scientific reports

Check for updates

OPEN

Genetic landscape and phenotypic spectrum of osteogenesis imperfecta in the Kazakhstani pediatric population

Mirgul Bayanova[1], Aigerim Abilova[1], Marzhan Rakhimzhanova[1], Assiya Bazenova[1], Lyazzat Nazarova[1], Dias Malik[1], Naanlep Matthew Tanko[1,2], Nursulu Altaeva[3] & Aidos Bolatov[1,3,4]✉

This study investigates the genetic landscape and phenotypic spectrum of osteogenesis imperfecta (OI) in the Kazakhstani pediatric population, focusing on 40 children diagnosed and treated at the "University Medical Center" Corporate Fund from July 2021 to June 2023. Genetic analysis was conducted using whole-genome sequencing for 22 participants at the "National Laboratory Astana" (Nazarbayev University, Astana, Kazakhstan) and whole-exome sequencing for 18 participants in private laboratories. Clinically significant genetic variants were found in 35 cases (87.5%). Mutations in the *COL1A1* and *COL1A2* genes were detected in 24 cases (68.6%), among them 5 variants were described for the first time. Among the rare cases of OI, variants in the *IFITM5* (n = 2), *SERPINF1* (n = 7), and *SERPINH1* (n = 1) genes were identified. At the same time, seven unrelated cases had identical variants in the *SERPINF1* gene (c.907C > T, 6 of which in the homozygous and 1 in the compound heterozygous state) and two cases in the *IFITM1* gene (c.-14C > T). Novel disease-causing variants were identified in 17% of cases, and a higher proportion of collagen defects were seen. The relatively high proportion of autosomal recessive inherited OI determined in the current study should be investigated at the population level in Kazakhstan and in the countries of Central Asia. Moreover, this study described the genotype–phenotype correlation, which complements and expands the existing knowledge about the OI.

Keywords Osteogenesis imperfecta, Genetics, Whole-genome sequencing, Whole-exome sequencing, Next-generation sequencing

Osteogenesis imperfecta (OI or brittle bone disease) is a phenotypically and molecularly heterogeneous group of inherited systemic connective tissue disorders characterized by low bone mass, bone fragility, and deformity[1,2]. OI is a fairly common rare disorder with prevalence ranges from about 1:15,000 to 1:20,000 births[3]. Type I is the most common type of OI (population frequencies range between 2.35 to 4.7 in 100,000), and the incidence of OI type II ranges between 1.4 to 2.5 in 100,000 live births. The exact incidence of types III and IV OI is unknown (Subramanian, Anastasopoulou, & Viswanathan,[4]). At the same time, the overall prevalence and incidence of OI in Kazakhstan have not been studied.

In the last 15–20 years, genetic studies have expanded our understanding of the causative mechanisms that underlie OI, and next-generation sequencing technology enabled cost-effective and timely diagnosis via expanded gene panels and exome or genome sequencing. The majority of OI (about 85–90% of cases) is associated with autosomal dominantly (AD)-inherited pathogenic variants in *COL1A1* or *COL1A2* genes, encoding the α1 and α2 chains of collagen type I, and are classified as having classic OI types I to IV[2,5,6]. Other OI described variants affect collagen post-translational modification and folding (autosomal recessively (AR)-inherited *CRTAP*, *P3H1*, and *PPIB*), collagen folding and crosslinking (AR-inherited *SERPINH1*, *FKBP10*, and *BMP1*), ossification and mineralization (AD-inherited *IFITM5* and AR-inherited *SERPINF1*), osteoblast differentiation and function (AR-inherited *SP7*, *TMEM38B*, *WNT1*, *CREB3L1*, *SPARC*, and XR-inherited *MBTPS2*)[1,5]. *TENT5A*-mediated

[1]"University Medical Center" Corporate Fund, Turan Ave. 38, 010000 Astana, Kazakhstan. [2]Department of Biomedical Sciences, School of Medicine, Nazarbayev University, Astana 010000, Kazakhstan. [3]Astana Medical University, Beybitshilik St. 49A, 010000 Astana, Kazakhstan. [4]Shenzhen University Medical School, Shenzhen University, Shenzhen 518060, Guangdong, China. ✉email: bolatovaidos@gmail.com

regulation of collagen expression[7], involved in ossification and protein folding *MESD* gene (Gene, NCBI[8,9]), *CCDC134*-mediated regulation MAPK signaling pathway[10], *KDELR2*-mediated disrupting fiber formation[11] were also described as causative OI genes.

Although the diagnosis of OI is mainly based on clinical and radiological features, genetic tests can determine the exact cause of the disease and provide useful information in unclear cases[12–14]. Molecular diagnostics provide information on recurrent risk and assists in the genetic counseling of families for the identification of affected members and pregnancy/family planning[13,15,16]. Moreover, the development of new treatment options and therapy schemas for OI predominantly depends on a better understanding of the genetic background and molecular mechanisms of this disease and may based on genetic engineering and molecular chaperones usage[3,17–19].

The Central Asian region remains a dark spot in understanding the genetic epidemiology of many hereditary diseases, including OI. In the current study, we aim to investigate the genetic epidemiology of OI using NGS technology and expand our understanding of the genotype–phenotype correlations.

Materials and methods

Study participants

The National Research Center for Maternal and Child Health of the "University Medical Center" Corporate Fund (UMC CF) is a specialized national-level hospital in Kazakhstan, where children with orphan diseases, including OI, are diagnosed and treated. For 2023, 107 cases of OI have been registered in the "University Medical Center" CF. The study involved 40 children from unrelated marriages, hospitalized in the Clinical Department of Pediatrics from July 2021 to June 2023. Among them, 22 participants underwent whole-genome sequencing (WGS) at the "National Laboratory Astana" PI (NLA, Nazarbayev University, Astana, Kazakhstan), and 18 participants underwent whole-exome sequencing (WES) in private laboratories.

Biological samples and DNA isolation

Peripheral blood samples (2 mL each) were collected from 22 participants and sent to the NLA for DNA extraction. DNA isolation was performed using the Promega™ kit (USA), following the manufacturer's protocols to ensure high-quality genomic DNA for downstream analysis. Additionally, 18 participants provided their blood samples in private laboratories, where DNA isolation was carried out using standard procedures specific to each facility. All samples were processed under stringent conditions to maintain sample integrity and ensure reliable genetic analysis.

