Whole-Exome Sequencing of Pakistani Consanguineous Families Identified Pathogenic Variants in Genes of Intellectual Disability

Maria Asif [1,2,†], **Maryam Anayat** [3,†], **Faiza Tariq** [3], **Tanzeela Noureen** [3], **Ghulam Naseer Ud Din** [3], **Christian Becker** [1], **Kerstin Becker** [1], **Holger Thiele** [1], **Ehtisham ul Haq Makhdoom** [4,5], **Pakeeza Arzoo Shaiq** [3], **Shahid M. Baig** [5,6,7], **Peter Nürnberg** [1,2], **Muhammad Sajid Hussain** [1,2,*], **Ghazala Kaukab Raja** [3] and **Uzma Abdullah** [3,*]

1   Cologne Center for Genomics (CCG), University of Cologne, Faculty of Medicine and University Hospital Cologne, 50931 Cologne, Germany
2   Center for Molecular Medicine Cologne (CMMC), University of Cologne, Faculty of Medicine and University Hospital Cologne, 50931 Cologne, Germany
3   University Institute of Biochemistry and Biotechnology (UIBB), PMAS-Arid Agriculture University Rawalpindi (PMAS-AAUR), Rawalpindi 46300, Pakistan
4   Neurochemical Biology and Genetics Laboratory (NGL), Department of Physiology, Faculty of Life Sciences, Government College University, Faisalabad 38000, Pakistan
5   Human Molecular Genetics Laboratory, Health Biotechnology Division, National Institute for Biotechnology and Genetic Engineering (NIBGE) College, PIEAS, Faisalabad 38000, Pakistan
6   Department of Biological and Biomedical Sciences, The Aga Khan University, Karachi 74800, Pakistan
7   Pakistan Science Foundation (PSF), Islamabad 44050, Pakistan
*   Correspondence: mhussain@uni-koeln.de (M.S.H.); uzma.abdullah@uaar.edu.pk (U.A.)
†   These authors contributed equally to this work.

Abstract: Intellectual disability (ID) is a condition of significant limitation of cognitive functioning and adaptive behavior, with 50% of etiology attributed to genetic predisposition. We recruited two consanguineous Pakistani families manifesting severe ID and developmental delay. The probands were subjected to whole exome sequencing (WES) and variants were further prioritized based on population frequency, predicted pathogenicity and functional relevance. The WES data analysis identified homozygous pathogenic variants in genes *MBOAT7* and *TRAPPC9*. The pathogenicity of the variants was supported by co-segregation analysis and in silico tool. The findings of this study expand mutation spectrum and provide additional evidence to the role of MBOAT7 and TRAPPC9 in causation of ID.

Keywords: intellectual disability; whole-exome sequencing; consanguineous; *MBOAT7*; *TRAPPC9*

1. Introduction

Intellectual disability (ID) is a condition of significant limitation of cognitive functioning and adaptive behavior (intelligent quotient lower than 70) manifesting before the age of 18 years old [1]. ID has a very heterogenous clinical presentation (syndromic and non-syndromic) and its etiology is attributed to both environmental and inherited aberrations [2]. The global incidence of ID ranges from 1 to 3% [3]. Patients need lifetime care that, in addition to the psychological toll it takes, casts serious financial strain on the families and healthcare system alike [4]. ID is observed in all patterns of mendelian inheritance and with a genetic etiology as diverse as large chromosomal abnormalities, submicroscopic copy number variants, and monogenic forms due to pathogenic variants in single genes [5–7]. The autosomal recessive form of ID (ARID) has been rare and therefore the least explored [8]. With the increasing feasibility of next-generation sequencing, more cases of ARID are being explored, which is revealing its ever increasing genetic heterogeneity [8]. According to SysNDD (https://sysndd.dbmr.unibe.ch/Entities (accessed on 30

October 2022)), 995 genes have been definitely linked to ARID, though the magnitude of candidate genes greatly exceeds this number [9] and is expected to be around 3000 [8]. This is reflected in a consistently low diagnostic yield of genetic diagnostic for ID, which has not exceeded 50%, at best [10,11]. In addition, an increasing adaption of next-generation sequencing for prenatal diagnosis and stronger evidence for pathogenicity of reported variants are required for accurate risk estimate; it urges the need for investigation of more informative pedigrees and deep phenotyping.

