HGG Advances

REPORT

A complex structural variant near *SOX3* causes X-linked split-hand/foot malformation

Elke de Boer,[1,2] Carlo Marcelis,[1] Kornelia Neveling,[1,2] Ellen van Beusekom,[1] Alexander Hoischen,[1,3,4] Willemijn M. Klein,[5] Nicole de Leeuw,[1] Tuomo Mantere,[1,6] Uirá S. Melo,[7,8] Jeroen van Reeuwijk,[1,2] Dominique Smeets,[1] Malte Spielmann,[7,9,10] Tjitske Kleefstra,[1,2,11] Hans van Bokhoven,[1,2,12,13] and Lisenka E.L.M. Vissers[1,2,13,*]

Summary

Split-hand/foot malformation (SHFM) is a congenital limb defect most typically presenting with median clefts in hands and/or feet, that can occur in a syndromic context as well as in isolated form. SHFM is caused by failure to maintain normal apical ectodermal ridge function during limb development. Although several genes and contiguous gene syndromes are implicated in the monogenic etiology of isolated SHFM, the disorder remains genetically unexplained for many families and associated genetic loci. We describe a family with isolated X-linked SHFM, for which the causative variant could be detected after a diagnostic journey of 20 years. We combined well-established approaches including microarray-based copy number variant analysis and fluorescence *in situ* hybridization coupled with optical genome mapping and whole genome sequencing. This strategy identified a complex structural variant (SV) comprising a 165-kb gain of 15q26.3 material ([GRCh37/hg19] chr15:99795320-99960362dup) inserted in inverted position at the site of a 38-kb deletion on Xq27.1 ([GRCh37/hg19] chrX:139481061-139518989del). *In silico* analysis suggested that the SV disrupts the regulatory framework on the X chromosome and may lead to *SOX3* misexpression. We hypothesize that *SOX3* dysregulation in the developing limb disturbed the fine balance between morphogens required for maintaining AER function, resulting in SHFM in this family.

Split-hand/foot malformation (SHFM) refers to a heterogeneous group of rare congenital limb anomalies, characterized by median clefts in hands and/or feet, syndactyly, and aplasia or hypoplasia of the metacarpal, metatarsal, and phalangeal bones, mostly affecting the central rays.[1,2] It arises during embryonic development as a result of failure to maintain normal function of the apical ectodermal ridge (AER), a critical signaling structure that directs morphogenesis of the developing limb along the proximal-distal axis.[3–7] SHFM is a rare disease with a prevalence of 1 per 90,000 live births.[2] The severity of clinical features varies substantially and ranges from syndactyly in mildly affected individuals to monodactyly or aphalangia in its most severe forms.[1] This clinical variability can be observed between members of the same family but even between limbs of a single affected individual.[1,6]

SHFM can be observed in a syndromic context as well as in isolated (non-syndromic) form. To date, OMIM describes over 50 syndromes that involve SHFM, with a range of associated (congenital) anomalies, including neurodevelopmental delay, growth retardation, hearing loss, and craniofacial, ectodermal, and internal organ abnormalities.[2,8,9] For the non-syndromic forms, various clinical criteria have been proposed to assist in its classification, which depend on anatomical and radiographic findings, such as position of the cleft and thumb web, unilateral or bilateral occurrence, involvement of the long bones, and clinical severity.[1] Based on the associated genetic locus, 12 distinct forms of non-syndromic SHFM can be recognized, with an autosomal dominant, autosomal recessive or X-linked mode of inheritance, often showing variable expressivity and incomplete penetrance.[2,10] Contiguous gene syndromes have been implicated in the etiology of SHFM, and for five of these 12 loci, the associated genes have been identified.[2] These include *DLX5* and *DLX6* in SHFM1 (MIM#183600 and MIM#220600), *TP63* in SHFM4 (MIM#605289), *WNT10B* in SHFM6 (MIM#225 300), *ZAK* in split-foot malformation with mesoaxial polydactyly (SFMMP; MIM#616890), and *EPS15L1* in SHFM8 (Table 1).[2,11] Additionally, a combinatorial effect of ectopic misexpression of multiple genes was shown in the etiology of SHFM3.[12]

[1]Department of Human Genetics, Radboudumc University Medical Center, Nijmegen, the Netherlands; [2]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands; [3]Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, the Netherlands; [4]Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, the Netherlands; [5]Department of Medical Imaging, Radiology, Radboud University Medical Center, Nijmegen, the Netherlands; [6]Laboratory of Cancer Genetics and Tumor Biology, Cancer and Translational Medicine Research Unit and Biocenter Oulu, University of Oulu, Oulu, Finland; [7]Max Planck Institute for Molecular Genetics, RG Development & Disease, Berlin, Germany; [8]Institute for Medical and Human Genetics, Charité Universitätsmedizin Berlin, Berlin, Germany; [9]Institute of Human Genetics, University Hospitals Schleswig-Holstein, University of Lübeck and Kiel University, 23562 Lübeck, Kiel, Germany; [10]DZHK (German Centre for Cardiovascular Research), Partner Site Hamburg/Lübeck/Kiel, Lübeck, Germany; [11]Center of Excellence for Neuropsychiatry, Vincent van Gogh Institute for Psychiatry, Venray, the Netherlands; [12]Department of Cognitive Neuroscience, Radboudumc, Nijmegen, the Netherlands
[13]These authors contributed equally
*Correspondence: lisenka.vissers@radboudumc.nl
https://doi.org/10.1016/j.xhgg.2023.100200.
© 2023 The Author(s). This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Table 1. Genetic aberrations in 12 non-syndromic forms of SHFM

Type	MIM/PMID	Locus	Inheritance	Genetic aberrations	Causative gene
SHFM1	#183600	7q21.2q21.3	AD	Duplication/deletion/rearrangement involving *DLX5*, *DLX6*, *DSS1* or possible regulatory elements; SNVs/indels of *DLX5* and *DLX6*	*DLX5*, *DLX6*
SHFM1D	#220600	7q21	AR	SNVs/indels in *DLX5*	*DLX5*
SHFM2	%313350	Xq26.3	XL	Unknown (linkage between DXS1114 and DXS1192, chrX:133,295,286-138,368,235)	Unknown
SHFM3	#246560	10q24	AD	(Micro)duplications and complex rearrangements, including *FGF8, LBX1, BTRC, POLL* and *FBXW4* (=*DACTYLIN*)	Unknown
SHFM4	#605289	3q28	AD	SNVs in *TP63*	*TP63*
SHFM5	%606708	2q31	AD	Unknown, possibly haploinsufficiency of HOXD gene cluster	Unknown
SHFM6	#225300	12q13.12	AR	SNVs/indels in *WNT10B*	*WNT10B*
SFMMP	#616890	2q31.1	AR	Intragenic deletions or SNVs of *ZAK*	*ZAK*
SHFM8	PMID:29023680 PMID:32021595	19p13.11	AR	Indel in *EPS15L1*	*EPS15L1*
		8q21.11q22.3	AR	Unknown	Unknown
SHFLD1	%119100	1q42.2q43	AD	Unknown	Unknown
SHFLD2	%610685	6q14.1	AD	Unknown	Unknown
SHFLD3	#612576	17p13.3	AD	(Micro)duplications involving *BHLHA9*	Unknown

AD, autosomal dominant; AR, autosomal recessive; XL, X-linked.

Finding genetic causes for SHFM is complicated by the rarity of the phenotype, the large number of morphogens associated with limb development, their complex interactions, including intertwining of signaling pathways acting in different spatial dimensions, and the presumed involvement of regulatory elements.[6,13] SHFM2 (MIM%313350) is the only mapped form of non-syndromic SHFM with X-linked inheritance.[14–17] Although X-linked inheritance for SHFM was already suggested in 1978,[17] SHFM2 is merely based on a single, large consanguineous family reported first in 1987.[14] In this family, 36 individuals in seven generations were affected by monodactyly or split-hand and split-foot, with full expression of the trait in hemizygous males and presumed homozygous females, whereas heterozygous females were either unaffected or showed milder phenotypes.[14] Linkage analysis in this large family defined a 5.1 Mb region on Xq26.3 (between DXS1114 and DXS1192, (GRCh37/hg19) chrX:133,295,286–138,368,235).[15,16] The exons and exon-intron boundaries of 19 candidate genes in the linkage region were sequenced, but no relevant variants were identified,[16] suggesting that regulatory elements must be at play in the pathophysiology of SHFM2.

Here, we report on a family in which five individuals are affected by non-syndromic SHFM, consistent with an X-linked inheritance pattern. We applied a combined approach of microarray-based copy number variant (CNV) analysis, fluorescence *in situ* hybridization (FISH), optical genome mapping (OGM), and whole genome sequencing (WGS). The combination of these technologies enabled us to identify a complex structural variant (SV), consisting of a 165-kb duplicated fragment originating from chromosome 15, to be inserted in inverted orientation at the site of a 38-kb deletion on the X chromosome near the previously described SHFM2 locus and *SOX3*. We hypothesize that perturbations of regulatory elements lead to dysregulated *SOX3* expression affecting AER maintenance in the limb bud and causing SHFM in this family.

Individuals from the reported family first received genetic counseling at a university medical center in the Netherlands over 20 years ago, and were followed up at low frequency thenceforth. In this four-generation kindred (Figure 1A), a total of five individuals were affected by congenital unilateral or bilateral non-syndromic split-hand malformation, but without clinical abnormalities of the feet. Evaluation of the pedigree indicated an X-linked mode of inheritance, or less likely, autosomal dominant inheritance with reduced penetrance. The Radboudumc review board approved the study (2019–5554) and affected individuals provided written informed consent.

Individual I-1 was a male affected by bilateral split-hand malformation (Figure 1B), who deceased at 82 years of age. He had three sons and six daughters from a non-consanguineous relationship, all of whom were clinically unaffected. Five of his nine children had offspring, but only his two oldest daughters (individual II-4 and II-7) had offspring with symptoms fitting the SHFM spectrum. Individual III-6, the 52-year-old daughter of individual II-4, showed a mild phenotype consisting of unilateral oligodactyly. She presented with four fingers in cleft-shape on

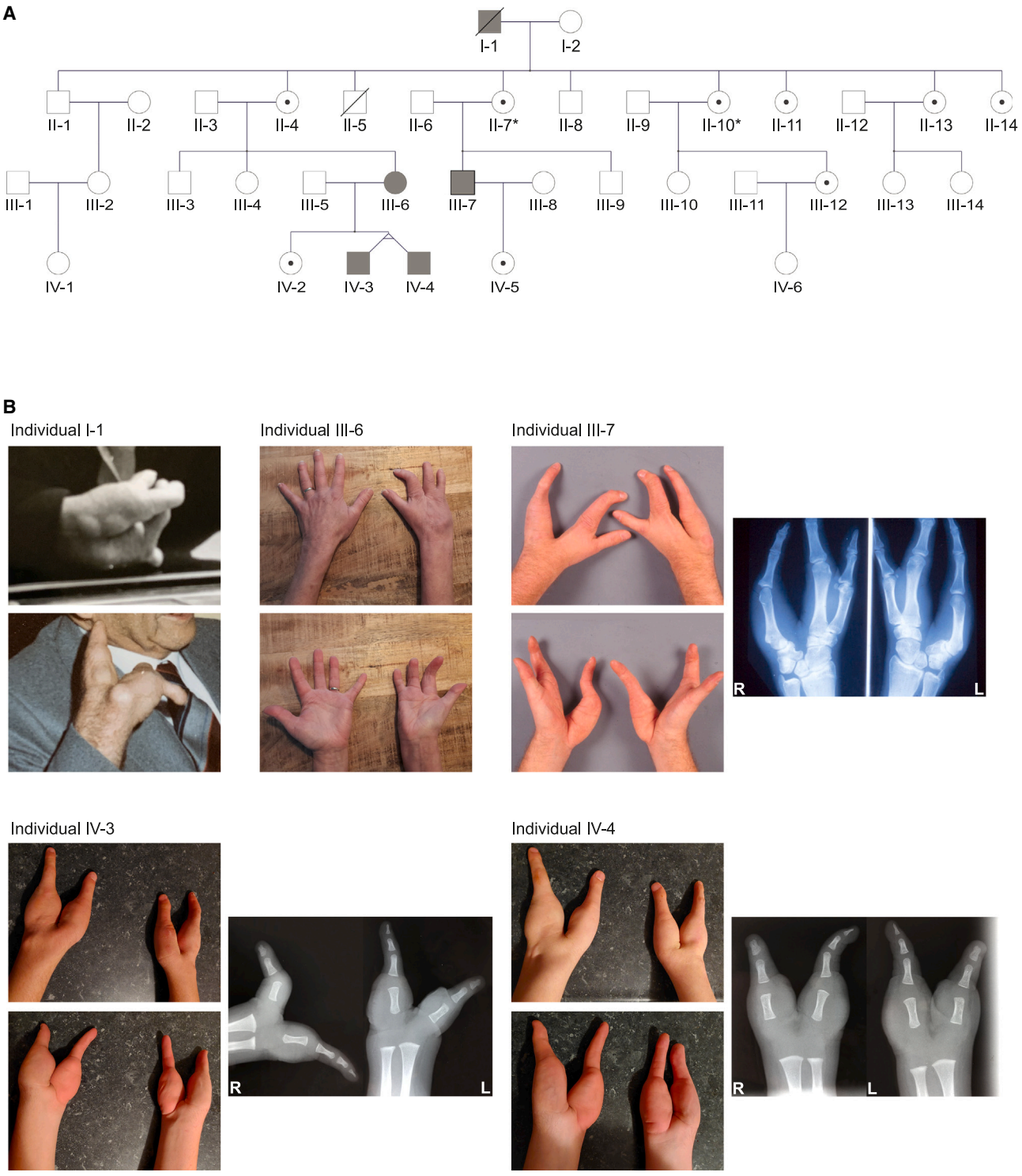


Figure 1. A family with X-linked split-hand malformation

(A) Pedigree of the family showing five affected individuals (marked in gray) in the first, third, and fourth generation, and unaffected carriers (indicated by a dot) in the second, third, and fourth generation, consistent with an X-linked inheritance pattern. Radiographic feet abnormalities are indicated by an asterisk.