Whole-genome sequencing

DNA libraries were prepared from 300 ng of genomic DNA using Illumina DNA PCR-Free Library Prep, Tagmentation protocol, with the IDT Indexes for Illumina DNA/RNA UD Set A. DNA libraries were validated using the Qubit ssDNA Assay on the Qubit Fluorometer according to the standard kit protocol. Sequencing was performed on the NovaSeq 6000 high throughput platform using the NovaSeq 6000 S4 Reagent Kit v1.5 (300 cycles) at the NLA.

Raw Data Preprocessing and Bioinformatics Analysis

The raw data files generated by the Illumina NovaSeq sequencing platform in binary base call (bcl) format were converted into FASTQ format using the bcl2fastq tool. The quality of the resulting sequences was assessed with FastQC v0.11.9 and summarized using MultiQC v1.12. Sequencing reads were mapped to the human reference genome (NCBI GRCh37, hg19) using the Burrows-Wheeler Aligner (BWA) v0.7.12. Duplicate reads were identified and marked, and the alignment files were sorted with Picard Tools v2.27.4. Variant calling was performed using the Genome Analysis Toolkit (GATK) v3.8, and the identified genomic variants were subsequently annotated using ANNOVAR[20].

Whole-exome sequencing

Whole-exome sequencing (WES) was performed in a private laboratory using the Illumina NovaSeq 6000 and Illumina NextSeq 550 platforms with paired-end sequencing (2 × 100 bp) at an average coverage depth of 70–100x. Library preparation was carried out using a selective capture approach targeting the coding regions of human genes. DNA enrichment was performed using Agilent SureSelect Human All Exon probes. Sequencing data processing included an automated pipeline involving alignment of sequencing reads to the human reference genome (hg19), post-alignment processing, variant calling, and quality-based variant filtering. Variants were annotated across all known gene transcripts from the RefSeq database.

Interpretation of genetic variants

Genetic variants were annotated following the nomenclature standards established by the Human Genome Variation Society (HGVS; http://www.hgvs.org/mutnomen). Variant interpretation was performed in accordance with the 5-level classification system recommended by the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP). Analysis was carried out using the Franklin platform (Franklin by Genoox, Genoox, USA). A comprehensive three-step approach was applied for variant evaluation:

1. Variants were filtered based on their frequency in population control databases (gnomAD, GenomeAsia, 1000 Genomes, ExAC, and ESP 6500).
2. Literature and database reviews were conducted to investigate the potential role of each variant in disease etiology and progression (HGMD, ClinVar, UniProt, LOVD, and OMIM).

3. Pathogenicity assessments were performed for each identified variant in the context of individual clinical cases (Revel, AlphaMissense, Varity, MutationAssessor, SIFT, MutationTaster, FATHMM, and DANN for missense variants, SpliceAI for splice-altering variants, and GenoCanyon and fitCons integrates functional assays).

Statistical analysis

The Shapiro–Wilk test was employed to assess the normality of continuous variables. Normally distributed continuous data was presented as mean and the standard deviation (SD), continuous data lacking normal distribution is presented as median and range. Comparisons of normally distributed continuous data were performed using the Student's t-test, whereas the Mann–Whitney U test was applied for non-normally distributed data. P-values less than 0.05 were considered to be statistically significant. All statistical analyses were performed using JASP (version 0.19.3) and Jamovi (version 2.6.17) software.

Ethical issues

The study was approved by the local commission on bioethics of the "University Medical Center" Corporate Fund (Astana, Kazakhstan), an extract from Protocol No. 1 dated 06/29/2021. All methods were performed by the relevant guidelines and regulations, including the Declaration of Helsinki. Written informed consent was obtained from all subjects involved in the study.

Results

Forty patients of the Clinical Department of Pediatrics of the UMC CF with OI were included in the study. Slightly more than half (n = 21, 52.5%) of the participants were female. The age of study participants at the time of recruitment ranged from 1 to 17 years, with an average age of 97 months (8 years and 1 month). Among the probands, a family history of OI was noted in a quarter of cases (10/40, 25%).

After the whole-genome sequencing and interpretation of whole-exome sequencing results from private laboratories, pathogenic, likely-pathogenic and variants of Unknown Significance (VUS) were identified in 35 cases (87.5%). The analysis identified the distribution of Osteogenesis Imperfecta (OI) types among the study participants as follows: 3 cases of OI type I associated with *COL1A1* variants, 14 cases of OI type III linked to *COL1A1*, and 7 cases of OI type IV, of which 5 were associated with *COL1A1* and 2 with *COL1A2*. Additionally, 2 cases of OI type V were linked to *IFITM5*, 8 cases of OI type VI were associated with *SERPINF1*, and 1 case of OI type X was attributed to a variant in *SERPINH1*. Clinical, phenotype, and genotype data of patients with OI are presented in Tables 1 and 2, and Fig. 1.

The median age at first fracture varies significantly among different OI types, reflecting the clinical heterogeneity of the disease. Patients with OI Type IV exhibited the highest median age at first fracture (Me: 3 years), indicating a milder disease course. Statistical comparisons revealed no significant differences in the age of the first fracture between OI types (Fig. 1B). The number of fractures also varied considerably across OI types, with a clear trend correlating with disease severity. Patients with OI type III had the highest median number of fractures (Me: 20 fractures), while those with OI Type I and type IV reported significantly fewer fractures (Me: 6 and 5 fractures, respectively, $p < 0.05$, Fig. 1C).

Genetic variants in the *COL1A1* and *COL1A2* genes, inherited in an autosomal-dominant manner and corresponding to OI types I-IV, were observed in 68.6% of cases (24 of 35). Among the identified genetic variants in the COL1A1 gene (n = 22), the majority were missense variants (n = 12). Notably, 10 (83%) of these missense variants resulted in the substitution of the amino acid glycine (Gly) within the collagen alpha I chain triple helical domain, a region critical for structural stability. Additionally, seven variants in COL1A1 were found to affect splicing. In the COL1A2 gene, two genetic variants were identified, both of which were missense variants (Table 2, Fig. 1D,E). One of these resulted in the substitution of glycine (Gly) within the collagen alpha I chain triple helical domain, similar to the observations in COL1A1.

AD-inherited *IFITM5* causing OI type V was detected in 2 cases (5.7%). At the same time, the proportion of patients with variants in the *SERPINH1* (n = 1) and *SERPINF1* (n = 8) genes inherited in an autosomal recessive manner amounted to about a third of all cases (25.7%). Variants in the SERPINF1 gene were predominantly nonsense variants (Fig1D,E). Moreover, *SERPINF1*:c.907C > T (n = 7, six in homozygous and one in compound heterozygous state) and *IFITM5*:c.-14C > T (n = 2, heterozygous state) variants were observed in more than one unrelated cases.

