the reference sequence using Sequencher software (Gene Codes Corporation, Ann Arbor, MI, USA).

## 2.7 | Ethic statement

This study was approved by the Institutional Review Board of Samsung Seoul Hospital (IRB no. 2014-07-001-004). Written informed consent was obtained from the parents of all pediatric patients included in this study.

## 3 | Results

### 3.1 | Clinical phenotypes

Among 40 patients (female:male = 22:18) with microcephaly, 34 patients presented at birth with primary microcephaly and the remaining six were identified as cases of secondary microcephaly based on the gradual development of a small HC after birth (Table 1). With proportionate microcephaly, identified in 15 patients (37.5%), all body gauges were below the sex and age-matched 3<sup>rd</sup> centile value. The other 25 patients (62.5%) showed disproportionate microcephaly, with height and/or weight in the normal range for sex and age. All patients with short stature ( $n = 26$ ) displayed type 3 failure to thrive during their infancy. Common features observed in all patients were infantile hypotonia and developmental delay. The average age at the time of participation in this study was 5.28  $\pm$  4.60 years old. Thirty-two patients showed noticeable developmental delay or intellectual disability. Of the remaining patients, one showed mild intellectual disabilities and seven displayed a delay (5 months below age) in reaching developmental milestones. Preschool-aged patients ( $n = 25$ ) also lagged in achieving developmental milestones, and eight school-aged patients received special education. Four patients were treated for autistic spectrum disorder ( $n = 3$ ) or behavioral disorder ( $n = 1$ ) in a neuropsychiatric clinic. Seven patients showed frequent dystonia or other involuntary movements like chorea, athetoid movements, or tremor.

Seventeen patients (42.5%) showed various accompanying anomalies, including congenital heart anomaly ( $n = 8$ ), skeletal anomaly ( $n = 4$ ), and cleft lip and/or palate ( $n = 5$ ) (Table 1). Three patients developed sensorineural hearing loss without a discernible etiology. Dysmorphic features were evident in more than half of patients ( $n = 28$ , 70%), these features were clear in 17 patients (42.5%) and notably vague in 11 patients (27.5%).

Twenty-seven patients (67.5%) had seizures. One of these patients had only one seizure episode provoked by febrile illness. Among the 26 patients with epilepsy, 11 showed intractable epilepsy, with having frequent seizures despite multiple antiepileptic medications. There were eight patients diagnosed with epileptic encephalopathy, including six patients with Lennox-Gastaut syndrome (LGS). Half of patients with LGS showed infantile spasms, and one patient had a neonatal seizure.

Brain MRIs were performed in all patients, and abnormal structural lesions were found in six patients. Three patients had agenesis or hypoplasia of the corpus callosum with microcephaly. In another three patients, holoprosencephaly, subependymal heterotopia, and hypoplasia of the pons and cerebellum were found, respectively. Thirty-four patients showed normal brain structure by MRI.

### 3.2 | Genetic evaluation in all patients with microcephaly

A molecular diagnosis was established in 19 patients (47.5%) from 39 families. In 12 patients, including one sibling pair (30%), 11 PVs or LPVs were identified by WES (Table 2). The inheritance patterns of

these mutations were autosomal dominant ( $n = 9$ ), autosomal recessive ( $n = 1$ ), and X-linked recessive ( $n = 1$ ). Another seven patients (17.5%) showed a pathogenic CNV by CMA or LD-WGS (Table 3). Using WES, PVs or LPVs were detected in nine genes previously associated with microcephaly: *GNB1*, *GNAO1*, *TCF4*, *ASXL1*, *SMC1A*, *KMT2A*, *VPS13B*, *ACTG1*, *EP300*, and *KMT2D*. Five PVs or LPVs were novel variants, and nine PVs or LPVs were *de novo* variants. Identified variants except BV and LBV were described in Supplementary Table 2.

