Novel *MYO1D* Missense Variant Identified Through Whole Exome Sequencing and Computational Biology Analysis Expands the Spectrum of Causal Genes of Laterality Defects

Rabab Said Alsafwani[1†], Khalidah K. Nasser[1,2†], Thoraia Shinawi[1], Babajan Banaganapalli[2,3], Hanan Abdelhalim ElSokary[2], Zhaher F. Zaher[4,5], Noor Ahmad Shaik[2,3,6], Gaser Abdelmohsen[4,7], Jumana Yousuf Al-Aama[2,3], Adam J. Shapiro[8], Osman O. Al-Radi[9*], Ramu Elango[2,3*] and Turki Alahmadi[4,10*]

[1] Department of Medical Laboratory Technology, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, [2] Princess Al-Jawhara Center of Excellence in Research of Hereditary Disorders, King Abdulaziz University, Jeddah, Saudi Arabia, [3] Department of Genetic Medicine, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia, [4] Department of Pediatrics, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia, [5] Pediatric Cardiac Center of Excellence, King Abdulaziz University Hospital, King Abdulaziz University, Jeddah, Saudi Arabia, [6] Department of Genetics, Al Borg Medical Laboratories, Jeddah, Saudi Arabia, [7] Pediatric Cardiology Division, Department of Pediatrics, Cairo University, Kasr Al Ainy Faculty of Medicine, Cairo, Egypt, [8] Division of Pediatric Respiratory Medicine, McGill University Health Centre Research Institute, Montreal Children's Hospital, Montreal, QC, Canada, [9] Department of Surgery Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia, [10] Pediatric Department, Faculty of Medicine in Rabigh, King Abdulaziz University, Jeddah, Saudi Arabia

Laterality defects (LDs) or asymmetrically positioned organs are a group of rare developmental disorders caused by environmental and/or genetic factors. However, the exact molecular pathophysiology of LD is not yet fully characterised. In this context, studying Arab population presents an ideal opportunity to discover the novel molecular basis of diseases owing to the high rate of consanguinity and genetic disorders. Therefore, in the present study, we studied the molecular basis of LD in Arab patients, using next-generation sequencing method. We discovered an extremely rare novel missense variant in *MYO1D* gene (Pro765Ser) presenting with visceral heterotaxy and left isomerism with polysplenia syndrome. The proband in this index family has inherited this homozygous variant from her heterozygous parents following the autosomal recessive pattern. This is the first report to show *MYO1D* genetic variant causing left–right axis defects in humans, besides previous known evidence from zebrafish, frog and *Drosophila* models. Moreover, our multilevel bioinformatics-based structural (protein variant structural modelling, divergence, and stability) analysis has suggested that Ser765 causes minor structural drifts and stability changes, potentially affecting the biophysical

and functional properties of *MYO1D* protein like calmodulin binding and microfilament motor activities. Functional bioinformatics analysis has shown that *MYO1D* is ubiquitously expressed across several human tissues and is reported to induce severe phenotypes in knockout mouse models. In conclusion, our findings show the expanded genetic spectrum of LD, which could potentially pave way for the novel drug target identification and development of personalised medicine for high-risk families.

INTRODUCTION

Laterality defects (LDs) are a group of developmental diseases that affect internal organ positioning in the body. In general, human LDs can be divided into three categories: (1) *situs solitus* (SS) with normally expected organ arrangement; (2) *situs inversus* (SI) characterised by complete mirror image of organs; and (3) *situs ambiguus* (SA) with organ arrangement falling along a spectrum of various anomalies between SS and SI, including congenital heart defects (CHDs). Within SA, a subgroup of patients presents a severe and complex form of congenital heart disease, which is commonly known as *heterotaxy* (1). Defective left–right (LR) patterning of internal organs is associated with multiple congenital diseases affecting the cardiovascular system, kidneys, liver, and biliary tract (2, 3). According to the National Birth Defects Prevention Study (4), the estimated prevalence of LD is 1.1 per 10,000 in the United States. Despite the rare likelihood of LD, its incidence is excepted to be higher among the Arab population due to their high rate of consanguinity and genetic disorders (5).

The aetiology of LD is complex and includes both environment (5–7) and genetic factors (8, 9). Disease-causing genetic variations are found in <20% of LD cases; and the remaining 80% of cases are due to unidentifiable causes (10, 11). Up to now, known LD genes were mostly associated with NODAL/TGFβ signalling (*NODAL*, *CFC1*, *ACVR2B*, *LEFTYB*, *GDF1*, *TGFBR2*, and *FOXH1*), SHH signalling (*ZIC3* and *LZTFL1*), and monocilia function (*NPHP2*, *NPHP3*, *NPHP4*, *PKD2*, and *TTC8*) (10). Other genetic alterations associated with early cardiac development (*NKX2-5*, *CRELD1*, *MMP2*1, and *PKD1L1*) were also implicated in LD development (10). The main functional roles of these genes were demonstrated in LR axis determination, controlling cardiac looping direction, nodal activity regulation in embryogenesis, protein interaction of primary cilia, and signalling involved in morphogenesis cascade (12–19). Hence, LDs occur in a variety of different diseases, affecting various cardiac, respiratory, and gastrointestinal organs, reflecting the complex genes involved in signalling pathways of organogenesis and ciliary function.

Genetic testing and molecular diagnostics are now regarded as an useful approach to discover molecular causes underlying the LD development. Whole-exome sequencing (WES) analysis proved to be a successful method to uncover novel candidate genes or novel variants in known candidate genes (10). To this end, there is an increasing need to study the rare developmental disorders across different ethnic populations

due to its potential in expanding the genetic spectrum of the disease. However, literature searches reveal sparse data on the Arab LD patients. We hypothesise that genetically investigating LD patients from a consanguineous Arab society will offer new insights into disease pathogenesis by identifying novel genes or novel variants in known genes, as demonstrated in other complex diseases (20). Therefore, the objective of the present study was to identify the genetic cause of LD in Arab patients, using WES and multilevel bioinformatics-based structural (protein variant structural modelling, divergence, and stability) and functional (gene expression and knockout mouse model) analysis approaches.

MATERIALS AND METHODS

The ethical approval for the present study was obtained from the institutional ethics committee of King Abdulaziz University Hospital (KAUH), Jeddah, Saudi Arabia. Informed consent forms were collected from both adult parents and their children (parental consent and children's assent for those <18 years old) prior to blood sample collection and genetic testing. As per the National Birth Defects Prevention Study, LD participants were selected based on the following clinical criteria: *situs inversus*, CHDs (heterotaxy), isomerism of the lungs (bilateral two lobes/left-sidedness and bilateral three lobes/right-sidedness), abdominal situs abnormality (*abdominal situs inversus* and SA), and spleen abnormality (asplenia and polysplenia) (4). One LD index family composed of the proband (affected child) and both parents was recruited to paediatric cardiology, surgery, and pulmonology clinics in KAUH. Clinical, laboratory, and radiological results were independently assessed by both paediatric cardiology and pulmonology consultants. Family pedigree was drawn by interviewing the parents. Samples from two other LD families were screened for the presence of the identified candidate variants.

Molecular Testing
Clinical Sampling
A total of 5 ml of whole EDTA peripheral blood samples was collected from each study participant.

DNA Extraction
The genomic DNA was isolated from circulating lymphocytes using QIAamp DNA blood Kit as per the manufacturer's protocol and quantified using a NanoDrop 2000 spectrophotometer. DNA integrity was checked on 2% agarose gel electrophoresis.

