

Heterogeneous Inheritance in Autism Genes Shared Across Neurodevelopmental and Neuromuscular Disorders in Consanguineous Singlets

Mohammad-Reza Ghasemi

Department of Medical Genetics, Faculty of Medicine, Shahid Beheshti University of Medical Sciences

Hossein Sadeghi

Genomic Research Center, Shahid Beheshti University of Medical Sciences

Farzad Hashemi-Gorji

Genomic Research Center, Shahid Beheshti University of Medical Sciences

Reza Mirfakhraie

Department of Medical Genetics, Faculty of Medicine, Shahid Beheshti University of Medical Sciences

Vijay Gupta

Hamad Bin Khalifa University

Afif Ben-Mahmoud

Hamad Bin Khalifa University

Katayoon Razjouyan

Shahid Beheshti University of Medical Sciences

Shadab Salehpour

Loghman Hakim Hospital, Shahid Beheshti University of Medical Sciences

Seyed Hassan Tonekaboni

Mofid Children's Hospital, Shahid Beheshti University of Medical Sciences

Mehdi Dianatpour

Shiraz University of Medical Sciences

Davood Omrani

Department of Medical Genetics, Faculty of Medicine, Shahid Beheshti University of Medical Sciences

Lawrence C. Layman

Augusta University

Hyung-Goo Kim (✉ hkim@hbku.edu.qa)

Hamad Bin Khalifa University

Mohammad Miryounesi

Shahid Beheshti University of Medical Sciences

Research Article

Keywords: syndromic autism, Exome sequencing, neurodevelopmental disorder, muscular dystrophy, de novo variants, consanguineous singlet family

Posted Date: March 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1411351/v1>

Abstract

Background

Autosomal recessive genetic disorders are leading health concerns in countries with consanguineous marriage. However, previously there was no strong association of consanguinity with autosomal dominant or X-linked inheritance patterns. Whether this would be the case in singlet consanguineous families remained unexplored.

Methods

The current study aimed to identify the disease genes and their inheritance patterns in 12 independent Iranian consanguineous singlet families with syndromic autism.

Results

Exome sequencing revealed two missense and two nonsense variants (one in each gene: *IQSEC2*, *FOXG1*, *DMD*, and *CHKB*) in four out of 12 families, thereby offering an aetiologic diagnosis. Autism, intellectual disability, learning disability, developmental delay, and behavioral problems were common features in these four families, and three individuals presented with impaired motor skills such as abnormal gait. Muscular dystrophy coupled with hypotonia were noted in two individuals, whereas in the remaining two families ADHD combined with language/speech delay were present. Along with the two X-linked (*IQSEC2* and *DMD*), one autosomal dominant (*FOXG1*) and one autosomal recessive genes (*CHKB*) were found in these four consanguineous families. Two genes-*DMD* and *CHKB* are known to be mainly involved in muscular dystrophy.

Limitation:

our study significance is limited by the small sample size, 12 individuals, and lack of functional experiment on a VUS on the *DMD* variant.

Conclusion

Our findings illustrate the mixed inheritance pattern in consanguineous singlet families and highlight the association of autism spectrum disorder with muscular dystrophy, both of which are seemingly not related.

Background

Autism spectrum disorder (ASD) is a group of neurodevelopmental conditions characterized by the triad of communication deficit, abnormal social interest, and restricted/repetitive behavior [1, 2]. It can be classified as syndromic and non-syndromic types depending on the presence or absence of other neurobehavioral phenotypes and dysmorphic features, respectively. Whereas non-syndromic autism describes the cases where autism is the primary diagnosis, syndromic autism manifests additional clinical features, such as intellectual disability, epilepsy, craniofacial anomalies or muscular dystrophy, which are usually associated with chromosomal abnormalities or single-gene mutations [3].

The heterogeneity of ASD includes not just clinical features, which involves neurobehavioral comorbidity, but also genetic diversity [4]. This significant level of genetic heterogeneity presents direct challenges to identifying causative variants and interpreting genetic test results in a clinical setting [3].

The most comprehensive genetic progress has been made in gene discovery of syndromic autism with more than 1,000 genes or gene loci involved thus far [5]. Although these genes represent a minority of all autism cases, studies demonstrated that syndromic and non-syndromic forms share common biological pathways [5]. Gene discovery in individuals with autism will provide tremendous insight into underlying pathogenic mechanisms. The latest discoveries have also identified the significance of *de novo* mutations (DNMs) underlying syndromic autism. The risk of autism is therefore due to large numbers of rare *de novo* and inherited variants, rather than small numbers of common variants [6].

Single gene testing in autism faces one of the largest challenges with such extreme heterogeneous genetic loci and clinical spectrum in the diverse group of disorders. Subsequently, with the advent of exome sequencing (ES) in human genetics, recent literature has supported that ES provides an efficient molecular tool for studying individuals suffering from a broad spectrum of the previously undiagnosed syndromic form [7], by identifying a likely pathogenic variant in approximately 15 to 30% of cases [8, 9]. Syndromic autism opens up new opportunities to provide a deep insight into the molecular heterogeneity involved in generating the phenotypic heterogeneity. In the present study, we analysed the genetic aetiology of syndromic autism by a phenotype first approach in the context of a clinically well-defined syndrome [3]. Our goal was to discover the inheritance pattern and interrelationship of co-existing clinical features in 12 singlet consanguineous Iranian families by identifying disease genes associated with diverse phenotypes including muscular dystrophy.

Methods

Human Subjects

The study sample encompassed twelve Iranian families with ASD diagnoses that were referred to the outpatient clinical genetic department by psychiatric specialists for ES. Prior to ES, karyotype analysis and microarray have been performed in twelve individuals from these twelve families to identify chromosome anomalies and CNVs. The all results were negative. All individuals were evaluated as having syndromic autism by multi-specialists in a joint clinic. After obtaining written informed consent from the families, complete clinical assessments and pedigree information were characterized prior to ES. Four out of twelve families have a positive WES result that is describe in this study (Fig. 1).

Clinical Description

Subject 1

This individual is a 21-year-old Iranian female from a consanguineous marriage (first cousins) who was diagnosed with a Rett-like phenotype with multiple comorbidities including intellectual disability, ASD, motor delay, and speech delay. She was born full-term by spontaneous vaginal delivery with ordinary biometric parameters. Her perinatal and neonatal periods were uneventful. Her parents reported that she did not walk, sit, or speak until 22 months old. She spoke her first word very late and has not spoken full sentences thus far, but currently can express some words. Her motor milestones were not normal but bolting (running away) is a frequent problem and causes injury. The first neuropsychological testing showed that her Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) full-scale score was extremely low (Composite Score Range is 69 and below). She scored poorly across visual-spatial, fluid reasoning, processing speed, working memory, and the verbal comprehension indices. She continues to struggle with independent planning and organization, as well as becoming frustrated with challenging tasks. There were neurobehavioral concerns when

repetitive movements and behavior were displayed. At 2 years 2 months, she was also diagnosed with ASD. She has no friends, exhibits no social skills, and escapes from a crowd. A progressively significant anxiety disorder was detected as well. Five fundamental deficits were considered: Acrophobia, Attention deficit hyperactivity disorder (ADHD), astasia, asthenia, and fatigability (unable to keep a muscular force against a resistance).

Subject 2

This proband is the only child of consanguineous healthy parents (first cousins). This individual is an 18-year-old Iranian female with a history of ASD, intellectual disability, obesity, and seizures. She was born full-term by spontaneous vaginal delivery with the average biometric parameters. From the first year of life, four fundamental deficits were measured: hypotonia, astasia-abasia, speech delay, and ADHD. She spoke her first words in the second year of life and not full sentences until five years old. She speaks some ambiguous words currently. Her most recent neuropsychological testing demonstrated that her WISC-V full-scale score was very low (Composite Score Range is 70–79). She scored poorly across working memory, visual-spatial, fluid reasoning, processing speed indices, but performed in the low average range on the verbal comprehension index. Neurobehavioral concerns were diagnosed from the very beginning when she displayed repetitive movements and behaviours. She was diagnosed with Rett syndrome with multiple comorbidities (Table 1). Several events happened in childhood during the same year when she had seizures without fever. Complex partial seizures were diagnosed which have been well controlled with sodium valproate. Significant obesity is evident with a BMI of more than 30 for the past 5 years and this continues to progress due to an unhealthy lifestyle.