Here, we report two consanguineous families with severe intellectual disability that were analyzed with next-generation sequencing, segregating a novel and known mutation in the genes of ARID.

2. Materials and Methods

2.1. Subjects

The study was approved by "The Ethics Committee for the Use of Human Subjects" of Pir Mehr Ali Shah Arid Agriculture University Rawalpindi (PMAS-AAUR). We recruited two consanguineous families from the remote area of Punjab province of Pakistan. The families were educated about the aim and nature of the study in their native language and written informed consent was obtained as per the rules described in the Declaration of Helsinki. From all consenting individuals, peripheral blood was collected in EDTA-coated tubes and stored at 4 °C.

2.2. Whole-Exome Sequencing

We extracted genomic DNA from blood samples of all individuals using phenol chloroform extraction method. The DNA of affected individual IV-1 from family A and IV-4 from family B were subjected to whole-exome sequencing. For this purpose, we used Agilent SureSelectXT HS Human All Exon V8 enrichment kit and the samples were subsequently run on an Illumina HiSeq 2000 sequencing system (paired-end reads, 2 × 100 bp) and data were analyzed as described before [12]. For variant interpretation, we used VARBANK, our in-house database and analysis platform (https://varbank.ccg.uni-koeln.de/varbank2/ (accessed on 1 July 2022)) of the Cologne Center of Genomics (CCG), University of Cologne, Germany.

2.3. Co-Segregation Analysis

The candidate variants were further analyzed through bidirectional Sanger sequencing of variant-spanning genomic fragments in all family members. The reference sequences were retrieved from the UCSC Human genome browser and sequences were analyzed through UGENE (Version 45.0).

2.4. In Silico Analysis

To assess the likely pathogenicity of the variants, multiple in silico tools were employed, using web interface of Variant effect predictor (VEP) available at Ensemble (https://asia.ensembl.org/info/docs/tools/vep/index.html (accessed on 15 August 2022)) and Varsom (https://varsome.com/ (accessed on 14 October 2022)).