(B) Photographs of upper limbs of individuals I-1, III-6, III-7, IV-3, and IV-4. Individual I-1 exhibited bilateral V-shaped hands with median clefts. Individual III-6 had right-sided split-hand with four fingers on the affected hand. Individual III-7 showed bilateral split-hand malformation, with abnormal carpal bone and absence of metacarpal and phalangeal bones of the third and fourth rays as seen on radiography. Monozygotic twins IV-3 and IV-4 are affected by bilateral bidactylous split-hand, characterized by deep V-shaped median clefts. Radiography of the hands at age 1 year and 9 months showed age-related absence of calcification of the carpal bones (therefore, the formation cannot be assessed based on these images) and complete absence of metacarpals and phalanges of the second, third, and fourth digits. The thumb consisted of a normal metacarpal and proximal phalangeal bone with a bifid distal phalangeal bone, and the fifth ray showed normal metacarpal and triphalangeal formation positioned in abnormal ulnar rotation.

the right hand (Figure 1B), requiring surgical intervention at age 18 months. Her brother (III-3), sister (III-4), and daughter (IV-2) were unaffected, but both her monozygotic twin sons (individual IV-3 and IV-4) exhibited bilateral split-hand malformation. The 22-year-old twins had very similarly affected hands characterized by absence of the three central digits, leading to a deep median cleft between the first and fifth rays, with the fifth ray in ulnar rotation. Radiography of hands in the posterior-anterior direction at age 1 year and 9 months showed a deep V-shaped cleft with complete absence of calcification of the carpals, as well as complete absence of metacarpals and phalanges of the second, third, and fourth rays. The thumb consisted of a normal metacarpal and proximal phalangeal bone, with a bifid top phalanx. The fifth ray comprised normal metacarpal, and normal proximal, mid and end phalangeal bones, positioned in abnormal ulnar rotation (Figure 1B). As this ulnar rotation interfered with grip and pinch function of hands, and herewith development of fine motor skills, the twins underwent a rotational osteotomy at the base of the fifth metacarpal of the right (dominant) hand at 2 years and 9 months to improve functional anatomy of the hand. They received hand therapy of a specialized occupational therapist. At age 3 years and 10 months, they could perform fine motor skills at a slow but age-appropriate level. Besides a diagnosis of attention-deficit/hyperactivity disorder in both, they had a normal development and there were no additional congenital abnormalities or medical problems. Their 50-year-old second cousin, individual III-7, son of individual II-7, showed a phenotype of bilateral split-hand malformation with absence of the third and fourth rays, including absence of metacarpal and phalangeal bones, as well as abnormalities of the carpal bones (large scaphoid, fused trapezoid and trapezium, absent capitate, rotated hamate, small triquetrum, pisiform and lunate), as seen on radiography (Figure 1B). His feet showed mild bilateral cutaneous 2-3-syndactyly (Figure S1). There were no other congenital anomalies or medical problems, and his brother (III-9) and daughter (IV-5) were unaffected. To understand segregation of the trait in this family, individuals II-1, II-4, II-7, II-10, II-11, III-4, and III-9 were screened for subclinical hand and feet abnormalities by radiography. For individual II-7, this showed bony coalition of talus and navicular bone and shortened second metatarsal bone of the right foot. Individual II-10 had bilateral shortened second metatarsal bones. All other imaging results were reported normal.

Over the course of the years, genetic testing in this family included chromosome X-specific exome sequencing, and targeted sequencing of *TP63*, associated with autosomal dominant SHFM4 with reports of remarkable non-penetrance,[4,10] but neither identified a causative variant. Additionally, linkage analysis failed to identify a locus with a significant logarithm of odds (LOD) score, although analysis with markers from the SHFM2 locus had revealed that linkage in the present family would be compatible with the SHFM2 locus identified in the previously published family.[14–16] Because of the family's ongoing search for a genetic diagnosis and CNVs being implicated in several forms of SHFM, we used DNA isolated from blood of individual III-7 for microarray-based CNV analysis (Affymetrix CytoScan HD 2.6M), using routine diagnostic procedures[18] and reporting based on genome build GRCh37/hg19. This analysis identified two rare CNVs, of which the first was an intragenic deletion of 175 kb in 10q26.2, arr [GRCh37/hg19] 10q26.2(129,012,678-129,186,196)x1, including *DOCK1*. The second CNV was an interstitial gain of 161 kb in 15q26.3, arr [GRCh37/hg19] 15q26.3(99,796,482-99,957,320)x3, encompassing the gene *LRRC28* and pseudogene *HSP90B2P*. These CNVs were neither reported before by our diagnostic laboratory, nor in online CNV databases of healthy or affected individuals.[19–21] In addition, no disease-gene associations were known for *DOCK1, LRRC28,* or *HSP90B2P* in OMIM.[8,9] Therefore, both CNVs were classified as variants of uncertain significance. To gain further insights in the clinical relevance of these CNVs, we continued with array-based CNV analysis in an additional affected relative (individual IV-4), who was, together with his monozygotic twin brother, the most distant affected individual in the pedigree (separated by five meiotic cell divisions from III-7). In individual IV-4, the 10q26.2 deletion was not present, but the 15q26.3 gain as found in individual III-7 was detected (Figure 2A). Based on this observation and the pedigree information, his affected twin brother (individual IV-3), affected mother (individual III-6), and affected great-grandfather (individual I-1) are likely carriers of the 15q26.3 gain, which herewith segregates with the SHFM phenotype. However, an aberration on one of the autosomes would not explain the classic X-linked inheritance in this family, as evaluated from the absence of SHFM in the entire second generation, although there are at least two obligate female carriers of the 15q26.3 gain (individual II-4 and II-7), together with the observation of a milder phenotype in the only affected female (individual III-6) compared with her male counterparts. We thus hypothesized that the observed 15q26.3 gain could be part of a more complex SV, possibly involving the X chromosome.

We continued with karyotyping and metaphase FISH experiments in duplicate on cultured Epstein-Barr virus immortalized lymphoblastoid cells (EBV-LCLs) of affected male individual III-7, using RP11-668P3 (Empire Genomics, Williamsville, NY, USA) as a region-specific probe for 15q26.3 and probe CEP-15 (Vysis, Abbott, Abbott Park, IL, USA) for the centromere of chromosome 15 or probe CEP-X (Vysis, Abbott, Abbott Park, IL, USA) for the centromere of chromosome X. Conventional karyotyping showed a normal male karyotype as the copy number gain of 15q26.3 is smaller than the detection limit of G-banding (5–10 Mb). FISH, however, showed an abnormal pattern with the duplicated segment of 15q26.3 resulting in a third signal on the long arm of the X chromosome (Figure 2B). The exact locus of insertion of the duplicated fragment

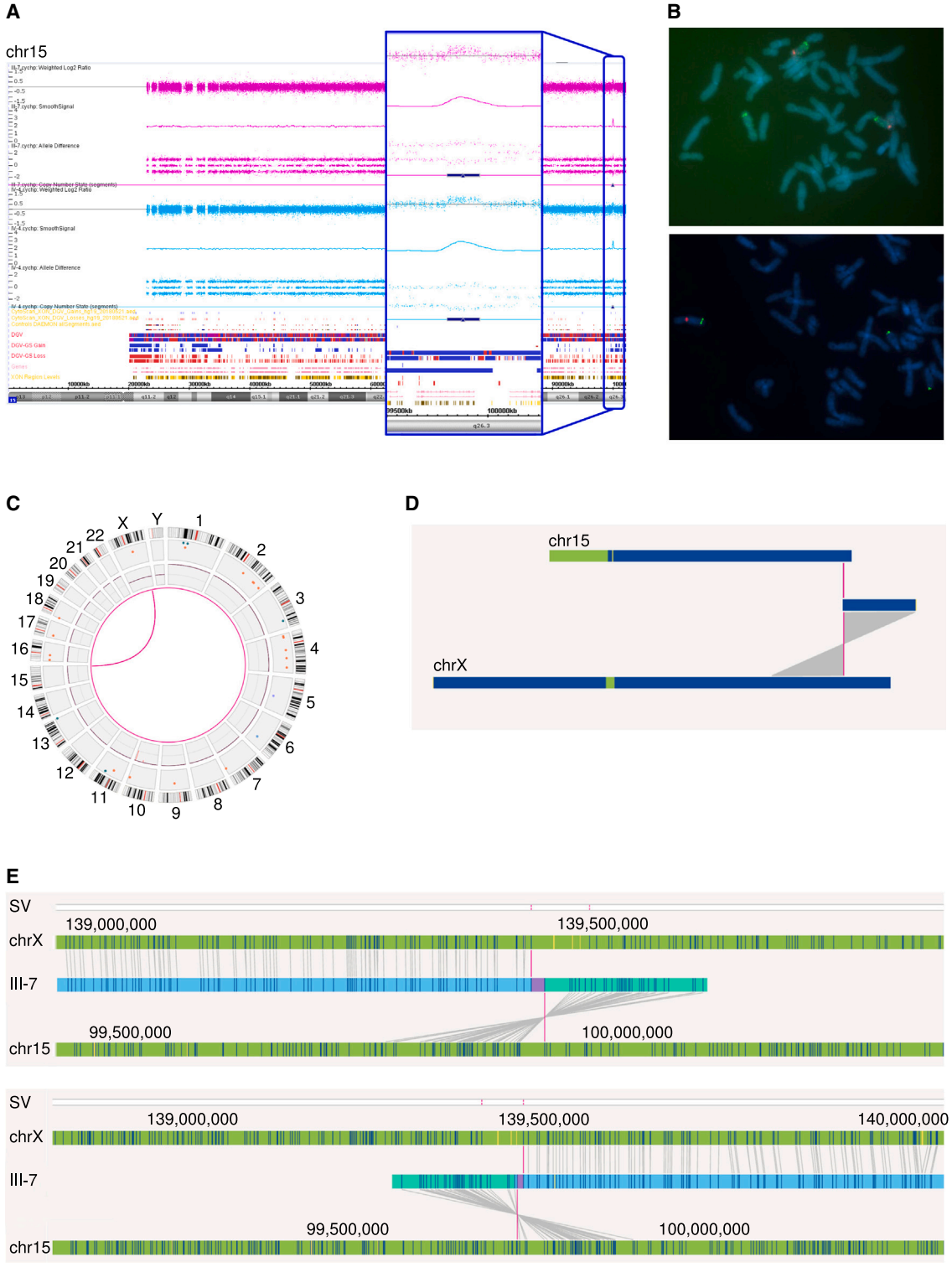


Figure 2. Microarray-based CNV analysis, FISH, and OGM detected an inverted 15q26.3 gain inserted in Xq27.1

(A) Microarray-based CNV analysis of individual III-7 and IV-4 shows a 15q26.3 gain of ~161 kb.

(B) FISH experiments show this 15q26.3 gain is inserted in the long arm of the X chromosome. Upper image: the centromeres of both chromosomes 15 (probe CEP 15, Vysis) are labeled in red. In green, probe RP11-668p3 (Empire Genomics) marks the duplicated region that is present on both chromosomes 15 with a third signal on Xq27, pinpointing that the gain of chromosome 15 is inserted into the distal long arm of the X chromosome. Lower image: metaphase FISH results with in red the centromere of chromosome X (probe CEP-X, Vysis) and in green probe RP11-668p3, that is present on both chromosomes 15 and again showing a third signal on the X chromosome.

(C) Circosplot showing chromosomes X and 15 connected with a pink line, representing a translocation event.

(D) Whole chromosome view from OGM, indicating a translocation between chromosome 15 and the X chromosome with the pink line, t(X; 15)t(q27.1; q26.3).

(legend continued on next page)

of chromosome 15 into chromosome X could not be determined, but based on the FISH experiments, this was most probably at Xq27 (46,der(X)ins(X; 15) (q2?7; q26.3q26.3), Y. ish der(X) (q2?7) (RP11-668P3+)).

Although FISH confirmed our hypothesis of a structural chromosome rearrangement involving not only chromosome 15 but also the X chromosome, it lacked the resolution to elucidate the exact genomic architecture underlying the variant. Therefore, we continued with a two-pronged approach consisting of parallel OGM and WGS to gain insights in the breakpoints and orientation of the individual elements of the SV. OGM technology (Bionano Genomics) generates ultra-long DNA fragments ($\geq$150 kb) with fluorescent labels targeting a 6-nucleotide motif occurring non-randomly throughout the genome, resulting in a *de novo* genome assembly with "barcode-like" visualization of the genome of interest, and a more than 10,000-fold higher resolution than conventional karyotyping.[22,23] Comparison of the occurrence of this 6-nucleotide motif between the genome of interest and a reference allows for detection of CNVs and (complex) SVs.[22,23] We isolated ultra-high molecular weight DNA from EBV-LCLs of individual III-7, followed by the previously published OGM workflow,[22,23] including *de novo* genome assembly and variant calling of CNVs and SVs. We obtained an N50 molecule length (of molecules larger than 150 kb) of 0.321 Mbp and a map rate of 85.9%, resulting in an effective coverage of 224-fold.[22] Variants were prioritized as previously described.[22] OGM readily revealed the variant of interest in more intricate detail: the gain originating from chromosome 15 ([GRCh37/hg19] chr15:99795709-99959476) inserted in an inverted fashion at the Xq27.1 locus between positions [GRCh37/hg19] chrX:139468438 and chrX:139527176 (Figures 2C–2E and S2, and Table S1A). Additionally, visual inspection of output from the CNV algorithm suggested the presence of a small deletion on Xq27.1 at the locus of the insertion, although not detected by the CNV calling software.