Subject 3

This individual is a 16-year-old male with a history of muscular dystrophy and autism. He was born full-term by spontaneous vaginal delivery with normal biometric parameters from consanguineous healthy parents (second cousins). Even though he had hypotonia in the first year of life; (specifically poor head and neck control), his motor milestones were standard during childhood, but have gradually diminished. He has been diagnosed with Becker muscular dystrophy and ASD since childhood. He had elevated serum creatine phosphokinase (CPK) level of 1068 U/L (normal range/male: 39–308 U/L) and LDH was 346 U/L (normal range/male:140–280 U/L). His WISC-V report revealed a non-progressive low average score (composite Score Range is 80–89). He exhibited all three types of specific learning disabilities: dyslexia (reading disabilities), dysgraphia (written communication disorder), and dyscalculia (mathematics problem). He seemed disappointed with puzzling tasks but performed in the average range on the verbal comprehension index. He has few close friends, exhibits immature social skills, and displays repetitive movements and behavior.

Subject 4

This proband is a 4-year-old Iranian boy with a history of global developmental delay, intellectual disability, muscular dystrophy, and ASD. He was born after two spontaneous abortions by Cesarean section with normal biometric parameters from consanguineous healthy parents (second cousins). He has hypotonia, specifically poor head and neck control. He sat independently at 18 months, walked at 30 months with an abnormal gait, and has abasia. His WISC-V score was low (Composite Score Range is 70–79) from the first test and that has not changed. All WISC-V structures such as fluid reasoning, visual-spatial, verbal comprehension, working memory, processing speed indices were reported low. Although his MRI of the brain/spine was normal, the neurobehavioral problems are demonstrated by the appearance of repetitive behavior. He has no friends and reveals almost no social skills. Her mother's thrombophilia test was abnormal during pregnancy (the report is unavailable). A muscle biopsy test of the proband demonstrated megaconical congenital muscular dystrophy.

Table 1
Clinical characteristics of the four index patients.

Case Number	Subject 1	Subject 2	Subject 3	Subject 4
Current age	21	18	16	4
Sex	F	F	M	M
Ethnicity	Iran	Iran	Iran	Iran
Method of detection	ES	ES	ES	ES
Developmental delay	+	+	+	+
Intellectual disability	+	+	+	+
Learning disability	+	+	+	+
Autism	+	+	+	+
Epilepsy	-	+	-	-
Language/speech delay	+	+	-	+
ADHD	+	+	-	-
Rett Syndrome	Rett like	+	-	-
Behavioral problems	+	+	+	+
Anxiety disorder	+	-	-	NA
muscular dystrophy	-	-	+	+
Hypotonia	-	-	+	+
astasia-abasia	astasia	astasia-abasia	-	abasia
Obesity	-	+	-	-
Acrophobia	+	-	-	-

DNA Extraction, ES and Bioinformatics Analysis

To find the genetic defects in the affected members, genomic DNA (gDNA) was extracted from peripheral blood samples of probands and their parents by using the salting out method. The concentration and quality of gDNA were assessed by NanoDrop 1000 (Thermo Fisher Scientific, Inc., Wilmington, DE, USA). Custom-designed Nimblegen chip capturing and next-generation sequencing were utilized to perform ES. Briefly, after shearing one μg of gDNA from the proband, the SureSelectXT2 V6 kit was used for the enrichment of the targeted sequence, following by paired-end sequencing on Illumina HiSeq4000 with read length 101bp and coverage 100X. Manufacturer's proprietary software generated read files (Fastq) from the sequencing platform. To remove the adaptor and low-quality reads; the Fastq file was trimmed and aligned to human reference genome hg19 by recruiting Burrows-Wheeler Aligner (BWA). Then, BAM files were evaluated for variant calling by using SAMtools and Picard – which were followed by variant annotation via ANNOVAR software [10–14].

Additional filtering was also performed by determining the variant frequency in public databases– gnomAD (<https://gnomad.broadinstitute.org/>), dbSNP138 (<https://www.ncbi.nlm.nih.gov/snp/>), 1000 Genome projects (<http://www.internationalgenome.org>), and the Iranome database [15]. Only low-frequency variants with minor allele

frequency (MAF) less than 0.01 were selected. Variants (rare and novel) were prioritized according to nonsynonymous, indel, and putative splice sites. To predict the pathogenicity of the candidate variant, predictor tools: including Polyphen2 (<http://genetics.bwh.harvard.edu/pph2/>), SIFT (<https://sift.bii.a-star.edu.sg/>), MutationTaster (<http://www.mutationtaster.org/>), and most importantly CADD software were employed for *in silico* evaluation [16–21]. The ultimate sequencing results were also analyzed via international mutation and polymorphism databases alongside the in-house database.

Sanger Confirmation and Segregation Analysis

To confirm the ES results, flanking PCR primers were designed for all variants, and subsequent PCR products were sequenced by Sanger sequencing. Segregation analysis using proband and parents' samples exhibited that just two heterozygous (*IQSEC2* and *FOXP1*) variants and one hemizygous variant in *DMD* in 3 families are *de novo*, and one homozygous variant in *CHKB* segregated in the family in an autosomal recessive fashion.

Three Dimensional Structure Predictions

Phyre2 and SWISS-MODEL were applied for protein modelling. PyMol software (PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC, USA) and MetaDome webserver [22] were used to visualize the 3D protein structure model and analysis of tolerance of point mutation along with structural effects in the proteins, respectively.

CADD score

Four validated sequence variants were annotated by using CADD (Combined Annotation Dependent Depletion) scores (C-scores) [23], a publicly available online tool that ranks multiple genetic variants to predict an estimation of deleteriousness of variants throughout the human genome reference assembly [24]. According to the CADD framework, a machine learning CADD model is trained by using more than 60 diverse annotations to classify variants as neutral versus deleterious. After model training, CADD scores, calculated for all ~9 billion potential SNVs of the human reference genome, were annotated by using the same features. Finally, 0 to 10 scores were scaled related to 90% (~7.7 billion) of all GRCh37/hg19 reference SNVs (~8.6 billion), while the next 9% (top 10% to top 1%, spanning ~774 million SNVs) were scored 10 to 20, etc. CADD scores reflect causal or fine-mapping variants in estimated pathogenicity relative to all possible SNVs in the reference genome [24]. The higher the CADD score, the greater the likelihood that the variant is deleterious.

RT-qPCR

Total RNA from the human brain and other tissues (Human Total RNA Master Panel II, Cat# 636643, Clontech), were used for RT-qPCR. The catalogue number of eight brain tissues from Clontech are as follows (dorsal root ganglion-636150, fetal brain total-636526, substantia nigra total-636560, cerebral cortex-636561, occipetal lobe-636570, parietal lobe-636571, post central gyrus-636573, hippocampus-636565). The cDNA synthesis was performed using 1 µg of total RNA using high-Capacity cDNA Reverse Transcription Kit and analyzed by RT-PCR on QuantStudio 6 Flex system using SYBR Green (ThermoFisher, Waltham, MA). Utilizing the comparative Δ Ct method, relative expression of each gene was calculated. In short, relative gene expression was calculated by the difference (Δ Ct) between the Ct value of the gene of interest and reference gene, GAPDH. After that the fold change ($2^{-\Delta$ Ct) was measured and relative expression was plotted as excel graphs.

Results

Identification of four novel ASD-related gene variants

Four novel variants were identified within *IQSEC2* (c.1637G > A, p.W546X), *FOXP1* (c.563C > A, p.A188E), *DMD* (c.631T > A, p.L211M), and *CHKB* (c.382G > T, p.E128X) in known syndromic autism-related genes in the four families, based on

ES (Table 2).

