# Biallelic *MFSD2A* variants associated with congenital microcephaly, developmental delay, and recognizable neuroimaging features.

**Running title: *MFSD2A*-related congenital microcephaly.**

The work was supported by National Research Foundation grants (NRF2016NRF-NRFI001-15); Biomedical Research Council of A*STAR; March of Dimes Research Grant; National Institute for Health Research University College London Hospitals Biomedical Research Centre.

Marcello Scala[1,2,3*], Geok Lin Chua[4*], Cheen Fei Chin[4], Hessa S Alsaif[5], Artem Borovikov[6], Saima Riazuddin[7], Sheikh Riazuddin[8,9], M. Chiara Manzini[10], Mariasavina Severino[11], Alvin Kuk[4], Hao Fan[12,13,14], Yalda Jamshidi[15], Mehran Beiraghi Toosi[16], Mohammad Doosti[16], Ehsan Ghayoor Karimiani[16], Vincenzo Salpietro[1,2], Elena Dadali[6], Galina Baydakova[6], Fedor Konovalov[17,18], Ekaterina Lozier[17,18], Emer O'Connor[1], Yasser Sabr[19], Abdullah Alfaifi[20], Farah Ashrafzadeh[21], Pasquale Striano[2,3], Federico Zara[2,22], Fowzan S Alkuraya[23,24], Henry Houlden[1], Reza Maroofian [1,15¶], David L. Silver[4¶].

[1] Department of Neuromuscular Disorders, Institute of Neurology, University College London, London, United Kingdom

[2] Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Genoa, Italy

[3] Pediatric Neurology and Muscular Diseases Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy

[4] Signature Research Program in Cardiovascular and Metabolic Disorders, Duke-NUS Medical School, Singapore, 169857, Singapore

[5] Department of Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia

[6] Research Centre for Medical Genetics, Moscow, Russia

[7] Department of Otorhinolaryngology Head & Neck Surgery, School of Medicine, University of Maryland, Baltimore, MD 21201, USA

[8] Center for Genetic Diseases, Shaheed Zulfiqar Ali Bhutto Medical University, Pakistan Institute of Medical Sciences, Islamabad, Pakistan

[9] National Centre of Excellence in Molecular Biology, University of the Punjab, Lahore 53700, Pakistan

[10] Department of Neuroscience and Cell Biology and Child Health Institute of New Jersey, Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ 08901, USA

[11] Neuroradiology Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy

[12] Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), 30 Biopolis St., Matrix No. 07-01, Singapore, 138671, Singapore

[13] Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore, 117543, Singapore.

[14] Centre for Computational Biology, DUKE-NUS Medical School, 8 College Road, Singapore, 169857, Singapore

[15] Genetics Research Centre, Molecular and Clinical Sciences Institute, St George's, University of London, Cranmer Terrace, London SW17 0RE, UK

[16] Department of pediatric diseases, Faculty of medicine, Mashhad University of Medical Sciences, Mashhad, Iran

[17] Independent Clinical Bioinformatics Laboratory, Moscow, Russia

[18] Genomed Ltd., Moscow, Russia

[19] Department of Obstetrics and Gynecology, King Saudi University, Riyadh, Saudi Arabia

[20] Pediatrics Department, Security Forces Hospital, Riyadh, Saudi Arabia

[21] Department of Pediatric Diseases, Mashhad University of Medical Sciences, Mashhad, Iran

[22] Unit of Medical Genetics, IRCCS Istituto Giannina Gaslini, Genova Italy

[23] Department of Genetics, King Faisal Specialist Hospital and Research Center, Saudi Arabia

[24] Department of Anatomy and Cell Biology, College of Medicine, Alfaisal University, Riyadh, Saudi Arabia

* These authors contributed equally to this work

¶ Correspondence should be addressed to R.M. (r.maroofian@ucl.ac.uk, +44 (0) 203448 4069 (Internal x84069)) and D.L.S. (david.silver@duke-nus.edu.sg, +65 6516 7666).

## Abstract

Major Facilitator Superfamily Domain containing 2a (*MFSD2A*) is an essential endothelial lipid transporter at the blood-brain barrier. Biallelic variants affecting function in *MFSD2A* cause autosomal recessive primary microcephaly 15 (MCPH15, OMIM# 616486). We sought to expand our knowledge of the phenotypic spectrum of MCPH15 and demonstrate the underlying mechanism of inactivation of the MFSD2A transporter. We carried out detailed analysis of the clinical and neuroradiological features of a series of 27 MCPH15 cases, including eight new individuals from seven unrelated families. Genetic investigation was performed through exome sequencing (ES). Structural insights on the human Mfsd2a model and in-vitro biochemical assays were used to investigate the functional impact of the identified variants. All patients had primary microcephaly and severe developmental delay. Brain MRI showed variable degrees of white matter reduction, ventricular enlargement, callosal hypodysgenesis, and pontine and vermian hypoplasia. ES led to the identification of six novel biallelic *MFSD2A* variants (NG_053084.1, NM_032793.5: c.556+1G>A, c.748G>T; p.(Val250Phe), c.750_753del; p.(Cys251SerfsTer3), c.977G>A; p.(Arg326His), c.1386_1435del; p.(Gln462HisfsTer17), and c.1478C>T; p.(Pro493Leu)) and two recurrent variants (NM_032793.5: c.593C>T; p.(Thr198Met) and c.476C>T; p.(Thr159Met)). All these variants and the previously reported NM_032793.5: c.490C>A; p.(Pro164Thr) resulted in either reduced MFSD2A expression and/or transport activity. Our study further delineates the phenotypic spectrum of MCPH15, refining its clinical and neuroradiological characterization and supporting that *MFSD2A* deficiency causes early prenatal brain

developmental disruption. We also show that poor MFSD2A expression despite normal transporter activity is a relevant pathomechanism in MCPH15.

**Keywords: *MFSD2A*; microcephaly; developmental delay; brain MRI.**

## Introduction

Major Facilitator Superfamily Domain containing 2a (*MFSD2A*) is a sodium-dependent lysophosphatidylcholine (LPC) transporter that is highly expressed at the endothelium of the blood-brain barrier (BBB).[1] Omega-3 fatty acids and other mono- and polyunsaturated fatty acids conjugated as LPCs are transported by MFSD2A, which plays a pivotal role in the supply of omega-3 fatty acids to the brain[1]. The essential role of *MFSD2A* in regulating lipogenesis in the developing brain has been recently demonstrated using loss-of-function mouse models.[2]

Five distinct homozygous loss-of-function *MFSD2A* variants have been reported in patients with neurodevelopmental abnormalities from seven consanguineous families. These patients showed developmental delay (DD), microcephaly, and neuroimaging abnormalities such as ventriculomegaly and hypoplasia of the corpus callosum, brainstem, and cerebellum. These observations underscored the fundamental role of LPC transport at the BBB for human brain development and clarified the structure-function relationships in the MFSD2A-mediated transport mechanism.[3-9]

In this study, we report seven new families with biallelic variants affecting function in *MFSD2A*, expanding the phenotype and defining the characteristic neuroimaging features of *MFSD2A*-related neurodevelopmental disorder, also known as Autosomal Recessive Microcephaly 15, (MCPH15, OMIM #616486). We provide clinical, genetic, and functional characterization of these novel variants and the previously reported NM_032793.5:c.593C>T; p.(Thr198Met) and c.490C>A; p.(Pro164Thr) variants on the transporter activity, which further substantiates the functional importance of LPC transport for human brain development.

## Materials and methods

### Patients ascertainment

Eight patients from seven unrelated families were locally referred for exome sequencing (ES) in the context of severe microcephaly and psychomotor delay. Patients were enrolled in accordance with the Declaration of Helsinki and informed consent was obtained for all of them in agreement with the requirements of Iranian, Pakistani, Russian, and Saudi bioethics laws. Subjects were examined by several geneticists, neurologists, and pediatricians with expertise in pediatric neurology. Detailed family history was collected for all families. Brain MRI were locally acquired with different protocols, but all included diffusion weighted images, T1 and T2-weighted, and FLAIR images on the 3 planes. Images were reviewed by an experienced pediatric neuroradiologist (MS) and a pediatrician with expertise in neurogenetics (MS) in consensus. Blood samples were obtained from patients and parents.