WGS (BGISeq500, PE100) was performed in parallel to OGM on DNA derived from blood of individual III-7. Prioritization of SNVs/indels, CNVs and SVs from WGS data did not reveal any other compelling variants,[24,25] but allowed us to refine the breakpoints of the SV at single nucleotide resolution. We obtained a mean coverage of 40-fold, and data were interpreted with a targeted approach, prioritizing CNVs (Control-FREEC[26] and Canvas[27]) and SVs (Manta[28]) on 15q26.3 and Xq27.1 using an ANNOVAR-based[29] annotation pipeline. Both CNV calling algorithms identified the rare 15q26.3 gain observed in previous experiments (Tables S1B and S1C). Additionally, both tools detected a rare ~38 kb Xq27.1 deletion (Tables S1B and S1C). Further analysis of the Manta output suggested the duplicated fragment of 15q26.3 to be inserted in inverted orientation at the locus of the Xq27.1 deletion, with inclusion of four nucleotides of unknown origin at the right breakpoint on the X chromosome (Table S1D). Visual inspection of WGS data in the Integrative Genomics Viewer (IGV)[30] (Figures 3A–3D, S3, and S4) and analysis of breakpoints using BLAT on DNA[31] confirmed the 165-kb duplicated fragment from chromosome 15 ([GRCh37/hg19] chr15:99795320-99960362dup) inserted at the site of a 38-kb deletion ([GRCh37/hg19] chrX:139481061-139518989del) on the X chromosome (Figure 3A). Both breakpoints were confirmed by PCR and Sanger sequencing. Segregation analysis by breakpoint-spanning PCRs, including all family members for whom DNA was available (23 individuals), confirmed the presence of the SV in all affected individuals (III-6, III-7, IV-3, and IV-4) and in all putative carriers (II-4 and II-7). In addition, it revealed seven additional clinically unaffected female carriers (II-10, II-11, II-13, II-14, III-12, IV-2, IV-5; Figure S5). From these data, we concluded that the SV segregates with the SHFM phenotype (Figure 1A) and that expression of clinical features is consistent with an X-linked inheritance pattern.

Of all female carriers, only individual III-6 showed a mild unilateral split-hand malformation, whereas the other nine did not. Additionally, individual II-7 and II-10 exhibited mild subclinical radiographic foot abnormalities, although it remains uncertain whether these foot abnormalities are implicated in SHFM2 or occur randomly. We hypothesized that differences in X-inactivation could play a role in phenotypic variability between female carriers and continued with X-inactivation studies by quantifying methylation of the human androgen receptor locus including all 10 confirmed female carriers (II-4, II-7, II-10, II-11, II-13, II-14, III-6, III-12, IV-2, and IV-5). However, this analysis did not yield any conclusive results to understand the difference in phenotypic presentation among female carriers (Figure S6; supplemental information).

Collectively, the proximity of the here described SV to the known SHFM2 locus, the absence of variation in this locus in control populations, and the segregation pattern observed in the family, strongly suggest that this variant is causative for the SHFM in this family. We therefore pursued further biological and functional understanding of this variant by focusing on three aspects, being (1) the genomic contents of the duplicated fragment of chromosome 15, (2) the genomic contents deleted on chromosome X, and (3) the disruption of regulatory context by the insertion of chromosome 15 material into the X chromosome.

(E) OGM results from individual III-7 shows the 15q26.3 gain inserts in an inverted fashion on the X chromosome. The upper image illustrates the left breakpoint, the lower image illustrates the right breakpoint. In both these images, the upper green bar indicates the X chromosome, the middle blue bar indicates the genome of individual III-7, and the lower bar indicates chromosome 15. Each of the lines in these bars represents the fluorescent labels targeting a 6-nucleotide motif that occurs randomly throughout the genome. The lines connecting labels found in individual III-7 with labels on the X chromosome and on chromosome 15 indicate the software recognizes these labels are the same.

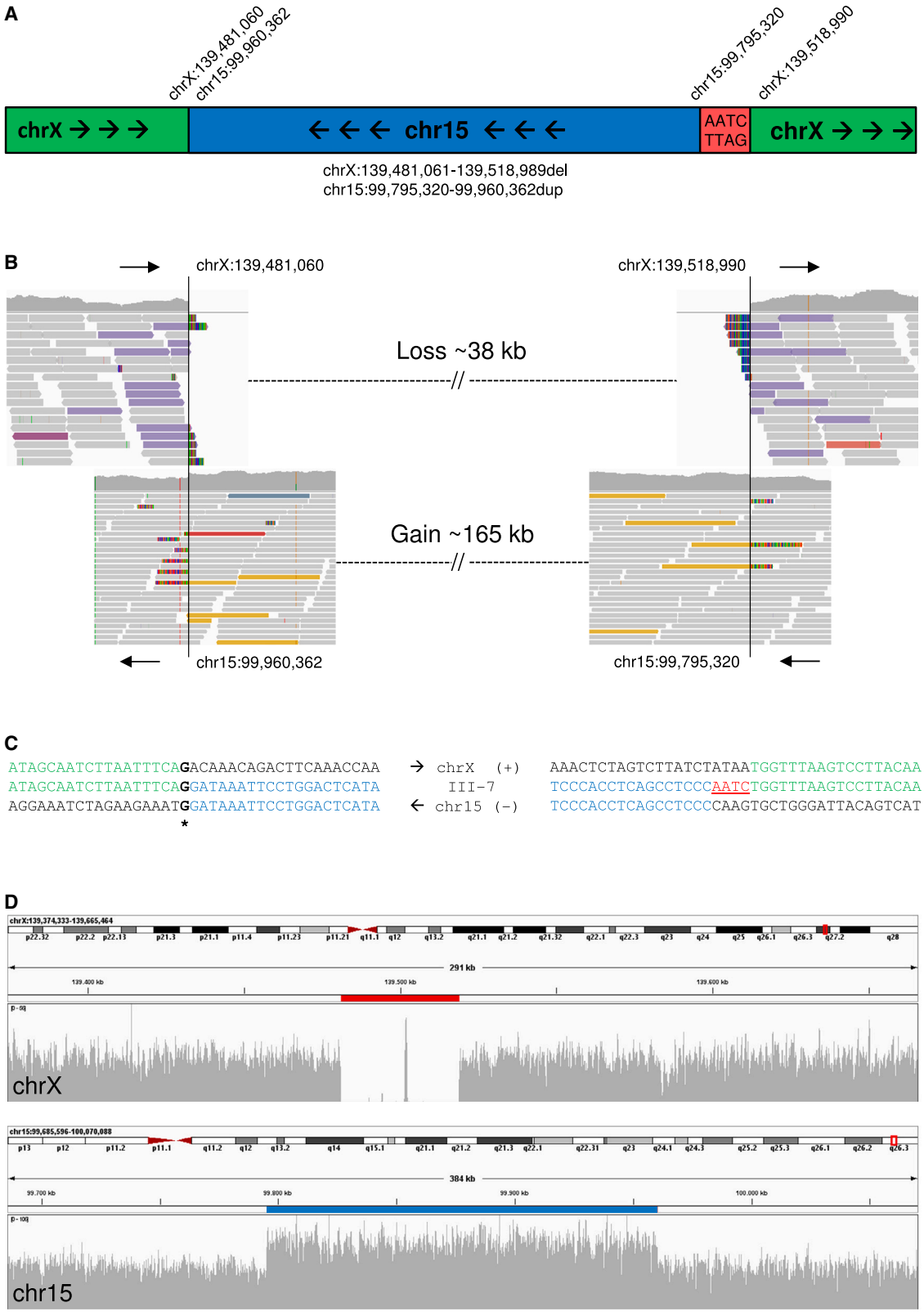


Figure 3. Detection of SV breakpoints at single nucleotide resolution with WGS

(A) Schematic representation of the structural variant, comprising an inverted 165-kb gain from 15q26.3 inserted at the site of a 38-kb deletion on Xq27.1.

(B) Visualization of alignment of reads from WGS in IGV. The proximal (left) breakpoint on the X chromosome attaches to the distal end of the duplicated fragment of chromosome 15, whereas the distal (right) breakpoint on the X chromosome is connected with the proximal end of the duplicated fragment of chromosome 15.

(legend continued on next page)

The duplicated segment of chromosome 15 contains part of *LRRC28* and pseudogene *HSP90B2P*. Despite *LRRC28* not being triplosensitive (pTriplo 0.07),[32] and the duplicated protein-coding sequencing only involving exons 2–10, we first hypothesized that this duplication could play a role in SHFM pathophysiology in this family (Figure S7A). *LRRC28* is ubiquitously expressed and encodes Leucine Rich Repeat Containing protein 28 (LRRC28), characterized by Leucine Rich Repeats (LRRs). Although its exact functions are largely uncharacterized, LRRC28 was reported to be involved in RAS-mediated signaling.[33,34] We measured *LRRC28* expression with quantitative real-time PCR experiments on RNA isolated from cultured EBV-LCLs of individuals III-6, III-7, IV-3, and IV-4, alongside three male and three female healthy unrelated controls, and found that expression of *LRRC28* is not affected in individuals carrying the variant (Figures S7B and S7C), suggesting that partial *LRRC28* duplication is, as anticipated, not the underlying molecular cause of SHFM.

The 38-kb deletion on Xq27.1 does not contain any (protein-coding) genes. Interestingly, Xq27.1 is prone to structural variation, and insertions occurring at or near a quasi-palindromic sequence (ChrX:139,502,865–139,503,044) are associated with nine other distinct disease phenotypes.[35] For two of these disease phenotypes, the underlying pathophysiological mechanism is shown to be dysregulation of a nearby gene, being *FGF13*[36] and *SOX3*,[37] respectively. We thus progressed to our hypothesis that the SV disrupts the genomic context, spatial organization and possibly regulation of genes in close proximity to the breakpoints. Assessing the region involved (300 kb up- and downstream of breakpoints) identified the protein-coding gene *SOX3* and two non-coding genes of which little is known, including the small nuclear RNA *U7* (*LOC124905265*; RF00066) and the long non-coding RNA *LINC00632*.[38] *SOX3*, located 67 kb downstream of the SV breakpoint encodes SRY-Box Transcription Factor 3 (SOX3), a member of the SOX family of transcription factors, which includes important regulators of cell fate during embryonic development.[39] *SOX3* is predominantly expressed in the fetal brain and spinal cord,[40] where it is implicated in central nervous system development,[41–44] by regulating gene expression in neural progenitor cells.[44] Duplication of *SOX3* has been associated with intellectual disability with isolated growth hormone deficiency (MIM#300123)[45] and panhypopituitarism (MIM#312000). Additionally, an SV located 67 kb downstream of *SOX3* is implicated in hypoparathyroidism,[46] and *SOX3* dysregulation caused by insertion of chromosome 1 material at Xq26.3 is implicated in 46,XX male sex reversal (MIM#300833).[37]

SOX3 is a plausible candidate in the etiology of SHFM2, as other members of the SOX family of transcription factors, such as *SOX5, SOX6, SOX8, SOX9,* and *SOX10*, are implicated in limb chondrogenesis.[47] In addition, induced ectopic expression of *SOX3* in the limb bud alters expression patterns of Sonic Hedgehog (SHH)-regulated genes,[48] and during early development, *SOX3* is co-expressed at the neural plate border with *TP63*,[49] the gene implicated in SHFM4 (MIM#605289).[4] SOX3 and TP63 share numerous transcription factor binding sites for neural plate genes, with TP63-dependent inhibition of SOX3 binding.[49] Together, these proteins are thought to define the distinction between surface ectoderm and neuroectoderm.[49] In the limb bud, TP63 is an important factor for AER maintenance, via regulation of *DLX5, DLX6,* and *FGF8* expression.[50,51] *Tp63*$^{-/-}$ mutant mice display defective AER maintenance, lack several limb components, and show reduced expression of *fgf8* in the limb bud,[52] a gene involved in AER maintenance and thought to regulate regionalized expression of *SOX3* (Figure 4A).[53]

Taking this information together, we hypothesized that the SV interferes with tissue-specific and time-dependent regulation of *SOX3* expression, although it is also a possibility that genes even more distant from the breakpoints are dysregulated.[35] Unfortunately, *SOX3* is neither expressed in any accessible tissue nor in cultured fibroblasts or EBV-LCLs, limiting the possibility to test this hypothesis. Additionally, it is uncertain whether measuring gene expression would accurately reflect the *in vivo* situation if not taking into account the relevant developmental time point and cell type. Regardless, given the observed *SOX3* misexpression in EBV-LCLs in 46,XX male sex reversal associated with an insertion near *SOX3*,[37] we measured *SOX3* expression with quantitative real-time PCR experiments on RNA isolated from cultured EBV-LCLs of individuals III-6, III-7, IV-3, IV-4, and three male and three female healthy unrelated controls, also taking along *FGF13*.[36] Expression levels of *SOX3* and *FGF13* in EBV-LCLs were too low in affected individuals as well as controls (data not shown), herewith not providing any further insights in our hypothesis. We therefore continued by *in silico* analysis reasoning that the SV would either disrupt topological associating domain (TAD) structures, and/or affect SOX3 transcription factor binding sites (TFBS). We found that the SV deletes four TFBSs on the X chromosome,[54–56] two of which have *SOX3* as a target gene. Additionally, four TFBSs locate in a 100-kb region proximal to the locus of the inserted fragment on Xq27.1, and for three of these, *SOX3* is the target gene. Potentially, the SV affects *SOX3* expression as a result of loss of the TFBS and/or increased distance or disturbed genomic 3D organization between target gene

(C) Sequence alignment at both breakpoints, showing insertion of four nucleotides of unknown origin at the distal breakpoint on the X chromosome.
(D) Coverage data obtained from WGS shows a 38-kb deletion on the X chromosome (upper image, red line) and a 165-kb duplication on chromosome 15 (lower image, blue line).

