Supplemental Data

American Journal of Human Geneteics, Volume *88*

Exome Sequencing Identifies Truncating Mutations in Human *SERPINF1* in Autosomal-Recessive Osteogenesis Imperfecta

Jutta Becker, Oliver Semler, Christian Gilissen, Yun Li, Hanno Jörn Bolz, Cecilia Giunta, Carsten Bergmann, Marianne Rohrbach, Friederike Koerber, Katharina Zimmermann, Petra de Vries, Brunhilde Wirth, Eckhard Schoenau, Bernd Wollnik, Joris A. Veltman, Alexander Hoischen, and Christian Netzer

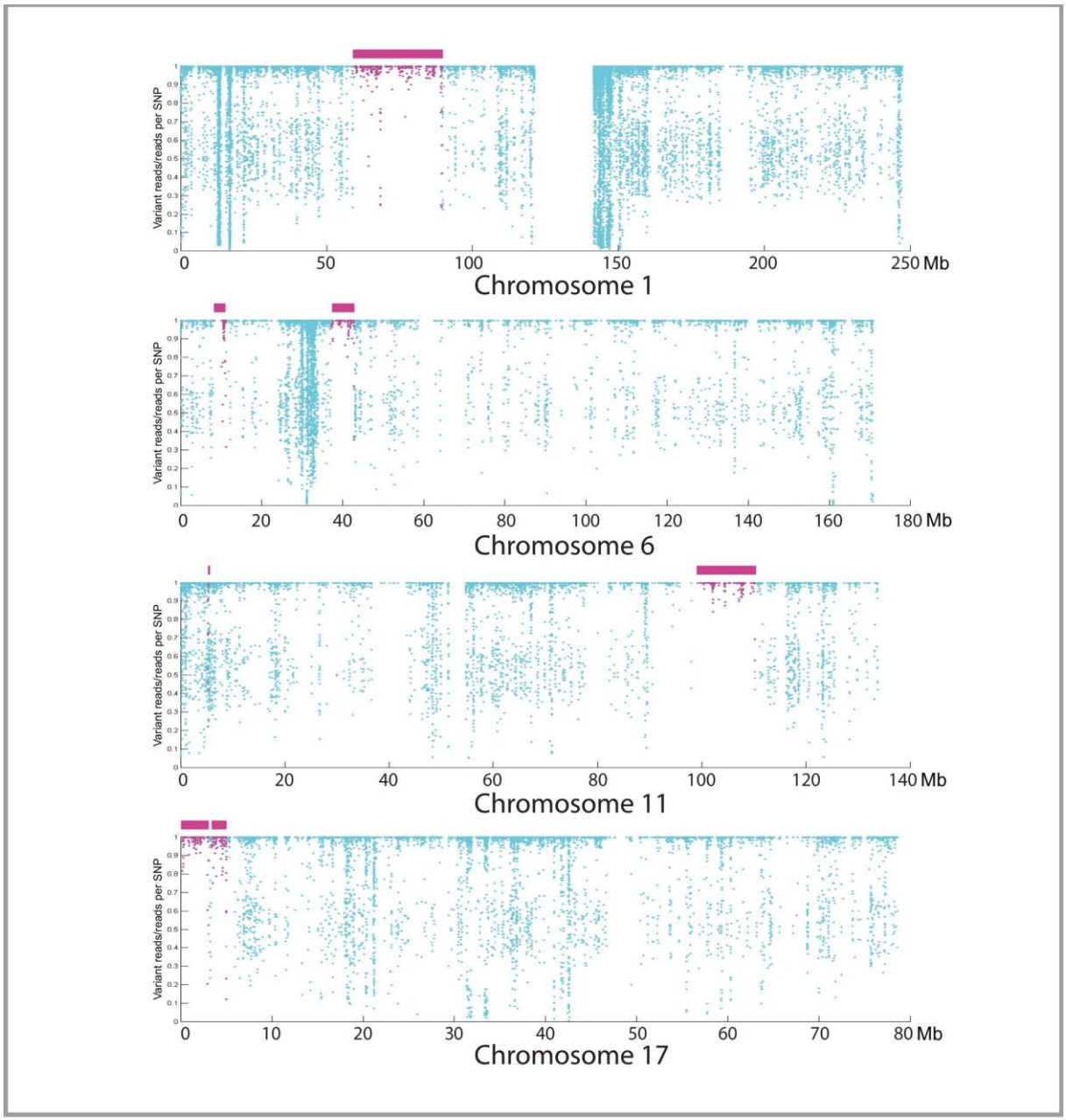
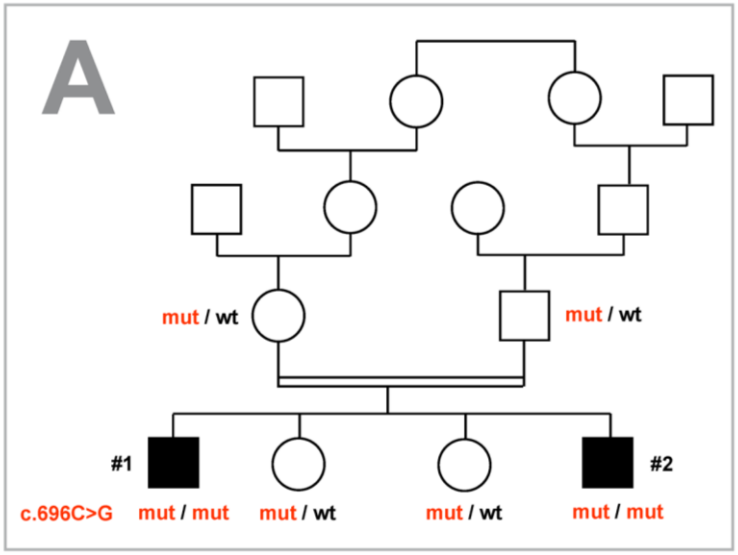
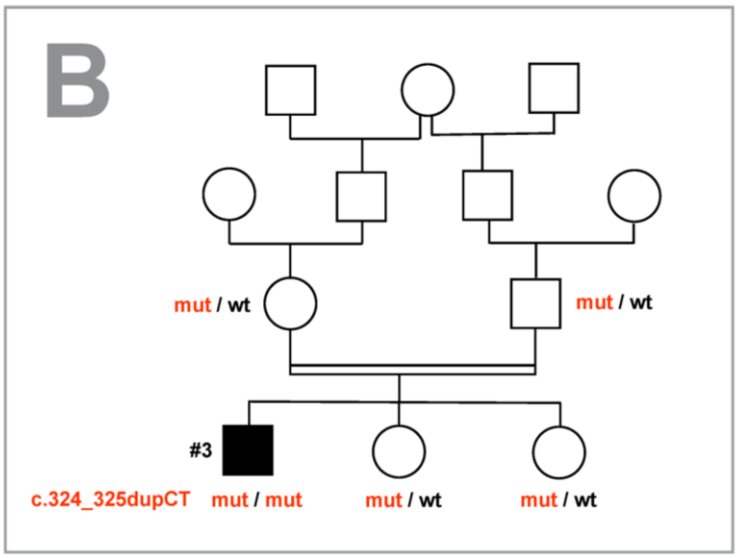