### 3.3 | Clinical phenotypes of the patients with a PV/LPV determined by WES

There was a 9-year-old girl (M-011) with an LPV in the *GNB1* gene. She presented with infantile spasms at the age of 4 months and was treated with antiepileptic medications. Her epileptic spasms disappeared during infancy, but, she failed to meet developmental milestones. Gradually, different type of seizures and involuntary movements developed, despite being administered a number of antiepileptic medications. A *de novo* LPV (c.284T>C;p.Leu95Pro) that was absent from the control databases was identified in the *GNB1* gene. This c.284T>C variant was previously associated with *de novo* severe neurodevelopmental disability, hypotonia, and seizures (Petrovski et al., 2016).

In a 4-year-old boy (M-020) suffering from intractable dystonia with severe growth failure and profound psychomotor retardation, a *de novo* LPV (c.607G>A;p.Gly203Arg) that was absent from the control databases was identified in the *GNAO1* gene. A c.607G>A change in *GNAO1* was previously reported many times in *de novo* in epileptic encephalopathy patients, suggesting that mutation hotspot (Arya, Spaeth, Gilbert, Leach, & Holland, 2017).

In patient M-039, a PV (c.3115C>T;p.Gln1039Ter) of the *ASXL1* gene was identified. She was admitted on the neonatal intensive care unit for congenital hypotonia. Because her respiration was unstable and weak, she needed a tracheostomy and home ventilator support. She showed severe feeding problems and profound failure to thrive. In addition, the patient suffered from intractable epilepsy and severe dystonia. The c.3115C>T variant in *ASXL1* was absent in the control population and occurred *de novo*. This c.3115C>T variant was previously reported mainly as a somatic variant in patients with myelodysplastic syndrome (Wang et al., 2013), but no germline variant has been reported.

In the *SMC1A* gene, known as a causative gene of Cornelia de Lange syndrome, an LPV (c.2368C>T;p.Arg790Trp) was identified in a male sibling pair (M-055-P, M-055-S) who were clinically suspected of Cornelia de Lange syndrome. They showed distinctive facial features, namely long eyelashes, synophrys, small nose, wide nasal bridge, low-set ears, and small chin. They could not speak a meaningful word and showed severe growth retardation in not only HC but also height and weight. The c.2368C>T variant in *SMC1A* was hemizygous. The patients' mother had this variant, in her case, it was heterozygous. There were two unaffected maternal uncles, who have not married. They displayed normal intelligence and facial feature, and did not been tested for the variant. This variant was previously reported to be associated with Cornelia de Lange syndrome (Ansari et al., 2014).

### 3.4 | Clinical phenotypes of patients with a pathogenic CNV

Of 40 patients, seven (17.5%) were found to have a pathogenic CNV by CMA or LD-WGS (Table 3). The size of duplication or deletion region was between 312 Kb and 28.43 Mb.

Patient M-004 presented with severe intellectual disability, autistic behaviors, epilepsy, and dysmorphic facial features, and showed an 801 Kb deletion at 7q11.21. This region contains the *AUTS2* gene, and the deletion of the *AUTS2* gene has been reported to be associated with autism, intellectual disability, short stature, and microcephaly (Beunders et al., 2013).

M-017 presented with primary microcephaly and congenital heart disease, and showed a 2,790 Kb deletion at 9q34.3. This region was previously associated with 9q subtelomeric syndrome characterized by severe developmental delay, microcephaly, congenital heart disease, and seizure (Stewart, & Kleefstra, 2007).

M-048 presented with primary microcephaly and congenital heart disease, and displayed a 28.43 Mb du-

plication at 16p13.3p11.2. This region was registered at OMIM as associated with chromosome 16p11.2 duplication syndrome (MIM 614671). Patients with a 16p11.2 duplication have been reported to have motor delays, congenital anomalies, seizures, and microcephaly (Shinawi et al., 2010).

M-074 presented with primary microcephaly, epilepsy, and cleft palate, and carried a 368 Kb duplication at 6q25.3. This region contains the *ARID1B* gene, which is a causative gene for Coffin-Siris syndrome, which is characterized by mental retardation and associated with coarse facial features, hypertrichosis, sparse scalp hair, and hypoplastic or absent fifth fingernails or toenails. Patients with a *ARID1B* duplication have been reported to show moderate developmental delays, hypotonia, and high-arched palate (Hoyer et al., 2012).