## Exome Sequencing

After standard DNA extraction from peripheral blood, proband-only ES was performed in all the families as previously described.[10-12] Variants were filtered out according to frequency, conservation, and predicted impact on protein function by several bioinformatic tools (SIFT, Polyphen-2, Mutation Taster). Candidate variants were subsequently validated through co-segregation studies by Sanger sequencing and submitted to the gene variant database LOVD at https://databases.lovd.nl/shared/genes/MFSD2A (Individual IDs 00276067, 00276070, 00276071, 00276074, 00276075, 00276076, 00276077). All the variants are reported according to the NM_032793.5 transcript. GeneMatcher was used for the distributed case-matching.[13] Further details available in the Supplementary Methods.

## Functional tests summary methods

Site-directed mutagenesis was used to create the Mfsd2a variants NM_032793.5:c.1478C>T; p.(Pro493Leu), c.593C>T; p.(Thr198Met), c.490C>A; p.(Pro164Thr), c.977G>A; p.(Arg326His), and c.748G>T; p.(Val250Phe) in a mammalian expression vector, which were used to determine the effects on transporter function in mammalian cells. The amino acid variants in Mfsd2a protein were modeled and visualized to understand the causative mechanism of transporter dysfunction. Further details are available in the Supplementary Methods.

# Results

## Clinical features

We present eight patients (Table 1) from seven unrelated families of varying ancestry (Saudi, Iranian, Pakistani, and Russian), including six consanguineous families (Families A, B, C, E, F, and G) (Fig. 1a, b).

Patient 1 (Family A) is a 4-year-old female born to consanguineous parents (first-cousins) of Iranian ancestry. Prenatal ultrasound revealed microcephaly. At birth, her occipital frontal circumference (OFC) was 28 cm (-4.6 SDS). At the age of 6 months, she had head-lag, was unable to roll over, and lacked babbling. At 1 year of age, she started to suffer from myoclonic seizures and failure to thrive (FTT) due to dysphagia. Physical examination at 4 years showed progressive microcephaly with an OFC of 41 cm (-5.6 SDS) and bilateral talipes equinovarus (TEV). She was unable to walk and neurological examination revealed spastic quadriparesis and hyperreflexia. Karyotyping and metabolic testing were normal.

Patient 2 (Family B) is 4-year-old Iranian male born to consanguineous parents. Family history revealed several previous miscarriages. His older brother was healthy. At birth, his OFC was 27 cm (-3.9 SDS). He was diagnosed with global DD during infancy and started to suffer from generalized tonic-clonic seizures since the age of 2 years. At 4 years, he was unable to sit and his language was very limited. Physical examination revealed bilateral TEV, progressive microcephaly with OFC of 37 cm (-8.8 SDS) and spastic quadriparesis.

Patient 3 and 4 (Family C) belong to a consanguineous family of Pakistani descent consisting of six siblings. Two males were reported to have microcephaly and died in the neonatal period due to a possible infection. Two males were healthy. The proband (patient 3), a 17-year-old female, and her sister (patient 4), currently 27 years old, presented with severe global DD and aggressive behavior during infancy. They had no seizure history. Physical evaluation revealed mild muscle weakness, language limited to few words, and severe microcephaly, with an OFC of 49 cm (-5.0 SDS) and 47 cm (-6.9 SDS) in patients 3 and 4, respectively.

Patient 5 (Family D) is the youngest of two siblings born to unrelated parents of Russian descent. Neonatal history was unremarkable except for microcephaly. The baby started to suffer from generalized tonic-clonic seizures at the age of 1 month. Global DD was subsequently diagnosed at 1 year of age as he was unable to sit without support and could not speak. At 5 years, the patient was unable to walk and nonverbal. He had microcephaly with OFC of 46 cm (-3.6 SDS), gross and fine motor impairment, and axial hypotonia. He also had dysphagia, excessive drooling, and some dysmorphic features, including wide nasal bridge and prominent epicanthal folds.

Patient 6 (Family E) is a 1-month-old Saudi female born to consanguineous parents. She was the youngest of four siblings. Her older brother had microcephaly but died during infancy. The patient was diagnosed with severe microcephaly at birth, with an OFC of 28.5 cm (-6.2 SDS). During the neonatal period she suffered from FTT due to severe dysphagia and physical examination further revealed generalized spasticity.

Patient 7 (Family F) is a 2-year-old male born to consanguineous parents from Saudi Arabia. During the neonatal period, he suffered from FTT and received percutaneous endoscopic gastrostomy (PEG) due to severe dysphagia. At 1 year of age, he started to suffer from recurrent seizures treated with phenobarbital and sodium valproate. Developmental milestones were severely delayed. The patient was also diagnosed with gastro-esophageal reflux. Physical examination showed microcephaly, bilateral TEV, generalized muscle weakness, and spasticity.

Patient 8 (Family G) is a 4-month-old female born to consanguineous Saudi parents. Prenatal ultrasound showed microcephaly and foetal echogenic bowel. Perinatal course was uneventful, but at the age of 1 week the baby was admitted to neonatal intensive care unit due to relevant feeding difficulties. At 4 months, she started to suffer from seizures requiring hospitalization. Physical examination showed microcephaly, generalized spasticity, bilateral hip dislocation, and left TEV.

**Neuroimaging**

Brain MRI revealed mild to severe white matter reduction with consequent ventricular dilatation in all subjects (Fig. 1c). In particular, the supratentorial white matter was markedly thinned with severe ventriculomegaly in 5/8 patients. The degree of myelination was appropriate for the age in all subjects. The cortical gyral pattern was mildly to severely simplified in all cases, without other associated cortical malformations. The thalami were small and the corpus callosum was abnormal in all patients. In particular, in 5 subjects the

corpus callosum was markedly thin and short, in 2 patients there was hypoplasia of the anterior portion of the corpus callosum, while in the remaining patient it was globally thin. Of note, the cingulate gyrus was present in all subjects. Finally, inferior vermian hypoplasia was observed in all cases, while pontine hypoplasia was present in 6/8 patients.

**Genetic findings**

After filtering for allele frequency, conservation, and predicted functional impact, biallelic *MFSD2A* variants were prioritized as candidate disease-causing variants. Eight different variants were identified (Fig. 1d), including three homozygous missense variants (c.1478C>T; p.(Pro493Leu) in patient 1; c.593C>T; p.(Thr198Met) in patient 3 and 4; c.476C>T; p.(Thr159Met) in patient 6), a homozygous splice site variant (patient 2: NG_053084.1(NM_032793.5): c.556+1G>A, NC_000001.11(NM_032793.5): c.556+1G>A, LRG_199t1), two homozygous frameshift variants (c.1386_1435del; p.(Gln462HisfsTer17) in patient 7; c.750_753del; p.(Cys251SerfsTer3) in patient 8), and two compound heterozygous missense variants (c.[748G>T];[977G>A], p.[(Val250Phe)];[(Arg326His)] in patient 5) (Table 2). Biparental segregation confirmed the autosomal recessive inheritance model. In Family C (Fig. 1a), unaffected individuals (II-1 and II-3) were heterozygous for the c.593C>T; p.(Thr198Met) variant in *MFSD2A*, whereas the DNA of the deceased individuals (II-2 and II-6) was not available due to their premature death. All the identified variants are absent in the homozygous state and extremely rare in the heterozygous state in the most common population databases

(including our database of 10,000 exomes, gnomAD, Greater Middle East Variome - GME, Iranome, and Ensembl). Missense variants were located at the amino acid residues with high levels of conservation, with a Genomic Evolutionary Rate Profiling (GERP) score between 5.49 to 5.94. The predicted effect on protein function was also consistent with a loss-of-function mechanism, with a Combined Annotation Dependent Depletion (CADD) score ranging from 24.4 to 34. The two frameshift variants are predicted to result in nonsense mediated mRNA decay, likely leading to a functional knock-out. All the identified variants are predicted to be damaging by several bioinformatic tools, such as SIFT, Polyphen-2, and Mutation Taster. The splicing variant c.556+1G>A is predicted to result in aberrant splicing through the alteration of the wildtype (WT) donor site by Human Splice Finder and Variant Effect Predictor.