The pathogenic *de novo* variant c.1637G > A in *IQSEC2* (IQ motif and sec7domain2, MIM 300522) in Subject 1 is located between IQ-like and Sec7 domains far before the critical residue for enzymatic activity at position 849. This nonsense variant p.W546X predicted a truncated protein lacking the SEC7, PH, and PDZ binding domains in *IQSEC2* if expressed (Fig. 2). This premature stop codon could prompt NMD since the stop codon is located 661 nucleotides (NTs) upstream of the junction between exons 5 and 6 (50–55 nucleotide rule) [25] and prevent the expression of the truncated protein. Even if a C-terminal truncated protein escapes from NMD to some degree, it would lack the most prominent Sec7 enzyme activity and both the pleckstrin homology domain (PH) and PDZ binding motif.

An additional observed alteration in Subject 2 was a *de novo* heterozygous missense mutation in *FOXP1* (Forkhead box G1, MIM 164874), c.563C > A/ p.A188E, in an affected female. A detected variant located on the forkhead binding domain (FBD, amino acids from 179 to 269), highly conserved (Fig. 3) in this single-exon gene, is absent in her unaffected parents.

The candidate variant (c.631T > A, p.L211M) in Subject 3 was a *de novo* hemizygous missense change in *DMD* (Dystrophin, MIM 300377) in an affected male because it was absent in his both unaffected parents.

The methionine residue at 211 of *DMD* misfolds this protein and leads to an alteration in the 3D structure and surface electrostatic potential in the protein modelling (Fig. 4). So, in Subject 3, not only does this variant act as a prominent cause of muscular dystrophy, but also is the prime candidate for autistic features. The phenotypic spectrum, as demonstrated in this case, illustrates the multifunctional role of dystrophin. Its CADD score scale is 15, where a score greater than 10 indicates the top of less than 10% of most deleteriousness; however, it is still most likely the number one candidate for the comorbid neuropsychiatric phenotype. Careful NGS data analysis and consideration of all modes of inheritance also suggest that it has an X-linked recessive pattern.

The variant (c.382G > T, p.E128X) in exon 3 in *CHKB* (Choline Kinase Beta, MIM 612395) was found in Subject 4 and is predicted to result in a truncated protein and to disrupt a major portion of the central domain and the entire *CHKB* C-terminal domain. *CHKB* comprises two chains according to PDB 2IG7 (Fig. 5). Protein absence was predicted due to a premature stop codon that could trigger NMD to degrade the mRNA because the stop codon is located 66 nucleotides (NTs) upstream of the junction of exons 3 and 4 (50–55 nucleotide rule) [25]. Even if a C-terminal truncated protein escapes NMD to some degree, *CHKB* dimers important for its choline kinase function would not be produced (Fig. 5).

Protein Modelling

In Subject 2 with an A188E variant, the residue Alanine at 188 is located in a highly intolerant region of *FOXP1* based on the MetaDome web server (Fig. 3C) and its substitution to the bulky Glutamic acid is likely to create spatial and steric hindrance leading to misfolding of *FOXP1* (Figs. 3A & 3B).

Homology modelling of amino acid 227 in *DMD* showed spatial inhibition for the mutated amino acid, Methionine, at position 211 in Subject 3 that may result in misfolding of the structure of *DMD* (Fig. 4B). Moreover, L211M leads to an alteration in the 3-dimensional structure as well as the surface electrostatic potential of *DMD* (Fig. 4C).

The negatively charged glutamic acid at position 128 is essential for the dimerization of the *CHKB* monomer. Figures 5A and C display the crucial role of E128 for the dimerization of *CHKB*. The nonsense variant p.E128X at this residue is, therefore, predicted to be highly deleterious in Subject 4.

The American College of Medical Genetics and Genomics (ACMG) classification

According to ACMG standards and guidelines [26], we interpreted the pathogenicity of genetic variants with specific standard terminology: “pathogenic,” “likely pathogenic,” “uncertain significance,” “likely benign,” and “benign”. Since our four novel variants found in Iranian patients are absent in publicly available population databases including gnomAD and 1000 Genome Project, this was considered a moderate piece of evidence for pathogenicity (PM2). Based on the rules combining criteria to classify sequence variants from ACMG, three variants have been classified as pathogenic, whereas one was likely pathogenic (Table 2).

3.4 CADD Scores

CADD scores of *IQSEC2* (c.1637G > A, NM_001111125.3), *FOXG1* (c.563C > A, NM_005249.5), and *CHKB* (c.382G > T, NM_005198.4) are more than 30 indicating the top 0.1% most deleterious, and the *DMD* (c.631T > A, NM_004010.3) score was 15, predicted to be in the top 10% most deleterious substitutions (Table 2). These CADD scores are very well harmonized with the ACMG classification of genetic variants.

Table 2
Clinical data, ES results and some prominent related data for defined novel mutations.

ID	Subject 1	Subject 2	Subject 3	Subject 4
Gene name	<i>IQSEC2</i>	<i>FOXG1</i>	<i>DMD</i>	<i>CHKB</i>
Exon number	exon 5 of 15	exon 1 of 1	exon 10 of 79	exon 3 of 11
Nucleotide change	c.1637G > A	c.563C > A	c.631T > A	c.382G > T
NM number	(NM_001111125.3)	(NM_005249.5)	(NM_004010.3)	(NM_005198.4)
genomic change (hg38)	g.53,250,939 C > T	g.28,767,842 C > A	g.32,645,113 A > T	g.50,581,814 C > A
Effect on protein	p.W546X	p.A188E	p.L211M	p.E128X
NP number	NP_001104595.1	NP_005240.3	NP_004001.1	NP_005189.2
Alteration Type	nonsense	missense	missense	nonsense
ACMG classification	Pathogenic PVS1/PS2/PM2/PP3	Pathogenic PS2/PM1/PM2/PM5/PP2/PP3	VUS PS2/PM2/PP3	Pathogenic PVS1/ PM2/PM3/PP1/PP3
CADD Score	36	42	15	32
Zygoty	Het	Het	Hemi	Hom
ClinVar	SCV001251727	SCV001337647.1	SCV000746310	SCV000923677
MAF	N/A	N/A	0.0000109	N/A
Inheritance	XLD / <i>de novo</i>	AD / <i>de novo</i>	XLR / <i>de novo</i>	AR
sex	F	F	M	M
Method of detection	ES	ES	ES	ES
Associated Syndrome (MIM)	Intellectual disability, X-linked 1 (MIM 309530) Rett-like syndrome	Rett syndrome, congenital variant (MIM 613454)	Becker Muscular Dystrophy (MIM 300376) dilated	Megaconial type of congenital muscular dystrophy (CMD) (MIM 602541)
<p>AF, autism family; MAF, minor allele frequency; IP, inheritance pattern; Hemi, hemizygous; Het, heterozygous; Hom, homozygous; Z, zygoty; AD, autosomal dominant; AR, autosomal recessive; XLR, X-linked recessive; XLD, X-linked Dominant; ES, exome sequencing; NA, not available; PVS, very strong; PM, Moderate; PP, supporting</p>				

Expression patterns of four ASD-related genes in the brain and other human tissues

We determined the transcript levels of *FOXG1*, *IQSEC2*, *DMD* and *CHKB* in various human tissues by RT-qPCR. Higher *FOXG1* and *IQSEC2* expressions were detected in the fetal brain compared to other tissues whereas *CHKB* and *DMD* expressions in the skeletal muscles were highest (Fig. 6A). All four genes *FOXG1*, *IQSEC2*, *DMD* and *CHKB* showed expressions in various parts of the brain with varying levels (Fig. 6B). High levels of expression of *FOXG1* and *IQSEC2* in the fetal brain tissues indicated their involvement in causing neurological phenotypes, and higher levels of *CHKB* and *DMD* expression in the skeletal muscles denote their role in muscle disorders.

Discussion

In consanguineous families with human developmental disorders, autosomal recessive genes are generally identified as the genetic aetiology. We have performed exome sequencing in 12 independent Iranian consanguineous singlet families with syndromic autism and identified four novel variants in *IQSEC2*, *FOXG1*, *DMD*, and *CHKB* in four families. Among them, only *CHKB* is an autosomal recessive gene, while *IQSEC2* and *DMD* are X-linked and *FOXG1* is an autosomal dominant gene. Interestingly, two of them, *DMD* and *CHKB*, are known muscular dystrophy genes. This emphasizes the striking association between autistic spectrum disorder and muscular dystrophy. Autosomal dominant disease genes with homozygous variants were reported in consanguineous Iranian families previously and both parents, who are heterozygous carriers in those families, were healthy [15, 27]. By contrast, Subject 2 with a heterozygous variant in an autosomal dominant gene, *FOXG1*, in our consanguineous singlet Iranian family was affected with syndromic autism.