Figure S1. Overview of all Identified Homozygous Regions. The graph shows the identified homozygous regions on chromosomes 1, 6, 11 and 17 (y-axis, ratio of variant reads/all reads per SNP position; x-axis, genomic position on the respective chromosome). Each dot represents one of the SNP markers that were included in the analysis. SNPs within a region that was called as homozygous are colored in purple. The position along the y-axis indicates whether all reads or just a fraction of reads contained a certain variant. To identify homozygous regions in an unbiased genome-wide fashion, all autosomal SNP genotypes were considered, not only variants identified in the course of the initial quality filtering step of the exome data analysis (i.e., differences from the reference genome). We extracted the genotypes of all autosomal SNPs annotated in dbSNP130 that had a read coverage of at least 15-fold. This resulted in a total of 299,494 genotypes (38,927 SNPs on chromosome 1, 21,275 on chromosome 6, 16,279 on chromosome 11 and 18,031 on chromosome 17). Variants displaying in ≥95% of all reads an identical SNP allele were considered to be homozygous, SNPs with 30% to 70% variation reads were considered to be heterozygous, and SNPs with <30% or with more than 70% and less than 95% variation reads were considered ambiguous. A genomic region was identified as a homozygous stretch if at least 500 consecutive SNP markers were called as homozygous, with a maximum tolerance of two heterozygous SNPs per 500 markers (to account for possible sequencing or mapping errors). The genomic positions of all homozygous regions are highlighted by purple bars. The two homozygous regions on chromosome 17 are just separated by a few heterozygous variants, possibly false-positives, i.e. mapping artifacts in a cluster of olfactory receptor genes (stretch of 111 SNPs with 14 heterozygous calls, chr17:3,091,589-3,299,082), and therefore likely represents one homozygous stretch of more than 4.5Mb.