Among the results presented, six variants (17.1%), namely c.4006-2A > T, c.1269_1277del, c.3278G > C, c.3659A > T, and c.662G > T in *COL1A1* gene, and *SERPINF1*:c.957_961del, were firstly described in the current research.

Blue sclera, dentinogenesis imperfecta, mobility impairment, scoliosis, vertebral compression fracture, and femoral fractures were found in both the collagen to non-collagen OI, and there was no significant difference in frequencies (Table 2, Fig. 1F). At the same time, such phenotypic signs as late fontanel closure (> 2 years) and cardiac abnormalities were noted only in patients with a mutation in the *COL1A1* gene, and joint hyperextension in the group of non-collagen OI (*IFITM5*, *SERPINF1*).

Discussion

This study represents the first comprehensive genetic analysis of osteogenesis imperfecta (OI) in the Kazakhstani pediatric population, utilizing next-generation sequencing (NGS) technology. Our findings significantly expand the understanding of the genetic basis and phenotypic spectrum of OI in Central Asia.

The identification of pathogenic and likely pathogenic variants in 87.5% of our cohort underscores the robustness of NGS in diagnosing OI. The following genes have been identified as OI etiology: *COL1A1* (Types

Case # (gender)	Classical Aillence classification	Gene (RefSeq)	Genetic variant	ACMG/ Clinical evidence	Age of the first fracture (years)	Phenotype	Additional information (number of fractures)	Family cases
1 (m)	Type I (AD)	*COL1A1* NM_000088.4	chr17-48,272,927C > T (c.1155 + 1G > A)	LP/-	3	Blue sclera, mild to moderate bone fragility, kyphoscoliosis	History of up to 10 fractures	The father, grandmother from father's side, and father's sibling have frequent fractures
2 (m)	Type I (AD)	*COL1A1* NM_000088.4	chr17-48,272,927 C > T (c.1155 + 1G > A)	LP/-	0	Blue sclera, opalescent dentin, deformity of the lower extremities	History of 1 fracture	Father, two siblings, grandmother from father's side, father's sibling have frequent fractures
3 (m)	Type I (AD)	*COL1A1* NM_000088.4	chr17-48,276,949 ACT > A (c.299_300del, p.Glu100ValfsTer68)	P/ClinVar-LP	2	Blue sclera, bone pain, periostitis	History of more than 6 fractures	(-)
4 (f)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,269,229 C > T (c.2047G > A, p.Gly683Ser)	LP/UniProt-P	0	Limb shortening and progressive deformities, frontal bossing, opalescent dentin, bone pain, short neck, pectus excavatum	The history of the first fracture was prenatal, and there are up to 20 fractures in total	The father has a history of frequent fractures
5 (m)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,273,524 C > T (c.994G > A, p.Gly332Arg)	P/ ClinVar-P, UniProt-P	4	Deformity of the lower extremities, triangular shaped face, opalescent dentin, short neck, narrow chest	History of up to 15 fractures	(-)
6 (m)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,272,201 C > T (c.1354-12G > A)	P/ ClinVar-P	2.5	Short stature, deformities of long bones, frontal bossing, long bone deformity, scoliosis	History of more than 20 fractures	The mother, grandfather on the mother's side has frequent fractures
7 (m)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,264,000 C > T (c.3814 + 1G > A)	P/ ClinVar-LP	3	Limb shortening and progressive deformities, frontal bossing, blue sclera, pectus excavatum	History of more than 8 fractures	Mother, grandfather on mother's side and mother's sibling have frequent fractures
8 (m)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,267,460 C > T (c.2461G > A, p.Gly821Ser)	P/ ClinVar-P, UniProt-P	2	Deformity of the lower extremities, opalescent dentin, scoliosis, knee joint contraction	History of more than 10 fractures	(-)
9 (m)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,272,190 C > T (c.1354-1G > A)	LP/-	2	Bowing of limbs due to multiple fractures, limb muscle atrophy, frontal bossing, opalescent dentin, short neck, broad chest, scoliosis	History of more than 20 fractures	The mother has a history of frequent fractures
10 (f)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,263,383 T > A (c.4006-2A > T)	LP/-	3	Limb shortening and progressive deformities, bowing of limb, opalescent dentin, blue sclera at birth becoming normal with age	History of more than 14 fractures	Father and sibling have frequent fractures
11 (m)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,272,614 AGGG CCGCCG > A (c.1269_1277del, p.Gly424_Pro426del)	VUS/-	1.5	Limb shortening and progressive deformities, frontal bossing, opalescent dentin, broad chest, muscle weakness, bowed legs	History of more than 10 fractures	(-)
12 (f)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,269,187 G > A (c.2089C > T, p.Arg697Ter)	P/ ClinVar-P	0	Limb shortening and progressive deformities, micrognathia, blue sclera at birth becoming normal with age, multiple fractures present at birth	History of more than 20fractures	(-)
13 (f)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,265,492 C > T (c.3226G > A, p.Gly1076Ser)	P/ ClinVar-P, UniProt-P	3	Deformity of the lower extremities, scoliosis, short deformed femurs, frontal bossing	History of more than 25 fractures	(-)
14 (f)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,265,328 C > G (c.3278G > C, p.Arg1093Pro)	VUS/-	3	Limb shortening and progressive deformities, triangular face, frontal bossing, thin gracile ribs, scoliosis	History of more than 20 fractures	(-)
15 (m)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,265,980 C > T (c.3118G > A, p.Gly1040Ser)	P/ ClinVar-P, UniProt-P	2	Short limb, wormian bones, long bone deformity, short deformed femurs, frontal bossing	History of more than 20 fractures	(-)