**Mfsd2a variants lead to loss-of-function and/or loss-of-expression**

Human Mfsd2a is a 530 amino acid glycosylated sodium-dependent MFS transporter composed of 12 conserved transmembrane domains.[7] To understand the consequence of the c.1478C>T; p.(Pro493Leu), c.490C>A; p.(Pro164Thr), c.593C>T; p.(Thr198Met), c.977G>A; p.(Arg326His), and c.748G>T; p.(Val250Phe) variants on the structure and function of Mfsd2a, we utilized a published structural model of human Mfsd2a to carry out bioinformatic predictions.[7] In the c.593C>T; p.(Thr198Met) mutant model, M198 faces the internal cavity of the transporter and forms more favorable hydrophobic interactions with neighboring residues such as F399 from helix X, in comparison to T198 in the WT model

that faces the membrane exterior (Fig. 1e). In the c.1478C>T; p.(Pro493Leu) mutant model, the proline-to-leucine amino acid change results in the extension of helix XII that is stabilized by a hydrophobic cluster formed by sidechains of L493 and three other residues Y294, L297, and F489 (Fig. 1e). In addition, multiple polar interactions observed in the WT model are absent in the c.1478C>T; p.(Pro493Leu) mutant model, including the hydrogen bonding interaction between Y294 and E497 as well as ionic locks between R498 and a negatively charged surface comprising D408, D411, and D412. These ionic locks were previously suggested to be important for the transporter function.[7] Taken together, we observed enhanced hydrophobic packing in both mutant models likely leading to increased structure rigidity and reduced mobility of the transporter, indirectly inactivating the transport of substrate. Additionally, the c.1478C>T; p.(Pro493Leu) mutant would be predicted to show a reduction in transport due to the partial loss of ionic locks.

We next utilized HEK293 cells, which do not endogenously express Mfsd2a, as an in vitro cell system to determine if Mfsd2a variants affect protein expression, localization, and transport function. Mock transfected and the sodium binding transporter inactive mutant p.(Asp97Ala) (p.(D97A)) served as negative controls,[1,7] while WT Mfsd2a served as a positive control. Western blot analysis of WT Mfsd2a showed the multiple protein bands similar to results previously reported for overexpression of Mfsd2a in HEK293 cells,[3,4,6] while all the five mutants c.1478C>T; p.(Pro493Leu), c.593C>T; p.(Thr198Met), c.490C>A; p.(Pro164Thr), c.977G>A; p.(Arg326His), and c.748G>T; p.(Val250Phe) were expressed at less than 30% of WT Mfsd2a (Fig.2a). This low level of protein expression of these five Mfsd2a mutants is consistent with predicted negative effects of these variants on protein

folding (Fig. 1e). Despite low level expression of all five Mfsd2a mutants, immunofluorescence microscopy indicated that all mutants were expressed at the plasma membrane similarly to WT (Fig. 2b).

To directly test the functional consequences of these five variants on LPC transport, we utilized an established transport assay that quantifies net transport of $^{14}$C-LPC-DHA in HEK293 cells. To directly compare transport activity between WT and the five mutants c.1478C>T; p.(Pro493Leu), c.593C>T; p.(Thr198Met), c.490C>A; p.(Pro164Thr), c.977G>A; p.(Arg326His), and c.748G>T; p.(Val250Phe), we first titrated down the amount of plasmid DNA for the transfection of WT Mfsd2a into cells to obtain a comparable expression level of WT to all five mutants. We found that 0.1 µg of WT yielded similarly low levels of expression as cells transfected with 2 µg of mutants (Fig. 2c). Surprisingly, at comparable protein expression levels of WT and mutants, four of the five mutants demonstrated comparable transport of 14C-LPC-DHA in HEK293 cells with c.593C>T; p.(Thr198Met) at 75%, c.490C>A; p.(Pro164Thr) at 82%, c.977G>A; p.(Arg326His) at 104%, and c.748G>T; p.(Val250Phe) at 80% of WT transport activity. Only P493L was similar to non-functional D97A negative control, indicating it is inactive (Fig. 2d).

Previously reported non-synonymous variants in Mfsd2a have been shown to affect transport function but not protein expression.[3,4,6] In our cases, five of the variants (c.593C>T; p.(Thr198Met), c.490C>A; p.(Pro164Thr), c.977G>A; p.(Arg326His), c.748G>T; p.(Val250Phe), and c.1478C>T; p.(Pro493Leu)) were extremely lowly expressed (Fig. 2a). Our findings indicate that poor expression of Mfsd2a, despite normal transporter activity, can also be an underlying cause for severe microcephaly and hypomyelination in these patients,

294 which further defines the etiology of Mfsd2a-related microcephaly.

295

## 296 Discussion

297 MFSD2A is a sodium-dependent 12-pass transmembrane protein belonging to the major
298 facilitator superfamily of secondary transporters. Mfsd2a plays a pivotal role at the BBB for
299 the transport of plasma-derived LPCs conjugated to polyunsaturated fatty acids such as the
300 omega-3 fatty acid docosahexaenoic acid (DHA) to the brain.[1,2,14] The deficiency of the DHA
301 in the brain of *Mfsd2a*-knockout mice is associated with a severe neurodevelopmental
302 phenotype characterized by microcephaly, cognitive impairment, ataxia, and severe
303 anxiety.[12] In particular, microcephaly is likely explained by the fact that LPC transport not
304 only provides accretion of DHA by the developing brain, but is also critical for providing
305 LPC as building blocks for neuron arborization and regulation of membrane phospholipid
306 composition.[2,5,15] The reports of loss-of-function *MFSD2A* variants in patients with a
307 progressive microcephaly syndrome with severe ID and neuroimaging abnormalities have
308 supported the relevant role of this lipid transporter in human brain development and
309 functioning.[3,4,9] The relevance of proper DHA metabolism for brain development and
310 functioning is further supported by CYP2U1 deficiency. This enzyme is a member of the
311 cytochrome P450 family 2 subfamily U and catalyzes the hydroxylation of arachidonic acid
312 (AA) and AA-related long-chain fatty acids, including DHA.[16] Biallelic loss-of-function
313 *CYP2U1* variants cause spastic paraplegia 56 (SPG56), a complex neurological condition
314 characterized by spasticity, cognitive impairment, and white matter abnormalities.[16]

Here, we present seven families with eight distinct loss-of-function variants in *MFSD2A*, including seven novel variants affecting function. Patient 4 was part of a large cohort of consanguineous families with recessive intellectual disability reported by Riazuddin et al.[8] Patients 6 and 7 were briefly described before by Shaheen et al. and Monies et al., respectively.[17,18] In line with previously reported cases, our patients showed a complex neurodevelopmental phenotype primarily characterized by severe progressive microcephaly, ID, spasticity, and speech delay (Table 1) (Fig. 1f).[3,4,6,8,9] Less common clinical features were also identified in our cohort, including axial hypotonia, increased deep tendon reflexes, and seizures (Fig. 1b).[3,4,6,8,9] Of note, none of our patients died prematurely, although some of their siblings who died prematurely were most likely affected by the same condition. The longest follow-up was 27 years (patient 4), allowing assessment of the progression of microcephaly over time. Language was delayed in most subjects and one patient was nonverbal. Four patients showed skeletal abnormalities consistent with TEV. Dysmorphic features were observed in patient 5 only.

In previously reported cases, brain MRI revealed a spectrum of abnormal findings, including ventricular enlargement secondary to white matter paucity and hypoplasia of the corpus callosum, cerebellum, and brainstem.[3,4] In our study, we provide further evidence that affected subjects present severe microcephaly with simplified gyral pattern, associated with variable degrees of white matter reduction leading to mild to severe ventricular dilatation. Of note, the myelination was always appropriate for patients' age in our series, ruling out a hypomyelinating disorder. Interestingly, the corpus callosum was always abnormal, with severe hypodysplasia in most subjects. However, the cingulate gyrus was present in the most

severe cases as well, indicating that the corpus callosum was initially formed. Finally, the inferior cerebellar vermis was small in all subjects while hypoplasia of the pons was noted in almost all of them. Taken together, these neuroimaging features are consistent with an early prenatal developmental disruption and likely suggest a relevant role of LPCs in the development of both the cerebral gray and white matter.

A clear correlation between the severity of the clinico-radiological phenotype and the variants affecting function in *MFSD2A* could not be observed. Despite the *MFSD2A* variants identified in the current study impair protein expression rather than the transporter function, no substantial difference between the phenotypes of previously reported affected individuals and patients from the current cohort was noticed (Table 1). This observation supports the loss of function as the main pathogenic mechanism in MCPH15, regardless of the specific underlying cause. All patients show a variable degree of progressive microcephaly and a comparable level of psychomotor delay, but some speculations on selected phenotypic features are possible. In fact, behavioural disturbances appeared to be more frequent in subjects carrying missense variants affecting the transporter function (c.1016C>T; p.(Ser339Leu), c.476C>T; p.(Thr159Met), and c.497C>T; p.(Ser166Leu)),[3,4] whereas skeletal abnormalities might be more common in patients carrying variants resulting in decreased MFSD2A expression, as showed by patients 1, 2, 7, and 8 from our cohort. Interestingly, extrapyramidal disorders have been associated with the previously reported variants c.1205C>A; p.(Pro402His) and c.490C>A; p.(Pro164Thr),[6,9] but were absent in our cases. As to the neuroimaging features, the degree of involvement of grey and white matter

structures is quite variable in the affected individuals and does not appear to be correlated to *MFSD2A* variant type.