A girl (M-075) showing microcephaly with pontine and cerebellar hypoplasia by brain MRI was found to have a 2.8 Mb deletion at Xp11.4p11.3 (Figure 2). This region contains the *CASK* gene, and deletion of the *CASK* gene has been reported to be associated with mental retardation and microcephaly with disproportionate pontine and cerebellar hypoplasia in females (Hayashi et al., 2012).

M-116 presented with intellectual disability and autistic features, midline hypoplasia of the face, hypertrichosis, and bush eyebrows, and showed a 312 Kb deletion at 2q23.1. This region contains the *MBD5* gene, and deletion of the *MBD5* gene has been reported to be associated with intellectual disability, seizures, significant speech impairment, and behavioral problems (Bonnet et al., 2013).

### 3.5 | Additional candidate microcephaly genes

An allele search in familial trio exome data revealed four additional candidate genes. In a patient (M-001) with primary microcephaly, we found a *de novo* variant (c.2123A>C;p.Tyr708Ser) in the *JARID2* gene. *JARID2* (jumonji and AT-rich interaction domain containing 2) is highly homologous to mouse Jumonji (Jmj). In mice, Jmj is required for normal morphogenesis of the neural tube (Takeuchi et al., 1995). This c.2123A>C variant in *JARID2* was absent from the control databases.

*RFX7* (regulatory factor X7) is a member of the regulatory factor X (RFX) family of transcription factors (Aftab, Semenec, Chu, & Chen, 2008). *RFX7* is required for the formation of cilia in the neural tube (Manojlovic, Earwood, Kato, Stefanovic, & Kato, 2014). In one patient (M-100), we found a *de novo* variant (c.2727\_2730dup;p.Gly911PhefsTer58) in the *RFX7* gene that was absent from the control databases.

*LMNB1* (lamin B1) encodes one of the two B-type lamin protein. Lamin B1, a key component of the nuclear lamina, plays an important role in brain development (Mahajani et al., 2017). A heterozygous tandem genomic duplication of *LMNB1* is known to be the cause of autosomal dominant adult-onset demyelinating leukodystrophy. In one patient (M-118), we found a *de novo* variant (c.1091T>C;p.Leu364Pro) in this gene that was absent from the control databases.

*HPRT1* (hypoxanthine guanine phosphoribosyltransferase 1) encodes hypoxanthine-guanine phosphoribosyl transferase, an enzyme involved in purine metabolism. An *HPRT1* gene mutation is the cause of Lesch-Nyhan syndrome, which is an X-linked recessive disorder (Gibbs, & Caskey, 1987). In one patient (M-122), we identified a hemizygous variant (c.151C>T;p.Arg51Ter) that was inherited from his mother. This c.151C>T variant was absent from the control databases and was previously reported in a Japanese patient with Lesch-Nyhan syndrome (Fujimori, Kamatani, Nishida, Ogasawara, & Akaoka, 1990). The patient in this study did not have the clinical characteristics of that syndrome.

## 4 | Discussion

This study used WES and CNV analyses to identify genetic causes of microcephaly in Korean patients. The overall diagnostic rate was 47.5%; with 12 cases (30%) identified by WES and seven cases by CNV analysis. The nine of all variants found by WES, including five novel ones, were *de novo* and autosomal dominant traits. Of the 40 patients included in this study, 34 patients (85%) had primary microcephaly. Six showed various structural anomalies by brain MRI, whereas the others showed normal brain structures despite microcephaly. Hypotonia and developmental delay were the most frequent accompanying symptoms. Seizure was the next most frequent symptom ( $n = 27$ , 65%), and 11 patients had intractable seizures.