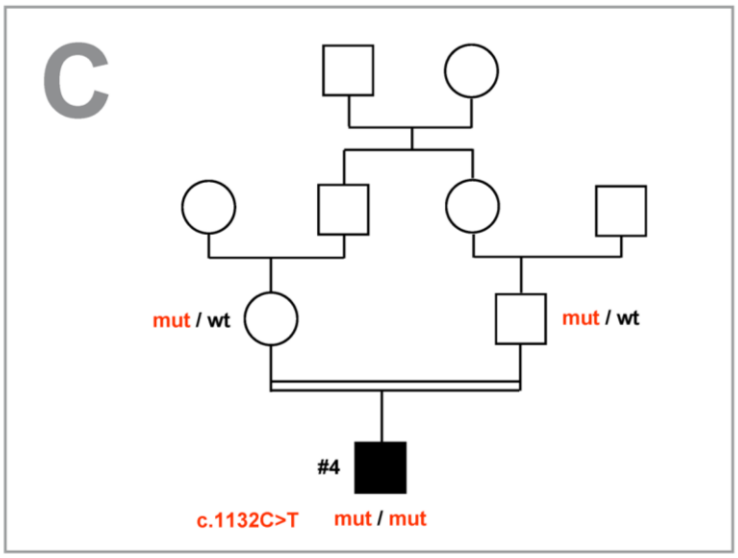


Figure S2. Pedigrees of Three OI Families with *SERPINF1* Mutations. Affected family members carry a number sign (#1 to #4). The genotypes of the patients, of their healthy siblings and of their parents are displayed, with mut indicating a mutated allele and wt indicating the wildtype allele. **(A)** Pedigree of the index family from the United Arab Emirates. The patients´ parents are second cousins. **(B)** Pedigree of patient #3. The family originates from Turkey. The patient´s parents are half first cousins, i. e., their fathers are half siblings. **(C)** Pedigree of patient #4. The family originates from Turkey. The patient´s parents are first cousins.