FIGURE 1 | (A) Family pedigree chart of laterality defect (LD) Arab family. Proband is indicated by the arrow. The affected proband (shaded circle) is homozygous of c.2293C>T mutation in *MYO1D* gene. Both parents were consanguineous (double horizontal lines) and heterozygous carriers of the identified mutation. **(B)** Chest X-rays showing left isomerism (heterotaxy).
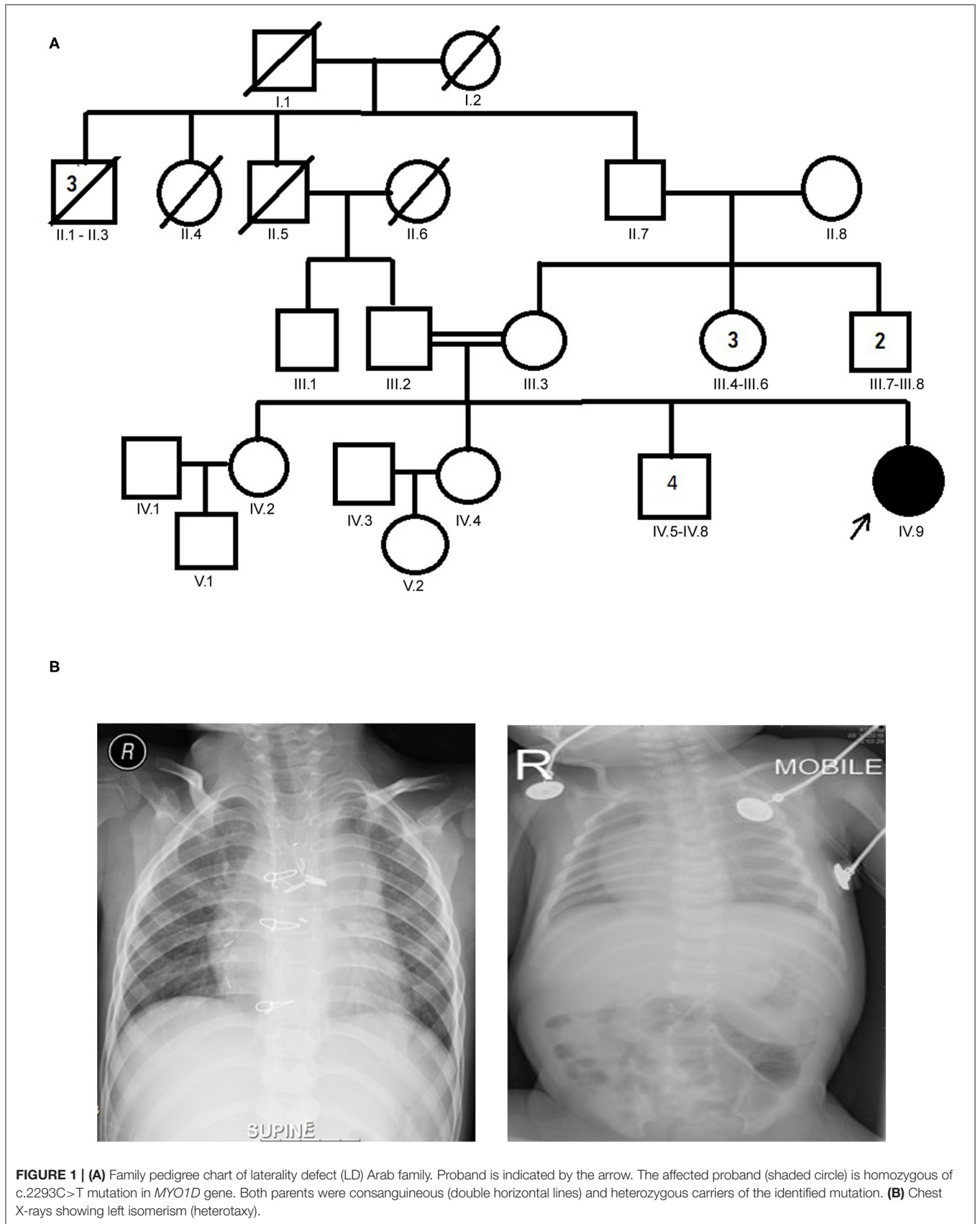
TABLE 1 | List of LD potential candidate variants that showed autosomal recessive inheritance pattern.

	Gene name	Chrom	ref	Alt	Effect	HGVS.c	HGVS.p	dbSNP151_ID	1000G_AF	ExAC_AF	gnomAD_exomes_AF
(1)	ZC3H12A	chr1	G	T	Missense	c.99G>T	p.Arg33Ser	rs116208741	0.00139776	0.002405	0.002499959
(2)	ZC3H12A	chr1	C	T	Missense	c.95C>T	p.Pro32Leu	rs115805535	0.00119808	0.002117	0.002117
(3)	TMEM18A	chr7	C	T	Missense	c.116G>A	p.Gly39Glu	rs183593116	0.00219649	0.004772	0.00515989
(4)	CPNE7	chr16	C	T	Missense	c.898C>T	p.Pro300Ser	rs150443459	0.000399361	0.0002966	0.0003931566
(5)	ASXL2	chr2	C	T	Missense	c.1489G>A	p.Ala497Thr	rs192716734	0.00139776	0.003805	0.004480134
(6)	MYO1D	chr17	G	A	Missense	c.2293C>T	p.Pro765Ser	rs7209106	0.0043	0.002759	0.002424917
(7)	EPB42	chr15	C	T	Missense	c.1477G>A	p.Gly493Ser	rs148871144	0.000199681	0.0007578	0.0007836991
(8)	FAM220A	chr7	G	A	Missense	c.437C>T	p.Pro146Leu	rs75910050	0.00139776	0.0009884	0.0009261591
(9)	OR10A2	chr11	C	A	Missense	c.741C>A	p.Phe247Leu	rs150322658	0.00119808	0.0014	0.001421724

LD, laterality defect.

TABLE 2 | Computational pathogenicity prediction scores of the LD candidate variants.

S. no	Variant	Gene	CADD	FATHMM	MetaLR	MutationTaster	PROVEAN	REVEL
(1)	rs115805535	ZC3H12A	16.06	0.22678	0.0431	0.08975	0.30964	0.027
(2)	rs116208741	ZC3H12A	–	–	–	–	–	–
(3)	rs150443459	CPNE7	–	–	–	–	–	–
(4)	rs192716734	ASXL2	14.94	0.18248	0.025	0.25126	0.24026	0.079
(5)	rs148871144	EPB42	0.135	0.78537	0.1928	0.08975	0.23156	0.121
(6)	rs75910050	FAM220A	0.524	0.06931	0.0102	0.08975	0.58248	0.049
(7)	rs150322658	OR10A2	–	–	–	–	–	–
(8)	rs183593116	TMEM184A	9.855	0.43279	0.0355	0.08975	0.07008	0.039
(9)	rs7209106	MYO1D	23.2	0.88298	0.6557	0.81001	0.80682	0.553

LD, laterality defect.
CADD: >25 = damaging; <25 = tolerated; FATHMM: >0.5 = damaging; <0.5 = tolerated; MetaLR: > 0.5 = damaging; <0.5 = tolerated; Mutation Tester: >0.5 = damaging; <0.5 = tolerated; PROVEAN: >0.5 = damaging; <0.5 = tolerated; REVEL: >0.5 = damaging; <0.5 = tolerated.

Whole-Exome Sequencing Analysis

The DNA library was prepared using Agilent Sure Select Target Enrichment Kit. DNA library was captured using ultralong 120 mer biotinylated cRNA baits. The library was sequenced using HiSeq2000 Next Generation Sequencer (Illumina, San Diego, CA, USA). The FASTQ format sequence was obtained, and reads were aligned using Burrows-Wheeler Aligner (BWA) software (Version bwa-0.7.12) against human genome reference sequence build 38 (GRCH38.p12). Variant calling was conducted using the genome analysis tool kit (GATK). The filtration pipeline was applied as follows: all coding variants that passed quality control (phred > 30 score) were included. All variants with a minor allele frequency (MAF) <0.015% were included. Known candidate gene variants were filtered based on the function of the gene and their role in LD development. Genes that are related to LD disease were collected through the Coremine Medical™ tool, National Center for Biotechnology Information (NCBI), OMIM, and literature review.