Microcephaly can present as non-syndromic or with various hereditary syndromic features (von der Hagen et al., 2014). According to the OMIM database, there are over 900 OMIM phenotype and genes related to microcephaly entered to date. In the present study, the diagnostic yield of WES (PVs and LPVs) was 30% and was comparable to that found in previous reports of genetic causes of microcephaly (Boonsawat et al., 2019; Rump et al., 2016; Shaheen et al., 2019). In general, the diagnostic yield of WES in neurodevelopmental disorders ranges from 17.5% to 29% (Marques Matos, Alonso, & Leao, 2019; Vissers et al., 2017). The yield depends on several factors such as phenotypic distinctiveness of the disease, genes associated with the phenotype, sequence read depth, bioinformatic filtering, and/or level of clinical medical review (Farwell et al., 2015). Some neurologic disorders are more closely associated with muscle disease, ataxia, and epilepsy, which are known to have heterogeneous genetic causes (Farwell et al., 2015; Marques Matos, Alonso, & Leao, 2019). Thus, WES should be the first-line test for complex diseases presumed extensive genetic causes, based on considerations of diagnostic yield and economic cost (Marques Matos, Alonso, & Leao, 2019; Vissers et al., 2017). Similar to those identified in previous reports, the variants found in this study were highly diverse. This study reconfirms the utility of WES in determining the genetic causes of microcephaly.

In this study, the CNV analysis increased the diagnostic rate from 30% to 47.5%. This rate was similar to that reported in a recent study by Boonsawat et al (2019). The diagnostic rate of pathogenic CNVs in neurodevelopmental disorders, including primary microcephaly, has been reported to be 13.2 – 20.8% (Jang et al., 2019; Shoukier et al., 2013). According to a previous study, combined WES and CNV analyses better enable diagnosis of rare neurologic disorders, with a 47.8% diagnostic rate comparing to that with WES alone (32.7%) (Jiao et al., 2019), which is consistent with the results of this study. A recent study showed 11 definite and 7 probable CNVs in 53 patients with primary microcephaly (Tsoutsou et al., 2017). Therefore, WES with CNV analysis may be the mostly effective approach to diagnosing the underlying causes of microcephaly. A diagnosis of microcephaly via WES and CNV analysis can help determine the genetic mechanism of the disease and predict the prognosis. In addition, it may provide information to patients for future reproductive decisions and genetic counseling, including opportunities for targeted therapies.

In the present study, most of the detected variants showed an autosomal dominant inheritance pattern. Two variants were associated with autosomal recessive and X-linked recessive inheritance respectively. Previous studies dealt with congenital microcephaly in consanguineous families and mainly focused on neurodevelopmental defects in the fetal period (Darvish et al., 2010; Sajid Hussain et al., 2013; Shaheen et al., 2019). They classified cases of microcephaly as autosomal recessive primary microcephaly (MCPH), and 18 loci in MCPH genes have been revealed (Jayaraman, Bae, & Walsh, 2018). In our patients, no mutations were identified in *MCPH1*, *WDR62*, *CDK5RAP2*, *CEP152*, *ASPM*, *CENPJ*, *STIL*, *CEP63*, *CEP135*, *CASC5*, *PHC1*, *CDK6*, *CENPE*, *SASS6*, *MFSD2A*, *ANKLE2*, *CIT*, or *WDFY3* genes. Our results suggest that autosomal dominant disorders are highly prevalent among Korean patients with microcephaly. The microcephaly-associated genes identified in this study were diverse, and no genes were found to be predominant causes of microcephaly. This suggests that various genes may lead to microcephaly, as reported previously (Shaheen et al., 2019). These genes are involved in neuronal myelination, neurotransmission, or regulation of neuronal excitability, rather than centrosome-related neurogenesis and DNA damage repair process (Woods, & Parker, 2013).

In this study, 34 patients showed normal brain structures despite small brain size. In the six patients (15%) with structural abnormalities observed by brain MRI, only one patient was found to have a causative genetic mutation in the *CASK* gene. The *CASK* gene encodes a membrane-associated guanylate kinase and is involved in neurotransmitter regulation, axon branching, and dendritic outgrowth. The patient with this gene variant showed a wide range of X-linked intellectual disabilities, including microcephaly with pontine and cerebellar hypoplasia (MICPCH) syndrome, nystagmus, dysmorphic faces, and hypotonia (Hayashi et al., 2017). Because this is located on the X chromosome, most cases of *CASK* mutation are females, and affected males are very rare and show severe manifestations (Cristofoli, Devriendt, Davis, Van Esch, & Vermeesch, 2018). We found a *CASK* mutation through CMA in a female patient with MICPCH syndrome. She was born at term without perinatal complications. Her HC at birth was 32.2 cm, and the fourth ventricle was extensively dilated according to a brain ultrasonogram performed on the second day of life. She showed

poor growth in all body gauges and dysmorphic facial features such as arched eyebrows, midline hypoplasia, long philtrum, and micrognathia. Her brain MRI was consistent with the criteria for microcephaly and pontocerebellar hypoplasia, and her achievement of developmental milestones was assessed at the five-month level at the age of one year (Figure 2).