Continued

Case # (gender)	Classical Aillence classification	Gene (RefSeq)	Genetic variant	ACMG/ Clinical evidence	Age of the first fracture (years)	Phenotype	Additional information (number of fractures)	Family cases
16 (f)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,275,309 C > T (c.642 + 1G > A)	LP/ ClinVar-P	3	Limb shortening and progressive deformities, triangular face, frontal bossing, multiple fractures, scoliosis, blurish sclera, hearing problems	History of more than 30 fractures	(-)
17 (f)	Type III (AD)	*COL1A1* NM_000088.4	chr17-48,265,492 C > T (c.3226G > A, p.Gly1076Ser)	P/ ClinVar-P, UniProt-P	0.5	Short stature, brachycephaly, blue sclera, short neck, long bone fractures, mild-moderate skeletal deformity, muscle weakness	History of more than 6 fractures	(-)
18 (m)	Type IV (AD)	*COL1A1* NM_000088.4	chr17-48,264,156 T > A (c.3659A > T, p.Asp1220Val)	LP/-	2.5	Short stature, pale blue sclera, mild-moderate skeletal deformity, scoliosis	History of more than 10 fractures	The father has a history of frequent fractures
19 (f)	Type IV (AD)	*COL1A1* NM_000088.4	chr17-48,274,031 C > T (c.805G > A, p.Gly269Ser)	P/ ClinVar-LP/P	3	Bowed legs and arms, muscle weakness, opalescent dentin, pectus excavatum, shortening of the right lower limb	History of more than 5 fractures	(-)
20 (f)	Type IV (AD)	*COL1A1* NM_000088.4	chr17-48,275,127 C > A (c.662G > T, p.Gly221Val)	LP/-	6	Bowed legs, mild-moderate skeletal deformity, kyphosis	History of more than 5 fractures	(-)
21 (m)	Type IV (AD)	*COL1A1* NM_000088.4	chr17-48,272,673 C > T (c.1219G > A, p.Gly407Ser)	LP/ClinVar-LP	5	Bowed legs, muscle weakness, opalescent dentin, mild-moderate skeletal deformity	History of more than 5 fractures	(-)
22 (m)	Type IV (AD)	*COL1A1* NM_000088.4	chr17-48,272,981 C > T (c.1102G > A, p.Gly368Ser)	LP/-	0	Macrocephaly, microstomy, opalescent dentin, long bone fractures, skeletal deformity	History of 2 fractures	(-)
23 (m)	Type IV (AD)	*COL1A2* NM_000089.4	chr7-94,058,603 G > T (c.3815G > T, p.Cys1272Phe)	LP/ ClinVar-P	5	Short stature, muscle weakness, frontal bossing, narrow chest, bone deformities	History of more than 10 fractures	(-)
24 (f)	Type IV (AD)	*COL1A2* NM_000089.4	chr7-94,037,168 G > C (c.604G > C, p.Gly202Arg)	LP/ ClinVar-P, UniProt-P	0	Blue sclera, ligamentous laxity, long bone fractures, mild-moderate skeletal deformity,	History of more than 10 fractures	Father, two siblings, grandfather from father's side, father's sibling, grandfather two siblings from father's side, have frequent fractures
25 (f)	Type V (AD)	*IFITM5* NM_001025295.3	chr11-299,504 G > A (c.-14C > T)	P/ ClinVar-P	2	Short stature, triangular face, moderate to severe bone fragility, hyperextensible joints, bluish sclera	History of more than 10 fractures	(-)
26 (f)	Type V (AD)	*IFITM5* NM_001025295.3	chr11-299,504 G > A (c.-14C > T)	P/ ClinVar-P	1.5	Ligamentous laxity, deformities of long bones, severe scoliosis, frontal bossing	History of more than 15 fractures	(-)
27 (m)	Type VI (AR)	*SERPINF1* NM_002615.6	chr17-1,679,946 C > T (c.907C > T, p.Arg303Ter)	P/ ClinVar-LP/P	1.5	Bowed legs and arms, brachycephaly, muscle weakness, short neck	History of more than 20 fractures	(-)
28 (f)	Type VI (AR)	*SERPINF1* NM_002615.6	chr17-1,679,946 C > T (c.907C > T, p.Arg303Ter)	P/ ClinVar-LP/P	2	Bowed legs and arms, bone pain, ligamentous laxity, blue sclera, short neck, scoliosis	History of more than 20 fractures	(-)
29 (f)	Type VI (AR)	*SERPINF1* NM_002615.6	chr17-1,680,559 G > A (c.1076G > A, p.Arg359Gln)	VUS/ ClinVar-VUS	1	Limb shortening and progressive deformities, frontal bossing, opalescent dentin, obesity	The history of the first fracture was prenatal, and there are up to 40 fractures in total	(-)
30 (m)	Type VI (AR)	*SERPINF1* NM_002615.6	chr17-1,679,946 C > T (c.907C > T, p.Arg303Ter)	P/ ClinVar-LP/P	1.5	Short stature, long bone fractures, bowing of upper extremities, blue sclera, normal hearing	History of more than 10 fractures	(-)
31 (m)	Type VI (AR)	*SERPINF1* NM_002615.6	chr17-1,679,946 C > T (c.907C > T, p.Arg303Ter) chr17-1,679,994 CTGAGT > C (c.957_961del, p.Ser320Ter)	P/ ClinVar-LP/P LP/-	1	Deformity of the lower and upper extremities, wormian bones, flattened vertebrae, scoliosis	History of more than 20 fractures	The father has a history of frequent fractures
32 (f)	Type VI (AR)	*SERPINF1* NM_002615.6	chr17-1,679,946 C > T (c.907C > T, p.Arg303Ter)	P/ ClinVar-LP/P	1.5	Short stature, bowed legs and arms, long bone fractures	History of more than 7 fractures	(-)

Continued

Case # (gender)	Classical Aillence classification	Gene (RefSeq)	Genetic variant	ACMG/ Clinical evidence	Age of the first fracture (years)	Phenotype	Additional information (number of fractures)	Family cases
33 (f)	Type VI (AR)	*SERPINF1* NM_002615.6	chr17-1,679,946 C > T (c.907C > T, p.Arg303Ter)	P/ ClinVar-LP/P	2	Short stature, long bone fractures, bowing of upper extremities	History of more than 8 fractures	(-)
34 (f)	Type VI (AR)	*SERPINF1* NM_002615.6	chr17-1,679,946 C > T (c.907C > T, p.Arg303Ter)	P/ ClinVar-LP/P	2.5	Bowed legs and arms, small pelvis, joint hypermobility	History of more than 5 fractures	(-)
35 (f)	Type X (AR)	*SERPINH1* NM_001235.5	chr11-75,277,449 G > A (c.55G > A, p.Ala19Thr)	VUS/ ClinVar-VUS	2	Deformity of the upper and lower extremities, triangular face, opalescent dentin,	History of more than 10 fractures	(-)

Table 1. Clinical, phenotype, and genotype data of patients with OI. M – male, F – female, AD – variant linked to an autosomal dominant condition, AR – variant linked to an autosomal recessive condition.