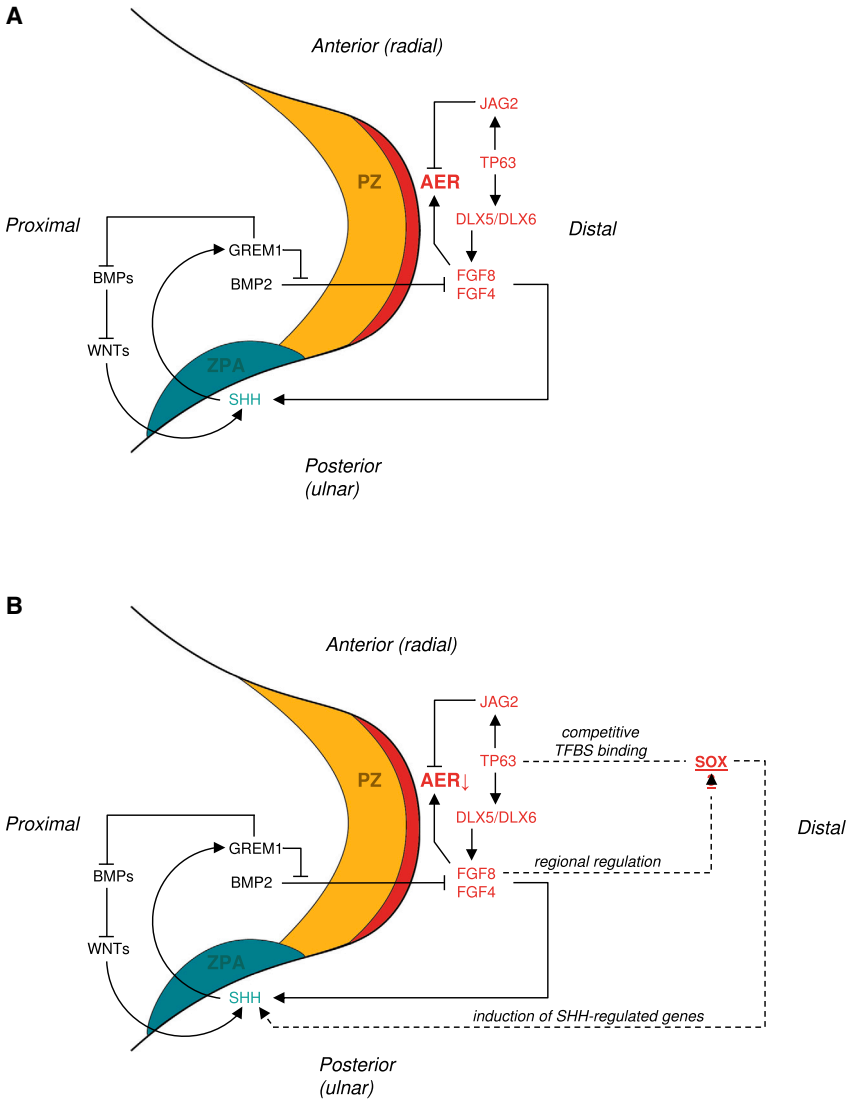


Figure 4. Misexpression of *SOX3* is hypothesized to disturb AER maintenance

(A) Schematic representation of the developing limb bud in early limb formation with simplified representation of genes involved in AER maintenance. The AER in red orchestrates outgrowth of the limb along the proximal-distal axis. Mesenchymal cells in the progress zone (PZ) in orange stay in an undifferentiated state under AER control. The zone of polarizing activity (ZPA) in turquoise regulates development along the anterior-posterior (radial-ulnar) axis. Factors determining dorsal-ventral growth are not shown. Signaling pathways associated with polarizing activity along the three axes are intertwined. Patterning along the anterior-posterior axis is mediated by SHH, also required for integrity of the AER. Patterning along the proximal-distal axis is regulated by FGFs, including *FGF4* and *FGF8*, maintaining AER integrity. *TP63*, the gene underlying SHFM4, stimulates expression of *FGF8*, via induction of *DLX5* and *DLX6* (implicated in SHFM1), herewith contributing to AER maintenance. *TP63* can also negatively regulate AER maintenance via induction of *JAG2*. (B) We hypothesize that *SOX3* misexpression in the developing limb bud disturbs the balance between signaling molecules needed for AER maintenance, for example by competitive TFBS binding with TP63 or by induction of SHH-regulated genes.

SOX3 and TFBS located proximal of the SV (Figures 5A and S8, and Table S2A). Assessing whether a similar SV is implicated in the etiology of SHFM2 in the previously published family[14] would be an interesting future prospect, given that the linkage region from literature[15,16] and the SV reported here are approximately 1.1 Mb apart but partly locate to the same TAD (Figure S8).

The duplicated 15q26.3 fragment inserted on the X chromosome also contains regulatory sequences, such as an enhancer (GH15J099385) that normally interacts with the *LRRC28* promoter located upstream on chromosome 15 (GH15J099248; Figures 5B and S9). This enhancer contains TFBS of *CTCF, HNF4A, FOSL2, JUND, JUNB,* and *SMARCA4* (Table S2B), of which *SMARCA4* is implicated in tissue-specific developmental gene regulation[57] and limb development in mice, with abnormal AER morphology and defective hindlimb development associated with *SMARCA4* ablation (MGI:3606859).[58] As enhancers are known for their functional autonomy, also when combined with heterologous promoters and genes,[59] one could speculate that the inserted enhancer may retain its transcription-activating function in the new genomic context and as a result might cause *SOX3* misexpression (Figure 5C). Additionally, the duplicated region contains 10 candidate *cis*-Regulatory Elements that are active in the human developing limb (Table S3).[60] Therefore, we hypothesize that misexpression of *SOX3* in the limb bud disturbs the fine balance required for AER maintenance as orchestrated by TP63 and FGF8, possibly resulting from competitive binding between SOX3 and TP63, or resulting from altered SHH-activity (Figure 4B),[5–7,48,50,51,53,61,62] leading to SHFM in this family.

In summary, we describe a family with X-linked non-syndromic SHFM. By combining older but well-established approaches with innovative genomics technologies, we could delineate the exact genetic aberration underlying SHFM in this family. The phenotype is caused by a complex SV, comprising an inverted 165-kb 15q26.3 gain inserted at the site of a 38-kb Xq27.1 deletion. As the SV does not alter expression of the partly duplicated protein-coding gene *LRRC28*, we anticipate that the SV influences regulatory functions of the affected region on the X chromosome, including *SOX3* as a plausible candidate gene. Hence, we hypothesize that the SV causes misexpression of *SOX3* in the developing limb bud, resulting in failure to maintain AER function.

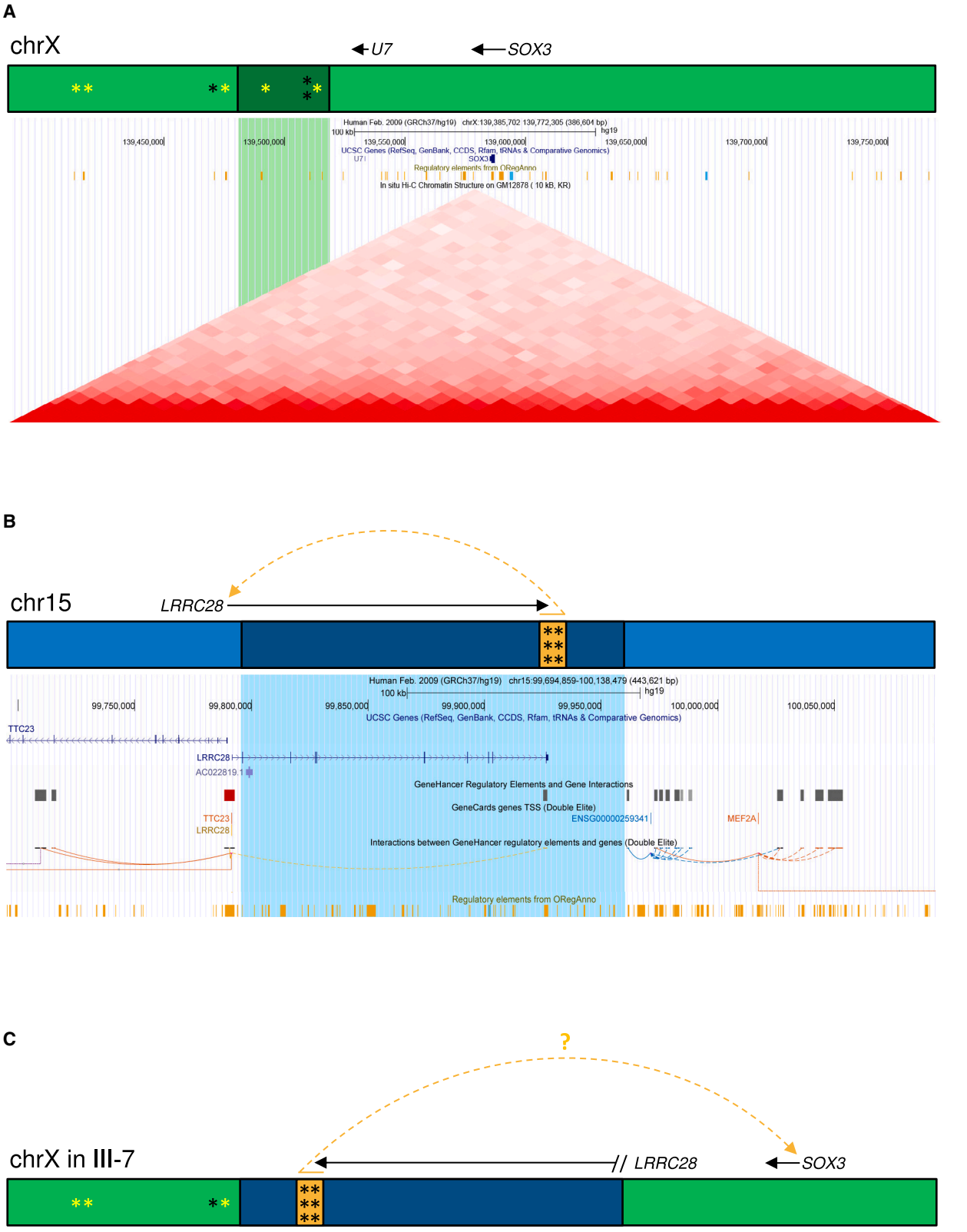


Figure 5. The SV is predicted to disturb regulation on the X chromosome

(A) Schematic representation of genes and regulatory elements on the X chromosome (upper image), and screenshot from the UCSC genome browser with the 38-kb deletion on the X chromosome in green (lower image). The deleted fragment of the X chromosome (dark green in the schematic representation) is located ~67 kb proximal of *SOX3* and contains four TFBSs (asterisk), of which two target *SOX3* (yellow asterisk; OREG1571551, TFBS of *FOXA1*; OREG1521521, TFBS of *ESR1*). In the 100-kb region proximal to the deleted fragment on the X chromosome, there are four TFBSs as well, of which three target *SOX3* (OREG1412659, TFBS of *E2F1*; OREG1412658 TFBS of *E2F1*; OREG1571552, TFBS of *FOXA1*).

(legend continued on next page)

Data and code availability

Disclosure of the full (raw and annotated) datasets are restricted by the level of consent for data sharing provided by the participants. Requests for more detailed analyses and (partial) data access should be discussed with the corresponding author.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.xhgg.2023.100200.

Acknowledgments

We are very grateful to all individuals from the family for their participation in this study. We would like to thank Joost Kummeling, Guillaume van de Zande, Martina Ruiterkamp-Versteeg, Ronald van Beek, Michael Kwint, Michiel Oorsprong, Ellen Kater-Baats, Jordi Corominas Galbany, Christian Gilissen, the Radboudumc Technology Center Genomics, and the Radboudumc Cell Culture Facility for their technical and bioinformatic support, and Musa M. Mhlanga and Lucas L. Boer for the fruitful discussions. This work was financially supported by Aspasia grants of the Dutch Research Council (015.014.036 to T.K. and 015.014.066 to L.V.) and Netherlands Organization for Health Research and Development (91718310 to T.K.). The collaborations in this study were facilitated by ERN ITHACA, one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States, co-funded by the European Commission. The aims of this study contribute to the Solve-RD project (E.d.B., A.H., T.K., L.V.), which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 779257.

Declaration of interests

The authors declare no competing interests.

Received: December 13, 2022
Accepted: April 13, 2023

Web resources

https://omim.org/.
http://www.oreganno.org/dump/.
https://screen.wenglab.org/.