ASD aetiology has consistently pointed to robust genetic components; however it has been complicated by extensive genetic heterogeneity. The current study disclosed four novel disease variants in *IQSEC2*, *FOXG1*, *DMD*, and *CHKB* in four syndromic singlet autism families from consanguineous marriages. Accordingly, two *de novo* heterozygous variants were found in *IQSEC2* (c.1637G > A, p.W546X) and *FOXG1* (c.563C > A, p.A188E), respectively, that associate with ASD and a neurodevelopmental disorder. A *de novo* variant in *DMD* (c.631T > A, p.L211M) and a novel homozygous variant in *CHKB* (c.382G > T, p.E128X) are likely involved in neuromuscular disorders and ASD.

The phenotypic spectrum of Subject 1 demonstrates that *IQSEC2* is an X-linked gene with a female-specific X-linked dominant phenotype, including intellectual disability, autism, developmental delay, speech delay, stereotypic hand movements, poor eye contact, acrophobia, sleep disorder, and ADHD. High expression of *IQSEC2* in the human fetal brain highlights the functional role of this gene in early human development, which is consistent with these clinical features. The truncating variant W546X in *IQSEC2* is likely to generate loss-of-function because it will eliminate three functional domains (Sec7, PH, and PDZ binding motif) affected, even if it escapes NMD. According to ACMG classification, this variant belongs to the category of PVS1/PS2/PM2/PP3 and combining these criteria classifies it as pathogenic (Table 2). Moreover, the CADD score is 36, and a score of greater than 30 designates the top of less than 0.1% of most deleteriousness. Careful NGS data analysis with consideration of all modes of inheritance suggests that it may be unrelated to shared recessive genes identified in consanguineous families.

Residue Ala188 in *FOXG1* is located in Forkhead domain, which is highly conserved and intolerant. Thus, the *de novo* variant Glu188 (c.563C > A) in Subject 2 will likely affect DNA binding of this domain. According to protein modeling, its replacement with the bulky Glutamic acid is predicted to cause spatial and steric hindrance, resulting in *FOXG1* misfolding. Development delay, ASD, seizures, and intellectual disability found in this 18-year-old Iranian girl can be explained by the high expression pattern of *FOXG1* in the human fetal brain (Fig. 6A). Based on ACMG classification, this belongs to the categories of PS2/PM1/PM2/PM5/PP2/PP3 and thus can be classified as pathogenic (Table 2). Moreover, The CADD score is 42 and greater than 40, indicates the top of less than 0.01% of most deleteriousness. A *de novo* variant also suggests autosomal dominant and unlikely to be an autosomal recessive pattern based on the consanguineous marriage of her parents.

The X-linked variant (c.631T > A, p.L211M) in Subject 3 was a *de novo* hemizygous missense change in *DMD* due to its absence in his unaffected parents. Two major phenotypes of a neurodevelopmental feature coupled with muscular dystrophy in this Iranian male could be explained by the highest expression of *DMD* in the skeletal muscle followed by the second highest expression in the human fetal brain (Fig. 6A). The Z score of a missense variant in *DMD* in the general population is -2.43 based on gnomAD data (<https://gnomad.broadinstitute.org/>), suggesting more tolerance to the missense variation. However, this does not preclude the possibility of missense mutations in this gene including L211M reported here.

There are multiple lines of evidence suggesting the pathogenicity of this *DMD* missense variant in this proband with development delay, autism, intellectual disability, muscular dystrophy, and hypotonia. Firstly, there are totally 5,159 mutations reported in *DMD* based on the HGMD website (Human Gene Mutation Database, <http://www.hgmd.cf.ac.uk/ac/index.php>). Among them, 989 are missense or nonsense mutations. Approximately one third of these are reported missense mutations, which suggests > 300 reported missense mutations in *DMD* in human patients. Secondly, L211M is absent in both parents and thus *de novo*. It is also absent in gnomAD and thus defined as PM2 based on ACMG guidelines [26]. Notably an elevated serum CPK level and the clinical details suggested dystrophinopathy as third evidence. Furthermore, *in silico* protein modeling of the residue L211M in Fig. 4 shows the disruption of the 3-dimensional structure in the mutant DMD. Taken together, pathogenicity has been determined as “likely pathogenic” in Table 2 by the entire body of evidence (PS2/PM2/PP3). Without our result in Subject 3, clinicians or geneticists would never have thought that an ASD male with muscular dystrophy from a consanguineous marriage could have a mutation in the X-linked *DMD* gene.

In Subject 4, the homozygous nonsense variant, c.382G > T (p.E128X) in exon 3 of *CHKB* was heterozygous in both healthy carrier parents. This premature stop codon was predicted to result in a truncated protein. Even if a C-terminal truncated protein escapes NMD to some extent, CHKB dimers, which are required for choline kinase function, would not be generated (Fig. 5). The CADD score scale was 32, which indicates the top of less than 0.1% of most deleteriousness along with ACMG classification criteria using combined criteria of PVS1/PM2/PM3/PP1/PP3 indicated that it is pathogenic (Table 2). Clearly, this variant follows an autosomal recessive inheritance pattern, based on the family history and in line with our hypothesis for consanguineous families. *CHKB* had higher levels of mRNA expression in skeletal muscles and brain tissues (Fig. 6A), which is consistent with its involvement in muscular dystrophy and autism.

In the remaining eight families, we identified six polymorphic variants in six families, among whom three heterozygous variants were found in one of the healthy parents, whereas another three heterozygous variants were found in both healthy parents in the other three families. In two families no candidate variants were found.

RTT and autism

Rett syndrome (RTT, MIM 312750) is a rare neurodevelopmental disorder, which affects females almost exclusively, occurring in 1 in 10,000–15,000 female live births, and which is responsible for around 10% of female severe intellectual disabilities of genetic origin [28]. Using a systematic gene screen approach from RTT families, in whom exclusion mapping studies mapped the disease gene locus to Xq28, *MeCP2* (methyl-CpG-binding protein 2) was identified as a disease gene. Clinical features of typical RTT appear to be enigmatic for the normal perinatal period until 6–18 months of age, followed by loss of speech and purposeful hand use, and developing microcephaly, seizures, and autism. *IQSEC2* mutations were identified in some individuals with either classic RTT or clinical symptoms similar to RTT [29]. whereas *FOXP1* mutations underlie atypical RTT [30]. So, from a genetic point of view, *IQSEC2* and *FOXP1* are clearly responsible for atypical and variant RTTs [31].

IQSEC2, identified in the families with intellectual disability by systematic X chromosome exome sequencing in 2010 [32], is an X-linked gene (Xp11.22) having been named for two protein domains, IQ-like and Sec7 domains (Fig. 2). *IQSEC2* is a member of GEF (guanine nucleotide exchange factor) for the ADP-Ribosylation Factor (ARF) family of small GTP-binding proteins (ARFGEF) [32]. According to the high expression level of *IQSEC2* in all brain regions (highest in the hippocampus), its disruption represents a significant cause of intellectual disability [33]. *IQSEC2* variants are also associated with seizures, behavioural and psychiatric abnormalities. A recent review reported 70 different types of *IQSEC2* mutations, in which autism presented in 25% of affected males and 30% of females [34]. Escaping X inactivation to some extent may explain the rather high prevalence of both intellectual disability and autism in heterozygous females. Due to this high level of comorbidity, it has been demonstrated that *IQSEC2*-associated intellectual disability and autism share common biochemical abnormalities [21]. Males have both missense and truncating pathogenic variants; however, females have mostly truncating variants like in Subject 1 of our study [35].