In conclusion, our observations expand the phenotypic spectrum of *MFSD2A*-related microcephaly syndrome and provide new insights into the underlying pathogenic mechanisms. Refining the neuroradiological characterization of MCPH15, we suggest that some neuroimaging clues can be extremely relevant for an early diagnosis. We also show that poor MFSD2A expression plays a relevant role in MCPH15 pathogenesis, further defining the etiology of this condition. A better understanding of the role of MFSD2A in brain physiology will foster the development of targeted therapies or specific metabolic supplementation regimens to bypass LPC transport deficiency. The identification and characterization of further patients harboring loss-of-function *MFSD2A* variants will support efforts to exploit LPCs as therapeutic lipids to improve DHA delivery and promote proper brain development in affected individuals.

## Acknowledgments

The work was supported in part by National Research Foundation and Ministry of Health grants, Singapore; by the Biomedical Research Council of A*STAR; by March of Dimes Research Grant; as part of the Queen Square Genomics group at University College London, supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

## Conflict of Interest

The authors declare no conflict of interest.


## Funding

National Research Foundation grants, Singapore NRF2016NRF-NRFI001-15 and OF-IRG MOH-000217 (to D.L.S.); Biomedical Research Council of A*STAR (to H.F.); The MRC (MR/S01165X/1, MR/S005021/1, G0601943), The National Institute for Health Research University College London Hospitals Biomedical Research Centre, Rosetree Trust, Ataxia UK, MSA Trust, Brain Research UK, Sparks GOSH Charity, Muscular Dystrophy UK (MDUK), Muscular Dystrophy Association (MDA USA), March of Dimes USA (to M.C.M.), The R01 RNS107428A by the National Institute of Neurological Disorders and Stroke/National Institutes of Health (NINDS/NIH).

## References


1. Nguyen LN, Ma D, Shui G, Wong P, Cazenave-Gassiot A, Zhang X, et al. Mfsd2a is a transporter for the essential omega-3 fatty acid docosahexaenoic acid. Nature. 2014;509:503-6.

2. Chan JP, Wong BH, Chin CF, Galam DLA, Foo JC, Wong LC, et al. The lysolipid transporter Mfsd2a regulates lipogenesis in the developing brain. PLoS Biol. 2018;16:e2006443.

3. Alakbarzade V, Hameed A, Quek DQ, Chioza BA, Baple EL, Cazenave-Gassiot A, et al. A partially inactivating mutation in the sodium-dependent lysophosphatidylcholine transporter MFSD2A causes a non-lethal microcephaly syndrome. Nat Genet. 2015;47:814-7.

4. Guemez-Gamboa A, Nguyen LN, Yang H, Zaki MS, Kara M, Ben-Omran T, et al. Inactivating mutations in MFSD2A, required for omega-3 fatty acid transport in brain, cause a lethal microcephaly syndrome. Nat Genet. 2015;47:809-13.

5. Guesnet P, Alessandri JM. Docosahexaenoic acid (DHA) and the developing central nervous system (CNS) - Implications for dietary recommendations. Biochimie. 2011;93:7-12.

6. Harel T, Quek DQY, Wong BH, Cazenave-Gassiot A, Wenk MR, Fan H, et al. Homozygous mutation in MFSD2A, encoding a lysolipid transporter for docosahexanoic acid, is associated with microcephaly and hypomyelination. Neurogenetics. 2018;19:227-35.

7. Quek DQ, Nguyen LN, Fan H, Silver DL. Structural insights into the transport mechanism of the human sodium-dependent lysophosphatidylcholine transporter Mfsd2a. J Biol Chem. 2016;291:9383-94.

8. Riazuddin S, Hussain M, Razzaq A, Iqbal Z, Shahzad M, Pollaet DL, et al. Exome sequencing of Pakistani consanguineous families identifies 30 novel candidate genes for recessive intellectual disability. Mol psychiatry. 2017;22:1604-14.

9. Hu H, Kahrizi K, Musante L, Fattahi Z, Herwig R, Hosseini M, et al. Genetics of intellectual

disability in consanguineous families. Mol Psychiatry. 2019;24:1027-39.

10. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754-60.

11. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics. 2013;43:11.10.1-11.10.33.

12. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38:e164.

13. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015;36:928-30.

14. Andreone BJ, Chow BW, Tata A, Lacoste B, Ben-Zvi A, Bullock K, et al. Blood-Brain Barrier Permeability Is Regulated by Lipid Transport-Dependent Suppression of Caveolae-Mediated Transcytosis. Neuron. 2017;94:581-94.

15. Ahmad A, Moriguchi T, Salem N. Decrease in neuron size in docosahexaenoic acid deficient brain. Pediatr Neurol. 2002;26:210-8.

16. Tesson C, Nawara M, Salih MA, Rossignol R, Zaki MS, Al Balwi M, et al. Alteration of fatty-acid-metabolizing enzymes affects mitochondrial form and function in hereditary spastic paraplegia. Am J Hum Genet. 2012;91:1051-64.

17. Shaheen R, Maddirevula S, Ewida N, Alsahli S, Abdel-Salam GMH, Zaki MS, et al. Genomic and phenotypic delineation of congenital microcephaly. Genet Med. 2019;21:545-52.

18. Monies D, Abouelhoda M, Assoum M, Moghrabi N, Rafiullah R, Almontashiri N, et al. Lessons Learned from Large-Scale, First-Tier Clinical Exome Sequencing in a Highly Consanguineous Population. Am J Hum Genet. 2019;104:1182-201.