References

1. Elliott, A.M., Scott, W.J., Jr., Chudley, A.E., Reed, M.H., and Evans, J.A. (2021). Classifications of split hand foot malformation (SHFM) should include transverse deficiencies: why Maisels was correct. Am. J. Med. Genet. *185*, 2809–2814.
2. Umair, M., and Hayat, A. (2020). Nonsyndromic split-hand/foot malformation: recent classification. Mol. Syndromol. *10*, 243–254.
3. Sowińska-Seidler, A., Socha, M., and Jamsheer, A. (2014). Split-hand/foot malformation - molecular cause and implications in genetic counseling. J. Appl. Genet. *55*, 105–115.
4. Ianakiev, P., Kilpatrick, M.W., Toudjarska, I., Basel, D., Beighton, P., and Tsipouras, P. (2000). Split-hand/split-foot malformation is caused by mutations in the p63 gene on 3q27. Am. J. Hum. Genet. *67*, 59–66.
5. Provot, S., Schipani, E., Wu, J., and Kronenberg, H. (2008). Osteoporosis. Chapter 10 - Development of the Skeleton, Third ed (Academic Press).
6. Duijf, P.H.G., van Bokhoven, H., and Brunner, H.G. (2003). Pathogenesis of split-hand/split-foot malformation. Hum. Mol. Genet. *12 Spec No 1*, R51–R60.
7. Pownall, M., and Isaacs, H. (2010). FGF Signalling in Vertebrate Development: San Rafael (CA) (Morgan & Claypool Life Sciences).
8. Amberger, J.S., Bocchini, C.A., Schiettecatte, F., Scott, A.F., and Hamosh, A. (2015). OMIM.org: online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. Nucleic Acids Res. *43*, D789–D798.
9. Amberger, J.S., Bocchini, C.A., Scott, A.F., and Hamosh, A. (2019). OMIM.org: leveraging knowledge across phenotype-gene relationships. Nucleic Acids Res. *47*, D1038–D1043.
10. Spranger, M., and Schapera, J. (1988). Anomalous inheritance in a kindred with split hand, split foot malformation. Eur. J. Pediatr. *147*, 202–205.
11. Umair, M., Ullah, A., Abbas, S., Ahmad, F., Basit, S., and Ahmad, W. (2018). First direct evidence of involvement of a homozygous loss-of-function variant in the EPS15L1 gene underlying split-hand/split-foot malformation. Clin. Genet. *93*, 699–702.
12. Cova, G., Glaser, J., Schöpflin, R., Ali, S., Prada-Medina, C.A., Franke, M., et al. (2022). Combinatorial effects on gene expression at the Lbx1/Fgf8 locus resolve Split-Hand/Foot Malformation type 3. Preprint at bioRxiv. https://doi.org/10.1101/2022.02.09.479724.
13. Bilal, M., Hayat, A., Umair, M., Ullah, A., Khawaja, S., Malik, E., Burmeister, M., Bibi, N., Umm-E-Kalsoom, Memon, M.I., et al. (2020). Sequence variants in the WNT10B and TP63 genes underlying isolated split-hand/split-foot malformation. Genet. Test. Mol. Biomarkers *24*, 600–607.
14. Ahmad, M., Abbas, H., Haque, S., and Flatz, G. (1987). X-chromosomally inherited split-hand/split-foot anomaly in a Pakistani kindred. Hum. Genet. *75*, 169–173.
15. Faiyaz ul Haque, M., Uhlhaas, S., Knapp, M., Schüler, H., Friedl, W., Ahmad, M., and Propping, P. (1993). Mapping of the gene for X-chromosomal split-hand/split-foot anomaly to Xq26-q26.1. Hum. Genet. *91*, 17–19.
16. Faiyaz-Ul-Haque, M., Zaidi, S.H.E., King, L.M., Haque, S., Patel, M., Ahmad, M., Siddique, T., Ahmad, W., Tsui, L.C., and Cohn, D.H. (2005). Fine mapping of the X-linked split-hand/split-foot malformation (SHFM2) locus to a 5.1-Mb

(B) Schematic representation of genes and regulatory elements on chromosome 15 (upper image), and screenshot from the UCSC genome browser with the 15q26.3 gain in blue (lower image). The duplicated 15q26.3 fragment (dark blue in the schematic representation) contains exon 2 to exon 10 of *LRRC28* and an enhancer (orange in the schematic representation), normally regulating *LRRC28* expression. This enhancer contains six TFBSs, associated with binding of *CTCF, HNF4A, FOSL2, JUND, JUNB,* and *SMARCA4*.
(C) Schematic representation of regulatory elements affected by the SV. Two TFBSs targeting *SOX3* are deleted, the distance between three *SOX3*-targeting TFBS and *SOX3* is increased, and an enhancer is inserted on the X chromosome, which may retain its transcription-activation functions in the new genomic context. All these factors may affect expression of nearby genes, such as *SOX3*.

region on Xq26.3 and analysis of candidate genes. Clin. Genet. *67*, 93–97.
17. van den Berghe, H., Dequeker, J., Fryns, J.P., and David, G. (1978). Familial occurrence of severe ulnar aplasia and lobster claw feet: a new syndrome. Hum. Genet. *42*, 109–113.
18. de Ligt, J., Boone, P.M., Pfundt, R., Vissers, L.E.L.M., de Leeuw, N., Shaw, C., Brunner, H.G., Lupski, J.R., Veltman, J.A., and Hehir-Kwa, J.Y. (2014). Platform comparison of detecting copy number variants with microarrays and whole-exome sequencing. Genom. Data *2*, 144–146.
19. Lappalainen, I., Lopez, J., Skipper, L., Hefferon, T., Spalding, J.D., Garner, J., Chen, C., Maguire, M., Corbett, M., Zhou, G., et al. (2013). DbVar and DGVa: public archives for genomic structural variation. Nucleic Acids Res. *41*, D936–D941.
20. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. Am. J. Hum. Genet. *84*, 524–533.
21. Rehm, H.L., Berg, J.S., Brooks, L.D., Bustamante, C.D., Evans, J.P., Landrum, M.J., Ledbetter, D.H., Maglott, D.R., Martin, C.L., Nussbaum, R.L., et al. (2015). ClinGen–the clinical genome resource. N. Engl. J. Med. *372*, 2235–2242.
22. Mantere, T., Neveling, K., Pebrel-Richard, C., Benoist, M., van der Zande, G., Kater-Baats, E., Baatout, I., van Beek, R., Yammine, T., Oorsprong, M., et al. (2021). Optical genome mapping enables constitutional chromosomal aberration detection. Am. J. Hum. Genet. *108*, 1409–1422.
23. Neveling, K., Mantere, T., Vermeulen, S., Oorsprong, M., van Beek, R., Kater-Baats, E., Pauper, M., van der Zande, G., Smeets, D., Weghuis, D.O., et al. (2021). Next-generation cytogenetics: comprehensive assessment of 52 hematological malignancy genomes by optical genome mapping. Am. J. Hum. Genet. *108*, 1423–1435.
24. Jacobsen, J.O.B., Kelly, C., Cipriani, V., Robinson, P.N., and Smedley, D. (2022). Evaluation of phenotype-driven gene prioritization methods for Mendelian diseases. Brief. Bioinform. *23*, bbac188.
25. Smedley, D., Jacobsen, J.O.B., Jäger, M., Köhler, S., Holtgrewe, M., Schubach, M., Siragusa, E., Zemojtel, T., Buske, O.J., Washington, N.L., et al. (2015). Next-generation diagnostics and disease-gene discovery with the Exomiser. Nat. Protoc. *10*, 2004–2015.
26. Boeva, V., Popova, T., Bleakley, K., Chiche, P., Cappo, J., Schleiermacher, G., Janoueix-Lerosey, I., Delattre, O., and Barillot, E. (2012). Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics *28*, 423–425.
27. Roller, E., Ivakhno, S., Lee, S., Royce, T., and Tanner, S. (2016). Canvas: versatile and scalable detection of copy number variants. Bioinformatics *32*, 2375–2377.
28. Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Källberg, M., Cox, A.J., Kruglyak, S., and Saunders, C.T. (2016). Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics *32*, 1220–1222.
29. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164.
30. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.
31. Kent, W.J. (2002). BLAT–the BLAST-like alignment tool. Genome Res. *12*, 656–664.
32. Collins, R.L., Glessner, J.T., Porcu, E., Lepamets, M., Brandon, R., Lauricella, C., Han, L., Morley, T., Niestroj, L.M., Ulirsch, J., et al. (2022). A cross-disorder dosage sensitivity map of the human genome. Cell *185*, 3041–3055.e25.
33. Piepoli, A., Palmieri, O., Maglietta, R., Panza, A., Cattaneo, E., Latiano, A., Laczko, E., Gentile, A., Carella, M., Mazzoccoli, G., et al. (2012). The expression of leucine-rich repeat gene family members in colorectal cancer. Exp. Biol. Med. *237*, 1123–1128.
34. Ng, A., and Xavier, R.J. (2011). Leucine-rich repeat (LRR) proteins: integrators of pattern recognition and signaling in immunity. Autophagy *7*, 1082–1084.
35. Boyling, A., Perez-Siles, G., and Kennerson, M.L. (2022). Structural variation at a disease mutation hotspot: strategies to investigate gene regulation and the 3D genome. Front. Genet. *13*, 842860.
36. DeStefano, G.M., Fantauzzo, K.A., Petukhova, L., Kurban, M., Tadin-Strapps, M., Levy, B., Warburton, D., Cirulli, E.T., Han, Y., Sun, X., et al. (2013). Position effect on FGF13 associated with X-linked congenital generalized hypertrichosis. Proc. Natl. Acad. Sci. USA *110*, 7790–7795.
37. Haines, B., Hughes, J., Corbett, M., Shaw, M., Innes, J., Patel, L., Gecz, J., Clayton-Smith, J., and Thomas, P. (2015). Interchromosomal insertional translocation at Xq26.3 alters SOX3 expression in an individual with XX male sex reversal. J. Clin. Endocrinol. Metab. *100*, E815–E820.
38. Barrett, S.P., Parker, K.R., Horn, C., Mata, M., and Salzman, J. (2017). ciRS-7 exonic sequence is embedded in a long noncoding RNA locus. PLoS Genet. *13*, e1007114.
39. Kamachi, Y., and Kondoh, H. (2013). Sox proteins: regulators of cell fate specification and differentiation. Development *140*, 4129–4144.
40. Wood, H.B., and Episkopou, V. (1999). Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. Mech. Dev. *86*, 197–201.
41. Bylund, M., Andersson, E., Novitch, B.G., and Muhr, J. (2003). Vertebrate neurogenesis is counteracted by Sox1-3 activity. Nat. Neurosci. *6*, 1162–1168.
42. Dee, C.T., Hirst, C.S., Shih, Y.H., Tripathi, V.B., Patient, R.K., and Scotting, P.J. (2008). Sox3 regulates both neural fate and differentiation in the zebrafish ectoderm. Dev. Biol. *320*, 289–301.
43. Bergsland, M., Ramsköld, D., Zaouter, C., Klum, S., Sandberg, R., and Muhr, J. (2011). Sequentially acting Sox transcription factors in neural lineage development. Genes Dev. *25*, 2453–2464.
44. McAninch, D., and Thomas, P. (2014). Identification of highly conserved putative developmental enhancers bound by SOX3 in neural progenitors using ChIP-Seq. PLoS One *9*, e113361.
45. Laumonnier, F., Ronce, N., Hamel, B.C.J., Thomas, P., Lespinasse, J., Raynaud, M., Paringaux, C., Van Bokhoven, H., Kalscheuer, V., Fryns, J.P., et al. (2002). Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. Am. J. Hum. Genet. *71*, 1450–1455.
46. Bowl, M.R., Nesbit, M.A., Harding, B., Levy, E., Jefferson, A., Volpi, E., Rizzoti, K., Lovell-Badge, R., Schlessinger, D., Whyte, M.P., and Thakker, R.V. (2005). An interstitial deletion-insertion involving chromosomes 2p25.3 and Xq27.1, near SOX3, causes X-linked recessive hypoparathyroidism. J. Clin. Invest. *115*, 2822–2831.

47. Chimal-Monroy, J., Rodriguez-Leon, J., Montero, J.A., Gañan, Y., Macias, D., Merino, R., and Hurle, J.M. (2003). Analysis of the molecular cascade responsible for mesodermal limb chondrogenesis: sox genes and BMP signaling. Dev. Biol. *257*, 292–301.
48. Oosterveen, T., Kurdija, S., Ensterö, M., Uhde, C.W., Bergsland, M., Sandberg, M., Sandberg, R., Muhr, J., and Ericson, J. (2013). SoxB1-driven transcriptional network underlies neural-specific interpretation of morphogen signals. Proc. Natl. Acad. Sci. USA *110*, 7330–7335.
49. Santos-Pereira, J.M., Gallardo-Fuentes, L., Neto, A., Acemel, R.D., and Tena, J.J. (2019). Pioneer and repressive functions of p63 during zebrafish embryonic ectoderm specification. Nat. Commun. *10*, 3049.
50. Guerrini, L., Costanzo, A., and Merlo, G.R. (2011). A symphony of regulations centered on p63 to control development of ectoderm-derived structures. J. Biomed. Biotechnol. *2011*, 864904.
51. Kawata, M., Taniguchi, Y., Mori, D., Yano, F., Ohba, S., Chung, U.I., Shimogori, T., Mills, A.A., Tanaka, S., and Saito, T. (2017). Different regulation of limb development by p63 transcript variants. PLoS One *12*, e0174122.
52. Yang, A., Schweitzer, R., Sun, D., Kaghad, M., Walker, N., Bronson, R.T., Tabin, C., Sharpe, A., Caput, D., Crum, C., and McKeon, F. (1999). p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. Nature *398*, 714–718.
53. Abelló, G., Khatri, S., Radosevic, M., Scotting, P.J., Giráldez, F., and Alsina, B. (2010). Independent regulation of Sox3 and Lmx1b by FGF and BMP signaling influences the neurogenic and non-neurogenic domains in the chick otic placode. Dev. Biol. *339*, 166–178.
54. Lesurf, R., Cotto, K.C., Wang, G., Griffith, M., Kasaian, K., Jones, S.J.M., Montgomery, S.B., Griffith, O.L.; and Open Regulatory Annotation Consortium (2016). ORegAnno 3.0: a community-driven resource for curated regulatory annotation. Nucleic Acids Res. *44*, D126–D132.
55. Portales-Casamar, E., Arenillas, D., Lim, J., Swanson, M.I., Jiang, S., McCallum, A., Kirov, S., and Wasserman, W.W. (2009). The PAZAR database of gene regulatory information coupled to the ORCA toolkit for the study of regulatory sequences. Nucleic Acids Res. *37*, D54–D60.
56. Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C.y., Chou, A., Ienasescu, H., et al. (2014). Jaspar 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res. *42*, D142–D147.
57. Attanasio, C., Nord, A.S., Zhu, Y., Blow, M.J., Biddie, S.C., Mendenhall, E.M., Dixon, J., Wright, C., Hosseini, R., Akiyama, J.A., et al. (2014). Tissue-specific SMARCA4 binding at active and repressed regulatory elements during embryogenesis. Genome Res. *24*, 920–929.
58. Indra, A.K., Dupé, V., Bornert, J.M., Messaddeq, N., Yaniv, M., Mark, M., Chambon, P., and Metzger, D. (2005). Temporally controlled targeted somatic mutagenesis in embryonic surface ectoderm and fetal epidermal keratinocytes unveils two distinct developmental functions of BRG1 in limb morphogenesis and skin barrier formation. Development *132*, 4533–4544.
59. Catarino, R.R., and Stark, A. (2018). Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. Genes Dev. *32*, 202–223.
60. ENCODE Project Consortium, Moore, J.E., Purcaro, M.J., Pratt, H.E., Epstein, C.B., Shoresh, N., Adrian, J., Kawli, T., Davis, C.A., Dobin, A., et al. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature *583*, 699–710.
61. Capdevila, J., and Izpisúa Belmonte, J.C. (2001). Patterning mechanisms controlling vertebrate limb development. Annu. Rev. Cell Dev. Biol. *17*, 87–132.
62. Mills, A.A., Zheng, B., Wang, X.J., Vogel, H., Roop, D.R., and Bradley, A. (1999). p63 is a p53 homologue required for limb and epidermal morphogenesis. Nature *398*, 708–713.