A highly regulated process of brain development encompasses the particular spatio-temporal activation of cell signaling clues [36]. Transcription factors play a fundamental role in this process by spreading information from external signaling to the genome. The encoded protein, Forkhead box protein G1, *FOXP1* (formerly brain factor 1 [BF-1]) primarily acts as a brain-specific transcriptional repressor, which is involved in brain development [37]. It is expressed in the evolving nervous system, where it plays a critical role in the establishment of the regional subdivision of the developing brain ranging from telencephalon specification, patterning and neuronal differentiation, maintenance, and survival of mature neurons [38, 39]. It is one of the earliest expressed transcription factors that can change the structure of chromatin and permit the binding of other factors. Consequently, it is considered a pioneer transcription factor.

Any alterations in the *FOXP1* dosage resulted in a complex group of cellular effects with key consequences for human diseases, including neurodevelopmental disorders. As an example, *FOXP1* interacts with global transcriptional co-repressors of the Groucho family as well as with the transcriptional repressor *JARID1B*. The interaction of these proteins has functional significance for early brain development. *FOXP1* also associates indirectly with the histone deacetylase 1 protein (*HDAC1*), as *MeCP2* also does [40].

Autism and muscular dystrophy

Muscular dystrophy has also been accompanied by neuropsychiatric disorders like ASD and association between autistic spectrum disorders and dystrophinopathy (including Duchenne and Becker muscular dystrophies) has recently been reported [41]. Duchenne muscular dystrophy and megaconial congenital muscular dystrophy manifest a broad spectrum of clinical severity with multisystem involvement. Based on the largest study, only 11 of 351 (3.1%) boys with muscular dystrophy were reported to have ASD [42], indicating that ASD in DMD is still a rare event.

As one of the first cloned human genes by positional cloning, the *DMD* gene attracts attention not only for its medical prominence but also for its exceptionally large size. It spans a genomic range of more than 2 Mb and encodes a large protein, dystrophin, containing an N-terminal actin-binding domain, and multiple spectrin repeats alongside a family of N-terminal truncated isoforms by activating independent promoters. Mutations, most commonly deletions, in *DMD*, located on the short arm of the X chromosome (Xp21.1 to Xp21. 2), lead to Duchenne and Becker muscular dystrophies (DMD and BMD). These have an impact not only on muscle but also on the brain [43]. With seven promoters, *DMD* has a full-length and shorter isoforms that may play a crucial role in synaptogenesis and axon guidance during brain development; hence, Becker Muscular Dystrophy is associated with neuropsychiatric disorders including ASD, which has been reported in two individuals with *DMD* mutations [44]. This is also supported by the highest expression of *DMD* in the skeletal muscle tissues compared to other human organ tissues (Fig. 6A).

Dystrophin and several components of the dystrophin-associated protein complex (DPC) can be found in the soma and dendrites of cortical and hippocampal neurons as well as cerebellar Purkinje cells, where it is associated with the

postsynaptic membrane of neurons and is present at high levels in the postsynaptic density (PSD); therefore, it may have a role in the synapse structure or function [45]. Additionally, the full-length isoforms are largely expressed neuronally and localized to punctate structures, which correspond to a group of GABAergic synapses in many brain regions (the cerebellum, cortex, and hippocampus) that are responsible for higher-order functions such as learning and memory. Therefore, cognitive impairment in individuals with *DMD* mutations is caused by a lack of dystrophin in the neuronal membrane.

The *CHKB* is located on 22q13.33 that comprises 11 exons, 10 of which (except exon 10) have been implicated in the pathogenesis of megaconial congenital muscular dystrophy (CMD). It encodes the choline kinase β protein, which catalyzes the first step of the *de novo* biosynthesis of phosphatidylethanolamine and phosphatidylcholine, an important lipid component of the mitochondrial membrane of skeletal muscles and the brain [46]. Then, it plays a vital role in phospholipid biosynthesis.

Autosomal recessive congenital megaconial muscular dystrophy (CMD, MIM 602541) is characterized by neonatal hypotonia, early-onset muscle wasting, intellectual disability, and developmental delays without brain malformation as specific clinical features, which were manifested in Subject 4. It is a rare inborn error of metabolism with a broad spectrum of clinical severity and multi system involvement in which some individuals presented with autism [47]. CMD is characterized by a distinct sparseness or absence of mitochondria in the centers of fibers and enlarged (megaconial) mitochondria in the periphery of muscle [48]. *CHKB* catalyzes the first step of phosphatidylcholine (PC) biosynthesis and its mutations result in loss of choline kinase activity and decreased levels of PC [47]. In relation to the phenotype-first approach, the clinical diagnosis of mitochondrial disease was primarily considered to be due to the multi-system nature of the condition. Laboratory testing such as pathognomonic features in the muscle biopsy and elevated serum creatine kinase also confirmed phenotypic compatibility with the identified variant.

We have found two muscular dystrophy genes, *DMD* and *CHKB*, mutated in our two syndromic autism families. These two autism genes overlap with non-cognitive phenotypes of muscular dystrophy, demonstrating significant genetic relation across neurodevelopmental and neuromuscular disorders. It clearly shows that some genes transcend the boundary of clinical categories, and this result could have broad clinical implications for diagnosis.

Our study shows that some children with autism may have a genetic defect that affects the muscles, and consanguineous singlet autism individuals might have autosomal dominant or X-linked disease genes. Additionally, future research consisting of animal models and functional studies of these four novel variants are essential to understand the pathological role of these variants.

Limitation

The major limitation of our study should be illustrated in the context of Functional experiment and sample size. However, our finding is mostly pathogenic due to ACMG classification and could be causative of the disorder, Functional studies are still important to understand the pathoethology of the novel variants, mainly a VUS variant in the *DMD* gene. Although we performed the first study in Iranian patients with ASD and concluded the novel clinical insight to ASD, our sample size is small, twelve families. Notwithstanding that recruiting suitable cases in one type of ASD (Syndromic) is critical with multi-center collaboration with mostly clinical implications and specialists. Due to untrusted behavior from other scientists that created confusion of the families suffering from ASD, and things of that nature, we missed the suitable cases, 30 out of 42. Moreover, it is worth noting that our study established to depict the genetic architecture of syndromic autism in Iranian patients and in the future cohort study we involved more cases.

Conclusion

In four Iranian consanguineous singlet autism families, we have identified one autosomal dominant, one autosomal recessive, and two X-linked autism variants in four different genes. Among four novel coding variants, three are *de novo* (two heterozygous mutations in *IQSEC2* as well as *FOXP1*, respectively, and one hemizygous variant in *DMD*). Two nonsense (*IQSEC2* and *CHKB*) and two missense variants (*FOXP1* and *DMD*) in four individuals are likely loss-of-function variants. Two missense variants in *FOXP1* and *DMD*, respectively, are predicted to affect protein folding, whereas two nonsense variants in *CHKB* and *IQSEC2*, respectively, would likely trigger *NMD*, preventing their protein expression. When expressed, the nonsense variant in *CHKB* affects its dimerization, which is important for its choline kinase function and the nonsense variant in *IQSEC2* would eliminate 3 functional domains. Since *DMD* and *CHKB* are generally known to be involved in muscular dystrophy, this suggests a link between ASD and muscular dystrophy.

Declarations

Ethics approval and consent to participate

All procedures performed in this study, was approved by the Ethical Committee, deputy of research affairs of shahid beheshti university of medical sciences, Tehran, Iran with approval ID: IR.SBMU.MSP.REC.1398.575, following with the 1964 helsinki declaration and its later amendments or comparable ethical standards.

Consent for publication

All families consented for the publication

Availability of data and materials

The data that support the findings of this study are available from the corresponding author, [M. Miryounesi], upon reasonable request.

Competing interests

The authors declare no conflict of interest.

Funding

This research was funded by “Research Department of the School of Medicine Shahid Beheshti University of Medical Sciences”, grant number (Pajooan code): 21827” and the startup funding of the Qatar Biomedical Research Institute at Hamad Bin Khalifa University (H.-G.K.).