Racially homogeneous Korean patients with microcephaly, with no consanguineous family members, were enrolled in this study. This is the first study delineating the genetic causes of microcephaly in East Asia. The patients in this study showed the same clinical characteristics that were reported in previous studies, including infantile hypotonia and developmental delay. Epilepsy was observed in 65% of the patients, and about half of these patients showed an intractable clinical course. In addition, movement disorders, behavioral problems, and autism spectrum disorder were noted, which was similar to features in a recent report (Boonsawat et al., 2019). However, our study showed different genetic peculiarities from those reported previously for patients in the Middle East or Europe, in which autosomal recessive inheritance was frequent. In our study, 75% of detected variants were autosomal dominant; only one sibling pair had an *SMC1A* gene variant with X-linked inheritance. These results suggest that the genetic spectrum of microcephaly may differ between ethnic groups, even if clinical characteristics are similar. Therefore, consideration of the racial background might be helpful in interpreting WES results.

This study had several limitations. First, the number of patients in our study was small to determine the detection rate for genetic causes of microcephaly. Second, a functional study of detected VUSs was not performed. Although not classified as PVs or LPVs, some of the VUSs found in this study may be useful in determining the genetic cause of microcephaly through additional studies to verify gene function. In addition, we did not identify cases of somatic mosaicism and balanced translocation, which it is difficult to confirm by WES or CMA. However, somatic mosaicism and balanced translocation have rarely been reported as genetic causes of microcephaly.

We conclude that conducting WES and CMA for patients with microcephaly is useful and time/cost-effective for genetic diagnosis and treatment considerations. Unlike other studies, autosomal recessive microcephaly was rare; autosomal dominant was predominant mode of inheritance in Korean patients. This study confirmed that microcephaly is a condition with genotypically and phenotypically heterogeneous causes. Interdisciplinary cooperation between molecular geneticists and clinicians is important to facilitate diagnosis of the underlying causes of microcephaly.

### Acknowledgments

This research was supported by Basic Science Research Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2014M3C9A2064619, NRF-2017R1A2B4005276). The funding body had no role in this study design, data collection, or analysis, decision to publish, or preparation of the manuscript.

### Data Availability Statement

The data that support the findings of this study are openly available in ClinVar at [<https://www.ncbi.nlm.nih.gov/clinvar/>], submission name [Korean Microcephaly].

### References

- Abuelo, D. (2007). Microcephaly syndromes. *Seminars in Pediatric Neurology* 14(3):118-127. <https://doi.org/10.1016/j.spen.2007.07.003>
- Aftab, S., Semene, L., Chu, J. S., & Chen, N. (2008). Identification and characterization of novel human tissue-specific RFX transcription factors. *BMC Evolutionary Biology* 8:226. <https://doi.org/10.1186/1471-2148-8-226>
- Ansari, M., Poke, G., Ferry, Q., Williamson, K., Aldridge, R., Meynert, A. M., ... FitzPatrick, D. R. (2014). Genetic heterogeneity in Cornelia de Lange syndrome (CdLS) and CdLS-like pheno-

types with observed and predicted levels of mosaicism. *Journal of Medical Genetics* 51(10):659-668. <https://doi.org/10.1136/jmedgenet-2014-102573>

Arya, R., Spaeth, C., Gilbert, D. L., Leach, J. L., & Holland, K. D. (2017). GNAO1-associated epileptic encephalopathy and movement disorders: c.607G>A variant represents a probable mutation hotspot with a distinct phenotype. *Epileptic Disorders* 19(1):67-75. <https://doi.org/10.1684/epd.2017.0888>

Beunders, G., Voorhoeve, E., Golzio, C., Pardo, L. M., Rosenfeld, J. A., Talkowski, M. E., ... Sistermans, E. A. (2013). Exonic deletions in AUTS2 cause a syndromic form of intellectual disability and suggest a critical role for the C terminus. *American Journal of Human Genetics* 92(2):210-220. <https://doi.org/10.1016/j.ajhg.2012.12.011>