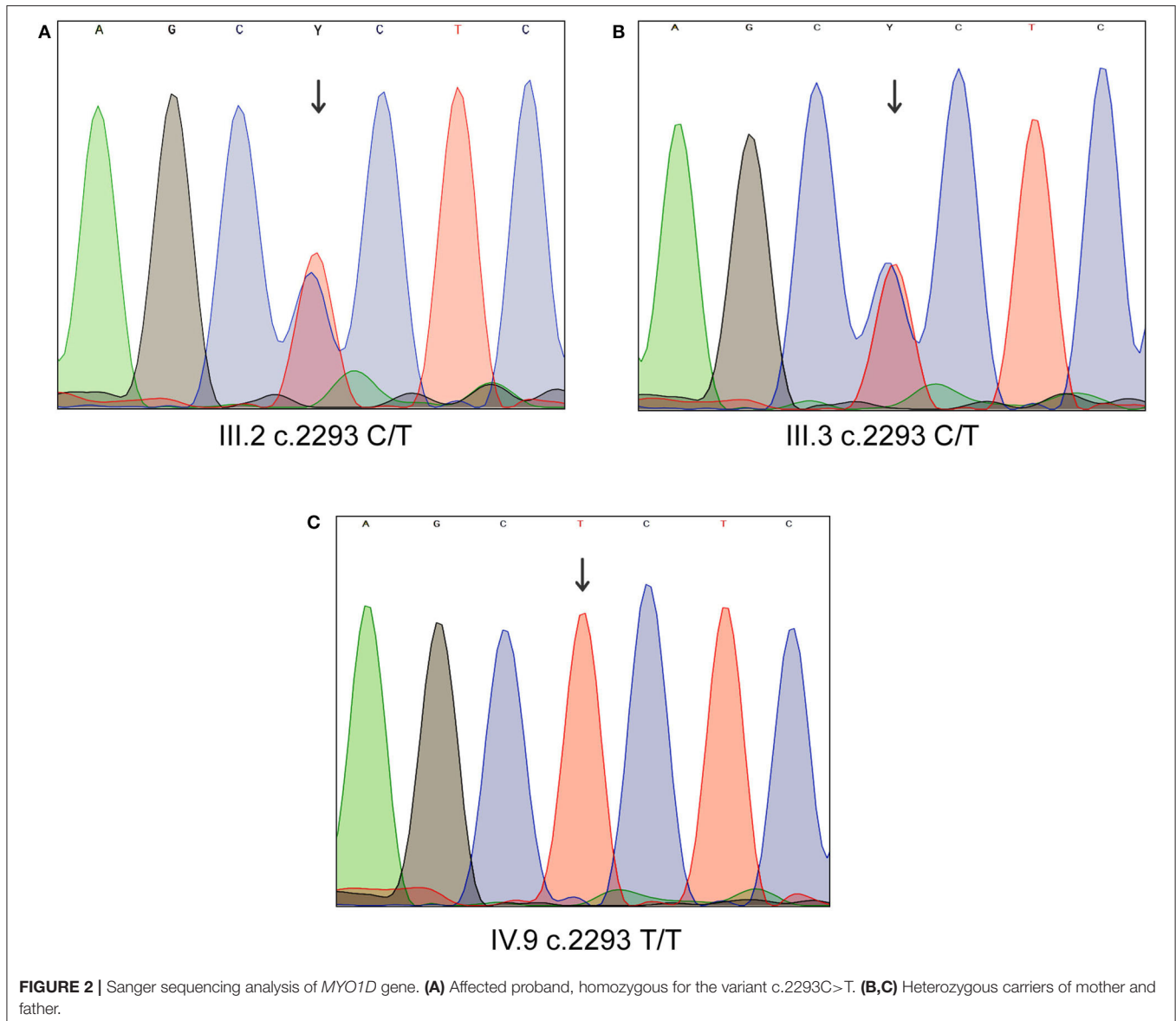## Variant Validation Using Sanger Sequencing Method

The potential LD candidate variant was validated using Sanger sequencer ABI 3500 Genetic Analyzer. The primers were designed using NCBI Primer Blast to capture targeted mutation in candidate gene. The sequence files (chromatogram) were analysed using BioEdit software.

Functional Analysis of Laterality Defect Variant Using Computational Methods
Amino Acid Conserved Domains

The functional relevance of LD candidate genetic variant on candidate proteins was predicted by searching the nucleotide and amino acid sequences against the functional domains of concerned protein as per the listing available in Conserved Domain Database (CDD). To estimate the sequence conservation characteristics of the functional domains in the candidate protein, CDD tool uses RPS-BLAST, which rapidly scans the query protein for pre-computed position-specific scoring matrices (PSSMs). The output file demonstrates the links between protein domains with annotations against the query input sequence together with imagining choices (21).

3D Protein Modelling

The pathogenic effects of amino acid variants on disease candidate proteins can be best understood when they are studied at the structural level. Therefore, the potential effect of LD variant on the tertiary structural features was explored through 3D simulation of the candidate protein. Based on the availability of the X-ray crystallographic structure of the query protein, either a combination of *ab initio* approaches or homology modelling approaches were followed (22, 23). The 3D simulated structure

FIGURE 2 | Sanger sequencing analysis of *MYO1D* gene. **(A)** Affected proband, homozygous for the variant c.2293C>T. **(B,C)** Heterozygous carriers of mother and father.

of the native protein was then used to construct the mutated version of candidate protein, which was then energy minimised and then analysed for structural deformities like amino acid or whole structure level deviations using YASARA software (24). The impact of candidate variant on the stability of protein structure was estimated using DUET webserver, which contains Protein Data Bank (PDB) structures of query proteins to predict the Gibbs free energy (G) values (25).