3. Results

3.1. Clinical Findings

The families, here designated as A and B (pedigrees provided in Figure 1A,B), consisted of two affected individuals each, born as a result of full-term, uneventful pregnancies to consanguineous couples. All four patients presented global developmental delay, severe intellectual disability and speech impairment. The facial features of the patients were evaluated by a clinical geneticist and confirmed to be normal, as shown in Figure 1C,D. The patients also presented variable features, such as microcephaly, reduced height, walking abnormality and seizures. The clinical features of each individual are detailed in Table 1.
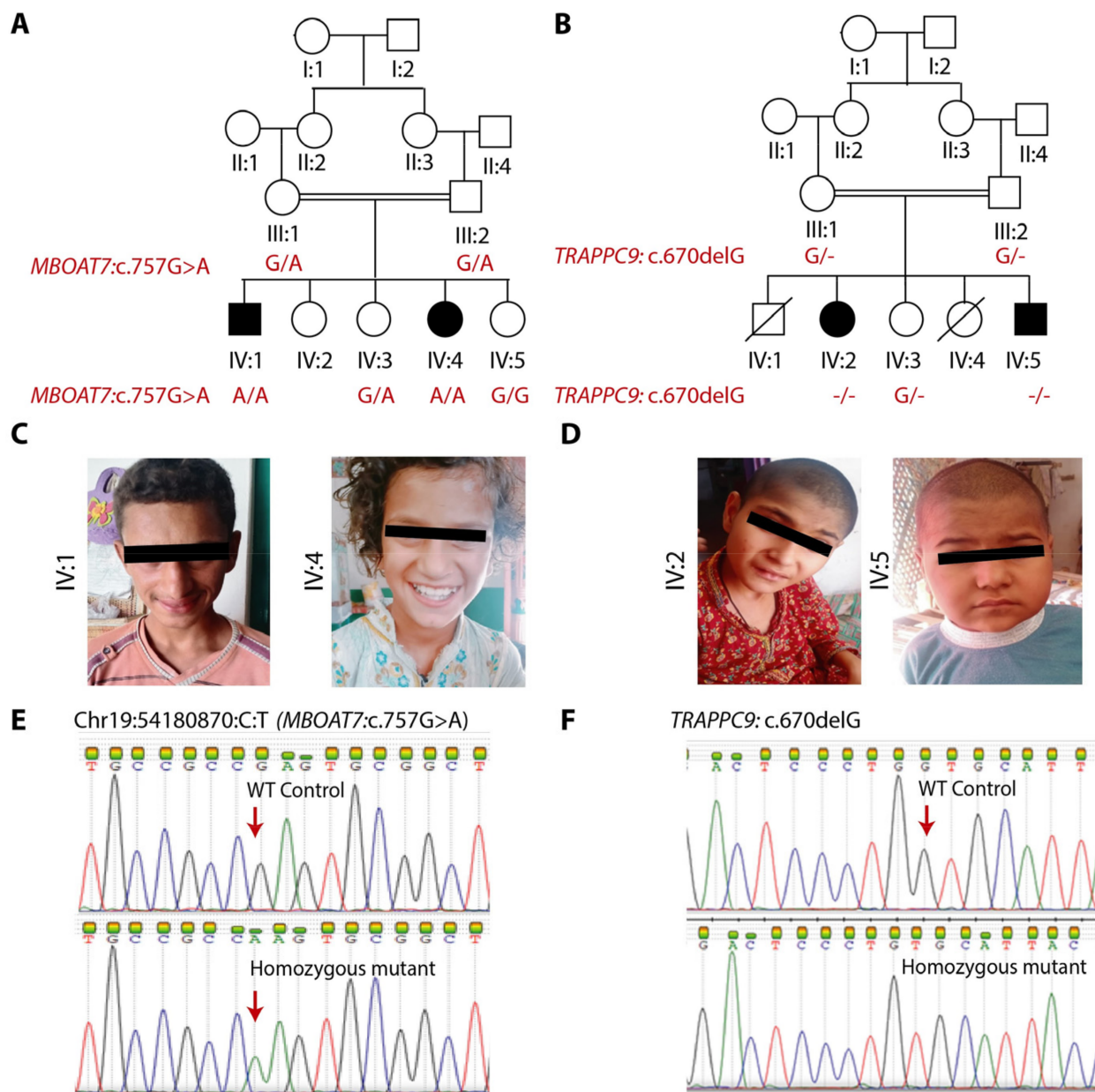
Figure 1. Clinical and genetic findings of families A and B (**A**) Pedigree of family A along with genotypes of candidate variant in *MBOAT7* showing its co-segregation. (**B**) Pedigree of family B along with genotypes of candidate variant in *TRAPPC9* (**C**) Frontal poses of patients of family A showing normal facial features. (**D**) Frontal poses of affected individuals of family B without any signs of facial deformities. (**E**) Sanger traces of the genomic region of *MBOAT7* showing identified variant in the affected member compared to the control. Arrow shows the position of the mutant nucleotide. (**F**) Sanger traces for mutation spanning region in of *TRAPPC9* showing wild-type sequence in a control individual in the upper panel, whereas it shows homozygous deletion c.670delG in the lower panel, as observed in both patients.

Table 1. Clinical information of the families A and B.

Family IDs	Family A		Family B	
Mutation	*MBOAT7* (NM_024298.4: c.757G>A; p.(Glu253Lys))		*TRAPPC9* (NM_001160372.3: c.670delG; p.(Val224Cysfs*13))	
Patient ID	IV:1	IV:4	IV:2	IV:5
Sex	Male	Female	Female	Male
Age at examination	16 years	10 years	10 years	4 years
Height	175.25 cm	137.20 cm (−0.3 SD)	121.90 cm (−3 SD)	86.35 cm (−4 SD)
HC	50.10 cm (−4 SD)	53.35 cm (+1 SD)	45.72 cm (−5 SD)	43.20 cm (−7 SD)
Intellectual disability	Severe	Severe	Severe	Severe
Self-care	No	No	No	No
Self-feed	No	No	No	No
Gestational history	Unremarkable	Unremarkable	Unremarkable	Unremarkable
Sitting	1 year	2 years	2 years	2 years
Walking	3 years	3.5 years	5 years	Not attained
Speech Impairment	Yes	Yes	Yes	Yes
Visual Impairment	No	No	No	No
Ambulation Difficulty	No	No	No	No
Facial Dysmorphism	No	No	No	No
Aggressive	Yes	Yes	N/A	N/A
Seizures	No	Yes, started at the age of 5 months	Yes, onset at the age of 4 years	Yes, onset at the age of 1.5 years
Skin anomalies	No	No	No	No