Bonnet, C., Ali Khan, A., Bresso, E., Vigouroux, C., Beri, M., Lejczak, S., ... Jonveaux, P. (2013). Extended spectrum of MBD5 mutations in neurodevelopmental disorders. *European Journal of Human Genetics* 21(12):1457-1461. <https://doi.org/10.1038/ejhg.2013.22>

Boonsawat, P., Joset, P., Steindl, K., Oneda, B., Gogoll, L., Azzarello-Burri, S., ... Rauch, A. (2019). Elucidation of the phenotypic spectrum and genetic landscape in primary and secondary microcephaly. *Genetics in Medicine* 21(9):2043-2058. <https://doi.org/10.1038/s41436-019-0464-7>

Cristofoli, F., Devriendt, K., Davis, E. E., Van Esch, H., & Vermeesch, J. R. (2018). Novel CASK mutations in cases with syndromic microcephaly. *Human Mutation* 39(7):993-1001. <https://doi.org/10.1002/humu.23536>

Darvish, H., Esmaeeli-Nieh, S., Monajemi, G. B., Mohseni, M., Ghasemi-Firouzabadi, S., Abedini, S. S., ... Najmabadi, H. (2010). A clinical and molecular genetic study of 112 Iranian families with primary microcephaly. *Journal of Medical Genetics* 47(12):823-828. <https://doi.org/10.1136/jmg.2009.076398>

Farwell, K. D., Shahmirzadi, L., El-Khechen, D., Powis, Z., Chao, E. C., Tippin Davis, B., ... Tang, S. (2015). Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. *Genetics in Medicine* 17(7):578-586. <https://doi.org/10.1038/gim.2014.154>

Fujimori, S., Kamatani, N., Nishida, Y., Ogasawara, N., & Akaoka, I. (1990). Hypoxanthine guanine phosphoribosyltransferase deficiency: nucleotide substitution causing Lesch-Nyhan syndrome identified for the first time among Japanese. *Human Genetics* 84(5):483-486. <https://doi.org/10.1007/bf00195826>

Gibbs, R. A., & Caskey, C. T. (1987). Identification and localization of mutations at the Lesch-Nyhan locus by ribonuclease A cleavage. *Science* 236(4799):303-305. <https://doi.org/10.1126/science.3563511>

Gilmore, E. C., & Walsh, C. A. (2013). Genetic causes of microcephaly and lessons for neuronal development. *Wiley Interdiscip Rev Dev Biol* 2(4):461-478. <https://doi.org/10.1002/wdev.89>

Hamdan, F. F., Srour, M., Capo-Chichi, J. M., Daoud, H., Nassif, C., Patry, L., ... Michaud, J. L. (2014). De novo mutations in moderate or severe intellectual disability. *Plos Genetics* 10(10):e1004772. <https://doi.org/10.1371/journal.pgen.1004772>

Hayashi, S., Okamoto, N., Chinen, Y., Takanashi, J., Makita, Y., Hata, A., ... Inazawa, J. (2012). Novel intragenic duplications and mutations of CASK in patients with mental retardation and microcephaly with pontine and cerebellar hypoplasia (MICPCH). *Human Genetics* 131(1):99-110. <https://doi.org/10.1007/s00439-011-1047-0>

Hayashi, S., Uehara, D. T., Tanimoto, K., Mizuno, S., Chinen, Y., Fukumura, S., ... Inazawa, J. (2017). Comprehensive investigation of CASK mutations and other genetic etiologies in 41 patients with intellectual disability and microcephaly with pontine and cerebellar hypoplasia (MICPCH). *PloS One* 12(8):e0181791. <https://doi.org/10.1371/journal.pone.0181791>