	COL1A1	COL1A2	IFITM5	SERPINF1	SERPINH1
Male/Female	13/9	2/0	0/2	3/5	0/1
Genetic features					
Variant type Nonsense substitution Missense substitution *(Gly substitution)* Inframe deletion Frameshift deletion Splice site Other: 5' UTR site	1 12 *(10)* 1 1 7 -	- 2 *(1)* - - - - -	- - - - - - 2	7 1 - - - - -	- 1 - - - - -
Clinical features					
Blue sclera	(±) 10/22	(±) 1/2	(±) 1/2	(±) 2/8	(-)
Dentinogenesis imperfecta	(±) 10/22	(-)	(-)	(±) 1/8	(+)
Hearing loss	(-)	(-)	(-)	(-)	(-)
Joint hyperextension	(-)	(±) 1/2	(+)	(±) 2/8	(-)
Late fontanel closure (> 2 years)	(±) 3/22	(-)	(-)	(-)	(-)
Mobility impairment	(±) 12/22	(-)	(±) 1/2	(±) 4/8	(-)
Radiological characteristic					
Scoliosis	(±) 8/22	(-)	(±) 1/2	(±) 2/8	(-)
Vertebral compression fracture	(±) 10/22	(-)	(±) 1/2	(±) 3/8	(-)
Femoral fracture	(±) 9/22	(±) 1/2	(±) 1/2	(-)	(-)
Comorbidity					
Mitral and tricuspid regurgitation	(±) 3/22	(-)	(-)	(-)	(-)

Table 2. Summary of clinical and radiological features of patients with OI.

I, III, and IV), *COL1A2* (Type IV), *IFITM5* (Type V), *SERPINF1* (Type VI), and *SERPINH1* (Type X). The predominance of mutations in the *COL1A1* and *COL1A2* genes (68.6%) aligns with global data, where these genes are implicated in the majority of OI cases. In other studies, the frequency of pathogenic variants in these two genes varied from 85–90%[6,8]. However, the relatively high proportion of autosomal-recessive (AR) variants (25.7%), particularly in the *SERPINF1* gene, suggests a unique genetic landscape in Kazakhstan that warrants further investigation. This finding is particularly notable given that AR forms of OI are less common globally, typically accounting for only 10–15% of cases.

The clinical variability among OI types observed in this study highlights the challenges in correlating genotype with phenotype. Patients with OI type IV, associated predominantly with *COL1A1* variants, exhibited the highest but not significant median age at first fracture (Me: 3 years), reflecting a relatively milder disease course. In contrast, patients with OI type III, also linked to *COL1A1*, had the highest median number of fractures (Me: 20), indicative of a more severe phenotype. These findings align with the general classification of OI severity but also underscore the variability that exists even within the same genetic and phenotypic categories[21].

Mutational analysis revealed that the majority of pathogenic variants in the *COL1A1* gene were missense variants (12/22), in seven cases there were variants disrupting the splicing process, and two variants were INDEL in nature. Among the identified 12 missense variants in *COL1A1*, 10 (83%) resulting in glycine substitutions. This finding is consistent with previous reports emphasizing the structural importance of glycine residues and their high mutational burden in OI[22]. Most of the identified variants in the *COL1A1* gene have already been described[23],Dalgleish, 2023). Whereas variants c.4006-2A > T, c.1269_1277del, c.3278G > C, c.3659A > T, and c.662G > T were first described in the current study. Among them, a variant c.3278G > C (p.Arg1093Pro) at the same point leading to another amino acid substitution was previously described (c.3278G > A (p.Arg1093His)

and was clinically classified as VUS and likely-benign[23,24],Dalgleish, 2023). One identified clinically significant variant in the *COL1A2* gene that had a missense nature has already been described[25]. These discoveries are crucial for improving diagnostic accuracy and for the potential development of targeted therapies. For instance, understanding the impact of these specific mutations on protein function could lead to novel therapeutic strategies that address the underlying molecular defects.

Variants in the *IFITM5* gene were identified in unrelated two cases with a similar recurrent variant linked to an autosomal dominant condition c.-14C > T. The mutation consists of a C > T transition at position -14 of the 5' untranslated region (UTR) of *IFITM5*, leading to the addition of five amino acids at the N-terminus of the BRIL protein, and affecting bone mineralization[26–28]. An earlier study of OI rare cases in China revealed that the mutation in the *IFITM5* gene accounted for 62.1% of all cases, while all cases had variant c.-14C > T[29].

Previously, OI was known as an autosomal dominant disorder, moreover, variants linked to an autosomal dominant condition in collagen type I are generally stated to be responsible for 90% of cases[1,30]. At the same time, in the cohorts of Chinese, Indians, Malaysians, and Brazilians, the proportion of mutations associated with a type 1 collagen defect was detected in 49–73% of cases[31], Mohd et al.,[32–34]. In this study, the prevalence of OI inherited in a recessive manner was 25.7%. The major cause of autosomal recessive OI was variants in *SERPINF1* (89%, 8/9), in one case clinically significant variant was found in SERPINH1. Among the described 8 cases with mutations in the *SERPINF1* gene, in 7 unrelated cases, variant c.907C > T was detected in the homozygous (n = 6) and compound heterozygous (n = 1) state. This variant leads to the termination of translation in 7 out of 8 exons (p.Arg303Ter) of the Serpin protein and early described as a single case in a large study conducted in China, in which 668 OI patients from 378 Chinese families were received[35].

In our study, we identified Variants of Unknown Significance (VUS) in 4 out of the 35 diagnosed patients (*COL1A1*, *SERPINF1*, and *SERPINH1* genes). According to the ACMG guidelines, VUS are variants with insufficient evidence to be classified as either pathogenic or benign. The interpretation of such findings presents several challenges, primarily due to the lack of functional studies validating their clinical significance. While these variants were detected in genes known to be associated with OI, their contribution to the disease phenotype remains uncertain. This highlights a critical limitation in the current state of genetic diagnostics, where the absence of functional validation complicates the interpretation of genetic findings. Further functional studies are necessary to evaluate the potential pathogenicity of these VUS and their role in the manifestation of OI in the affected individuals. Until then, clinical management decisions involving these variants should be approached with caution, considering both the genetic evidence and clinical presentation.

Our study also provides valuable insights into the genotype–phenotype correlations in OI. While classical symptoms like blue sclera, dentinogenesis imperfecta, and bone fragility were prevalent across different genotypes, certain phenotypic traits, such as late fontanel closure and joint hyperextension, were more common in specific genetic subtypes (Fig. 1F). This highlights the heterogeneity of OI and the importance of personalized medical management for affected individuals.