HGGA, Volume *4*

Supplemental information

A complex structural variant near *SOX3* causes X-linked split-hand/foot malformation

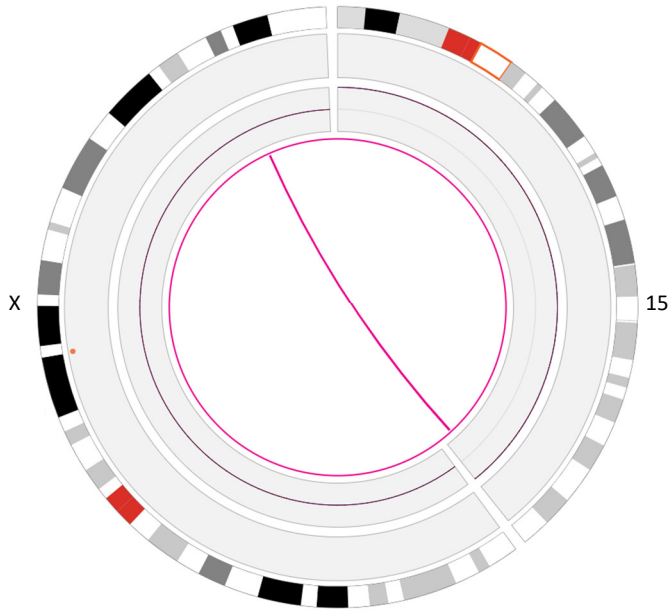
Elke de Boer, Carlo Marcelis, Kornelia Neveling, Ellen van Beusekom, Alexander Hoischen, Willemijn M. Klein, Nicole de Leeuw, Tuomo Mantere, Uirá S. Melo, Jeroen van Reeuwijk, Dominique Smeets, Malte Spielmann, Tjitske Kleefstra, Hans van Bokhoven, and Lisenka E.L.M. Vissers

Figure S1: Feet of individual III-7 showing mild cutaneous 2-3-syndactyly

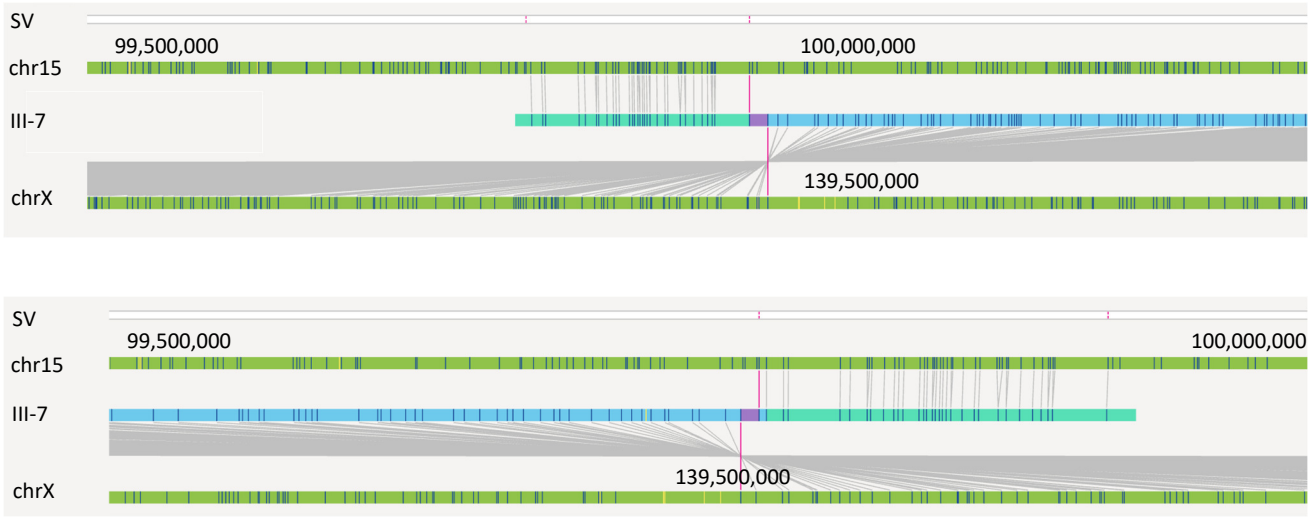
The feet of individual III-7 show mild cutaneous 2-3-syndactyly. Although syndactyly is common in SHFM, cutaneous 2-3-syndactyly of the toes is a non-specific clinical finding and relatively common in the general population.

Figure S2: OGM detects an inverted 15q26.3 gain inserted on Xq27.1

A.



B.



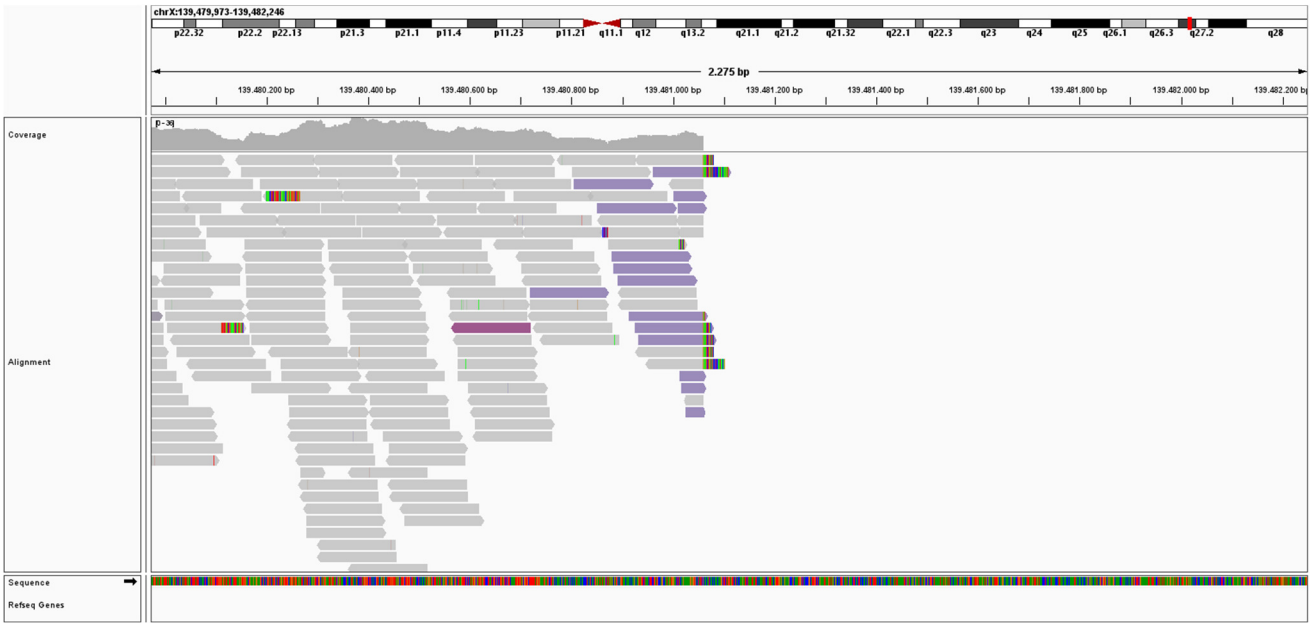
A) Circosplot showing a translocation between chromosome 15 and the X-chromosome. B) OGM results from individual III-7 indicates that the 15q26.3 gain inserts in an inverted fashion on the X-chromosome. The upper image illustrates the proximal breakpoint (distal end of gain from chromosome 15), the lower image illustrates the distal breakpoint (proximal end of gain from chromosome 15).

Figure S3: Visualization of the proximal breakpoint from WGS data

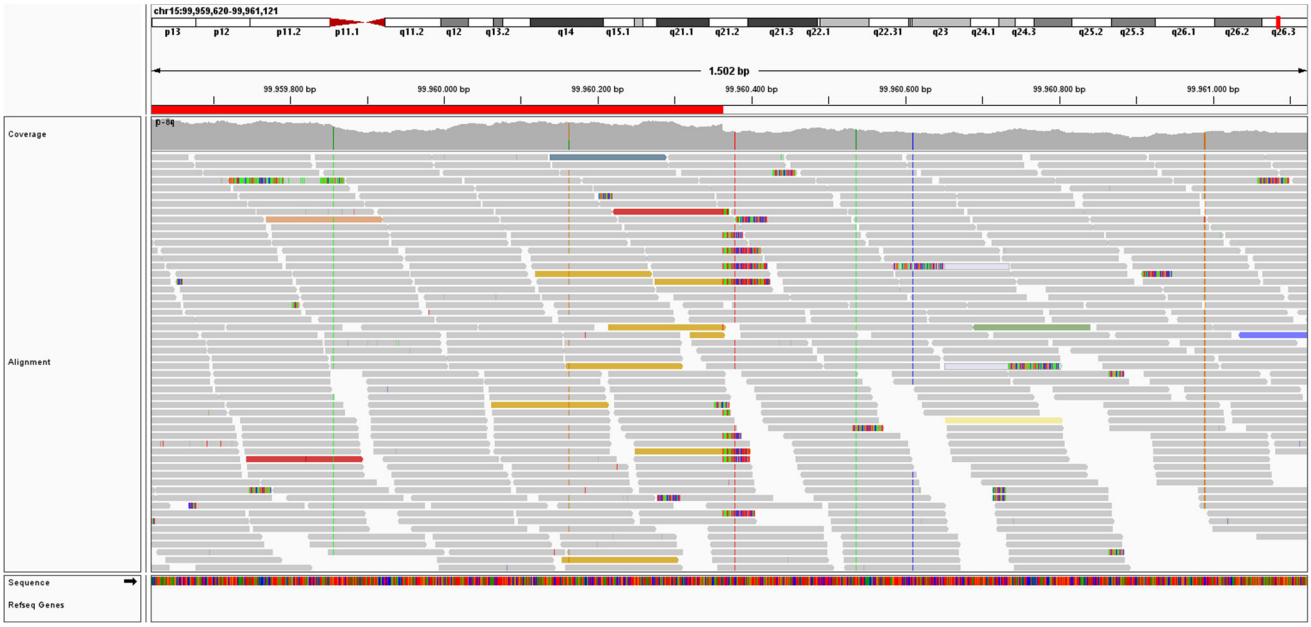
A.

```
→ Chromosome X (+)      GGATAGCAATCTTAATTTCAGACAAACAGACTTCAAACCAA
Individual III-7        GGATAGCAATCTTAATTTCAGGATAAATTCCTGGACTCATA
← Chromosome 15 (-)     CTAGGAAATCTAGAAGAAATGGATAAATTCCTGGACTCATA
```

B.



C.



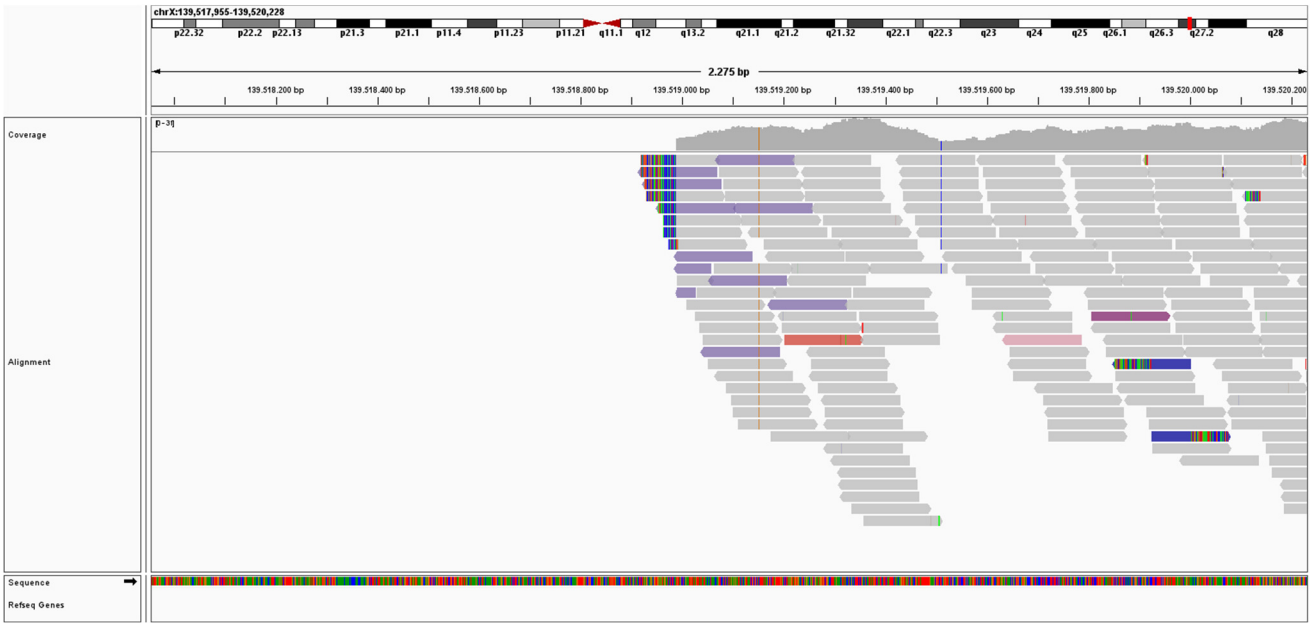
A) Alignment of the sequence at the proximal breakpoint. B) IGV screenshot of the proximal end of the deletion on the X-chromosome. C) IGV screenshot of the distal end of the duplication on chromosome 15.