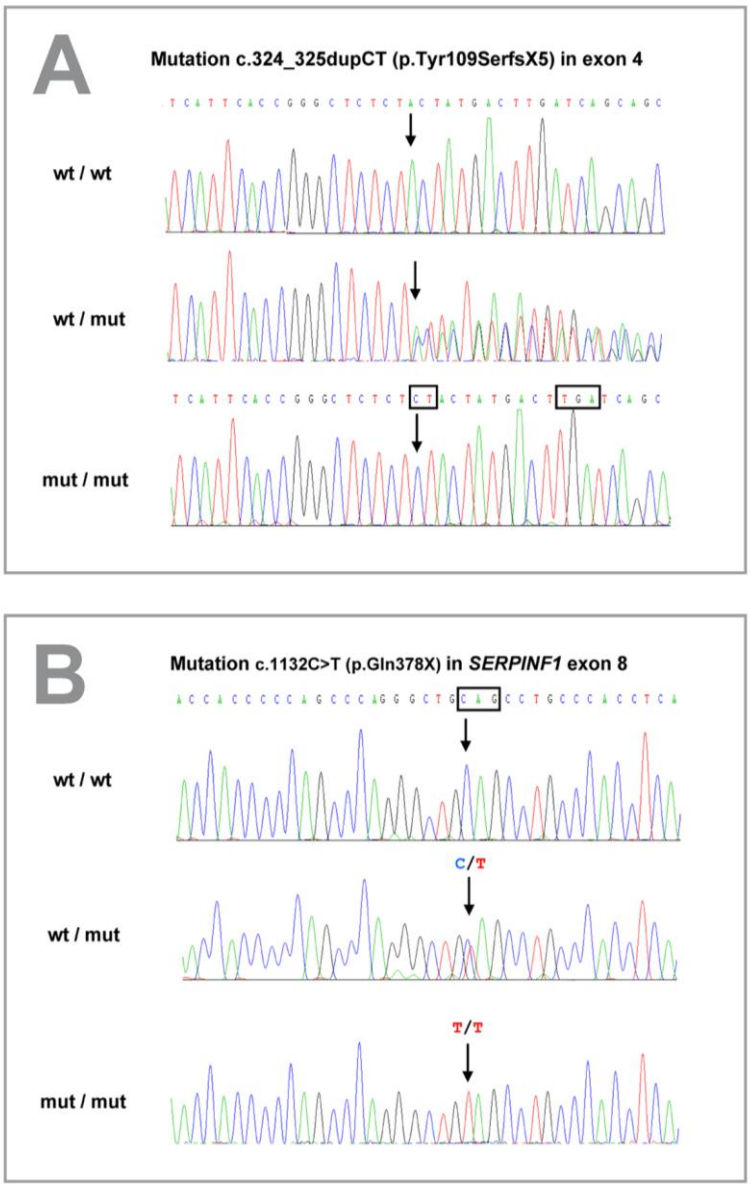


Figure S3. Electropherograms of *SERPINF1* Mutations Obtained by Sanger Sequencing. The coding exons of *SERPINF1* (NM_002615.4) were PCR-amplified from genomic DNA and analyzed by direct sequencing on an ABI3100 capillary sequencer. Primer sequences can be obtained upon request. (**A)** Electropherograms of individuals which carried the wild-type allele (wt) or the mutated allele c.324_325dupCT (mut) on one or both copies of *SERPINF1* exon 4. The position of the mutation is marked with an arrow. The two inserted nucleotides and the premature stop codon resulting from the frame-shift are framed in black. Note the "double" sequence downstream of the insertion in the heterozygous mutation carrier. This mutation was identified in patient #3. (**B)** Electropherograms of individuals which carried the wild-type allele (wt) or the mutated allele c.1132C>T (mut) on one or both copies of *SERPINF1* exon 8. The position of the mutation is marked with an arrow. The affected codon is framed in black. This mutation was identified in patient #4.

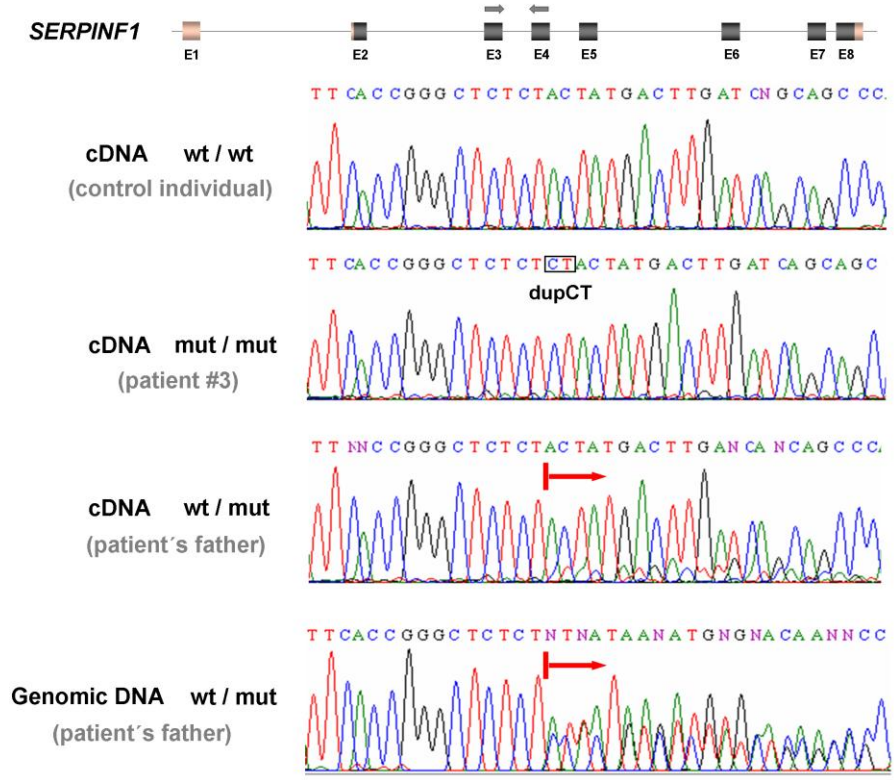


Figure S4. *SERPINF1* cDNA Analysis. RNA was isolated from whole blood of patient #3, his father and a control individual using standard methods. First strand cDNA synthesis was performed with the ReverseAid^{TM}-Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA). In the upper panel, the exon/intron structure of *SERPINF1* is depicted and the locations of the forward and reverse primer used for the subsequent cDNA amplification are given. Primer sequences can be obtained upon request. The resulting cDNA fragment encompasses the frameshift mutation c.324_325dupCT in exon 4 of *SERPINF1*. This cDNA fragment could be amplified from RNA of the control individual, of patient #3 and of the patient´s father. Sanger sequencing revealed, as expected, that the control individual carried the wildtype sequence (wt / wt) and that patient #3 was homozygous for the CT duplication (mut / mut). In the father of patient #3, the cDNA sequence of the mutated allele was only detectable as a faint background signal behind the wildtype sequence. The red arrow marks the beginning of the overlay or "double" sequence that can be observed in heterozygous carriers of insertion mutations. Sanger sequencing on genomic DNA derived from the patient´s father displayed equally strong signals for the mutated and the wildtype allele (lower panel). This observation suggests that the number of mutated transcripts serving as templates for the RT-PCR reaction is strongly reduced compared to the number of transcripts with the wildtype sequence, most certainly as a result of nonsense-mediated mRNA decay.