3.2. Genetic Findings

WES analysis identified a presumable homozygous loss-of-function variant *MBOAT7*, NM_024298.4:c.757G>A (Figure 1E). It is a missense variant and is predicted to substitute a highly conserved amino acid residue (p.(Glu253Lys)). Sanger sequences confirmed the homozygosity of the variant in both patients, whereas parents were heterozygous carrier (Figure 1A,E). For family B, a homozygous variant in another gene, *TRAPPC9*, NM_001160372.3,c.670delG, was identified. It is predicted to cause frameshift followed by premature termination p.(Val224Cysfs*13). The Sanger sequencing (Figure 1F) showed co-segregation of the variant, confirming the autosomal recessive mode of inheritance, as shown in Figure 1B.

3.3. In Silico Findings

Both of the candidate variants showed strong support for predicted pathogenicity from tools. The findings of in silico tools for the variant MOBAT7, NM_024298.5: c.757G>A, p.(Glu253Lys) are provided in Table 2. The variant is highly conserved across the animal kingdom (the conservation in some representative species is shown in Supplementary Figure S1). The *TRAPPC9*, NM_001160372.3, c.670delG is predicted to be likely pathogenic as per ACMG classification (9 points = 9 P − 0 B) and its conservation score on phyloP100 is 5.888.

Table 2. In silico analyses for pathogenicity of *MBOAT7*, NM_024298.5: c.757G>A, p.(Glu253Lys).

Prediction Tool	Prediction and Score
gnomAD frequency	1/152, 184 (0.000006571)
CADD	27.1
PolyPhen-2	Probably damaging (0.999)
DANN	Pathogenic (0.9993)
EIGEN	Pathogenic (0.7627)
SIFT	Deleterious
Mutation Taster	Disease-causing (56)
Polyphen-2	Probably damaging (1.000)
Conservation Scores	phyloP100: 4.065
EIGEN PC	Pathogenic (0.7029)
PROVEAN	Uncertain (3.28, 3.46, 3.05, 3.03)
ACMG Classification	Uncertain significance (PM1, PP3 and BP1)
SNAP^2	Neutral −41 (72%)
DEOGEN2	Uncertain (Damaging, Tolerated) 0.5828, 0.4976
MetaRNN	Pathogenic
FATHMM-MKL	Pathogenic (0.9838)
FATHMM-XF	Uncertain (0.8913)
MutPred	Pathogenic (0.84)
PANTHER	(Probably damaging) (0.74)
REVEL	Pathogenic (0.7459)
BayesDel addAF	Pathogenic (addAF score 0.2137)

4. Discussion

The study reports the clinical and genotypic findings of two consanguineous families with representative symptoms of intellectual disability. Although ID is a well-characterized disorder, the genetic and clinical heterogeneity of ID in general and its autosomal recessive form are partially documented [8]. The human brain receives the majority of genomic expression and many of these genes are underexplored. The genomic dissection of ARID holds the potential to generate a wealth of data that can contribute to a better understanding of the development and function of the human brain [8]. With the evolving application and down pricing of next-generation sequencing, it is increasingly possible to analyze smaller pedigrees, thus increasing the possibility of novel findings.