Figure S4: Visualization of the distal breakpoint from WGS data

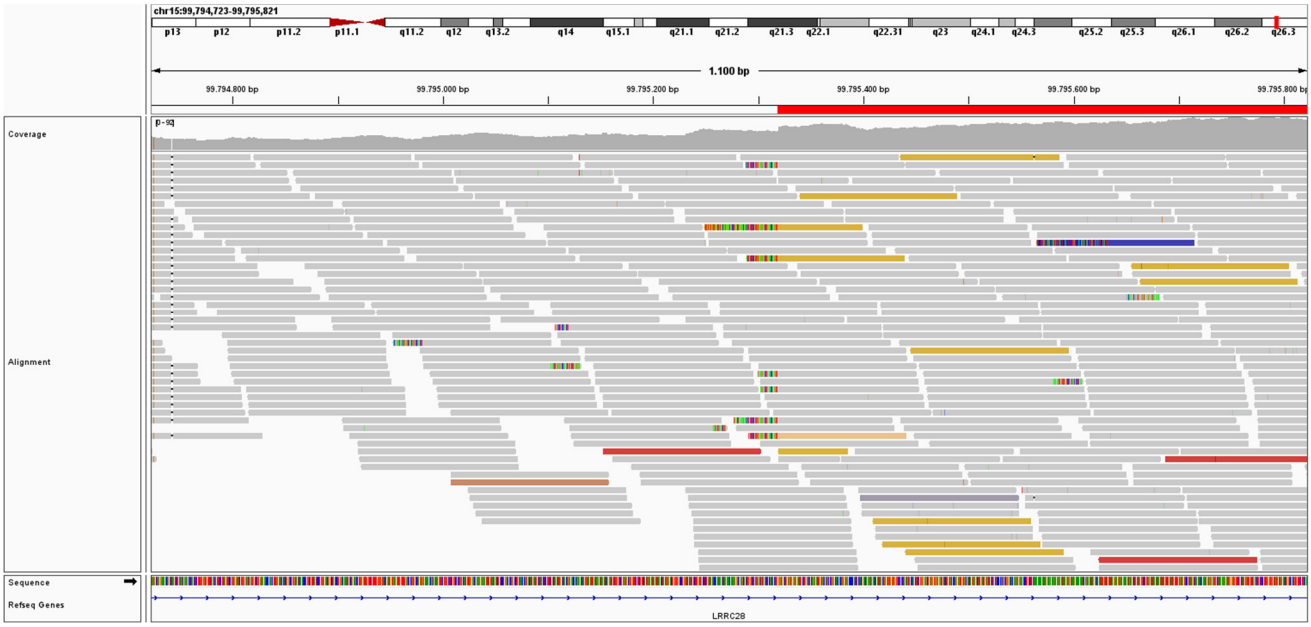
A.

```
→ Chromosome X (+)      GGAAACTCTAGTCTTATCTATAATGGTTTAAGTCCTTACAA
Individual III-7        CCTCCCACCTCAGCCTCCCAATCTGGTTTAAGTCCTTACAA
← Chromosome 15 (-)     CCTCCCACCTCAGCCTCCCCAAGTGCTGGGATTACAGTCAT
```

B.



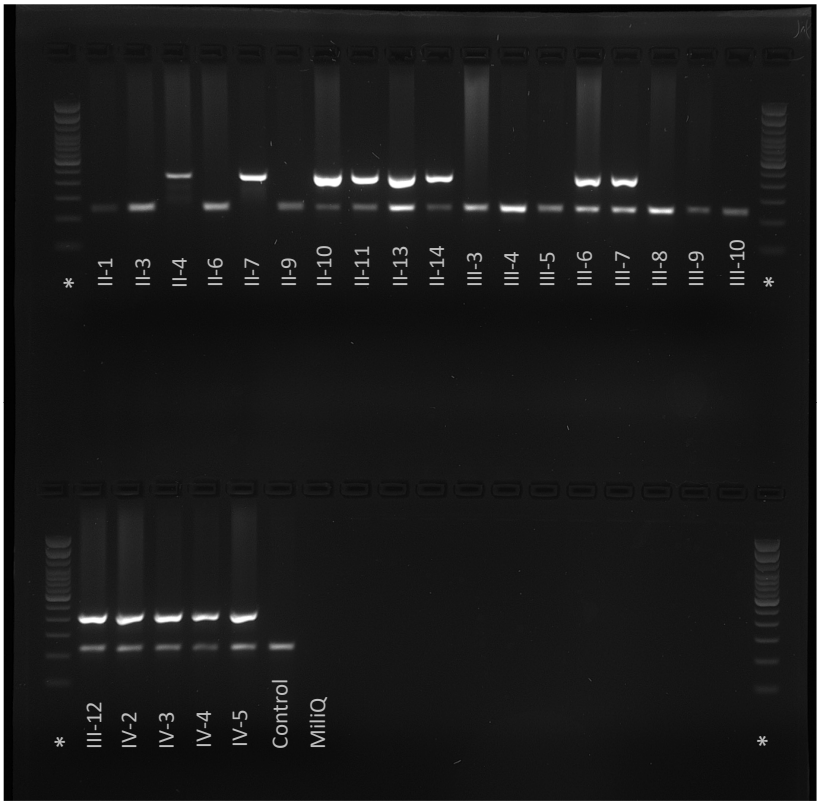
C.



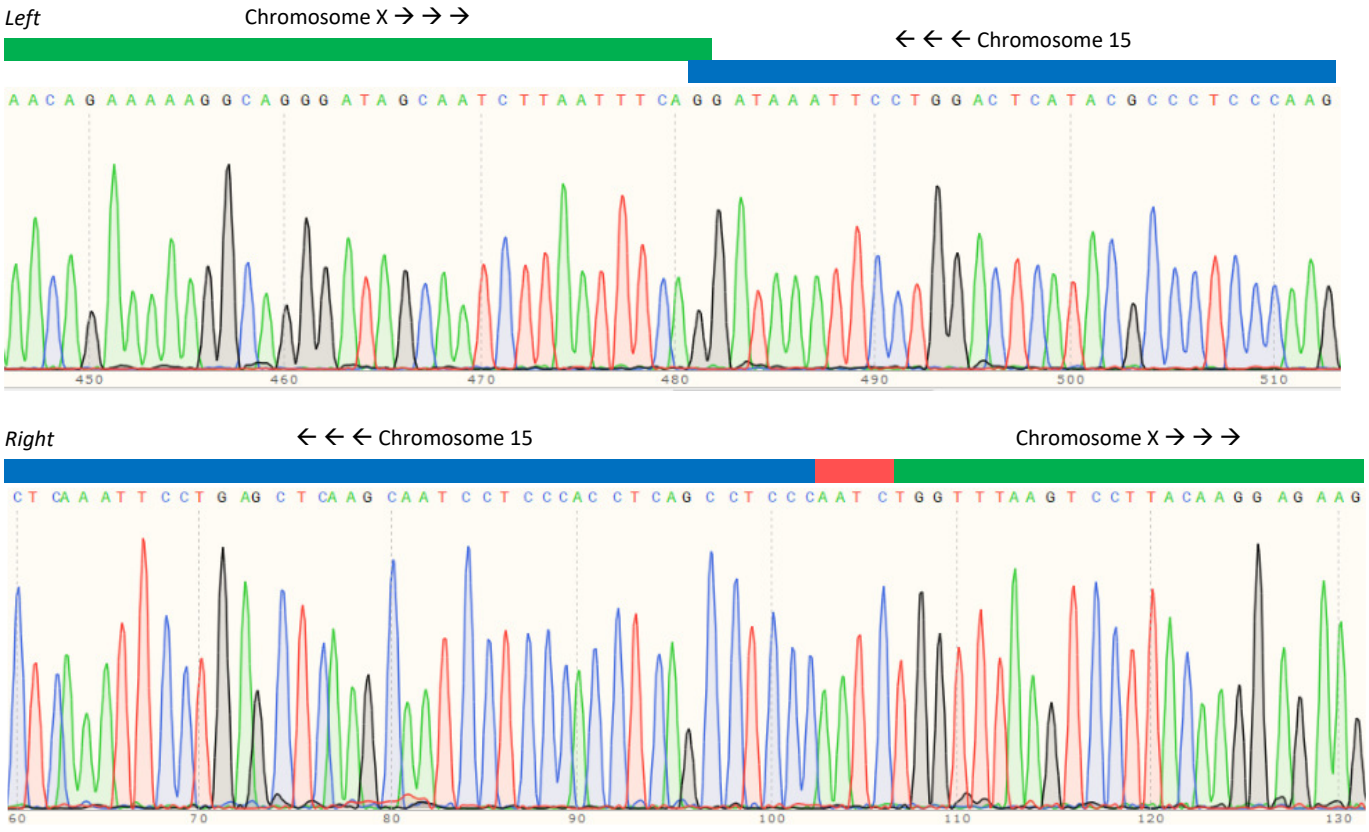
A) Alignment of the sequence at the distal breakpoint, including four nucleotides of unknown origin. B) IGV screenshot of the distal end of the deletion on the X-chromosome. C) IGV screenshot of the proximal end of the duplication on chromosome 15.

Figure S5: Validation and segregation of breakpoint junctions

A.

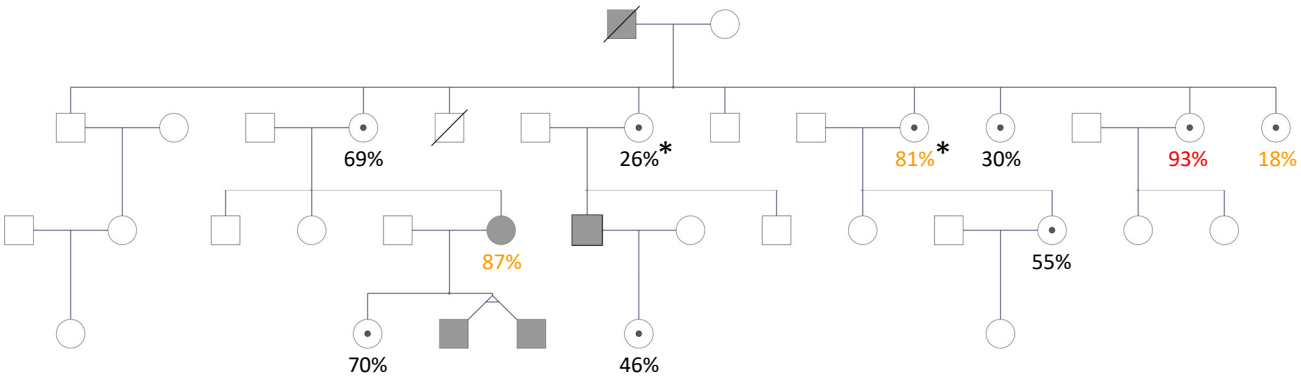


B.



A) PCR and gel electrophoresis of distal breakpoint (upper band), taking along exon 3 of B3GALNT2 as PCR control (lower band). The asterisks indicate the 100 bp DNA ladders. B) Sanger sequencing covering the proximal (upper) and distal (lower) breakpoints.

Figure S6: X-inactivation studies are not conclusive in explaining phenotypic variability between female carriers

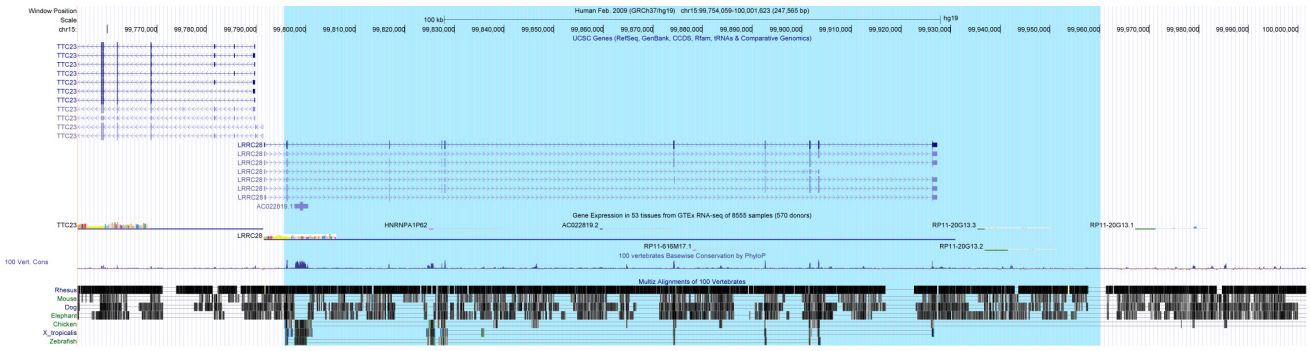


Random X-inactivation	21-79%
Skewed X-inactivation	11-20% and 80-89%
Extremely skewed X-inactivation	≤10% and ≥90%

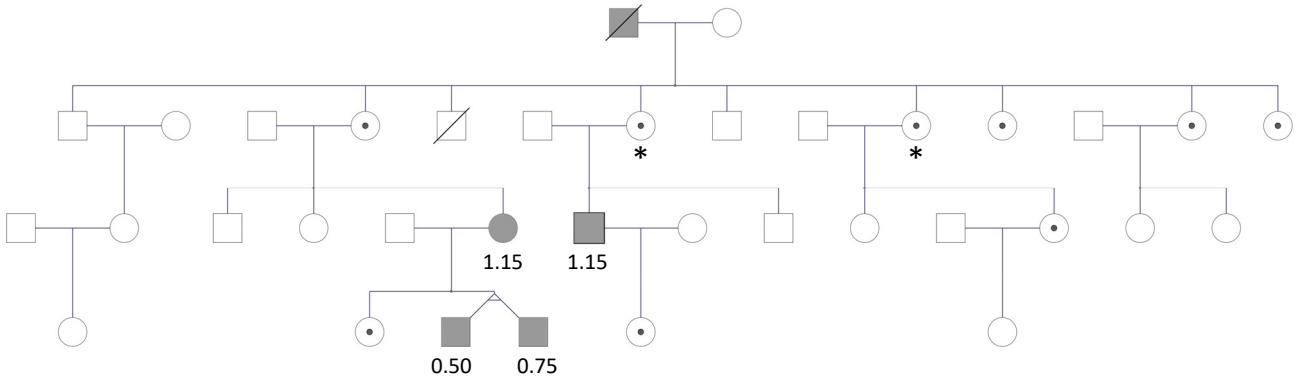
Pedigree of the family showing the degree of X-inactivation in all female carriers. This analysis indicates random X-inactivation in the majority of female carriers, skewed X-inactivation in three female relatives, including the only mildly affected female (III-6) and two putatively unaffected females, and extreme skewing in one putatively unaffected female carrier. Grey shading indicates individuals are affected, a dot indicates carriership and the asterisk indicates radiographic feet abnormalities.