RNA Expression, Gene Ontology, and Mouse Gene Knockout Model

The Human Protein Atlas (HPA) (https://www.proteinatlas.org/) database was used to determine the RNA expression status of the LD candidate gene. This database provides the expression profile of the query gene or protein based on primary antibody staining data in a series of immunohistochemistry pictures of clinical specimens. The functional enrichment analysis of the potential LD candidate gene was done using gene ontology (GO) webtool hosted in Ensembl web browser. Moreover, Mouse Genome Informatics (MGI) database (http://www.informatics.jax.org/) was used to better understand the functional role of potential LD gene on phenotype characteristics of knockout mouse models. The MGI resource provides a comprehensive set of data, tools, and analysis designed specifically for use in mouse laboratory model. It accepts input data in the form of a gene symbol and provides output corresponding to the physiological condition of knockout mice.
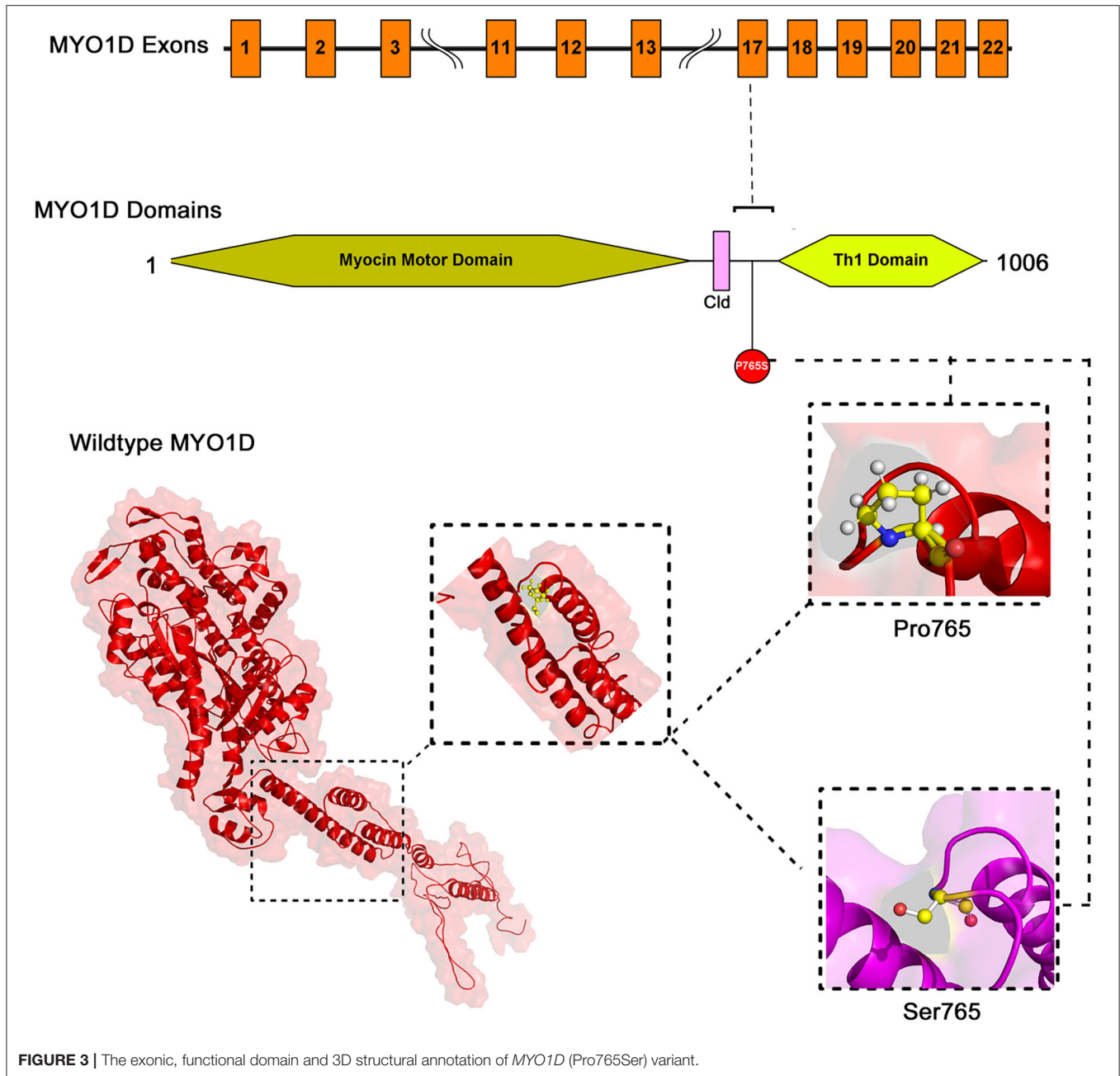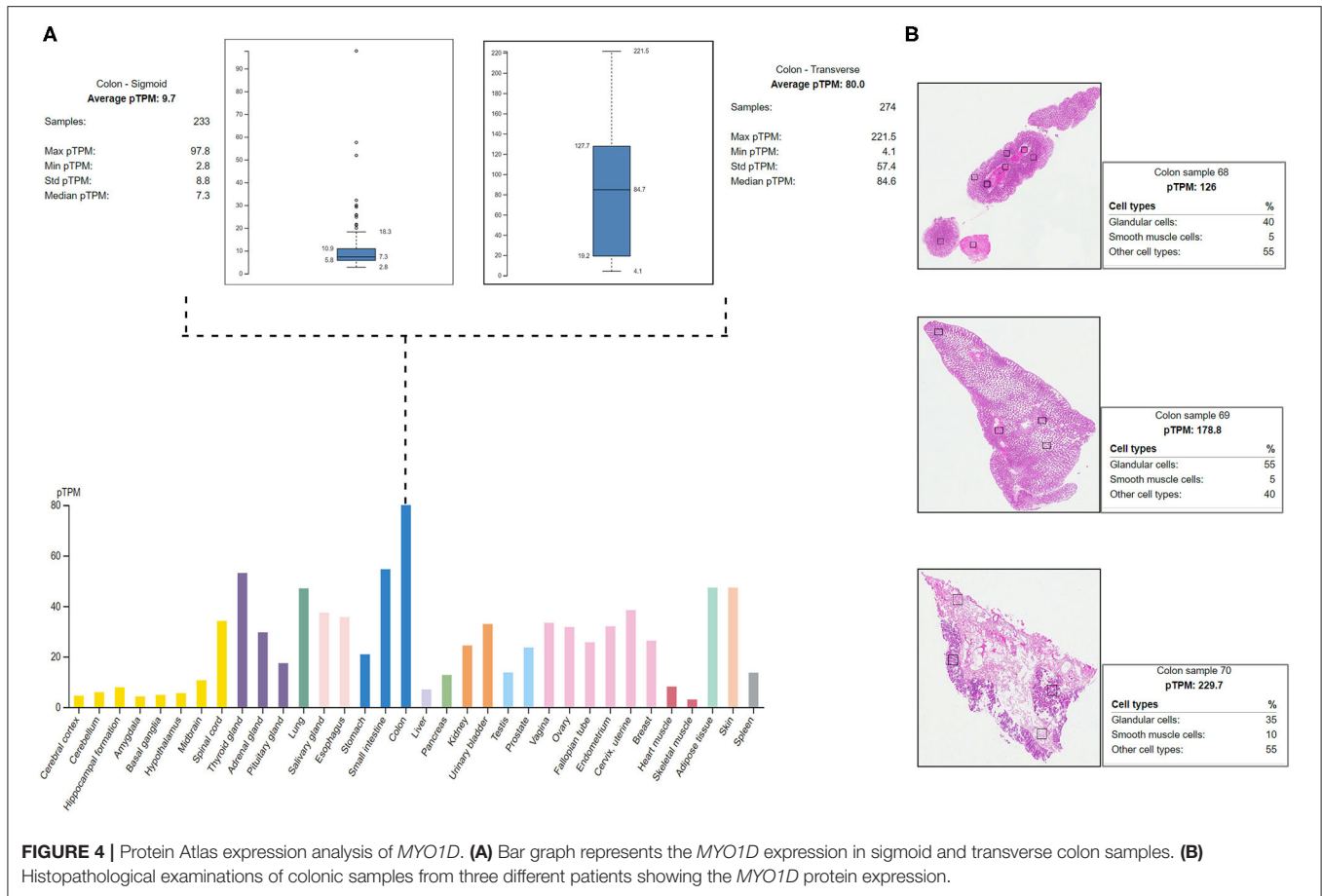
FIGURE 3 | The exonic, functional domain and 3D structural annotation of *MYO1D* (Pro765Ser) variant.

RESULTS

Clinical Assessment

The proband aged 4 years 6 months at the time of clinical diagnosis was born to an apparently healthy consanguineous parents of Arab origin (**Figure 1A**). The proband exhibited a spectrum of phenotypes including visceral heterotaxy (abnormal arrangements of thoracoabdominal organs) (**Figure 1B**), congenital cyanotic heart disease in the form of single ventricle physiology, left isomerism with polysplenia syndrome, double inlet atrioventricular connection (a heart defect that affects the valves and chambers), pulmonary atresia, interrupted inferior vena cava with absent supra-renal segment, and azygos continuation (a rare congenital abnormality often combined with cardiovascular and visceral malformations). At the age of 10 months, the proband underwent thorough palliative cardiac procedures in the form of ductal stenting in the neonatal period followed by Kawashima cavo-pulmonary shunt (a palliative surgical procedure performed in cases of left isomerism and azygos continuation of the inferior vena cava, and common atrioventricular valve with or without regurgitation and pulmonary stenosis) in addition to left pulmonary artery (LPA) balloon dilatation procedure. At 3 years of age, Fontan completion was performed *via* incorporation of hepatic veins to

FIGURE 4 | Protein Atlas expression analysis of *MYO1D*. **(A)** Bar graph represents the *MYO1D* expression in sigmoid and transverse colon samples. **(B)** Histopathological examinations of colonic samples from three different patients showing the *MYO1D* protein expression.
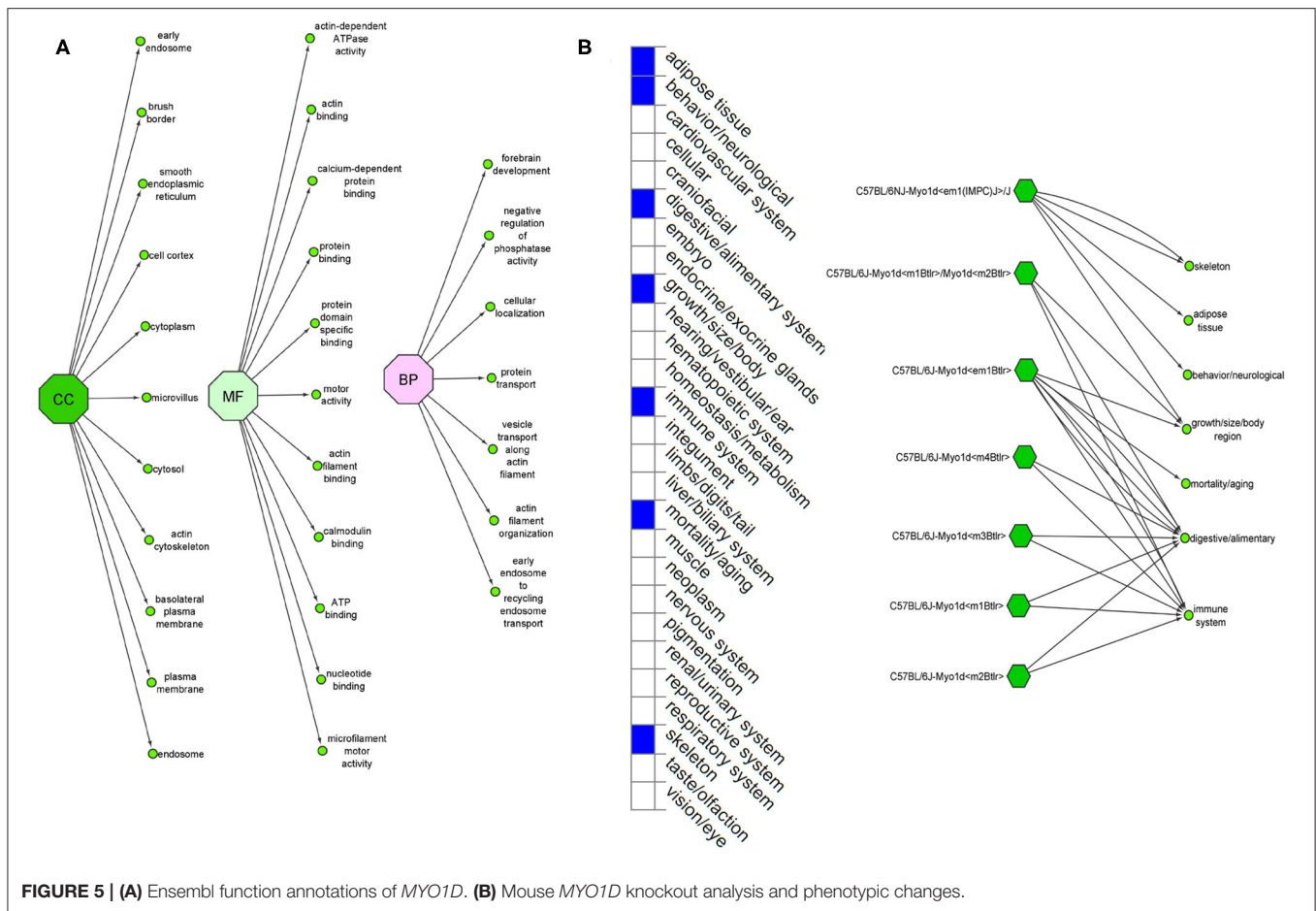
Genetic Analysis