Authors' contributions

Conceptualization; M-R.G.,M.D., H-G.K. and M.M.; Data curation, M-R.G, F.H-G, and M.D.; Formal analysis, M-R.G. and V.G.; Funding acquisition, M.M.; Investigation, M-R.G.,K.R., and M.M.; Methodology, M-R.G., M.D. and M.M.; Project administration, M.M.; Resources, M.M.; Software, M-R.G., F.H-G and M.D.; Supervision, H-G.K. and M.M.; Validation, H.S.,R. M., H-G.K. and M.M.; Visualization, M-R.G., K.R., S.S., and S-H.T.; Writing – original draft, M-R.G. and H-G K; Writing – review & editing, H.S., R.M., K.R., S.S., S-H.T.,A.B-M. D.O., M.D., L.C.L., V.G, A.B-M, and M.M. All authors have read and agreed to the published the manuscript.

Acknowledgements

Sincere gratitude to the families for their participation in this study and thanks to Lynn Chorich for proofreading the article. This study has been extracted from the Ph.D. dissertation written by Mr. Mohammad-Reza Ghasemi in School of

Medicine, Shahid Beheshti University of Medical Sciences (Registration no: 493). We also thank the support of funding provided by Caroline Jones-Carrick and Collin Carrick.

References

1. Association, A.P., *Diagnostic and statistical manual of mental disorders (DSM-5®)*. 2013: American Psychiatric Pub.
2. Lord, C., et al., Autism spectrum disorder. 2020. **6**(1): p. 1–23.
3. Fernandez, B.A. and S.W.J.D.i.c.n. Scherer, *Syndromic autism spectrum disorders: moving from a clinically defined to a molecularly defined approach*. 2017. **19**(4): p. 353.
4. Carter, M.T. and S.W.J.C.g. Scherer, *Autism spectrum disorder in the genetics clinic: a review*. 2013. **83**(5): p. 399–407.
5. Ziats, C.A., W.G. Patterson, and M.J.P.N. Friez, *Syndromic autism revisited: review of the literature and lessons learned*. 2020.
6. Iossifov, I., et al., *The contribution of de novo coding mutations to autism spectrum disorder*. 2014. **515**(7526): p. 216–221.
7. Satterstrom, F.K., et al., *Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism*. 2020. **180**(3): p. 568–584. e23.
8. Rossi, M., et al., *Outcomes of diagnostic exome sequencing in patients with diagnosed or suspected autism spectrum disorders*. 2017. **70**: p. 34–43. e2.
9. Arteché-López, A., et al., *Towards a Change in the Diagnostic Algorithm of Autism Spectrum Disorders: Evidence Supporting Whole Exome Sequencing as a First-Tier Test*. 2021. **12**(4): p. 560.
10. Cingolani, P., et al., *A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3*. 2012. **6**(2): p. 80–92.
11. Li, H. and R.J.B. Durbin, *Fast and accurate long-read alignment with Burrows–Wheeler transform*. 2010. **26**(5): p. 589–595.
12. Li, H., et al., *The sequence alignment/map format and SAMtools*. 2009. **25**(16): p. 2078–2079.
13. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data*. 2010. **20**(9): p. 1297–1303.
14. Nature, G.P.C.J., *A map of human genome variation from population-scale sequencing*. 2010. **467**(7319): p. 1061.
15. Hu, H., et al., *Genetics of intellectual disability in consanguineous families*. 2019. **24**(7): p. 1027–1039.
16. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations*. 2010. **7**(4): p. 248–249.
17. Keshava Prasad, T.t., et al., *Human protein reference database—2009 update*. 2009. **37**(suppl_1): p. D767-D772.
18. Kumar, P., S. Henikoff, and P.C.J.N.p. Ng, *Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm*. 2009. **4**(7): p. 1073.
19. Salgado, D., et al., *UMD-predictor: a high-throughput sequencing compliant system for pathogenicity prediction of any human cDNA substitution*. 2016. **37**(5): p. 439–446.
20. Schwarz, J.M., et al., *MutationTaster2: mutation prediction for the deep-sequencing age*. 2014. **11**(4): p. 361–362.
21. Tejada, M.I. and N.J.C.g. Ibarluzea, *Non-syndromic X linked intellectual disability: Current knowledge in light of the recent advances in molecular and functional studies*. 2020. **97**(5): p. 677–687.
22. Wiel, L., et al., *MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains*. 2019. **40**(8): p. 1030–1038.

23. Kircher, M., et al., *A general framework for estimating the relative pathogenicity of human genetic variants*. 2014. **46**(3): p. 310–315.
24. Rentzsch, P., et al., *CADD: predicting the deleteriousness of variants throughout the human genome*. 2019. **47**(D1): p. D886-D894.
25. Zhang, Z., et al., *Noisy splicing, more than expression regulation, explains why some exons are subject to nonsense-mediated mRNA decay*. 2009. **7**(1): p. 1–13.
26. Richards, S., et al., *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*. 2015. **17**(5): p. 405–423.
27. Thangarajh, M., et al., *Relationships between DMD mutations and neurodevelopment in dystrophinopathy*. 2019. **93**(17): p. e1597-e1604.
28. Aldosary, M., et al., *Rett syndrome, a neurodevelopmental disorder, whole-transcriptome, and mitochondrial genome multiomics analyses identify novel variations and disease pathways*. 2020. **24**(3): p. 160–171.
29. Lopergolo, D., et al., *IQSEC2 disorder: a new disease entity or a Rett spectrum continuum?* 2020.
30. Allou, L., et al., *14q12 and severe Rett-like phenotypes: new clinical insights and physical mapping of FOXP1-regulatory elements*. 2012. **20**(12): p. 1216–1223.
31. Allou, L., et al., *Rett-like phenotypes: Expanding the genetic heterogeneity to the KCNA2 gene and first familial case of CDKL5-related disease*. 2017. **91**(3): p. 431–440.
32. Shoubridge, C., et al., *Mutations in the guanine nucleotide exchange factor gene IQSEC2 cause nonsyndromic intellectual disability*. 2010. **42**(6): p. 486–488.
33. Tarpey, P.S., et al., *A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation*. 2009. **41**(5): p. 535–543.
34. Shoubridge, C., R.J. Harvey, and T.J.H.m. Dudding-Byth, *IQSEC2 mutation update and review of the female-specific phenotype spectrum including intellectual disability and epilepsy*. 2019. **40**(1): p. 5–24.
35. Mignot, C., et al., *IQSEC2-related encephalopathy in males and females: a comparative study including 37 novel patients*. 2019. **21**(4): p. 837–849.
36. Rakic, P.J.N.R.N., *Evolution of the neocortex: a perspective from developmental biology*. 2009. **10**(10): p. 724–735.
37. Hou, P.-S., et al., *Transcription and beyond: delineating FOXP1 function in cortical development and disorders*. 2020. **14**: p. 35.
38. Ariani, F., et al., *FOXP1 is responsible for the congenital variant of Rett syndrome*. 2008. **83**(1): p. 89–93.
39. Kumamoto, T., C.J.D. Hanashima, *growth, and differentiation, Evolutionary conservation and conversion of Foxg1 function in brain development*. 2017. **59**(4): p. 258–269.
40. Chahrour, M. and H.Y.J.N. Zoghbi, *The story of Rett syndrome: from clinic to neurobiology*. 2007. **56**(3): p. 422–437.
41. Parisi, L., et al., *Autism spectrum disorders in children affected by Duchenne muscular dystrophy*. 2018. **70**(3): p. 233–239.
42. Pagnamenta, A.T., et al., *A family with autism and rare copy number variants disrupting the Duchenne/Becker muscular dystrophy gene DMD and TRPM3*. 2011. **3**(2): p. 124–131.
43. Ricotti, V., et al., *Dystrophin and the brain*. 2011. **53**(1): p. 12.
44. Mori-Yoshimura, M., et al., *Psychiatric and neurodevelopmental aspects of Becker muscular dystrophy*. 2019. **29**(12): p. 930–939.
45. Lidov, H.G., et al., *Localization of dystrophin to postsynaptic regions of central nervous system cortical neurons*. 1990. **348**(6303): p. 725–728.