Family A segregated missense mutation in *MBOAT7*. The gene encodes membrane-bound O-acyltransferase domain-containing 7 (MBOAT7) protein, which is lysophospholipid acyltransferase 7 enzyme, functions for the incorporation of arachidonic acid into glycophosphatidylinositol and is involved in phosphatidylinositol acyl-chain remodeling in the Lands cycle [13]. The process is highly regulated: deficiency of MBOAT7 leads to elevated levels of arachidonic acid that compromise cellular physiology [14], including that of neurons. In 2016, biallelic loss-of-function mutations in this gene were identified to cause ARID by Johansen et al. [15], and since then, 19 different mutations have been identified in 51 patients, causing overlapping symptoms of developmental delay, intellectual disability, seizures and with variable reports of structural brain anomalies assessed by magnetic resonance imaging (MRI). As detailed in Supplementary Table S1, the patients with biallelic mutations in *MBOAT7* also present the most frequent co-segregation of seizures (40/51), speech impairment (45/51), delayed acquisition of motor milestones (23/51), occasionally autistic features (16/51), and microcephaly (01/51). The mutation observed in family A,

MOBAT7: NM_024298.5: c.757G>A; p.(Glu253Lys), has been previously reported in an unrelated Pakistani family [16]. The patients of the reported family manifested generalized tonic colonic seizures with the latest age of onset being 5 years, severe global developmental delay, severe intellectual disability and mild atrophy of the cerebellar cortex [16]. The patients of family A presented symptoms overlapping with the aforementioned reported patients, including severe intellectual disability, developmental delay, and speech impairment/deficiency. The patients IV:1, however, never developed epilepsy, even at the age of 16 years. The patients also had a normal walk and the disease appears to be static.

MBOAT family proteins, including MBOAT7, adopt a conserved helical core structure, comprising an extracellular funnel and intracellular tunnel [17]. Multiple in silico tools for the modeling of the mutation indicated that the substitution of p.(Glu253Lys) may cause the mutant protein to form ectopic hydrogen bonds that restrict its conformational flexibility of funnel, ultimately compromising its function [16], leading to the observed phenotype.

Family B shows a novel homozygous mutation in *TRAPPC9*. Its encoded protein is a subunit of the TRAPP complex (TRAPPC), a guanine nucleotide exchange factor for rab proteins that operate in secretory, endocytic, and autophagic pathways [18]. TRAPPC9 is the main component of TRAPP complex type II, also knows as TRAPII, which facilitates endoplasmic reticulum (ER)-to-Golgi vesicular transport, the transport within the Golgi, and the transport between Golgi and endosomes [19]. Biallelic loss-of-function mutations have been observed with intellectual disability, along with variable clinical co-segregation of obesity, hypoplastic corpus callosum, microcephaly, behavioral changes, and dysmorphic features [19–21]. The clinical features of the reported families have been summarized in Supplementary Table 2. As apparent in the table, the most frequent co-segregating features include microcephaly (39/44), delay or lack of speech acquisition (33/44), and delay of motor milestone or walking abnormality (12/44), whereas less frequently observed features included abnormalities of white matter or thinning of corpus callosum (18/44), more rarely autistic features (4/44) and seizures (2/44). Trappc9-deficient mice have also been observed to have behavioral deficits such as global delay of intellectual development, postnatal microcephaly, thin corpus callosum with imbalance in dopamine D1 and D2 receptor containing neurons [22]. *Trappc9*-null mice also showed decreased contents of neural stem cells in the subventricular zone of the lateral ventricle and the sub granular zone of the dentate gyrus [23], which likely explains the structural abnormalities of the brain observed in patients too.

The longest isoform of wild-type TRAPPC9 consists of 1148 amino acid residues. The mutation identified in family B is NM_001160372.3:c.670delG, which is predicted to shift the wild-type frame that will run into premature stop codon, making a polypeptide chain of 337 amino acids only p.(Val224Cysfs*13). Due to such an early termination, the mutant transcript is likely to be subjected to non-sense mediated decay. Thus, it will function as a null mutation, causing the complete absence of the protein.

Various types of homozygous null mutations of *TRAPPC9* that lead to ARID along with microcephaly and occasionally seizure (Supplementary Table S2) have been reported. Our patients exhibited similar symptoms of microcephaly, short stature, seizures and intellectual disability.