Whole-Exome Sequencing Variant Filtering and Novel Gene Identification

The sequencing of the index case generated approximately 98,000 variants, including 12,150 synonymous variants, 13,000 missense variants, and 11,500 indels. Variant filtration was based on its rare frequency, deleterious potential, autosomal recessive mode of inheritance, and functional relevance to disease (LD, primary ciliary dyskinesia (PCD), congenital heart disease, and heterotaxy). Nine genetic variants were identified as potential candidates (**Table 1**). Among these variants, only one missense variant (rs7209106: NM_015194.2:c.2293C>T; p.Pro765Ser) in *MYO1D* novel gene has survived our variant filtration criteria. This allele is absent in local databases like GME (Greater Middle East) (http://igm.ucsd.edu/gme/), DALIA (Disease Alleles in Arabs) (http://clingen.igib.res.in/dalia/index), and Saudi Human Genome Program (SHGP) (https://shgp.kacst.edu.sa/index.en.html#home). The MAF of this variant in international databases like 1,000 Genomes and gnomAD databases is 0.005 and 0.002, respectively. Although it has an allele frequency of 0.013 in the African population, only eight individuals are

reported as homozygous for this variant in the gnomAD. But their clinical details are not provided in the gnomAD database. In the index family studied here, both parents were heterozygous and do not have any symptoms associated with LD, confirming the autosomal recessive inheritance pattern as we initially deduced from their pedigree analysis. Moreover, more than 80% (5/6; 83.34) of the computational prediction methods like CADD, FATHMM, MetaLR, Mutation Taster, PROVEAN, and REVEL have attributed pathogenicity scores to this *MYO1D* (p.Pro765Ser) variant (**Table 2**). Functional biology data available from model organisms like *Drosophila*, zebrafish, and frog have proved the functional role of *MYO1D* gene in LDs.

Sanger Sequencing Validation

Sanger sequencing analysis confirmed that the LD patient is homozygous for c.2293C>T variant in *MYO1D* gene (1V.9, **Figures 1A**, **2**), whereas the mother and father were heterozygous carriers (111.2, 111.3, **Figures 1A**, **2**). This variant was absent in apparently healthy siblings and were homozygous for the T allele (1V.2, 1V.4, and 1V.5–8; **Figure 1A**). Two additional clinically diagnosed LD families were screened for this variant, and none carries this mutation, suggesting that *MYO1D* (p.Pro765Ser) variant is a rare private mutation in this family.

The pulmonary artery correcting thereby blood flow from the lower body parts directly to the lungs.

FIGURE 5 | (A) Ensembl function annotations of *MYO1D*. (B) Mouse *MYO1D* knockout analysis and phenotypic changes.

Computational Functional Analysis

Variant Mapping on MYOID1 Protein Domain

The mapping of conserved amino acid domains is a vital step in deducing the association between the nucleotide sequence, protein structure, and function of disease-causing proteins. The CDD analysis showed that *MYO1D* protein is made up of three domains, namely, motor (11–682 amino acids), IQ calmodulin-binding motif (699–719 amino acids), and *Myosin TH1* (803–1,000 amino acids) domains. The Pro765Ser variant is located between the *Myosin TH1* and calmodulin-binding domains (**Figure 3**).

MYO1D 3D Model Construction

The PDB database search revealed the availability of partial 707 aa (between 10 and 717 of 1,006 AA long *MYO1D* protein) X-ray crystal protein model (4L79) with 2.3-Å resolution. Hence, the remaining 306 aa long chain was simulated with iterative threading assembly modification (I-TASSER) webserver following an *ab initio* approach. From the I-TASSER output, the best *MYO1D* model was chosen based on its polypeptide prediction quality scores like confidence score (C = −1.52), template modelling (TM = 0.53 ± 0.15), and root-mean-square deviation (RMSD) (12.7 ± 4.3 Å) scores. These quality metrics indicate the very good structural similarity between the query

and template proteins (**Figure 3**). The stereochemical evaluation of the energy-minimised *MYO1D* protein model revealed that 96.2% of the amino acids are in the allowed portion of the protein, whereas 3.8% are in the non-allowed region. As per the above outlined processes, the native MYO1D model was used as a template to create a mutant variant by manually substituting proline for serine at the 765th position. PyMOL was used to depict native and mutant proteins.

Structural Deviation and Stability Findings

We have used YASARA tool to analyze Cα-atom coordinates of native and mutant *MYO1D* 3D structures to evaluate their structural drifts (in terms of RMSD) at residue and whole structure levels. RMSD value is used to quantitatively measure the structural similarity between two atomic coordinates when superimposed on each other. The impact of substitution mutations on amino acid structures can be calculated when there is a divergence at the polypeptide chain level. We noticed minor structural drifts in *MYO1D* structure only at 765th residue position due to the RMSD value difference (2.28) induced by the substitution of proline with serine (**Figure 3**). The DUET analysis of the *MYO1D* (P765S) variant predicted Gibbs free energy ($\Delta\Delta G$) alterations shifting the energy equilibrium to negative value, i.e., −0.959 kcal/mol, suggesting that the queried

TABLE 3 | Phenotypes and genetic data of LD and/or PCD among Arabs.

Nationality	Situs inversus (SI)	PCD	Dextrocardia	Heterotaxy of abdominal organs	Recurrent chest infections	Chronic rhinitis	Bronchiectasis	Chronic or recurrent otitis media	Gene	Reference
Yemeni	+	–	+	Visceral heterotaxy, polysplenia syndrome	–	–	–	–	MYO1D	Present study
Saudi (6 patients)	+6/6	–	NR	NR	–	NR	NR	NR	DNAH5, PKD1L1, DNAAF5/CYP21A2, DNAI1	(26)
Arab	+	NR	+	NR	NR	NR	NR	NR	GDF1	(27)
Saudi	+	NR	+	Abdominal ultrasound revealed that the liver and gallbladder were located in the left hypochondrium, spleen on the right side	NR	NR	NR	NR	NR	(28)
Arab	+	NR	+	+	NR	NR	NR	NR	HES7	(29)
Saudi (75 patients)	+	+73/75	NR	NR	+15/75	NR	NR	NR	CCDC151, CCDC39, CCDC40, DNAH11, GOLGA3, RSPH9, CCNO, RSPH9, ITCH, MCIDAS, RSPH4A, DNAH5, CEP164, GOLGA3	(26)
Morocco	+	+	+	+	+	+	+	–	NR	(30)
Saudi	+	+	+	NR	NR	+	+	+	DNAH1	(31)
Kuwaiti	+	+		NR	NR	+	NR	NR	DNAH5	(32),
Saudi	NR	+	NR	NR	+	+	+	NR	CCDC151	(33)
Palestine	NR	+	+	NR	NR	NR	+	+	LRRC6	(34)
Saudi	NR	+	NR	NR	NR	NR	+	NR	RSPH9	(35)
UAE	NR	+	NR	NR	+	NR	+	NR	RSPH9	(36)
UAE	–	+	NR	NR	+	+	+	+	RSPH9	(37)
Saudi	+	+	NR	NR	NR	NR	+	NR	19q13.3	(27)
Saudi	+	+	–	Situs inversus of cardiac shadow and gastric air bubble	+	NR	+	–	NR	(38)

+, presence of symptoms; –, absence of symptoms; NR, not reported; PCD, primary ciliary dyskinesia.

variant is potentially deleterious to the protein stability owing to its destabilising behaviour.