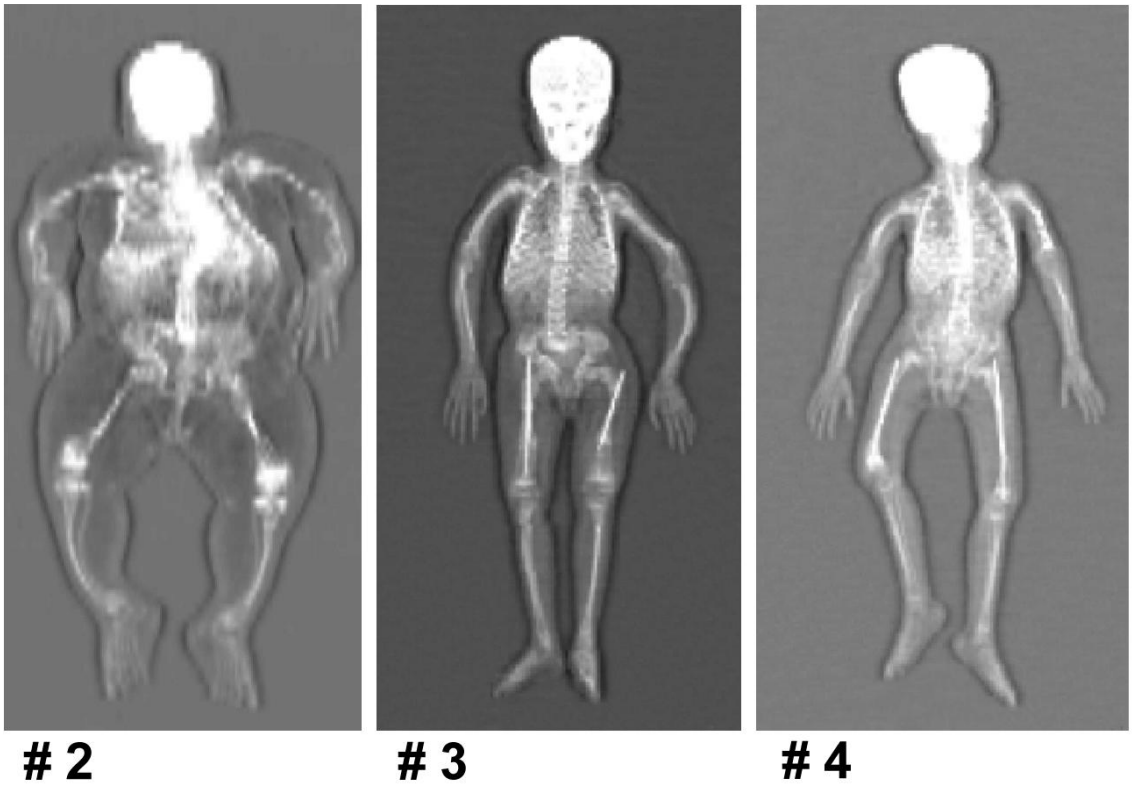
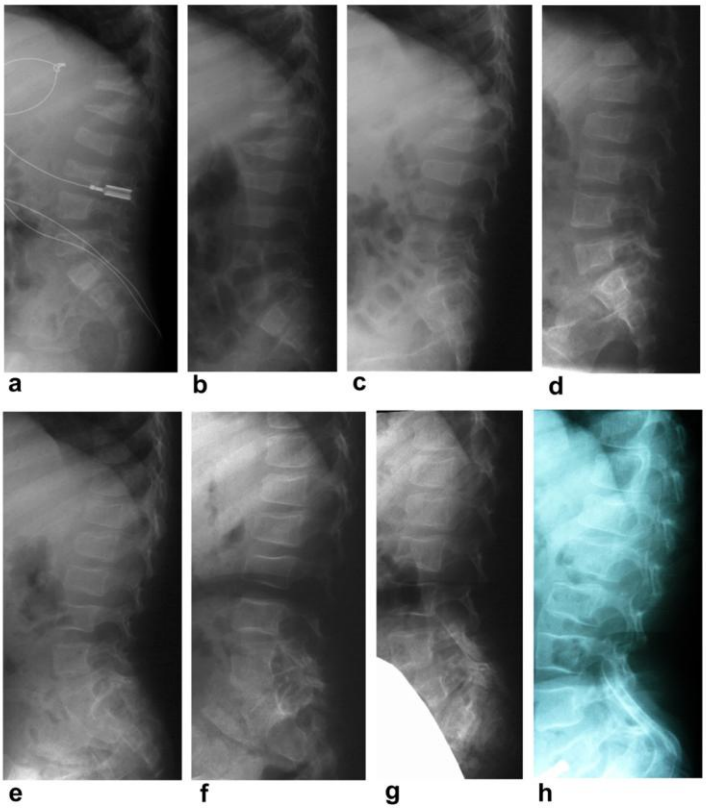
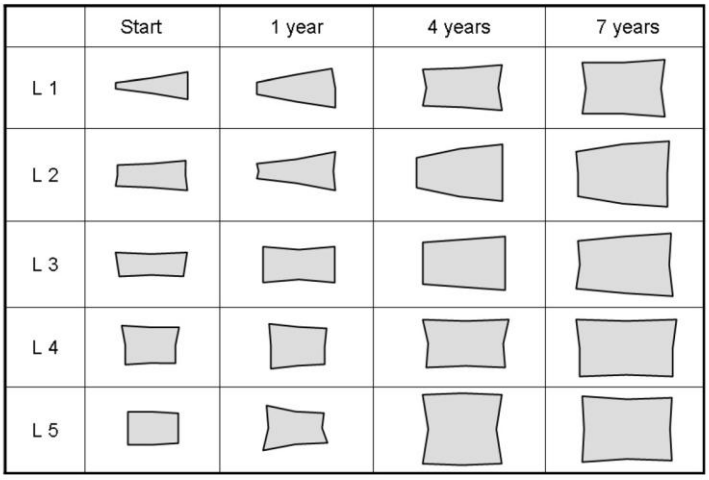