In conclusion, the study adds to the phenotypic and genetic spectrum of *MBOAT7*- and *TRAPPC9*-related intellectual disability. This knowledge adds evidence to the role of these two genes in the nervous system's development. It also adds to the genetic spectrum of ID in the Pakistani population.

Author Contributions: Conceptualization P.A.S., S.M.B., G.K.R.; methodology, M.A. (Maria Asif); validation, T.N., U.A. and M.S.H.; formal analysis, T.N., C.B. and K.B.; investigation, F.T., G.N.U.D.,

References

1. Schalock, R.L.; Borthwick-Duffy, S.A.; Bradley, V.J.; Buntinx, W.H.; Coulter, D.L.; Craig, E.M.; Gomez, S.C.; Lachapelle, Y.; Luckasson, R.; Reeve, A. *Intellectual Disability: Definition, Classification, and Systems of Supports*, 11th ed.; AAIDD: Washington, DC, USA, 2010.
2. Schalock, R.L.; Luckasson, R. A systematic approach to subgroup classification in intellectual disability. *Intellect. Dev. Disabil.* **2015**, *53*, 358–366. [CrossRef] [PubMed]
3. Leonard, H.; Wen, X. The epidemiology of mental retardation: Challenges and opportunities in the new millennium. *Ment. Retard. Dev. Disabil. Res. Rev.* **2002**, *8*, 117–134. [CrossRef]
4. Genereaux, D.; van Karnebeek, C.D.; Birch, P.H. Costs of caring for children with an intellectual developmental disorder. *Disabil. Health J.* **2015**, *8*, 646–651. [CrossRef]
5. Shaw-Smith, C.; Redon, R.; Rickman, L.; Rio, M.; Willatt, L.; Fiegler, H.; Firth, H.; Sanlaville, D.; Winter, R.; Colleaux, L. Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *J. Med. Genet.* **2004**, *41*, 241–248. [CrossRef] [PubMed]
6. Gilissen, C.; Hehir-Kwa, J.Y.; Thung, D.T.; van de Vorst, M.; van Bon, B.W.; Willemsen, M.H.; Kwint, M.; Janssen, I.M.; Hoischen, A.; Schenck, A. Genome sequencing identifies major causes of severe intellectual disability. *Nature* **2014**, *511*, 344–347. [CrossRef] [PubMed]
7. Vissers, L.E.; de Ligt, J.; Gilissen, C.; Janssen, I.; Steehouwer, M.; de Vries, P.; van Lier, B.; Arts, P.; Wieskamp, N.; del Rosario, M. A de novo paradigm for mental retardation. *Nat. Genet.* **2010**, *42*, 1109–1112. [CrossRef]
8. Hu, H.; Kahrizi, K.; Musante, L.; Fattahi, Z.; Herwig, R.; Hosseini, M.; Oppitz, C.; Abedini, S.S.; Suckow, V.; Larti, F. Genetics of intellectual disability in consanguineous families. *Mol. Psychiatry* **2019**, *24*, 1027–1039. [CrossRef]
9. Vissers, L.E.; Gilissen, C.; Veltman, J.A. Genetic studies in intellectual disability and related disorders. *Nat. Rev. Genet.* **2016**, *17*, 9. [CrossRef]
10. Leite, A.J.d.C.; Pinto, I.P.; Leijsten, N.; Ruiterkamp-Versteeg, M.; Pfundt, R.; de Leeuw, N.; da Cruz, A.D.; Minasi, L.B. Diagnostic yield of patients with undiagnosed intellectual disability, global developmental delay and multiples congenital anomalies using karyotype, microarray analysis, whole exome sequencing from Central Brazil. *PLoS ONE* **2022**, *17*, e0266493. [CrossRef]
11. Han, J.Y.; Lee, I.G. Genetic tests by next-generation sequencing in children with developmental delay and/or intellectual disability. *Clin. Exp. Pediatr.* **2020**, *63*, 195. [CrossRef]
12. Hussain, M.S.; Baig, S.M.; Neumann, S.; Nürnberg, G.; Farooq, M.; Ahmad, I.; Alef, T.; Hennies, H.C.; Technau, M.; Altmüller, J. A truncating mutation of CEP135 causes primary microcephaly and disturbed centrosomal function. *Am. J. Hum. Genet.* **2012**, *90*, 871–878. [CrossRef]
13. Caddeo, A.; Jamialahmadi, O.; Solinas, G.; Pujia, A.; Mancina, R.M.; Pingitore, P.; Romeo, S. MBOAT7 is anchored to endomembranes by six transmembrane domains. *J. Struct. Biol.* **2019**, *206*, 349–360. [CrossRef]