**Table 1. Genetic, clinical, and neuroradiological features of *MFSD2A* patients.**

| *Families* | *A (Pt 1)* | *B (Pt 2)* | *C (Pt 3)* | *C (Pt 4)#* | *D (Pt 5)* | *E (Pt 6)##* | *F (Pt 7)###* | *G (Pt 8)* | *Alakbarzade, 2015 (10 pts)* | *Guemez-Gamboa, 2015 (4 pts)†* | *Harel, 2018 (2 pts)* | *Hu, 2019 (3 pts)* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***Age (last FU), sex*** | 4 y, F | 4 y, M | 17 y, F | 27 y, F | 5 y, M | 1 mo, F | 2 y, M | 4 mo, F | Mean 12.6 y M/F = 2.3 | Mean N/A M/F = 0.3 | Mean 4.9 y M/F = 1 | Mean 22 y M/F = 0.5 |
| ***Origin*** | Iran | Iran | Pakistan | Pakistan | Russia | Saudi | Saudi | Saudi | Pakistan | Libya, Egypt | Jewish Moroccan | Iran |
| ***Consanguinity*** | + | + | + | + | - | + | + | + | + | + | + | + |
| ***MFSD2A variant [NM_032793.5]*** | c.[1478C>T];[1478C>T], p.[(Pro493Leu)];[(Pro493Leu)] | c.[556+1G>A];[556+1G>A]‡ | c.[593C>T];[593C>T], p.[(Thr198Met)];[(Thr198Met)] | c.[593C>T];[593C>T], p.[(Thr198Met)];[(Thr198Met)] | c.[748G>T];[c.977G>A], p.[(Val250Phe)];[(Arg326His)] | c.[476C>T];[476C>T], p.[(Thr159Met)];[(Thr159Met)] | c.[1386_1435del];[1386_1435del], p.[(Gln462HisfsTer17)];[(Gln462HisfsTer17)] | c.[750_753del];[750_753del], p.[(Cys251SerfsTer3)];[(Cys251SerfsTer3)] | c.[1016C>T];[1016C>T], p.[(Ser339Leu)];[(Ser339Leu)] | Fam 1825 c.[476C>T];[476C>T]; p.[(Thr159Met)];[(Thr159Met)] Fam 1422 c.[497C>T];[497C>T], p.[(Ser166Leu)];[(Ser166Leu)] | c.[1205C>A];[1205C>A], p.[(Pro402His)];[(Pro402His)] | c.[490C>A];[c.490C>A], p.[(Pro164Thr)];[(Pro164Thr)] |
| ***OFC at birth*** | 28 cm (-4.6 SDS) | 27 cm (-3.9 SDS) | N/A | N/A | N/A | 28.5 cm (-3.6 SDS) | 25.5 cm (-6 SDS) | 30.5 cm (-2.4 SDS) | N/A | Mean -1.3 SDS | Mean -2.5 SDS | N/A |
| ***OFC at FU*** | 41 cm (-5.6 SDS) | 37 cm (-8.8 SDS) | 49 cm (-5.0 SDS) | 47 cm (-6.9 SDS) | 46 cm (-3.6 SDS) | N/A | 36 cm (-8.9 SDS) | 36 cm (-3.9 SDS) | </= -3 SDS | Mean -5 SDS | Mean -3.25 SDS | Mean -4.3 SDS |
| ***GDD*** | + | + | + | + | + | + | + | + | + | + (3/3) | + (2/2) | + (3/3) |
| *Sitting* | - | - | + | + | + | - | - | + | N/A | - (2/3) | + (2/2) | + (3/3) |
| *Walking* | - | - | + | + | - | - | - | - | N/A | - | - | + (3/3) |
| *Speech* | Non-verbal | Severely Delayed | Severely Delayed | Severely Delayed | Non-verbal | Non-verbal | Non-verbal | Non-verbal | Absent/ limited (10/10) | Non-verbal (3/3) | Severely Delayed (2/2) | Non-verbal (2/3) |
| ***ID*** | N/A | N/A | Severe | Severe | Severe | Severe | Severe | Severe | Severe (10/10) | + (3/3) | + (2/2) | + (3/3) Mod-severe |
| ***Behavioural abnormalities*** | - | - | Aggressive | Aggressive | - | - | - | - | ASD (10/10) | ASD (3/3) | - | - |
| ***Appendicular spasticity*** | + | + | + | + | - | + | + | + | + (3/10) | + (3/3) | +, with dystonia (2/2) | -, but ataxia (3/3) |
| ***Axial hypotonia*** | + | - | - | - | + | - | - | - | N/A | + (3/3) | + (2/2) | - |
| ***Seizures*** | + | + | - | - | + | + | + | + | - | + (3/3) | - | - |
| ***Dysphagia*** | + | - | - | - | + | + | + | + | N/A | + (2/3) | - | - |
| ***Skeletal abnormalities*** | TEV | TEV | - | - | - | - | TEV | TEV, Bilateral DDH | N/A | TEV (2/3) | - | - |
| ***Premature death*** | - | - | - | - | - | - | - | - | N/A | + (mean 3 y) | - | - |
| ***MRI findings*** | | | | | | | | | | | | |
| *WM thinning with ventricular dilatation* | Severe | Severe | Moderate | Moderate | Mild | Severe | Severe | Severe | + | + (3/3) | + (2/2) | N/A |
| *Simplified gyral pattern* | Severe | Severe | Mild | Mild | Mild | Severe | Severe | Severe | N/A | N/A | N/A | N/A |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Corpus callosum hypoplasia* | *Severe* | *Severe* | *Mild* | *Mild* | *Mild* | *Severe* | *Severe* | *Severe* | *N/A* | *+ (3/3)* | *N/A* | *N/A* |
| *Inferior vermian hypoplasia* | + | + | + | + | + | + | + | + | N/A | + (3/3) | N/A | N/A |
| *Pontine hypoplasia* | + | + | - | - | - | + | + | + | N/A | N/A | N/A | N/A |

ASD Autism spectrum disorder Comp Het Compound heterozygous, DDH Developmental dysplasia of the hip F female, Fam Family, Hom Homozygous, FU Follow-up, M male, mo months, Mod moderate, N/A Not Applicable, OFC Occipito-frontal circumference, Pt Patient, TEV Talipes Equinovarus, y years. ‡ NG_053084.1(NM_032793.5): c.556+1G>A, NC_000001.11(NM_032793.5): c.556+1G>A, LRG_199t1). † Data available for 3 out of 4 patients. # PMID: 27457812. ## PMID: 30214071. ### PMID: 31585110.

**Table 2. Frequency, conservation, and predicted functional impact of *MFSD2A* variants.**

| ***MFSD2A* variant [NM_032793.5]** | **g. (hg19)** | **LOVD (ID)** | **Internal database ‡** | **ExAC/ gnomAD** | **GME** | **Iranome** | **Ensembl** | **ClinVar** | **SIFT** | **Mutation Taster** | **HSF/ VEP** | **GERP score** | **CADD score** | **ACMG class** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **c.476C>T (p.Thr159Met)** | chr1:40431005 C>T | 002760 75 | - | 0.000003 978 (1 het) | - | - | rs1057517 688 | Pathogenic | Damaging (score 0) | Disease causing | - | 5.75 | 34 | Likely pathogenic (PS3, PM2, PP3, PP4, PP5) |
| **c.593C>T (p.Thr198Met)** | chr1:40431565 C>T | 002760 71 | - | 0.000003 977 (1 het) | - | - | rs7564670 73 | - | Damaging (score 0.003) | Disease causing | - | 5.94 | 28.2 | Likely pathogenic (PS3, PM2, PP3, PP4) |
| **c.556+1G>A** | chr1:40431222 G>A | 002760 70 | - | 0.000003 978 (1 het) | - | - | rs7589530 00 | - | - | Disease causing | WT donor site alteration | 5.56 | 29.2 | Pathogenic (PVS1, PM2, PP3, PP4) |
| **c.750_753del (p.Cys251SerfsTer3)** | chr1:40432304 TTGTC>T | 002760 77 | - | 0.000003 982 (1 het) | - | - | - | - | - | Disease causing | - | - | - | Pathogenic (PVS1, PM2, PP4) |
| **c.748G>T (p.Val250Phe)** | chr1:40432306 G>T | 002760 74 | - | - | - | - | - | - | Damaging (score 0) | Disease causing | - | 5.79 | 33 | Likely pathogenic (PS3, PM2, PP3, PP4) |
| **c.977G>A (p.Arg326His)** | chr1:40432807 G>A | 002760 74 | - | 0.000007 956 (2 het) | - | - | rs7767413 31 | - | Tolerated (0.37 score) | Disease causing | - | 5.52 | 24.4 | Likely pathogenic (PS3, PM2, PP3, PP4) |
| **c.1386_1435del (p.Gln462HisfsTer17)** | chr1:40434271GCAG CCGGAACGTGTCA AGTTTACACTGAA CATGCTCGTGACC ATGGCTCC>G | 002760 76 | - | - | - | - | - | - | - | Disease causing | - | - | - | Pathogenic (PVS1, PM2, PP4) |
| **c.1478C>T (p.Pro493Leu)** | chr1:40434366 C>T | 002760 67 | - | - | - | - | - | - | Damaging | Disease causing | - | 5.49 | 32 | Likely pathogenic (PS3, PM2, PP3, PP4) |

*ACMG* American College of Medical Genetics and Genomics, *CADD* Combined Annotation Dependent Depletion, *GERP* Genomic Evolutionary Rate Profiling, *GME* Greater Middle East Variome Project, *HSF* Human Splice Finder, *LOVD-ID* Leiden Open Variation Database Identifier, *PVS* pathogenic very strong, *PS* pathogenic strong, *PM* pathogenic moderate, *PP* pathogenic supporting, *SIFT* Sorting Intolerant From Tolerant, *VEP* Variant Effect Predictor, *VUS* variant of unknown significance.