RNA Expression Analysis

The HPA shows the positive expression status of *MYO1D* gene in different tissues and organs of the human body like the colon, lungs, and thyroid gland. In particular, the highest expression was seen in the digestive system with the colon, where the 274 transverse colon samples showed a maximum of 221.5 protein transcripts per million (pTPM) and 233 sigmoid colon samples showed a maximum of 97.8 pTPM. The RNA-Seq analysis of immunohistochemistry tissue specimens from three control specimens showed that glandular cells showed the highest pTPM status of *MYO1D* gene when compared with smooth muscle cells and other cell types in the colon (**Figure 4**).

Gene Ontology and Gene Knockout Analysis

The Ensembl GO analysis of *MYO1D* gene showed its involvement in 43 GO terms, including those connected to biological processes (seven GO terms), molecular functions (11 GO terms), and cellular component (25 GO terms) (**Supplementary Table 1**). All the annotations collectively highlight that *MYO1D* is localised in the cytoplasm (cellular component) and plays an important role in actin filament organisation (biological process) as well as microfilament motor activity function (molecular function). **Supplementary Table 2** shows details of disease phenotypes corresponding to the *MYO1D* genetic background in different knockout mice models.

DISCUSSION

Recent evolution and easy accessibility of next-generation sequencing resulted in accurate molecular diagnosis of variety of genetic diseases from around the globe (39, 40). Owing to the genetic heterogeneity, identification of specific molecular cause of LD is very challenging in up to 80% of the cases. Majority of known LD causative genes are structural proteins of the cilia and are known for their involvement in NODAL/TGFβ or SHH signalling pathways. During embryonic development, these genes play an important role in symmetrical LR positioning of the organs (10).

MYO1D gene consists of 27 exons mapped to chromosome 17q11.2. Myosin heavy chain class 1 is a member of the myosin superfamily, playing essential roles in cytoskeletal structure, mechanical signal-transduction membrane dynamics (41), and endosome processing (42). In the present study, we identified the first case of a homozygous missense (c.2293C>T) mutation in *MYO1D* gene causing LD in the proband of an Arab consanguineous family. Further screening of LD participants from two additional Arab families did not reveal any mutations in this gene. In Middle Eastern Arab databases (with more than 10,000 exome data combined) such as SHGP, GMC, and DALIA, not a single case was recorded for this variant. Also, this variant was extremely rare (<0.005) in 1,000 Genomes and 0.013 in gnomAD across all ethnic groups. In gnomAD, only eight individuals were reported as homozygous (six males and two females), but no clinical data were available.

Various studies suggested the role of *MYO1D* gene in laterality disease in *Drosophila*, zebrafish, and frog (43–47). *MYO1D* has a role in organ asymmetry in *Drosophila*, which lacks cilia and nodal pathway while developing LD by using polar cell polarity (PCP), *MYO1D*, and *HOX* gene Abd-B (48, 49). Another study demonstrated the function of *MYO1D* in *Xenopus laevis* and influence the orientation of the cilia on the LR organiser (LRO) through planar cell polarity pathway as implicated in *Drosophila* (45). In zebrafish, *MYO1D* plays a fundamental role in the LR organisation (43, 50). Aside from the above initial reports, our understanding of *MYO1D* function in the context of human LR patterning remains largely unexplored. The *situs inversus* (SI) phenotype is reported in approximately 50% of PCD cases with congenital cardiac defects (51, 52). Approximately 3–7% of LD patients have CHDs (53). Many studies (54, 55) in a variety of species failed to identify a unifying mechanism for LR patterning. However, recent studies (43, 50) provided the first evidence of a shared origin of laterality in both arthropods and chordates through *MYO1D* gene. The clinical features are consistent with previous observations in zebrafish and *Drosophila*, indicating that MYO1D has an important role in LR patterning during embryogenesis. Moreover, asymmetric clustering of cilia was disrupted in ependymal cells of *MYO1D* KO rat models, consistent with LD (56).

MYO1D gene knockout in different mouse strains (seven different strains) is shown to demonstrate a variety of phenotypes like decreased body fat amount (adipose tissue), decreased startle reflex (behaviour/neurological), increased susceptibility to colitis (digestive/alimentary), decreased bone mineral content (skeleton), and increased susceptibility to weight loss (growth size/body weight and immune system and mortality or ageing) (**Figure 5**). Our study is the first one to report the association of defective *MYO1D* to LD in humans, confirming that the function is evolutionarily conserved from *Drosophila*, to zebrafish, to frog, to humans. Thus, *MYO1D* gene can be considered as the new causal gene for LD in humans. Though the *Drosophila* and zebrafish models clearly showed the visceral heterotaxy, mouse KO phenotypes were surprisingly not showing any LD-related phenotypes.

Extensive computational analysis of the protein structure and function adds the supporting evidence for *MYO1D* in LD. *MYO1D* protein is 1,006 aa long with a molecular weight of 116 kDa. It consists of a large, highly conserved Myocin Motor Domain (671 aa), short calmodulin-binding motif (20 aa), and a basic C-terminal tail homology-1 (TH1) domain (197 aa). The amino acid residue level structural deviation observed with the variant serine (Pro765Ser) in *MYO1D* is likely to disturb the primary, secondary, tertiary, and quaternary structural features in the protein. Numerous studies have shown the strong correlation between deviations in residue level RMSD score and structural properties for the disease-causing variants (23, 57–59). Disease causative pathogenic mutations have often changed the energy equilibrium, which is required to maintain the protein stability (60). Given the close physical proximity of P765S variant between calmodulin-binding motif and TH1 domains, the conformational and stability changes in *MYO1D* protein are likely to impact its main biological functions such as calmodulin

binding, actin-dependent ATPase activity, calcium-dependent protein binding, and microfilament motor activities (61).

LD is a complex disease, and its clinical phenotype presentations often overlap with PCD symptoms. A recent study from Saudi Arabia reported the overlapping clinical symptoms between PCD and LD patients (26). This report investigated a total of 81 patients, including 58 patients with sinopulmonary infections (SPIs), 15 patients with combined LD with SPIs, and six patients with LD alone. They reported mutations in the known PCD genes as follows: *RSPH9*, *CCNO*, DNAAF5, *RSPH4A*, *MCIDAS*, and *CCDC40* gene mutations in PCD patients with SPIs; *CCDC151*, *DNAH11*, *CCDC40*, *DNAH5*, and *CCDC39* gene mutations in LD patients with SPIs; *PKD1L1* and *DNAAF5* gene mutations in LD patients; and *RSPH9* and *MCIDAS* gene mutations in neonatal respiratory distress. Additionally, they have also identified gene mutations in *ITCH* and *CEP164* in two patients, demonstrating ITCH-related syndrome and Bardet–Biedl syndrome. Sparrow et al. (29) reported *HES7* as a cause of spondylocostal dysostosis with SI and dextrocardia. Molecular diagnosis of LD and PCD in Arab patients has revealed a spectrum of mutations in many genes with variable clinical presentations (35), which is summarised in **Table 3**.

CONCLUSION

In conclusion, we discovered missense mutation in *MYO1D* gene (c.2293C>T) in an Arab patient presenting with visceral heterotaxy and left isomerism with polysplenia syndrome by using higher-throughput WES technology. This is the first report to establish the relationship between *MYO1D* variants and LD, supporting the previous findings in *Drosophila zebrafish*, and frog. This exciting finding may support the critical role of *MYO1D* gene for LR patterning in humans. This study has some sincere limitations, as this is the first case identified with *MYO1D* mutation potentially contributing to LD phenotypes, and there are no reported cases with *MYO1D* variants to compare our data with. Therefore, testing *MYO1D* variants for LD patients in large cohort studies is recommended to verify our findings. Future functional studies are also recommended to investigate the specific molecular role and therapeutic prospects of targeting *MYO1D* genetic variants in patients demonstrating LD phenotypes.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because (a) participants' refusal to store or distribute the genomic data in the public domain and (b) as per the local Institutional Ethics committee approval and Saudi national policy on genomic data sharing in the public domain outside the country. Requests to access the data should be directed to RE or TS.

ETHICS STATEMENT

The Ethical approval for the present study was obtained from institutional Ethics Committee of King Abdulaziz University Hospital (KAUH), Jeddah, Saudi Arabia and informed consent forms were collected from both adult parents and their children (guardian consent for Language evaluation).