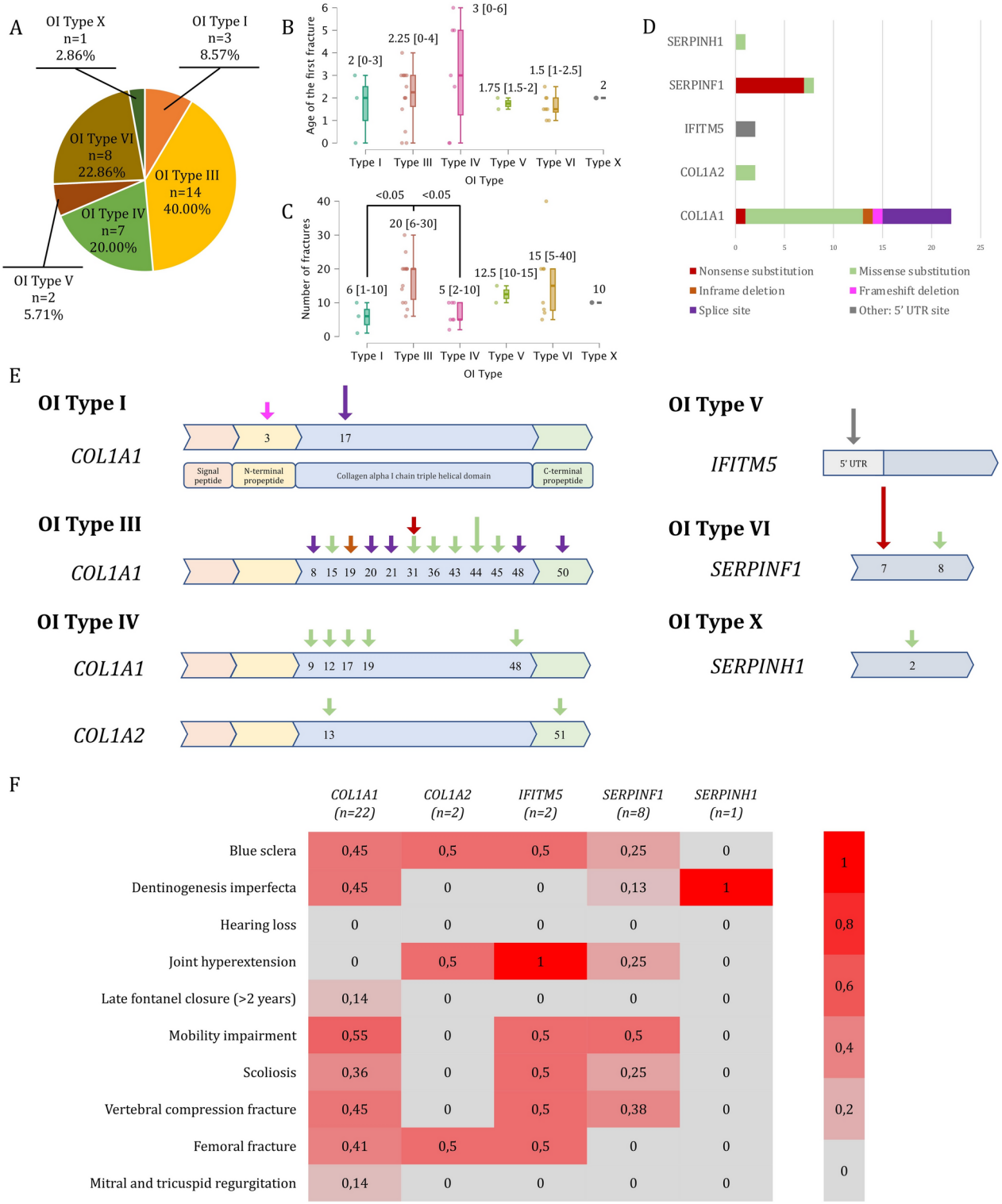
OI treatment goals include decreasing fracture incidence, improving pain, and promoting growth, mobility, and functional independence provided by multidisciplinary care including medical genetics[2]. Without a genetic cure for OI, management of the disease is aimed only at symptom reduction[36]. For instance, the antiresorptive agent denosumab decreased bone turnover markers and increased in areal BMD among children with *SERPINF1*, *COL1A1*, and *COL1A2* mutations[12]. Other treatment option based on anti-TGF-β therapy was well-tolerated and associated with increases in lumbar spine areal bone mineral density in participants with OI type IV, whereas participants with OI type III and VIII had unchanged or decreased this parameter[37]. Thus, the identification of molecular genetic etiology is important not only for the diagnosis and understanding of the mechanisms of OI, but also in the development of pathogenetic treatment and the selection of therapy.

The relatively high prevalence of OI with autosomal recessive inheritance observed in our cohort is intriguing and suggests possible founder effects or high rates of consanguinity in the population. Further, population-based studies are necessary to elucidate these factors and to understand the broader epidemiological trends in Central Asia. Additionally, expanding the sample size and including more diverse populations will be critical for validating our findings and for ensuring their generalizability. Since this disease involves a genetic etiology, and in this study, there were cases without genetic confirmation of the diagnosis by the NGS method, further research is needed aimed at studying the genetic mechanisms of OI and also expanded technological approaches in diagnosis.

One limitation of our study is the relatively small sample size, which may not capture the full spectrum of genetic variability in OI. Furthermore, while NGS provides comprehensive genetic data, the interpretation of some variants, particularly those classified as variants of uncertain significance (VUS), remains challenging. Future studies should incorporate functional assays and longitudinal clinical follow-ups to better characterize these variants and their impact on the OI phenotype.

Conclusion

In conclusion, this study highlights the utility of next-generation sequencing in unraveling the genetic basis of osteogenesis imperfecta in the Kazakhstani pediatric population. The identification of both known and novel pathogenic variants enriches the current understanding of OI and underscores the genetic diversity of this condition in Central Asia. Our findings emphasize the importance of integrating genetic diagnostics into routine clinical practice for the accurate diagnosis, management, and genetic counseling of OI patients.



The significant proportion of autosomal recessive cases identified calls for further research to explore the genetic epidemiology of OI in this region and to investigate potential founder effects. Future studies should focus on larger, more diverse cohorts and incorporate advanced functional analyses to fully elucidate the pathogenic mechanisms of newly discovered variants.

Overall, our study provides a foundational framework for future genetic and clinical research on OI in Central Asia and highlights the potential for personalized medicine approaches in the management of rare genetic disorders.

◀**Fig. 1**. Clinical and genetic characteristics of the Kazakhstani pediatric cohort with osteogenesis imperfecta (OI). (**A**). Distribution of OI types in the study population: a pie chart illustrating the percentage distribution of OI subtypes in the studied cohort (N = 35). (**B**). Age of first fracture by OI yype: box plots comparing the age at first fracture among different OI subtypes. (**C**). Number of fractures by OI type: box plots showing the total number of fractures experienced by patients across OI subtypes. Statistical comparisons between groups are annotated with p-values (< 0.05). (**D**). Types of genetic variants across OI genes: a bar graph showing the distribution of variants types (e.g., missense substitution, nonsense substitution, splice site mutations, in-frame deletions) in five genes associated with OI (COL1A1, COL1A2, IFITM5, SERPINF1, and SERPINH1). (**E**). Location of the identified genetic variants in OI-associated genes: schematic representations of the genes COL1A1, COL1A2, IFITM5, SERPINF1, and SERPINH1, indicating the positions and types of mutations identified in each OI subtype. (**F**). Heatmap visualizing phenotypic expression across different genes. The color intensity represents the frequency of each phenotype within the corresponding gene.