- Hoyer, J., Ekici, A. B., Endeley, S., Popp, B., Zweier, C., Wiesener, A., ... Reis, A. (2012). Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *American Journal of Human Genetics* 90(3):565-572. <https://doi.org/10.1016/j.ajhg.2012.02.007>
- Jang, W., Kim, Y., Han, E., Park, J., Chae, H., Kwon, A., ... Park, J. H. (2019). Chromosomal Microarray Analysis as a First-Tier Clinical Diagnostic Test in Patients With Developmental Delay/Intellectual Disability, Autism Spectrum Disorders, and Multiple Congenital Anomalies: A Prospective Multicenter Study in Korea. *Annals of Laboratory Medicine* 39(3):299-310. <https://doi.org/10.3343/alm.2019.39.3.299>
- Jayaraman, D., Bae, B. I., & Walsh, C. A. (2018). The Genetics of Primary Microcephaly. *Annual Review of Genomics and Human Genetics* 19:177-200. <https://doi.org/10.1146/annurev-genom-083117-021441>
- Jiao, Q., Sun, H., Zhang, H., Wang, R., Li, S., Sun, D., ... Jin, Y. (2019). The combination of whole-exome sequencing and copy number variation sequencing enables the diagnosis of rare neurological disorders. *Clinical Genetics* 96(2):140-150. <https://doi.org/10.1111/cge.13548>
- Mahajani, S., Giacomini, C., Marinaro, F., De Pietri Tonelli, D., Contestabile, A., & Gasparini, L. (2017). Lamin B1 levels modulate differentiation into neurons during embryonic corticogenesis. *Scientific Reports* 7(1):4897. <https://doi.org/10.1038/s41598-017-05078-6>
- Manojlovic, Z., Earwood, R., Kato, A., Stefanovic, B., & Kato, Y. (2014). RFX7 is required for the formation of cilia in the neural tube. *Mechanisms of Development* 132:28-37. <https://doi.org/10.1016/j.mod.2014.02.001>
- Marques Matos, C., Alonso, I., & Leao, M. (2019). Diagnostic yield of next-generation sequencing applied to neurological disorders. *Journal of Clinical Neuroscience* 67:14-18. <https://doi.org/10.1016/j.jocn.2019.06.041>
- Najmabadi, H., Hu, H., Garshasbi, M., Zemojtel, T., Abedini, S. S., Chen, W., ... Ropers, H. H. (2011). Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* 478(7367):57-63. <https://doi.org/10.1038/nature10423>
- Petrovski, S., Kury, S., Myers, C. T., Anyane-Yeboah, K., Cogne, B., Bialer, M., ... Goldstein, D. B. (2016). Germline De Novo Mutations in GNB1 Cause Severe Neurodevelopmental Disability, Hypotonia, and Seizures. *American Journal of Human Genetics* 98(5):1001-1010. <https://doi.org/10.1016/j.ajhg.2016.03.011>
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., ... Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine* 17(5):405-424. <https://doi.org/10.1038/gim.2015.30>
- Rump, P., Jazayeri, O., van Dijk-Bos, K. K., Johansson, L. F., van Essen, A. J., Verheij, J. B., ... Sikkema-Raddatz, B. (2016). Whole-exome sequencing is a powerful approach for establishing the etiological diagnosis in patients with intellectual disability and microcephaly. *BMC Medical Genomics* 9:7. <https://doi.org/10.1186/s12920-016-0167-8>
- Sajid Hussain, M., Marriam Bakhtiar, S., Farooq, M., Anjum, I., Janzen, E., Reza Toliat, M., ... Hansen, L. (2013). Genetic heterogeneity in Pakistani microcephaly families. *Clinical Genetics* 83(5):446-451. <https://doi.org/10.1111/j.1399-0004.2012.01932.x>
- Shaheen, R., Maddirevula, S., Ewida, N., Alsahli, S., Abdel-Salam, G. M. H., Zaki, M. S., ... Alkuraya, F. S. (2019). Genomic and phenotypic delineation of congenital microcephaly. *Genetics in Medicine* 21(3):545-552. <https://doi.org/10.1038/s41436-018-0140-3>
- Shinawi, M., Liu, P., Kang, S. H., Shen, J., Belmont, J. W., Scott, D. A., ... Lupski, J. R. (2010). Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *Journal of Medical Genetics* 47(5):332-341. <https://doi.org/10.1136/jmg.2009.073015>