CONSENT TO PARTICIPATE

Informed consent was obtained from all subjects involved in the study.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2021.724826/full#supplementary-material

REFERENCES

1. Shapiro AJ, Davis SD, Ferkol T, Dell SD, Rosenfeld M, Olivier KN, et al. Laterality defects other than situs inversus totalis in primary ciliary dyskinesia: insights into situs ambiguus and heterotaxy. *Chest.* (2014) 146:1176–86. doi: 10.1378/chest.13-1704

2. Kosaki K, Casey B. Genetics of human left-right axis malformations. *Semin Cell Dev Biol.* (1998) 9:89–99. doi: 10.1006/scdb.1997.0187

3. Lopez KN, Li AH, Hanchard N, Azamian M, Lalani S, Dickerson H, et al. Whole exome sequencing in congenital heart disease reveals variants in left-right patterning genes previously associated with heterotaxy syndrome and primary ciliary dyskinesia. *J Am Coll Cardiol.* (2015) 65(10 Suppl.):A490. doi: 10.1016/S0735-1097(15)60490-9

4. Lin AE, Krikov S, Riehle-Colarusso T, Frías JL, Belmont J, Anderka M, et al. Laterality defects in the national birth defects prevention study (1998-2007): birth prevalence and descriptive epidemiology. *Am J Med Genet A.* (2014) 164A:2581–91. doi: 10.1002/ajmg.a.36695

5. Best S, Shoemark A, Rubbo B, Patel MP, Fassad MR, Dixon M, et al. Risk factors for situs defects and congenital heart disease in primary ciliary dyskinesia. *Thorax.* (2019) 74:203–5. doi: 10.1136/thoraxjnl-2018-212104

6. Kuehl KS, Loffredo C. Risk factors for heart disease associated with abnormal sidedness. *Teratology.* (2002) 66:242–8. doi: 10.1002/tera.10099

7. Kuehl KS, Loffredo CA. Population-based study of l-transposition of the great arteries: possible associations with environmental factors. *Birth Defects Res A Clin Mol Teratol.* (2003) 67:162–7. doi: 10.1002/bdra.10015

8. Sutherland MJ, Ware SM. Disorders of left-right asymmetry: heterotaxy and situs inversus. *Am J Med Genet C.* (2009) 151C:307–17. doi: 10.1002/ajmg.c.30228

9. Deng H, Xia H, Deng S. Genetic basis of human left-right asymmetry disorders. *Expert Rev Mol Med.* (2015) 16:e19. doi: 10.1017/erm.2014.22

10. Li AH, Hanchard NA, Azamian M, D'alessandro LCA, Coban-Akdemir Z, Lopez KN, et al. Genetic architecture of laterality defects revealed by whole exome sequencing. *Eur J Hum Genet.* (2019) 27:563–73. doi: 10.1038/s41431-018-0307-z

11. Ware SM, Jefferies JL. New genetic insights into congenital heart disease. *J Clin Exp Cardiol.* (2012) S8:003. doi: 10.4172/2155-9880.S8-003

12. Collignon J, Varlet I, Robertson EJ. Relationship between asymmetric nodal expression and the direction of embryonic turning. *Nature.* (1996) 381:155–8. doi: 10.1038/381155a0

13. Olson EN, Srivastava D. Molecular pathways controlling heart development. *Science.* (1996) 272:671–6. doi: 10.1126/science.272.5262.671

14. Iratni R, Yan YT, Chen C, Ding J, Zhang Y, Price SM, et al. Inhibition of excess nodal signaling during mouse gastrulation by the transcriptional corepressor DRAP1. *Science.* (2002) 298:1996–9. doi: 10.1126/science.1073405

15. Otto EA, Schermer B, Obara T, O'toole JF, Hiller KS, Mueller AM, et al. Mutations in INVS encoding inversin cause nephronophthisis type 2, linking renal cystic disease to the function of primary cilia and left-right axis determination. *Nat Genet.* (2003) 34:413–20. doi: 10.1038/ng1217

16. Qian F, Germino FJ, Cai Y, Zhang X, Somlo S, Germino GG. PKD1 interacts with PKD2 through a probable coiled-coil domain. *Nat Genet.* (1997) 16:179–83. doi: 10.1038/ng0697-179

17. Yoshiba S, Shiratori H, Kuo IY, Kawasumi A, Shinohara K, Nonaka S, et al. Cilia at the node of mouse embryos sense fluid flow for left-right determination via Pkd2. *Science.* (2012) 338:226–31. doi: 10.1126/science.1222538

18. Calvet JP. Ciliary signaling goes down the tubes. *Nat Genet.* (2003) 33:113–4. doi: 10.1038/ng1078

19. Grimes DT, Keynton JL, Buenavista MT, Jin X, Patel SH, Kyosuke S, et al. Genetic analysis reveals a hierarchy of interactions between polycystin-encoding genes and genes controlling cilia function during left-right determination. *PLoS Genet.* (2016) 12:e1006070. doi: 10.1371/journal.pgen.1006070

20. Saadah OI, Banaganapalli B, Kamal NM, Sahly AN, Alsufyani HA, Mohammed A, et al. Identification of a rare exon 19 skipping mutation in ALMS1 gene in alström syndrome patients from two unrelated saudi families. *Front Pediatr.* (2021) 9:652011. doi: 10.3389/fped.2021.652011

21. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, Deweese-Scott C, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* (2010) 39:D225–D9. doi: 10.1093/nar/gkq1189

22. Ahmed Awan Z, Bima A, Rashidi OM, Jamil K, Khan IA, Almukadi HS, et al. Low resolution protein mapping and KB-R7943 drug-protein molecular interaction analysis of long-QT syndrome linked KCNH2 mutations. *All Life.* (2020) 13:183–93. doi: 10.1080/26895293.2020.1737249

23. Awan ZA, Bahattab R, Kutbi HI, Noor AOJ, Al-Nasser MS, Shaik NA, et al. Structural and molecular interaction studies on familial

hypercholesterolemia causative PCSK9 functional domain mutations reveals binding affinity alterations with LDLR. *Int J Pept Res Ther.* (2021) 27:719–33. doi: 10.1007/s10989-020-10121-8

24. Shaik NA, Bokhari HA, Masoodi TA, Shetty PJ, Ajabnoor GMA, Elango R, et al. Molecular modelling and dynamics of CA2 missense mutations causative to carbonic anhydrase 2 deficiency syndrome. *J Biomol Struct Dyn.* (2020) 38:4067–80. doi: 10.1080/07391102.2019.1671899

25. Pires DE, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res.* (2014) 42:W314–319. doi: 10.1093/nar/gku411

26. Shamseldin HE, Al Mogarri I, Alqwaiee MM, Alharbi AS, Baqais K, Alsaadi M, et al. An exome-first approach to aid in the diagnosis of primary ciliary dyskinesia. *Hum Genet.* (2020) 139:1273–83. doi: 10.1007/s00439-020-02170-2

27. Marek-Yagel D, Bolkier Y, Barel O, Vardi A, Mishali D, Katz U, et al. A founder truncating variant in GDF1 causes autosomal-recessive right isomerism and associated congenital heart defects in multiplex Arab kindreds. *Am J Med Genet A.* (2020) 182:987–93. doi: 10.1002/ajmg.a.61509

28. Saleh A. Dextrocardia with situs inversus totalis in a boy: a case report. *Int J Contemp Pediatr.* (2016) 1096–11098. doi: 10.18203/2349-3291.ijcp20162398

29. Sparrow DB, Faqeih EA, Sallout B, Alswaid A, Ababneh F, Al-Sayed M, et al. Mutation of HES7 in a large extended family with spondylocostal dysostosis and dextrocardia with situs inversus. *Am J Med Genet A.* (2013) 161A:2244–9. doi: 10.1002/ajmg.a.36073

30. Raoufi M, Sator H, Lahma J, El Ayoubi A, Nitassi S, Oujilal A, et al. A case of Kartagener syndrome with rhinolalia clausa. *Pan Afr Med J.* (2016) 23:159. doi: 10.11604/pamj.2016.23.159.8664

31. Imtiaz F, Allam R, Ramzan K, Al-Sayed M. Variation in DNAH1 may contribute to primary ciliary dyskinesia. *BMC Med Genet.* (2015) 16:14. doi: 10.1186/s12881-015-0162-5

32. Marafie MJ, Al Suliman IS, Redha AM, Alshati AM. Primary ciliary dyskinesia: Kartagener syndrome in a family with a novel DNAH5 gene mutation and variable phenotypes. *Egypt J Med Hum Genet.* (2015) 16:95–9. doi: 10.1016/j.ejmhg.2014.08.001