Figure S5. Whole Body DXA Scans of Three Patients with *SERPINF1* Mutations. The scans were performed with the DXA-machine *GE Lunar Prodigy* (GE Healthcare, Pollards Wood, United Kingdom). The pictures show patient #2 at the age of 15.9 years, patient #3 at 5.9 years, and patient #4 at 7.8 years. All patients display deformities of their humeri, which is a rather rare finding in patients with OI. Note that all received intramedullar rodding of their femura due to fractures and deformities. **Patient #2** (left picture) is severely adipose, a fact that decreases the quality of the DXA picture. He had multiple fractures of his forearms, resulting in severe deformities, and bowing of his tibiae. Additionally, he is affected by a severe scoliosis. Bisphosphonate and physical treatment started considerably later (at age 11) than in the other patients. **Patient #3** (middle) received additional rodding of his left tibia. He has only a mild scoliosis, but he displays severe bowing of his forearms, especially on the left side. **Patient #4** (right picture) underwent additional surgery of his left humerus with remaining osteosynthetic material. He also has a mild scoliosis.





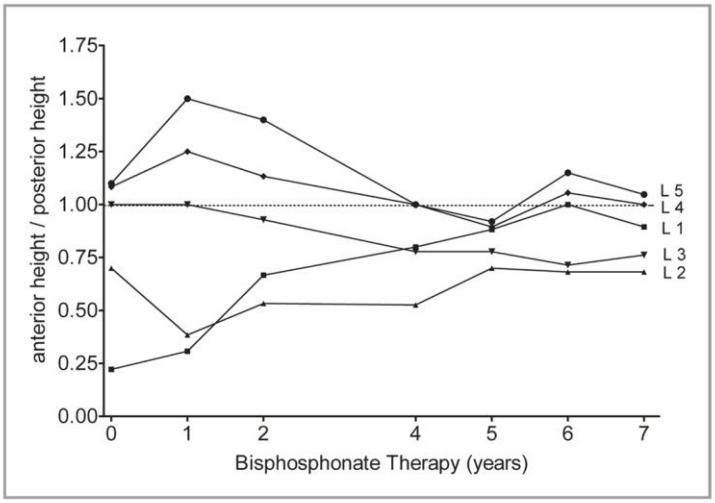


Figure S6. Development of Vertebral Deformities in Patient #4. Upper Panel: (A) Lateral spine x-ray picture of patient #4 at treatment start with i.v. bisphosphonates at the age of 7 months. **(B-H)** Yearly x-ray follow-up over a treatment period of 7 years. The lumbar vertebrae increased in size during this time and showed a reshaping close to the normal cuboid form. These radiographs are representative for the development in the other three OI patients with *SERPINF1* mutations, all of which received i.v. neridronate treatment. Patient #4 was selected for presentation as there is the longest follow-up available. **Middle Panel:** Schematic illustration of vertebral deformities of patient #4 during the treatment period with i.v. bisphosphonates. Vertebrae L1 – L5 on the lateral spine radiograms shown above were analyzed with an established 8-point measurement method and later converted into the schematic view. **Lower panel:** Anterior-posterior index of lumbar vertebra of patient #4 over a period of 7 years. The index is calculated by dividing the measured anterior height (in mm) by the posterior height (in mm). This method is established to quantify the reshaping of vertebral compression fractures in children with OI.[1; 2] A normalization of the anterior-posterior index close to the ratio of 1.0 can be observed over the treatment period.