14. Pérez-Chacón, G.; Astudillo, A.M.; Balgoma, D.; Balboa, M.A.; Balsinde, J. Control of free arachidonic acid levels by phospholipases A2 and lysophospholipid acyltransferases. *Biochim. Biophys. Acta BBA Mol. Cell Biol. Lipids* **2009**, *1791*, 1103–1113. [CrossRef]

15. Johansen, A.; Rosti, R.O.; Musaev, D.; Sticca, E.; Harripaul, R.; Zaki, M.; Çağlayan, A.O.; Azam, M.; Sultan, T.; Froukh, T. Mutations in MBOAT7, encoding lysophosphatidylinositol acyltransferase I, lead to intellectual disability accompanied by epilepsy and autistic features. *Am. J. Hum. Genet.* **2016**, *99*, 912–916. [CrossRef]

16. Lee, J.; Shamim, A.; Park, J.; Jang, J.; Kim, J.H.; Kwon, J.; Kim, J.-W.; Kim, K.K.; Lee, J. Functional and structural changes in the membrane-bound Oacyltransferase family member 7 (MBOAT7) protein. *Front. Neurol.* **2022**, *13*, 836954. [CrossRef]

17. Ma, D.; Wang, Z.; Merrikh, C.N.; Lang, K.S.; Lu, P.; Li, X.; Merrikh, H.; Rao, Z.; Xu, W. Crystal structure of a membrane-bound O-acyltransferase. *Nature* **2018**, *562*, 286–290. [CrossRef]

18. Barrowman, J.; Bhandari, D.; Reinisch, K.; Ferro-Novick, S. TRAPP complexes in membrane traffic: Convergence through a common Rab. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 759–763. [CrossRef]

19. Radenkovic, S.; Martinelli, D.; Zhang, Y.; Preston, G.J.; Maiorana, A.; Terracciano, A.; Dentici, M.L.; Pisaneschi, E.; Novelli, A.; Ranatunga, W. TRAPPC9-CDG: A novel congenital disorder of glycosylation with dysmorphic features and intellectual disability. *Genet. Med.* **2022**, *24*, 894–904. [CrossRef]

20. Bolat, H.; Ünsel-Bolat, G.; Derin, H.; Şen, A.; Ceylaner, S. Distinct Autism Spectrum Disorder Phenotype and Hand-Flapping Stereotypes: Two Siblings with Novel Homozygous Mutation in TRAPPC9 Gene and Literature Review. *Mol. Syndromol.* **2022**, *13*, 263–269. [CrossRef]

21. Alvarez-Mora, M.I.; Corominas, J.; Gilissen, C.; Sanchez, A.; Madrigal, I.; Rodriguez-Revenga, L. Novel Compound Heterozygous Mutation in TRAPPC9 Gene: The relevance of whole genome sequencing. *Genes* **2021**, *12*, 557. [CrossRef]

22. Ke, Y.; Weng, M.; Chhetri, G.; Usman, M.; Li, Y.; Yu, Q.; Ding, Y.; Wang, Z.; Wang, X.; Sultana, P. Trappc9 deficiency in mice impairs learning and memory by causing imbalance of dopamine D1 and D2 neurons. *Sci. Adv.* **2020**, *6*, eabb7781. [CrossRef] [PubMed]

23. Usman, M.; Li, Y.; Ke, Y.; Chhetri, G.; Islam, M.A.; Wang, Z.; Li, X. Trappc9 Deficiency Impairs the Plasticity of Stem Cells. *Int. J. Mol. Sci.* **2022**, *23*, 4900. [CrossRef] [PubMed]