## Legends

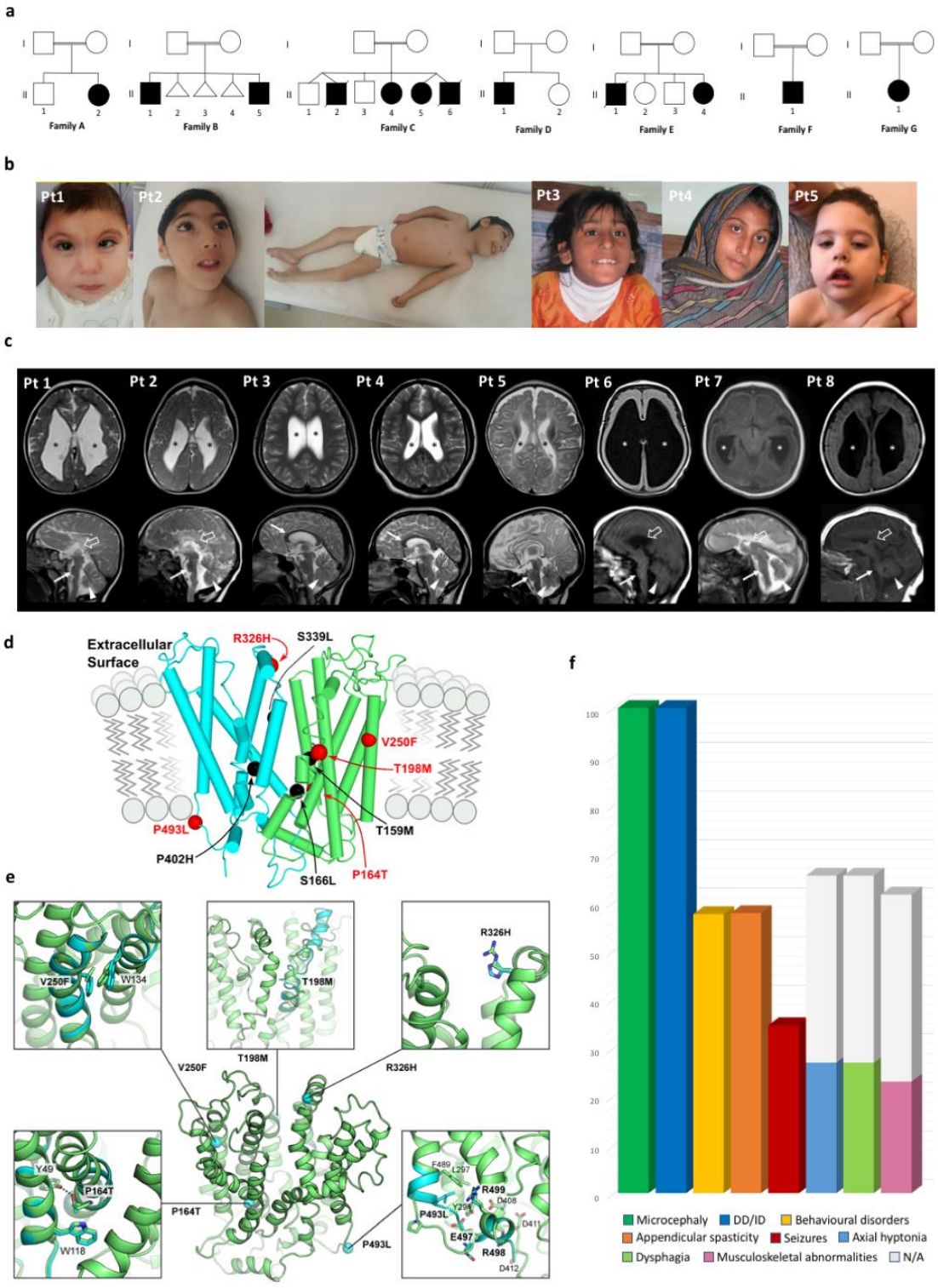



**Fig. 1 Clinical characterization, neuroimaging features, genetic findings and predicted**

**consequences of *MFSD2A* variants.** (**a**) Pedigrees of the seven reported families. (**b**) Main clinical features include severe microcephaly, axial hypotonia, talipes equinovarus, and minor dysmorphic features (e.g., epicanthal folds and broad nasal bridge in patient 5). (**c**) Brain MRI of affected subjects performed at 3 years (Pt 1), 1 year (Pt 2), 17 years (Pt 3), 27 years (Pt 4), 2 months (Pt 5), 1 month (Pt 6), 2 years (Pt 7), and 4 months of age (Pt 8). First row: axial T2, FLAIR or T1-weighted images of the patients. Second row: corresponding sagittal T2 or T1-weighted images. There is severe microcephaly with mildly to severely simplified gyral pattern in all subjects. The cerebral white matter is reduced with consequent ventricular dilatation (asterisks), especially in patients 1, 2, 6, 7, and 8. The corpus callosum is barely visible and markedly short in patients 1, 2, 6, 7, and 8 (empty arrows), while it is diffusely hypoplastic in Patient 5. Hypoplasia of the anterior portion of the corpus callosum is visible in patients 3 and 4 (arrows). Note that in all subjects the cingulate gyrus is present. The inferior portion of the vermis is small in all subjects (arrowheads), with associated pontine hypoplasia in patients 1, 2, 5, 6, 7, and 8. (**d**) 3D structural model of Mfsd2a (based on Quek DQ et al., 2016; Supplementary References) indicating the locations of previously reported variants (in black) and the variants identified in this study (in red). The N-terminus is indicated in green and C-terminus in cyan. (**e**) 3D structural models of the Mfsd2a variants. Positions of variants in the human Mfsd2a protein. Variants (cyan) were mapped to the published homology model of Mfsd2a (green). R326 is located at the putative extracellular gate and the R326H substitution might disrupt gate closure. V250 and P164 are both located in helical bundles. Their substitution by larger amino acids (V250F and P164T) might perturb protein folding by steric clash with neighboring sidechains (e.g., W134, W118). P164T might also form a hydrogen bond with Y49 that is not seen in canonical Mfsd2a. Variants T198M and P493L are predicted to alter the local protein structure. (**f**) Percentage distribution of the main clinical features

of *MFSD2A* patients. *DD* developmental delay; *ID* intellectual disability; *N/A* not applicable; *Pt* patient.