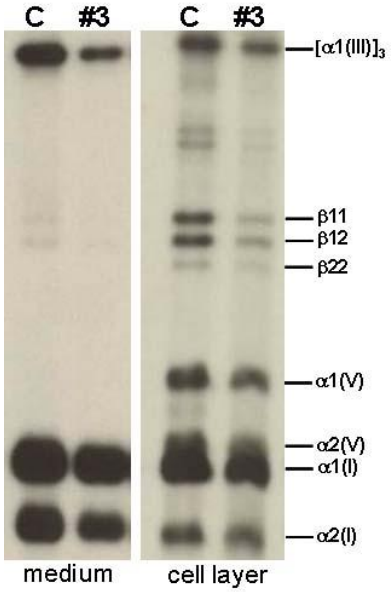


Figure S7. Collagen analysis. Pepsin-treated procollagens from medium and cell layers produced by fibroblasts from patient #3 and a control (C) were electrophoresed on a 5% SDS-polyacrylamide gel, processed for fluorography, and exposed to X-ray films. The α1(I), α2(I), α1(V), α2(V), α3(V), and $[\alpha1(III)]_3$ chains, and β components are indicated. The patient's collagens migrate similarly to those from the control.

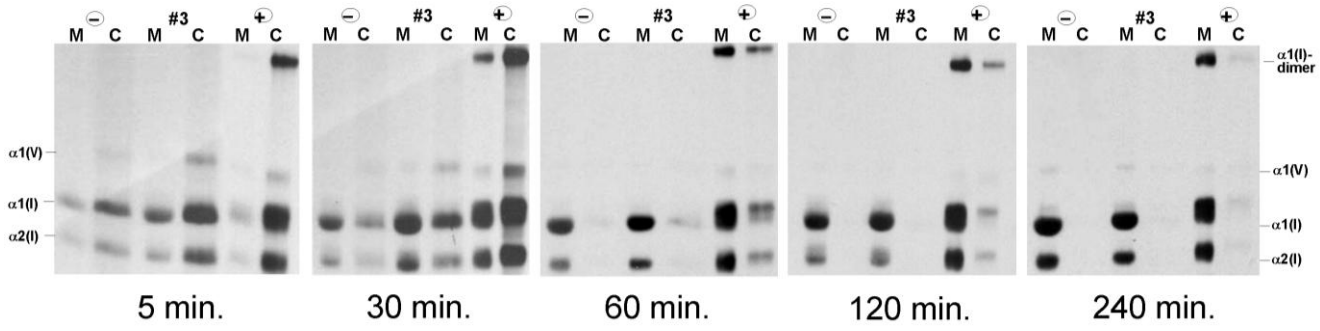


Figure S8. Collagen Synthesis and Secretion Analysis. Pulse-chase experiments were performed as follows: Fibroblasts of a healthy control individual (= "-" in the figure), of patient #3 from this study, and of a type II OI patient (= "+" in the figure, serves as a positive control for delayed collagen secretion due to a p.Gly988Cys mutation in the α1(I) chain[3]) were grown to confluence in DMEM and 10% FCS, and then plated ($2.5x10^5$ cells) in 35-mm dishes. Cells were stimulated overnight with ascorbic acid supplementation (50 μg/ml), and before labelling, medium was changed and DMEM (lacking serum, proline and glycine), supplemented with ascorbic acid (50μg/ml), was added. Cells were labelled 4 h with 20 μCi [^{3}H]-proline and 20 μCi [^{3}H]-glycine, and then chased with fresh DMEM supplemented with 2 mM cold proline and 2 mM cold glycine. The medium (= "M" in the figure) and cell (= "C" in the figure) layers were harvested separately for each sample at 5 min, 30 min, 1h, 2h and 4h chases, and digested with pepsin (50μg/ml) overnight at 4°C. Purified collagens were separated on 5% SDS-PAGE. Note that the fibroblasts of the healthy control individual and of our patient #3 almost completely secreted the α1(I) and α2(I) chains within the first hour, whereas fibroblasts from the positive control still retained collagens intracellularly, a portion of which formed α1(I) di-sulfid bonded dimers.[3]