Data availability

The raw genomic data supporting the findings of this study are not publicly available due to concerns regarding genetic data privacy and the sensitive nature of genomic information, particularly involving children. Instead, the data can be accessed upon reasonable request from the corresponding author.

Received: 13 June 2024; Accepted: 24 March 2025
Published online: 02 April 2025

References

1. Forlino, A. & Marini, J. C. Osteogenesis imperfecta. *Lancet (London, England)* **387**(10028), 1657–1671. https://doi.org/10.1016/S0140-6736(15)00728-X (2016).
2. Rossi, V., Lee, B. & Marom, R. Osteogenesis imperfecta: advancements in genetics and treatment. *Curr. Opin. Pediatr.* **31**(6), 708–715. https://doi.org/10.1097/MOP.0000000000000813 (2019).
3. Botor, M. et al. Osteogenesis Imperfecta: Current and Prospective Therapies. *Biomolecules* **11**(10), 1493. https://doi.org/10.3390/biom11101493 (2021).
4. Subramanian S, Anastasopoulou C, Viswanathan VK. Osteogenesis Imperfecta. [Updated 2023 Feb 6]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK536957/
5. Garibaldi, N. et al. Dissecting the phenotypic variability of osteogenesis imperfecta. *Disease models Mech.* https://doi.org/10.1242/dmm.049398 (2022).
6. Marini, J. C. et al. Osteogenesis imperfecta. *Nat. Rev. Dis. Primers* **3**, 17052. https://doi.org/10.1038/nrdp.2017.52 (2017).
7. Gewartowska, O. et al. Cytoplasmic polyadenylation by TENT5A is required for proper bone formation. *Cell Rep.* **35**(3), 109015. https://doi.org/10.1016/j.celrep.2021.109015 (2021).
8. MESD mesoderm development LRP chaperone [Homo sapiens (human)] - Gene - NCBI. (n.d.). https://www.ncbi.nlm.nih.gov/gene/23184
9. Moosa, S., Yamamoto, G. L., Garbes, L., Keupp, K., Beleza-Meireles, A., Moreno, C. A., Valadares, E. R., de Sousa, S. B., Maia, S., Saraiva, J., Honjo, R. S., Kim, C. A., Cabral de Menezes, H., Lausch, E., Lorini, P. V., Lamounier, A., Jr, Carniero, T. C. B., Giunta, C., Rohrbach, M., Janner, M., … Netzer, C. Autosomal-Recessive Mutations in MESD Cause Osteogenesis Imperfecta. American journal of human genetics, 105(4), 836–843. https://doi.org/10.1016/j.ajhg.2019.08.008. (2019).
10. Dubail, J. et al. Homozygous Loss-of-Function Mutations in CCDC134 Are Responsible for a Severe Form of Osteogenesis Imperfecta. *J. Bone Mineral Res.* **35**(8), 1470–1480. https://doi.org/10.1002/jbmr.4011 (2020).
11. van Dijk, F. S., Semler, O., Etich, J., Köhler, A., Jimenez-Estrada, J. A., Bravenboer, N., Claeys, L., Riesebos, E., Gegic, S., Piersma, S. R., Jimenez, C. R., Waisfisz, Q., Flores, C. L., Nevado, J., Harsevoort, A. J., Janus, G. J. M., Franken, A. A. M., van der Sar, A. M., Meijers-Heijboer, H., Heath, K. E., … Micha, D. (2020). Interaction between KDELR2 and HSP47 as a Key Determinant in Osteogenesis Imperfecta Caused by Bi-allelic Variants in KDELR2. American journal of human genetics, 107(5), 989–999. https://doi.org/10.1016/j.ajhg.2020.09.009
12. Bardai, G. et al. Molecular diagnosis in children with fractures but no extraskeletal signs of osteogenesis imperfecta. *Osteoporosis Inter.* **28**(7), 2095–2101. https://doi.org/10.1007/s00198-017-4031-2 (2017).
13. Palomo, T., Vilaça, T. & Lazaretti-Castro, M. Osteogenesis imperfecta: diagnosis and treatment. *Curr. Opin. Endocrinol. Diabetes Obes.* **24**(6), 381–388. https://doi.org/10.1097/MED.0000000000000367 (2017).
14. Shapiro, J. R. & Sponsellor, P. D. Osteogenesis imperfecta: questions and answers. *Curr. Opin. Pediatr.* **21**(6), 709–716. https://doi.org/10.1097/MOP.0b013e328332c68f (2009).
15. Deguchi, M. et al. Current Overview of Osteogenesis Imperfecta. *Medicina (Kaunas)* **57**(5), 464. https://doi.org/10.3390/medicina57050464 (2021).
16. Trejo, P. & Rauch, F. Osteogenesis imperfecta in children and adolescents-new developments in diagnosis and treatment. *Osteoporosis Inter.* **27**(12), 3427–3437. https://doi.org/10.1007/s00198-016-3723-3 (2016).
17. Deyle, D. R. et al. Normal collagen and bone production by gene-targeted human osteogenesis imperfecta iPSCs. *Mol. Ther. J. Am. Soc. Gene Therapy* **20**(1), 204–213. https://doi.org/10.1038/mt.2011.209 (2012).
18. Morello, R. Osteogenesis imperfecta and therapeutics. *Matrix Biol.* **71–72**, 294–312. https://doi.org/10.1016/j.matbio.2018.03.010 (2018).
19. Rousseau, J. et al. Allele-specific Col1a1 silencing reduces mutant collagen in fibroblasts from Brtl mouse, a model for classical osteogenesis imperfecta. *Eur. J. Hum. Genet.* **22**(5), 667–674. https://doi.org/10.1038/ejhg.2013.198 (2014).
20. Bayanova, M. et al. Whole-Genome Sequencing Among Kazakhstani Children with Early-Onset Epilepsy Revealed New Gene Variants and Phenotypic Variability. *Mol. Neurobiol.* **60**(8), 4324–4335. https://doi.org/10.1007/s12035-023-03346-3 (2023).
21. Zhytnik, L. et al. COL1A1/2 Pathogenic Variants and Phenotype Characteristics in Ukrainian Osteogenesis Imperfecta Patients. *Front. Genet.* **10**, 722. https://doi.org/10.3389/fgene.2019.00722 (2019).
22. Maioli, M. et al. Genotype-phenotype correlation study in 364 osteogenesis imperfecta Italian patients. *Eur. J. Human Genet.* **27**(7), 1090–1100. https://doi.org/10.1038/s41431-019-0373-x (2019).
23. Dalgleish, R. The Human Collagen Mutation Database 1998. *Nucl. Acids Res.* **26**(1), 253–255. https://doi.org/10.1093/nar/26.1.253 (1998).
24. Marshall, C. J. et al. Diagnostic conundrums in antenatal presentation of a skeletal dysplasia with description of a heterozygous C-propeptide mutation in COL1A1 associated with a severe presentation of osteogenesis imperfecta. *Am. J. Med. Genet. A* **170**(12), 3303–3307. https://doi.org/10.1002/ajmg.a.37943 (2016).