33. Hjeij R, Onoufriadis A, Watson CM, Slagle CE, Klena NT, Dougherty GW, et al. CCDC151 mutations cause primary ciliary dyskinesia by disruption of the outer dynein arm docking complex formation. *Am J Hum Genet.* (2014) 95:257–74. doi: 10.1016/j.ajhg.2014.08.005

34. Horani A, Ferkol TW, Shoseyov D, Wasserman MG, Oren YS, Kerem B, et al. LRRC6 mutation causes primary ciliary dyskinesia with dynein arm defects. *PLoS One.* (2013) 8:e59436. doi: 10.1371/journal.pone.0059436

35. Alsaadi MM, Gaunt TR, Boustred CR, Guthrie PA, Liu X, Lenzi L, et al. From a single whole exome read to notions of clinical screening: primary ciliary dyskinesia and RSPH9 p.Lys268del in the Arabian Peninsula. *Ann Hum Genet.* (2012) 76:211–20. doi: 10.1111/j.1469-1809.2012.00704.x

36. Reish O, Slatkin M, Chapman-Shimshoni D, Elizur A, Chioza B, Castleman V, et al. Founder mutation(s) in the RSPH9 gene leading to primary ciliary dyskinesia in two inbred Bedouin families. *Ann Hum Genet.* (2010) 74:117–25. doi: 10.1111/j.1469-1809.2009.00559.x

37. Castleman VH, Romio L, Chodhari R, Hirst RA, De Castro SC, Parker KA, et al. Mutations in radial spoke head protein genes RSPH9 and RSPH4A cause primary ciliary dyskinesia with central-microtubular-pair abnormalities. *Am J Hum Genet.* (2009) 84:197–209. doi: 10.1016/j.ajhg.2009.01.011

38. Bashi S, Khan MA, Guirjis A, Joharjy IA, Abid MA. Immotile-cilia syndrome with azoospermia: a case report and review of the literature. *Br J Dis Chest.* (1988) 82:194–6. doi: 10.1016/0007-0971(88)90043-5

39. Bokhari HA, Shaik NA, Banaganapalli B, Nasser KK, Ageel HI, Al Shamrani AS, et al. Whole exome sequencing of a Saudi family and systems biology analysis identifies CPED1 as a putative causative gene to Celiac Disease. *Saudi J Biol Sci.* (2020) 27:1494–502. doi: 10.1016/j.sjbs.2020.04.011

40. Gaboon NEA, Banaganapalli B, Nasser K, Razeeth M, Alsaedi MS, Rashidi OM, et al. Exome sequencing and metabolomic analysis of a chronic kidney disease and hearing loss patient family revealed RMND1 mutation induced sphingolipid metabolism defects. *Saudi J Biol Sci.* (2020) 27:324–34. doi: 10.1016/j.sjbs.2019.10.001

41. Huber LA, Fialka I, Paiha K, Hunziker W, Sacks DB, Bähler M, et al. Both calmodulin and the unconventional myosin Myr4 regulate membrane

trafficking along the recycling pathway of MDCK cells. *Traffic.* (2000) 1:494–503. doi: 10.1034/j.1600-0854.2000.010607.x

42. De La Cruz EM, Ostap EM. Relating biochemistry and function in the myosin superfamily. *Curr Opin Cell Biol.* (2004) 16:61–7. doi: 10.1016/j.ceb.2003.11.011

43. Juan T, Geminard C, Coutelis JB, Cerezo D, Poles S, Noselli S, et al. Myosin1D is an evolutionarily conserved regulator of animal left-right asymmetry. *Nat Commun.* (2018) 9:1942. doi: 10.1038/s41467-018-04284-8

44. Lebreton G, Geminard C, Lapraz F, Pyrpassopoulos S, Cerezo D, Speder P, et al. Molecular to organismal chirality is induced by the conserved myosin 1D. *Science.* (2018) 362:949–52. doi: 10.1126/science.aat8642

45. Tingler M, Kurz S, Maerker M, Ott T, Fuhl F, Schweickert A, et al. A conserved role of the unconventional myosin 1D in laterality determination. *Curr Biol.* (2018) 28:810–16.e813. doi: 10.1016/j.cub.2018.01.075

46. Yuan S, Brueckner M. Left-right asymmetry: myosin 1D at the center. *Curr Biol.* (2018) 28:R567–R9. doi: 10.1016/j.cub.2018.03.01

47. Blum M, Ott T. Mechanical strain, novel genes and evolutionary insights: news from the frog left-right organizer. *Curr Opin Genet Dev.* (2019) 56:8–14. doi: 10.1016/j.gde.2019.05.005

48. Spéder P, Ádám G, Noselli S. Type ID unconventional myosin controls left-right asymmetry in Drosophila. *Nature.* (2006) 440:803–7. doi: 10.1038/nature04623

49. González-Morales N, Géminard C, Lebreton G, Cerezo D, Coutelis J-B, Noselli S. The atypical cadherin dachsous controls left-right asymmetry in Drosophila. *Dev Cell.* (2015) 33:675–89. doi: 10.1016/j.devcel.2015.04.026

50. Blum M, Ott T. Animal left-right asymmetry. *Curr Biol.* (2018) 28:R301–R4. doi: 10.1016/j.cub.2018.02.073

51. Mirra V, Werner C, Santamaria F. Primary ciliary dyskinesia: an update on clinical aspects, genetics, diagnosis, and future treatment strategies. *Front Pediatr.* (2017) 5:135. doi: 10.3389/fped.2017.00135

52. Boon M, Smits A, Cuppens H, Jaspers M, Proesmans M, Dupont LJ, et al. Primary ciliary dyskinesia: critical evaluation of clinical symptoms and diagnosis in patients with normal and abnormal ultrastructure. *Orphanet J Rare Dis.* (2014) 9:11. doi: 10.1186/1750-1172-9-11

53. Escobar-Diaz MC, Friedman K, Salem Y, Marx GR, Kalish BT, Lafranchi T, et al. Perinatal and infant outcomes of prenatal diagnosis of heterotaxy syndrome (asplenia and polysplenia). *Am J Cardiol.* (2014) 114:612–7. doi: 10.1016/j.amjcard.2014.05.042

54. Blum M, Feistel K, Thumberger T, Schweickert A. The evolution and conservation of left-right patterning mechanisms. *Development.* (2014) 141:1603–13. doi: 10.1242/dev.100560

55. Coutelis JB, González-Morales N, Géminard C, Noselli S. Diversity and convergence in the mechanisms establishing L/R asymmetry in metazoa. *EMBO Rep.* (2014) 15:926–37. doi: 10.15252/embr.2014 38972

56. Hegan PS, Ostertag E, Geurts AM, Mooseker MS. Myosin Id is required for planar cell polarity in ciliated tracheal and ependymal epithelial cells. *Cytoskeleton (Hoboken).* (2015) 72:503–16. doi: 10.1002/cm.21259

57. Nasser KK, Banaganapalli B, Shinawi T, Elango R, Shaik NA. Molecular profiling of lamellar ichthyosis pathogenic missense mutations on the structural and stability aspects of TGM1 protein. *J Biomol Struct Dyn.* (2020) 1–11. doi: 10.1080/07391102.2020.1782770

58. Ibrahim AZ, Thirumal Kumar D, Abunada T, Younes S, George Priya Doss C, Zaki OK, et al. Investigating the structural impacts of a novel missense variant identified with whole exome sequencing in an Egyptian patient with propionic acidemia. *Mol Genet Metab Rep.* (2020) 25:100645. doi: 10.1016/j.ymgmr.2020.100645

59. Thirumal Kumar D, Udhaya Kumar S, Nishaat Laeeque AS, Apurva Abhay S, Bithia R, Magesh R, et al. Computational model to analyze and characterize the functional mutations of NOD2 protein causing inflammatory disorder - Blau syndrome. *Adv Protein Chem Struct Biol.* (2020) 120:379–408. doi: 10.1016/bs.apcsb.2019.11.005

60. Shaik NA, Nasser KK, Alruwaili MM, Alallasi SR, Elango R, Banaganapalli B. Molecular modelling and dynamic simulations of sequestosome 1 (SQSTM1) missense mutations linked to Paget disease of bone. *J Biomol Struct Dyn.* (2021) 39:2873–84. doi: 10.1080/07391102.2020.1758212

61. Ko Y-S, Bae JA, Kim KY, Kim SJ, Sun EG, Lee KH, et al. MYO1D binds with kinase domain of the EGFR family to anchor them to plasma membrane before their activation and contributes carcinogenesis. *Oncogene.* (2019) 38:7416–32. doi: 10.1038/s41388-019-0954-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.