Table S1. Exome Sequencing Statistics

Total number of sequenced reads	94,888,765
Total number of mapped reads	77,281,481 (81.44%)
Total number of bases mapped	3,727,327,393
Total bases uniquely mapped	3,135,079,346 (84.11%)
Total bases mapping to targets	2,639,906,551 (70.83%)
% targets with 10x coverage	92.90%
Mean target coverage	65.47

Table S2. Prioritization Scheme for Exome Data Analysis

Type of prioritization filter	**Number of remaining variants**
All variants	26922
Coding and canonical splice site (SS) variants after quality filtering (>5 variant reads, >15% variation)	13487
Non-synonymous variants, SS variants	6298
Not in dbSNP 130	657
Not in in-house database	318
Homozygous variants (>80% variation reads) (of which autosomal)	20 (17)
Of which overlap homozygous regions	3

The in-house database consists of data from in-house exome resequencing projects of patients with rare syndromes (425,529 variants), from the 1000 genomes project, and of published data from various studies (3,097,878 variants in total).[4-6]

Table S3. Homozygous Regions Identified when Considering all SNP (dbSNP 130) Genotypes Covered >15-Fold by Exome Sequencing Data

Chromosome	Start Position [hg 18, bp position]	End Position [hg 18, bp position]	Number of SNP marker (tolerated mismatches)	Size [bp]
Chr1	58,869,327	89,371,625	1,918 (6)	30,502,298
Chr6	8,007,840	11,293,519	1,526 (4)	3,285,679
Chr6	37,039,286	42,774,142	698 (2)	5,734,856
Chr11	5,109,976	5,301,257	555 (2)	191,281
Chr11	97,942,525	110,164,199	969 (2)	12,221,674
Chr17	97,058	3,091,589	568 (2)	2,994,531
Chr17	3,299,188	4,983,922	663 (2)	1,684,734

Table S4. Three Novel Homozygous Variants that are Non-Synonymous and Overlap with Homozygous Regions

Chromosome	position	reference	mutation	reads	variation reads	% variation	Abberation	Gene name	Gene id	Reference Amino Acid	Mutation Amino Acid	phyloP	Grantham Score
chr1	62,399,185	C	T	39	39	100	substitution	INADL	NM_176877	T	M	-0.93714	81
chr1	82,181,925	A	G	103	103	100	substitution	LPHN2	NM_012302	Q	R	2.58217	43
chr17	1,625,154	C	G	18	18	100	substitution	SERPINF1	NM_002615	Y	X	0.994354	NA

Supplemental References

1. Land, C., Rauch, F., Munns, C.F., Sahebjam, S., and Glorieux, F.H. (2006). Vertebral morphometry in children and adolescents with osteogenesis imperfecta: effect of intravenous pamidronate treatment. Bone 39, 901-906.
2. Sumnik, Z., Land, C., Rieger-Wettengl, G., Korber, F., Stabrey, A., and Schoenau, E. (2004). Effect of pamidronate treatment on vertebral deformity in children with primary osteoporosis. A pilot study using radiographic morphometry. Horm Res 61, 137-142.
3. Steinmann, B., Rao, V.H., Vogel, A., Bruckner, P., Gitzelmann, R., and Byers, P.H. (1984). Cysteine in the triple-helical domain of one allelic product of the alpha 1(I) gene of type I collagen produces a lethal form of osteogenesis imperfecta. J Biol Chem 259, 11129-11138.
4. Durbin, R.M., Abecasis, G.R., Altshuler, D.L., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A. (2010). A map of human genome variation from population-scale sequencing. Nature 467, 1061-1073.
5. Pushkarev, D., Neff, N.F., and Quake, S.R. (2009). Single-molecule sequencing of an individual human genome. Nat Biotechnol 27, 847-852.
6. Wang, J., Wang, W., Li, R., Li, Y., Tian, G., Goodman, L., Fan, W., Zhang, J., Li, J., Zhang, J., et al. (2008). The diploid genome sequence of an Asian individual. Nature 456, 60-65.