25. ClinVar. (2023, September 3). NM_000089.4(COL1A2):c.3815G>T (p.Cys1272Phe) AND multiple conditions - ClinVar - NCBI. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/clinvar/RCV002241395/
26. Jovanovic, M., Guterman-Ram, G. & Marini, J. C. Osteogenesis Imperfecta: Mechanisms and Signaling Pathways Connecting Classical and Rare OI Types. *Endocr. Rev.* **43**(1), 61–90. https://doi.org/10.1210/endrev/bnab017 (2022).
27. Marini, J. C., Reich, A. & Smith, S. M. Osteogenesis imperfecta due to mutations in non-collagenous genes: lessons in the biology of bone formation. *Curr. Opin. Pediatr.* **26**(4), 500–507. https://doi.org/10.1097/MOP.0000000000000117 (2014).
28. Rauch, F., Moffatt, P., Cheung, M., Roughley, P., Lalic, L., Lund, A. M., Ramirez, N., Fahiminiya, S., Majewski, J., & Glorieux, F. H. Osteogenesis imperfecta type V: marked phenotypic variability despite the presence of the IFITM5 c.-14C>T mutation in all patients. Journal of medical genetics, 50(1), 21–24. https://doi.org/10.1136/jmedgenet-2012-101307, (2013).
29. Xi, L., Zhang, H. & Zhang, Z. L. Genotypic and Phenotypic Characteristics of 29 Patients With Rare Types of Osteogenesis Imperfecta: Average 5 Years of Follow-Up. *Front. Genet.* https://doi.org/10.3389/fgene.2021.622078 (2021).
30. Lindahl, K. et al. Genetic epidemiology, prevalence, and genotype-phenotype correlations in the Swedish population with osteogenesis imperfecta. *Eur. J. Human Genet.* **23**(8), 1042–1050. https://doi.org/10.1038/ejhg.2015.81 (2015).
31. Y Liu Ma Asan D., Lv, F., Xu, X., Wang, J., Xia, W., Jiang, Y., Wang, O., Xing, X., Yu, W., Wang, J., Sun, J., Song, L., Zhu, Y., Yang, H., Wang, J., & Li, M. Gene mutation spectrum and genotype-phenotype correlation in a cohort of Chinese osteogenesis imperfecta patients revealed by targeted next generation sequencing Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 28 10 2985 2995 https://doi.org/10.1007/s00198-017-4143-8, (2017).
32. Mohd Nawawi, N., Selveindran, N. M., Rasat, R., Chow, Y. P., Abdul Latiff, Z., Syed Zakaria, S. Z., Jamal, R., Abdul Murad, N. A., & Abd Aziz, B. B. (2018). Genotype-phenotype correlation among Malaysian patients with osteogenesis imperfecta. Clinica chimica acta; international journal of clinical chemistry. https://doi.org/10.1016/j.cca.2018.05.048, (2018).
33. Fernandes, A. M. et al. The molecular landscape of osteogenesis imperfecta in a Brazilian tertiary service cohort. *Osteoporosis Inter.* **31**(7), 1341–1352. https://doi.org/10.1007/s00198-020-05366-4 (2020).
34. Mrosk, J. et al. Diagnostic strategies and genotype-phenotype correlation in a large Indian cohort of osteogenesis imperfecta. *Bone* **110**, 368–377. https://doi.org/10.1016/j.bone.2018.02.029 (2018).
35. Li, L. et al. Genotypic and phenotypic characterization of Chinese patients with osteogenesis imperfecta. *HuM. Mutation* **40**(5), 588–600. https://doi.org/10.1002/humu.23718 (2019).
36. Constantino, C. S. et al. Effect of Bisphosphonates on Function and Mobility Among Children With Osteogenesis Imperfecta: A Systematic Review. *JBMR Plus* https://doi.org/10.1002/jbm4.10216 (2019).
37. Song, I. W., Nagamani, S. C., Nguyen, D., Grafe, I., Sutton, V. R., Gannon, F. H., Munivez, E., Jiang, M. M., Tran, A., Wallace, M., Esposito, P., Musaad, S., Strudthoff, E., McGuire, S., Thornton, M., Shenava, V., Rosenfeld, S., Huang, S., Shypailo, R., Orwoll, E., … Lee, B. (2022). Targeting TGF-β for treatment of osteogenesis imperfecta. The Journal of clinical investigation, 132(7), e152571. https://doi.org/10.1172/JCI152571

Acknowledgements
The authors thank the "University Medical Center" CF (Astana, Kazakhstan) for providing financial support for the study and the "National Laboratory Astana" PI (Nazarbayev University, Astana, Kazakhstan) for providing whole-genome sequencing.

Author contributions
Conceptualization: MB, LN, and MR; Methodology: MB, AA, and ABo; Data collection: AA, ABa, AM, LN, and MR; Formal analysis and Investigation: AA, ABa, ABo, DM, and NA; Writing—original draft preparation: AA and ABo; Writing—review and editing: MB, NA, and NMT; Supervision: MB.

Funding
This study was funded and carried out as part of a joint pilot project between the "University Medical Center" Corporate Fund and the "National Laboratory Astana" Private Institution (contract DPMRV-1986/161–20 dated 09.12.2021).

Declarations

Competing interests
The authors declare no competing interests.

Ethics approval
The study was approved by the local commission on bioethics of the "University Medical Center" Corporate Fund, an extract from Protocol No. 1 dated 29.06.2021.

Consent to participate and publication
Written informed consent and permission to publish data was obtained from parents for all research subjects. Consent for (include appropriate statements): Written informed consent and permission to publish.

Additional information
Correspondence and requests for materials should be addressed to A.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.