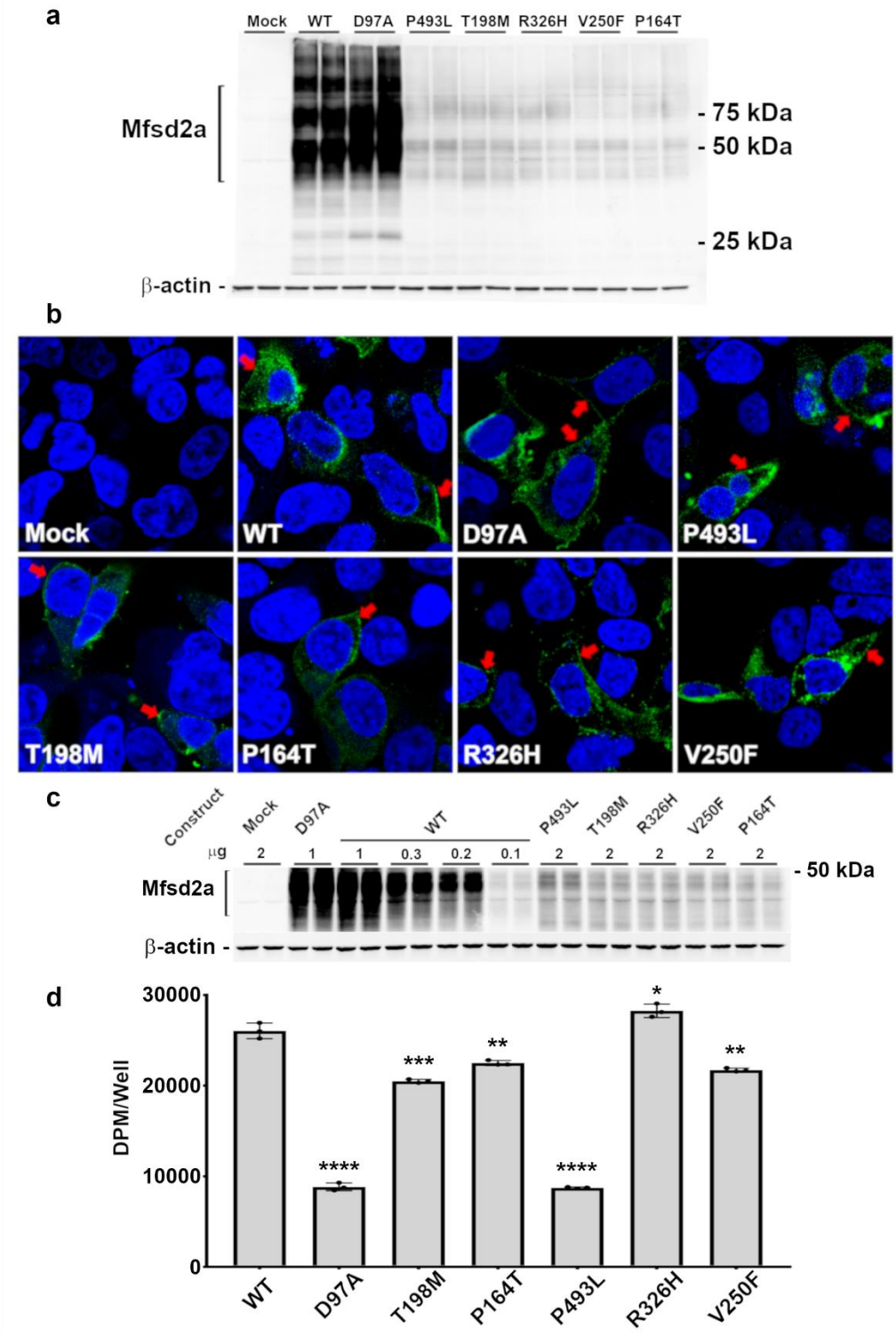


**Fig. 2 Biochemical analysis of Mfsd2a variants.** (**a**) Western blot probed for Mfsd2a and its mutants with β-actin used as loading control. (**b**) Confocal immunofluorescence micrographs of transiently transfected HEK293 cells with Mock, WT, D97A, P493L, T198M, P164T, R326H and V250F variants affecting function showing Mfsd2a localization in green cell nuclei in blue (Hoechst stain), red arrows pointing to the cell surface localization of Mfsd2a and its mutants. (**c**) Titration of varying amounts of WT Mfsd2a DNA (μg) to normalize the expression levels to determine the amount of WT Mfsd2a needed for comparable expression levels with cells transfected with 2 mg of mutant construct DNA. (**d**) Transport of 50 μM $^{14}$C LPC-DHA by comparable expression levels of MFSD2A in HEK293. Significance levels of difference compared with the transport activity of 0.1 μg of WT Mfsd2a (labeled WT on the graph). Transport activity are labeled with asterisks: **** representing P value < 0.0001, *** representing P value <0.001, ** representing P value < 0.01, * representing P value <0.1.

**Supplementary Material**

1. **Supplementary Methods**

2. **Supplementary Figure**

3. **Supplementary Table**

4. **Supplementary References**

## 1. Supplementary Methods

*Exome sequencing and variants analysis*

Genomic DNA was sent for whole exome sequencing at the Broad Institute Genomic Services. Sequencing reads were aligned to reference genome hg19 using Burrows Wheeler Aligner (Li & Durbin, 2009). Exome coverage was 92.9% with a mean target coverage of 82 reads. Aligned reads were sorted and duplicates marked using Picard Tools (Broad Institute). The Genome Analysis Toolkit was used to call variants, recalibrate base quality scores, then recall variants based on the recalibration scores using the best practices protocol for variant analysis (Van der Auwera et al., 2013). We used Annovar to annotate variants, loaded the variants into an SQL database, and used custom SQL queries to identify rare, homozygous and compound heterozygous nonsynonymous or truncating variants (Wang, Li, & Hakonarson, 2010). Variant frequency of less than 1% was filtered using data from the Genome Aggregation Database (Lek et al., 2016), the Greater Middle East Variome Project (Scott et al., 2016) and Iranome (Akbari et al., 2017). Protein pathogenicity of variants was predicted using CADD (Kircher et al., 2014), SIFT (Ng & Henikoff, 2003), and Polyphen-2 (Adzhubei et al., 2010). Further annotation on the clinical significance of variants was gathered from the databases UCSC Genome Browser (Kent et al., 2002), Uniprot (Poux et al., 2017), Online Mendelian Inheritance of Man (McKusick-Nathans Institute of Genetic Medicine), and The Human Gene Mutation Database (Stenson et al., 2017). The methodology of exome sequencing and variant analysis for family PKMR97 (Thr198Met) has been reported in detail previously (Nguyen et al., 2014).

*Generation of human point mutations in human Mfsd2a*

564 The five human mutations of Mfsd2a, Pro493Leu (P493L), Thr198Met (T198M), Pro164Thr (P164T),
565 and compound heterozygote Arg326His (R326H) and Val250Phe (V250F) were individually generated
566 through the amplification of human Mfsd2a using gene-specific and site-specific mutagenic primers and
567 ligated into pcDNA3.1 after digestion with restriction enzymes EcoRV and XbaI.

568

569 *3D structural modeling of the T198M, P164T, P493L, R326H and V250F mutants*

570 Starting from the published 3D model of MFSD2A WT in the outward occluded state, single point
571 mutations T198M and P493L were generated independently by sidechain prediction using SCWRL (Quek
572 et al., 2016). This initial model of T198M or P493L was subjected to local structural optimization by loop
573 modeling implemented in MODELLER (Sali et al., 1993), resulting in 2500 models that were evaluated
574 by the DOPE (discrete optimized protein energy) score to select the best ranked model.[3] For P164T,
575 R326H, and V250F, point mutations were generated from the same starting model[1] using the mutagenesis
576 function followed by local sphere regularization with secondary structure restraints in COOT (Emsley et
577 al., 2010). Molecular graphics were created in PyMOL (The PyMOL Molecular Graphics System, 2002).

578

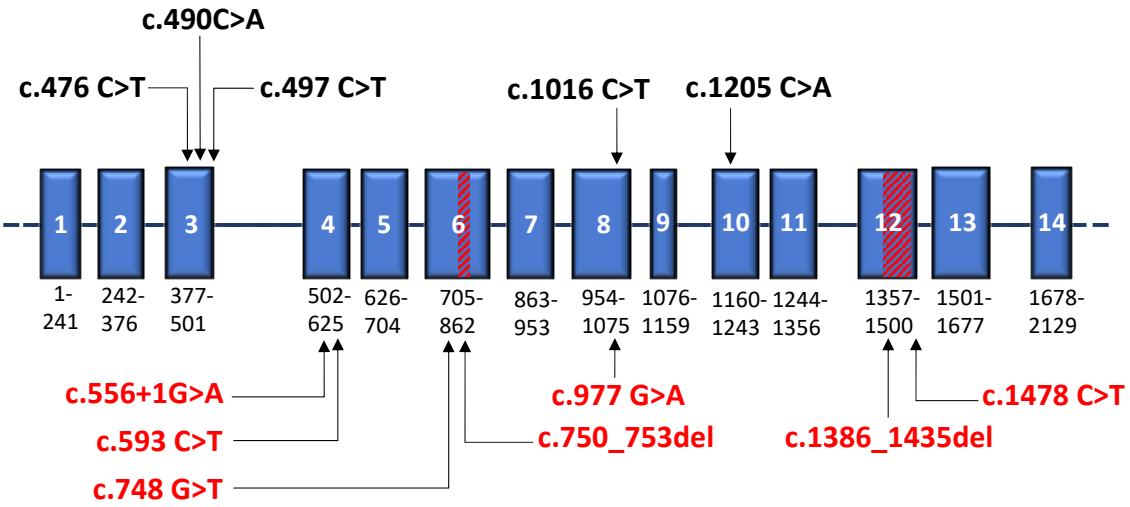
579 *Western Blot and Immunofluorescence analysis of mutant transiently transfected in HEK293 cells*

580 Cellular expression of the human mutants was compared with the wild-type (WT) Mfsd2a, and the non-
581 functional sodium binding mutant Asp97Ala (D97A) expression constructs by immunoblotting using a
582 rabbit polyclonal antibody against Mfsd2a on transiently transfected HEK293 cells (Chan et al., 2018).
583 Using the same antibody against MFSD2A, the cellular localizations of the mutants transiently transfected
584 into HEK293 were also visualized together with its WT Mfsd2a as a control using confocal
585 immunofluorescence microscopy (Zeiss). Cell transfected with an empty pcDNA3.1 was used as a
586 negative control. Details of these methods were previously described (Nguyen et al., 2014; Quek et al.,
587 2016).

*Mfsd2a Transport assays*

In vitro transport of the Mfsd2a ligand, $^{14}$C-Lysophosphatidylcholine-Docosahexaenoic acid (LPC-DHA) (ARC Radiochemicals), spiked into unlabeled 10 mM LPC-DHA (Vanteres Pte Ltd) was tested in HEK293 cells transiently transfected with wild-type (WT) Mfsd2a and mutants for 24 hours (Nguyen et al., 2014; Quek et al., 2016). Uptake activity of 14C-LPC-DHA for all constructs were measured after 30 minutes incubation with 50 mM LPC-DHA diluted in serum-free DMEM (Gibco). The cells were washed two times in serum-free DMEM (Gibco) containing 0.5% fatty-acid free bovine serum albumin and harvested with RIPA buffer into 4 ml of scintillation fluid (Ecolite, MP-biopharmaceuticals). Disintegrations Per Minute (DPM) of the incorporated LPC-DHA in each well of transfected HEK293 cells were counted using a scintillation counter (Tricarb, Perkin Elmer). All transport assays were carried out in triplicates using a 12-well plate.

## 2. Supplementary Figure




**Supplementary Figure 1.** Schematic drawing of *MFSD2A* with previously reported variants (in black) and the variants identified in this study (in red). Intragenic deletions are indicated by diagonal lines within the affected exon.