46. Mitsuhashi, S. and I.J.C.o.i.n. Nishino, *Megaconial congenital muscular dystrophy due to loss-of-function mutations in choline kinase β* . 2013. **26**(5): p. 536–543.
47. Chan, S.H., et al., *Megaconial congenital muscular dystrophy: same novel homozygous mutation in CHKB gene in two unrelated Chinese patients*. 2020. **30**(1): p. 47–53.
48. Mitsuhashi, S., et al., *A congenital muscular dystrophy with mitochondrial structural abnormalities caused by defective de novo phosphatidylcholine biosynthesis*. 2011. **88**(6): p. 845–851.

Figures

Figure 1

Pedigrees of the four families; solid circles and squares illustrated affected females and males, respectively; open circle/squares with black dots showed carrier mother/ father.

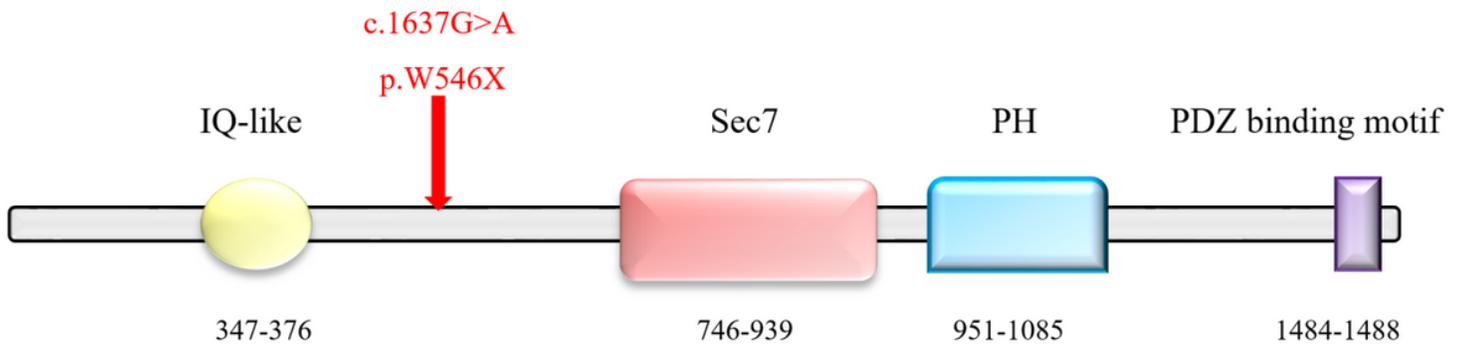


Figure 2

IQSEC2 structure and mutation. Human IQSEC2 protein Scheme (NP_001104595); illustrate the IQ-like domain, Sec7 domain, pleckstrin homology domain (PH), and PDZ binding motif.

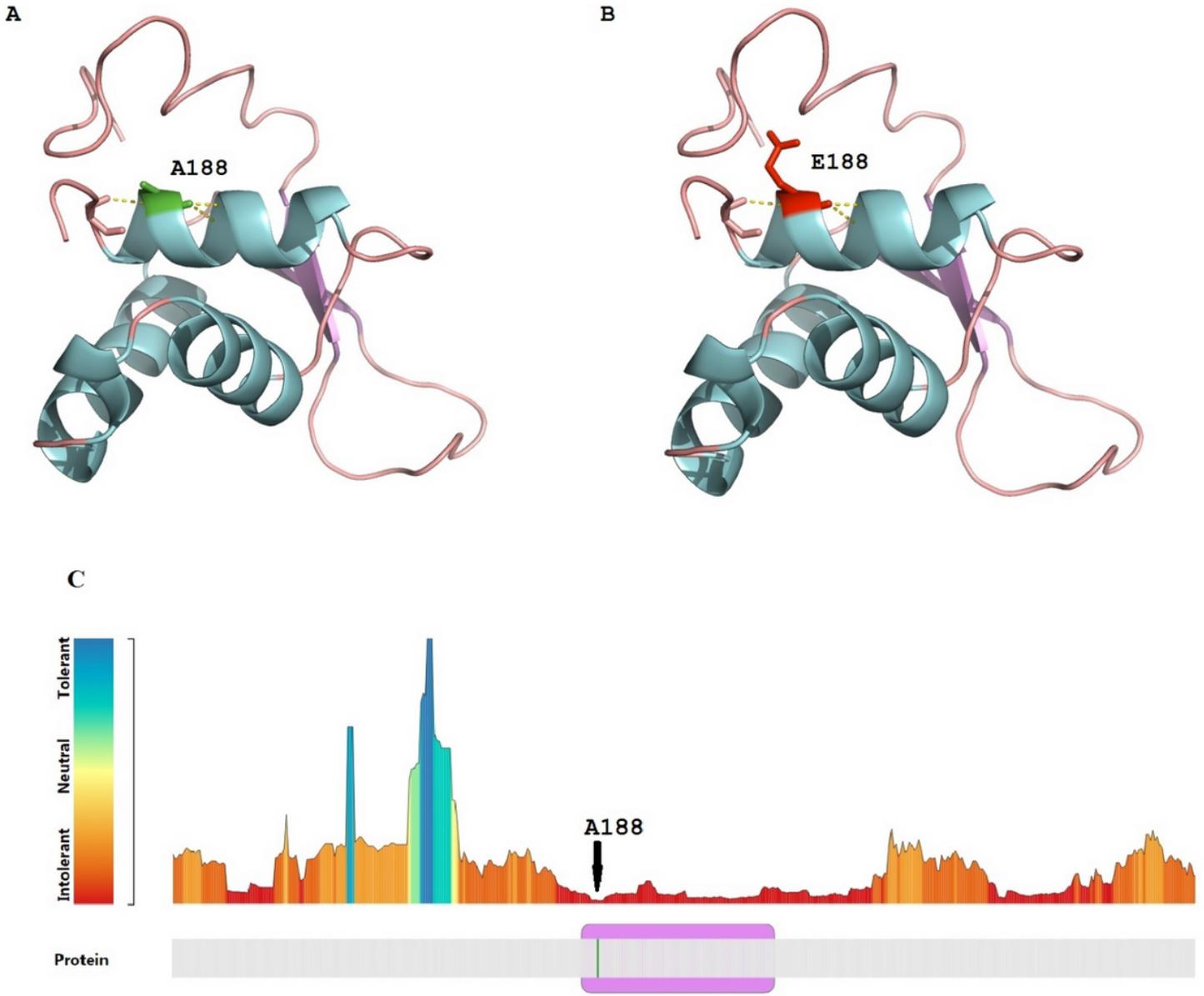


Figure 3

FOXG1; structure and location of *de novo* variant. The human model of the FOXG1 scheme was created from amino acid 168-280 (based on PDB 1vtn.1.C) using the SWISS model (with sequence identity 60.00% and coverage 0.86). 2A illustrates the overall cartoon structure of FOXG1 and the location of Alanine residue in green. 2B shows mutated residue Glutamic acid at position 188 in red. The bulky Glutamic acid residue in this position is likely to create spatial and steric hindrance leading to misfolding of the protein of FOXG1. 2C shows Ala 188 located in a highly intolerant region of FOXG1 based on the MetaDome web server.

Figure 4

Schematic representation of L211M substitution in the N-terminal of DMD. To 3D structure modeling of DMD, the first 227 amino acids at N-terminal utilized via Phyre 2 (86% of residues modeled at >90% confidence). PyMol molecular virtualization software [The PyMOL Molecular Graphics System, Version 1.7 Schrödinger, LLC] used for the assessment and visualization of the 3D protein structure. 3A represents the cartoon structure of wild-type DMD (in green) and mutated DMD (in cyan). It depicts that the Methionine residue in this position misfolds the DMD in the middle of the protein. A close-up view of the normal amino acid Leucine (blue) and mutated amino acid Methionine (red) at position 211 are shown in 3B. 3C demonstrates surface electrostaticity of the predicted wild-type DMD. In 4D, the Leu211Met mutation led to an alteration in the 3-dimensional structure, as well as the surface electrostatic potential of the mutated DMD.