Shoukier, M., Klein, N., Auber, B., Wickert, J., Schroder, J., Zoll, B., ... Zirn, B. (2013). Array CGH in patients with developmental delay or intellectual disability: are there phenotypic clues to pathogenic copy number variants? *Clinical Genetics* 83(1):53-65. <https://doi.org/10.1111/j.1399-0004.2012.01850.x>

Stewart, D. R., & Kleefstra, T. (2007). The chromosome 9q subtelomere deletion syndrome. *American Journal of Medical Genetics. Part C: Seminars in Medical Genetics* 145C(4):383-392. <https://doi.org/10.1002/ajmg.c.30148>

Takeuchi, T., Yamazaki, Y., Katoh-Fukui, Y., Tsuchiya, R., Kondo, S., Motoyama, J., & Higashinakagawa, T. (1995). Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation. *Genes and Development* 9(10):1211-1222. <https://doi.org/10.1101/gad.9.10.1211>

Thevenon, J., Duffourd, Y., Masurel-Paulet, A., Lefebvre, M., Feillet, F., El Chehadeh-Djebbar, S., ... Riviere, J. B. (2016). Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. *Clinical Genetics* 89(6):700-707. <https://doi.org/10.1111/cge.12732>

Tsoutsou, E., Tzetis, M., Giannikou, K., Braoudaki, M., Mitrakos, A., Amenta, S., ... Fryssira, H. (2017). Application of high-resolution array comparative genomic hybridization in children with unknown syndromic microcephaly. *Pediatric Research* 82(2):253-260. <https://doi.org/10.1038/pr.2017.65>

Vargas, J. E., Allred, E. N., Leviton, A., & Holmes, L. B. (2001). Congenital microcephaly: phenotypic features in a consecutive sample of newborn infants. *Journal of Pediatrics* 139(2):210-214. <https://doi.org/10.1067/mpd.2001.115314>

Visser, L., van Nimwegen, K. J. M., Schieving, J. H., Kamsteeg, E. J., Kleefstra, T., Yntema, H. G., ... Willemsen, M. (2017). A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. *Genetics in Medicine* 19(9):1055-1063. <https://doi.org/10.1038/gim.2017.1>

von der Hagen, M., Pivarcsi, M., Liebe, J., von Bernuth, H., Didonato, N., Hennermann, J. B., ... Kaindl, A. M. (2014). Diagnostic approach to microcephaly in childhood: a two-center study and review of the literature. *Developmental Medicine and Child Neurology* 56(8):732-741. <https://doi.org/10.1111/dmcn.12425>

Wang, J., Ai, X., Gale, R. P., Xu, Z., Qin, T., Fang, L., ... Xiao, Z. (2013). TET2, ASXL1 and EZH2 mutations in Chinese with myelodysplastic syndromes. *Leukemia Research* 37(3):305-311. <https://doi.org/10.1016/j.leukres.2012.10.004>

Woods, C. G., & Parker, A. (2013). Investigating microcephaly. *Archives of Disease in Childhood* 98(9):707-713. <https://doi.org/10.1136/archdischild-2012-302882>

## Figure Legends

**Figure 1** Generation process of a candidate variant list using a four-step strategy. Variants with less than 10x coverage were removed in the first stage and variants with low population frequency were selected according to the Genome Aggregation Database (gnomAD) (<http://gnomad.broadinstitute.org/>) or the Korean Reference Genome Database (KRGDB) (<http://coda.nih.go.kr/coda/KRGDB/index.jsp>). The third step was to prioritize variants causing missense, nonsense, frameshifts, and in-frame insertions/deletions variants, or changes affecting consensus splice site sequences. The last step was a gene-specific analysis with an in-silico gene panel composed of 903 genes, filtering for selected phenotype traits Human Phenotype Ontology (HPO)-terms for Microcephaly (HP:0000252) or Online Mendelian Inheritance in Man (OMIM) microcephaly phenotype genes.

**Figure 2** Chromosomal microarray (CMA) and brain magnetic resonance imaging (MRI) of patient M-075. (a) CMA shows a 2.8 Mb deletion at Xp11.4p11.3 containing the *CASK* gene. (b) Brain MRI of patient (M-075) demonstrates the microcephaly with pontine and cerebellar hypoplasia in the T2 weighted sagittal view.