## 3. Supplementary Table

**Table S1. Other potential causative variants in the reported *MFSD2A* families.**

| Families | Gene | Variant (hg19) | Status | gnomAD | GME | Iranome | ClinVar (ID) | SIFT | Mutation Taster | GERP score | CADD score | ACMG class |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | *MACF1* | chr1:39765977 C>A | hom | 0 | 0 | 0 | - | 0.238 | 0.977 | 5.81 | 16.5 | III (BP4, PM2, PP3) |
| | *SZT2* | chr1:43908592 C>T | hom | 0.00003 | 0 | 0 | - | 0.002 | 1 | 5.67 | 34 | III (BP1, PM2, PP3) |
| | *CACHD1* | chr1:65016278 G>A | hom | 0.00140 | 0.00302 | 0.00375 | - | 0.178 | 1 | 6.02 | 27.6 | II (BS1) |
| | *TTN* | chr2:179395282 G>C | hom | 0 | 0 | 0 | - | 1 | 1 | 5.23 | 13.1 | II (BP1, BP4, PM2) |
| | *PARD3B* | chr2:206057991 C>T | hom | 0.00003 | 0 | 0 | - | 0.102 | 0.999 | 5.63 | 22.5 | III (BP4) |
| | *ABCA12* | chr2:215802262 T>C | hom | 0 | 0 | 0 | - | 0.11 | 0.801 | 5.67 | 21.3 | II (BP1, BP4, PM2) |
| | *TBL1XR1* | chr3:176752064 T>C | hom | 0 | 0 | 0.00063 | - | 1 | 1 | 5.65 | 17.9 | III (PM2, PP2) |
| | *ALG3* | chr3:183960623 G>A | hom | 0.00007 | 0 | 0.00063 | - | 0.007 | 0.999 | 5.09 | 25 | III (PM2, PP2, PP3) |
| | *ATP13A5* | chr3:193039554 C>T | hom | 0.00020 | 0 | 0 | - | 0.592 | 1 | 5.82 | 5.8 | II (BS1, BP4) |
| | *LRRC15* | chr3:194081159 T>C | hom | 0.00020 | 0 | 0 | - | 0.029 | 1 | 5.02 | 17.5 | II (BS1, BP4) |
| | *RGS12* | chr4:3344267 T>C | hom | 0.00460 | 0 | 0.00625 | - | 0 | 0 | 1.49 | 1.8 | II (BS1, BP4) |
| | *ADAMTS8* | chr11:130289012 C>T | hom | 0.00040 | 0.00151 | 0.00438 | - | 0.041 | 1 | 5.62 | 13.8 | II (BS1, BP4) |
| | *MPP2* | chr17:41960701 G>C | hom | 0.00001 | 0 | 0 | - | 0.487 | 1 | 4.15 | 17.9 | III (BP4, PM2) |
| | *FAM187A* | chr17:42982324 C>T | hom | 0.00390 | 0 | 0.03062 | - | 0.465 | 1 | 5.54 | 11.3 | II (BS1, BP4) |
| | *STH* | chr17:44077019 C>G | hom | 0.00002 | 0 | 0.00063 | - | 0 | 1 | 2.03 | 34 | III (BP4, PM1, PM2) |
| | *ZDHHC8P1* | chr22:23742049 G>A | hom | 0 | 0 | 0.05882 | - | 0 | 0 | 1.82 | 4.7 | I (BA1, BP4) |
| | *CRYBB2P1* | chr22:25853368 T>C | hom | 0 | 0 | 0.1181 | - | 0 | 0 | 2.22 | 12.3 | I (BA1, BP4) |
| B† | - | - | - | - | - | - | - | - | - | - | - | - |
| C | *SCP2* | chr1:53393072 T>G | hom | 0.00005 | 0 | 0 | - | 0.778 | 0.885 | 3.14 | - | III (PM2, BP4) |
| | *TMCC2* | chr1:205241169 C>T | hom | 0.00006 | 0.00251 | 0.00437 | - | 0.492 | 0.999 | 5.18 | - | II (BS1, BP4) |
| | *NAGK* | chr2:71297921 G>C | hom | - | 0 | 0 | - | 0.26 | 0.995 | 4.95 | 20.7 | III (PM2, BP4) |
| | *NAGK* | chr2:71295842 G>T | hom | - | 0 | 0 | - | 0.002 | 1 | 5.11 | 36 | III (PM2, PP3) |
| | *MAP6* | chr11:75378664 C>T | hom | - | 0 | 0 | - | 0.438 | 0.999 | 4.5 | 6.9 | III (PM2, BP4) |
| | *POSTN* | chr13:38166262 C>T | hom | 0.00074 | 0 | 0.00063 | - | 0.331 | 0.999 | 5.18 | 22.2 | II (BS1, BP4) |
| D† | - | - | - | - | - | - | - | - | - | - | - | - |
| E | *CDKL5* | chrX:18668586 C>T | het | 0.00019 | 0 | 0 | RCV000475262 | 0 | 0.999 | -7.59 | 0 | I (BS1, BS2, BP4, BP6) |
| | *TUBB3* | chr16:90002195 G>A | het | 0.00035 | 0 | 0 | RCV000903349 | 0.001 | 1 | 4.66 | 16 | II (PP2, BS1, BP4, BP6) |
| F | *BRWD1* | chr21:40608526 T>C | hom | 0 | 0 | 0 | - | 0.154 | 0.899 | 5.44 | 15.4 | III (PM2, BP4) |
| G | *EXOSC8* | chr13:37583420 G>C | hom | 0.00385 | 0.00554 | 0.00875 | RCV000418794 | 0 | 1 | 5.85 | 14 | II (PP3, PP5, BS1, BP1) |
| | *ALDH5A1* | chr6:24495252 T>C | hom | 0.000096 | 0 | 0 | - | 0.212 | 0.999 | 1.27 | 4 | II (PM2, BP1, BP4) |

*ACMG* American College of Medical Genetics and Genomics, *BA* Benign stand alone, *BS* Benign Strong, *BP* Benign supporting, *CADD* Combined Annotation Dependent Depletion, *GERP* Genomic Evolutionary Rate Profiling, *GME* Greater Middle East Variome Project, *PM* Pathogenic Moderate, *PP* Pathogenic supporting, *SIFT* Sorting Intolerant From Tolerant, *VEP* Variant Effect Predictor, *VUS* variant of unknown significance. † In these two families, no other possible causative variant could be identified.

## 3. Supplementary References


Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. 2013;Chapter 7:Unit 7.20.

Chan JP, Wong BH, Chin CF, et al The lysolipid transporter Mfsd2a regulates lipogenesis in the developing brain. PLoS Biol. 2018;16:e2006443.

Emsley P, Lohkamp B, Scott WG, et al. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66:486-501.

Fattahi Z, Beheshtian M, Mohseni M, et al. Iranome: A catalog of genomic variations in the Iranian population. Hum Mutat. 2019;40:1968-1984.

Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. Genome Res. 2002;12:996-1006.

Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46:310-315.

Krivov GG, Shapovalov MV, Dunbrack RL Jr. Improved prediction of protein side-chain conformations with SCWRL4. Proteins. 2009;77:778-795.

Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536:285-291.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754-1760.

Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31:3812-3814.

Nguyen LN, Ma D, Shui G, et al. Mfsd2a is a transporter for the essential omega-3 fatty acid docosahexaenoic acid. Nature. 2014;509:503-506.

659 Poux S, Arighi CN, Magrane M, et al. On expert curation and scalability: UniProtKB/Swiss-Prot
660 as a case study. Bioinformatics. 2017;33:3454-3460.

661 Quek DQ, Nguyen LN, Fan H, Silver DL. Structural Insights into the Transport Mechanism of
662 the Human Sodium-dependent Lysophosphatidylcholine Transporter MFSD2A. J Biol
663 Chem. 2016;291:9383-9394.

664 Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol
665 Biol. 1993;234:779-815.

666 Scott EM, Halees A, Itan Y, et al. Characterization of Greater Middle Eastern genetic variation
667 for enhanced disease gene discovery. Nat Genet. 2016;48:1071-1076.

668 Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. Protein
669 Sci. 2006;15:2507-2524.

670 Stenson PD, Mort M, Ball EV, et al. The Human Gene Mutation Database: towards a comprehensive
671 repository of inherited mutation data for medical research, genetic diagnosis and next-generation
672 sequencing studies. Hum Genet. 2017;136:665-677.

673 The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC. 2002.

674 Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high
675 confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc
676 Bioinformatics. 2013;43:11.10.1-11.10.33.

677 Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-
678 throughput sequencing data. Nucleic Acids Res. 2010;38:e164.

679