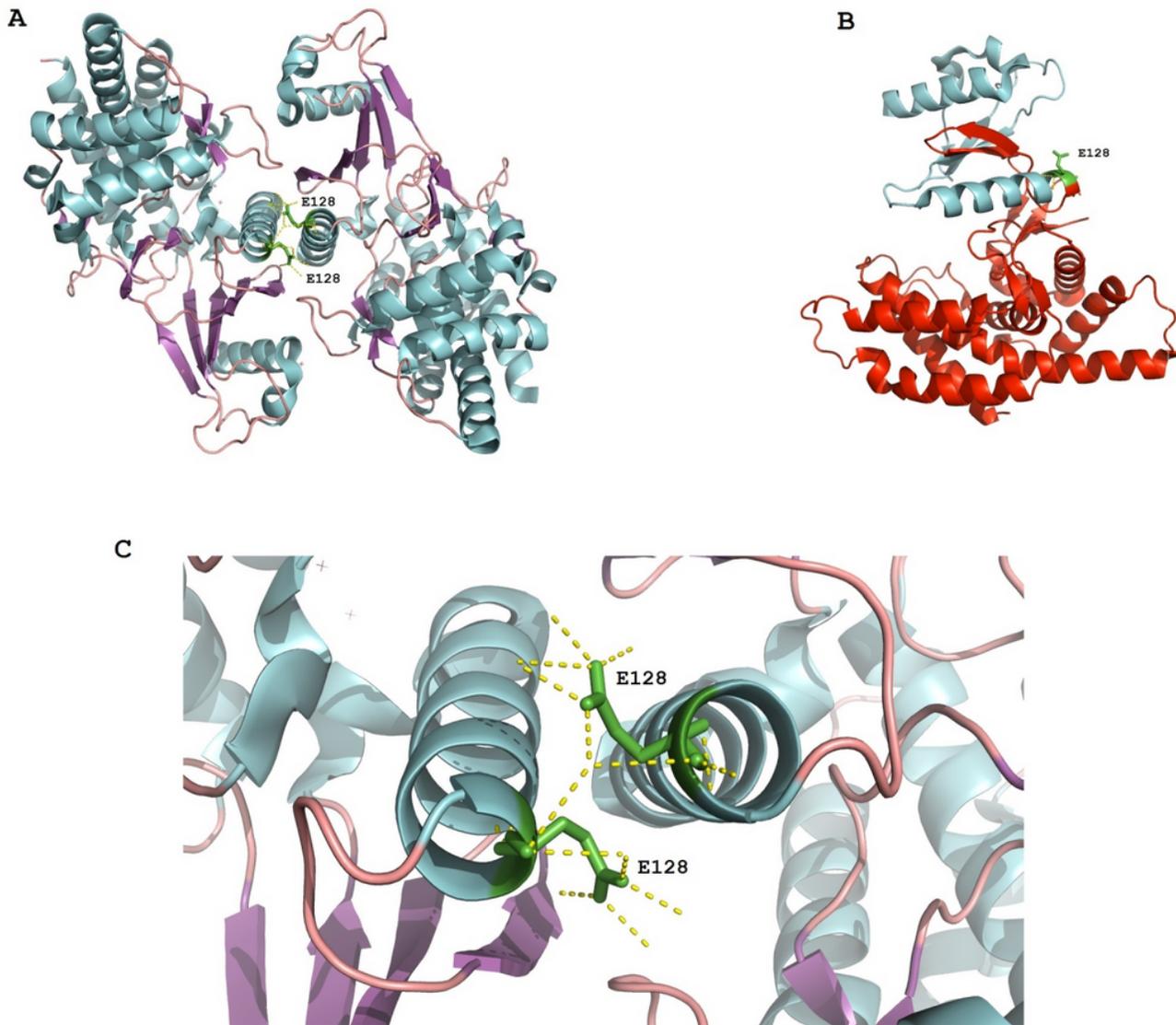


Figure 5

Schematic representation of the CHKB structure and location of glutamic acid at position 128. The human CHKB protein scheme illustrates Glutamic acid 128 (in green) as a negatively charged amino acid. This amino acid is essential for the dimerization of the CHKB monomer (4B). Figure 4A, C displays the dimer of CHKB and the crucial role of E128.

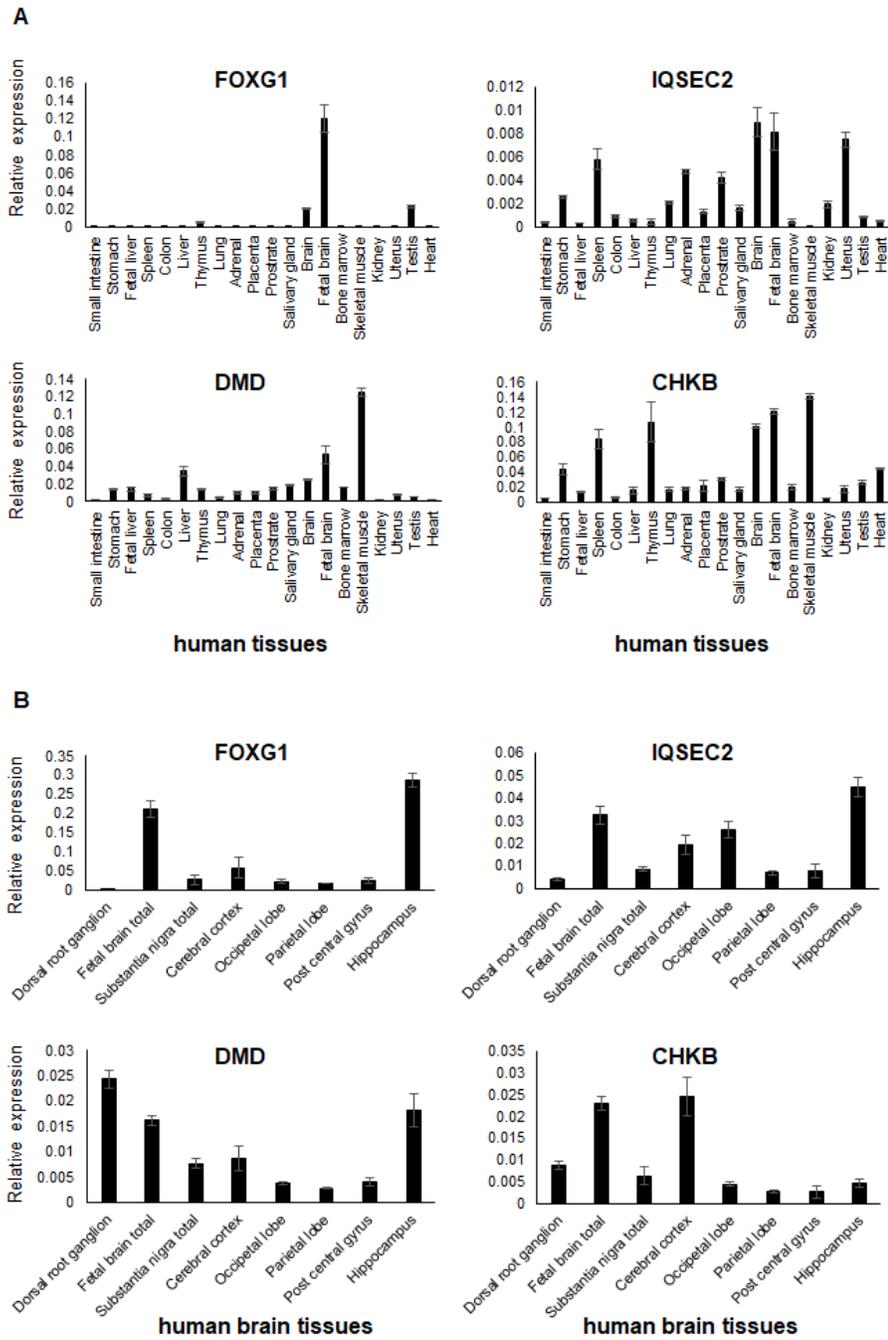


Figure 6

Transcript levels of *FOXG1*, *IQSEC2*, *DMD* and *CHKB* in various human tissues were determined by RT-qPCR. **(A)** Higher *FOXG1* and *IQSEC2* expression was detected in the fetal brain compared to other tissues, whereas *CHKB* and *DMD* expression in the skeletal muscles was highest. **(B)** All four genes *FOXG1*, *IQSEC2*, *DMD* and *CHKB* are expressed in the various parts of brain with varying levels.