Figure S7: The SV does not affect *LRRC28* expression

A.

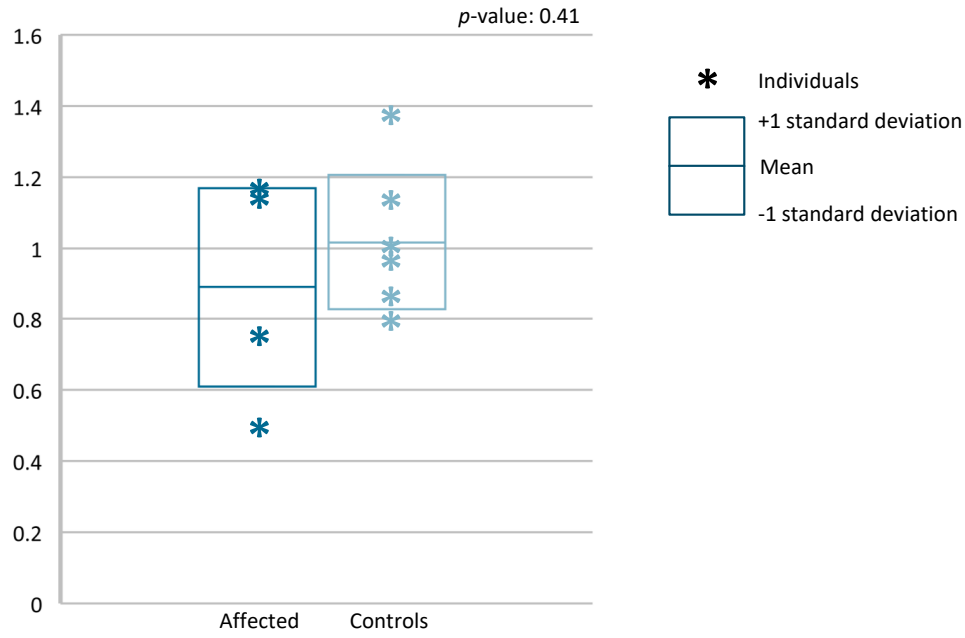


B.



Mean of 6 healthy unrelated controls: 1.02 (range: 0.80-1.38)

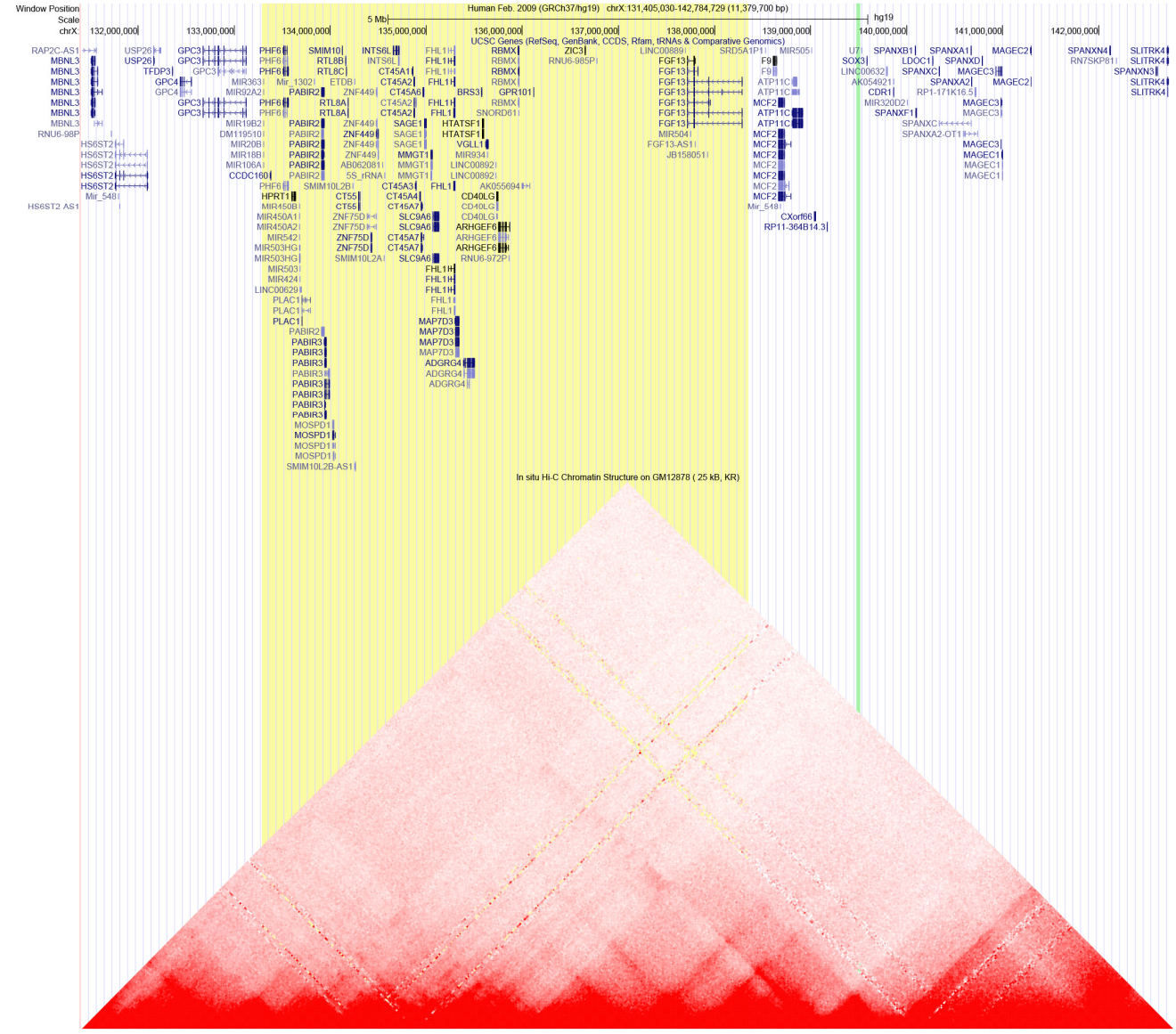
C.



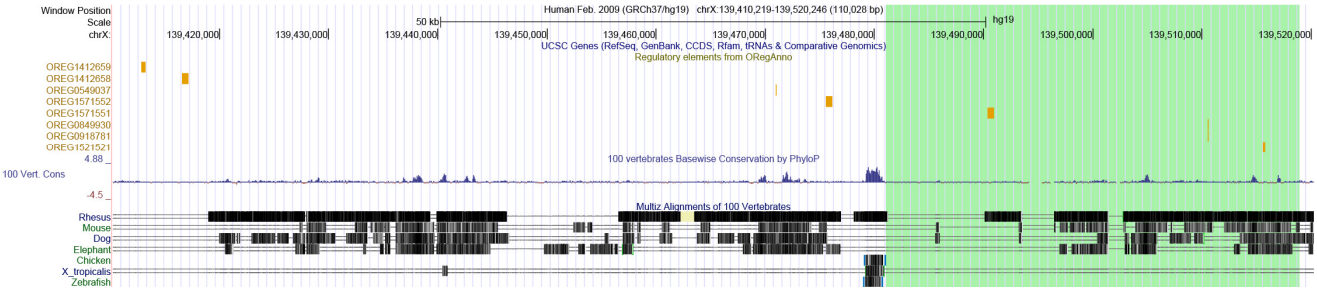
A) Screenshot from UCSC genome browser with gain of 15q26.3 material (blue), showing this gain contains exon 2-10 of LRRC28. B) Relative normalized expression of LRRC28 tested by qPCR on RNA from EBV-LCLs for all individuals for whom EBV-LCLs were available, alongside three healthy male and three female unrelated controls. Grey shading indicates individuals are affected, a dot indicates carriership and the asterisk indicates radiographic feet abnormalities. C) Relative expression of LRRC28 in affected individuals versus controls. The mean and standard deviation are visualized by the boxplot, with measurements per individual indicated by an asterisk.

Figure S8: Genomic 3D organization and TFBS at the locus of the SV and the SHFM2 linkage region on the X-chromosome

A.

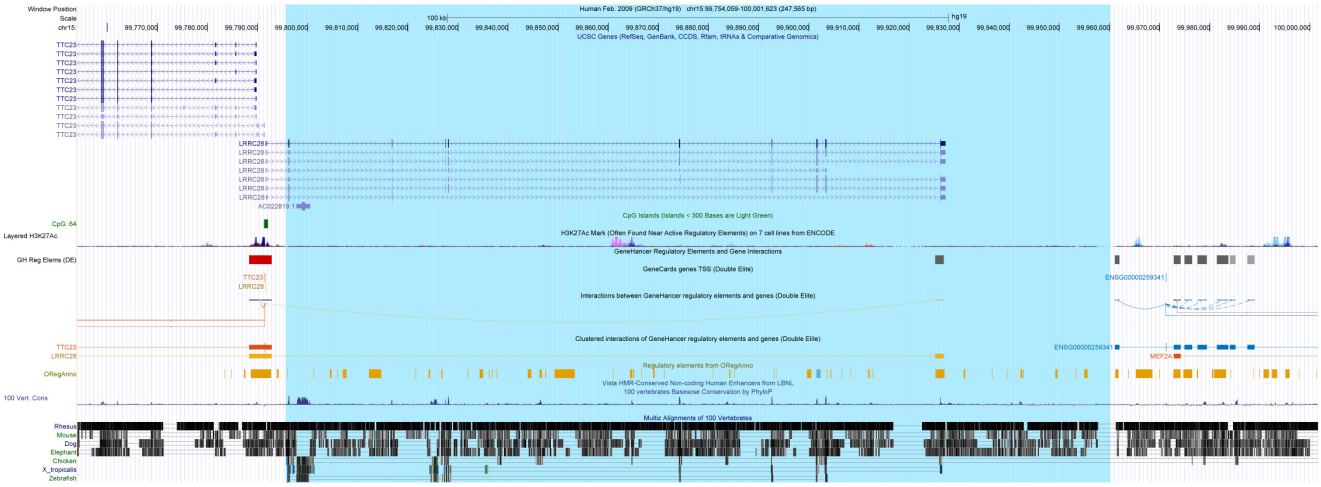


B.



A) Screenshot from the UCSC genome browser with the locus of the SV on the X-chromosome (green) and the SHFM2 linkage region from literature (yellow), showing SOX3 locates downstream of the deletion. The linkage region and the SV are approximately 1.1 Mb apart but partly locate to the same TAD. Based on the distance between the linkage region and the SV, identity by descent is deemed unlikely, although we could not formerly exclude this in the absence of molecular data for the previously published family. However, the fact that both the SV and the linkage region largely locate to the same TAD, suggests the phenotypes observed in the families might be caused by different variants with the same downstream effects. B) Screenshot from the UCSC genome browser showing the deletion of the X-chromosome (green) together with the 100 kb region proximal of the deletion. The deletion and the 100 kb proximal to the SV both contain four TFBS, of which several have SOX3 as target gene. These include OREG1412659 (TFBS of E2F1), OREG1412658 (TFBS of E2F1), OREG1571552 (TFBS of FOXA1), OREG1571551 (TFBS of FOXA1), OREG1521521 (TFBS of ESR1).

Figure S9: An enhancer is included in the duplicated chromosome 15 region



Screenshot from the UCSC genome browser showing the duplicated fragment of chromosome 15 (blue). The gain contains an enhancer that normally interacts with the promoter of LRRC28.

Supplemental information

Materials and methods

Individuals and consent

All affected individuals provided written informed consent to be included in this study. For publication of clinical photographs and radiological imaging, additional consent for photo publication was obtained. All consent procedures are in accordance with both the local ethical guidelines, and the Declaration of Helsinki. Clinical characterization was performed by reviewing the medical files and radiological imaging data and/or by revising the phenotypes of the individuals in the clinic. The study was approved by the Radboudumc local ethics board (2019-5554).

Microarray analysis

Microarray analysis was performed with the Affymetrix CytoScan HD (2.6M) array platform following the manufacturer's specification (Thermo Fisher Scientific, Waltham, USA). Experiments and interpretation of results were performed in the diagnostic workflow with an estimated average resolution of 20 kb on genome build GRCh37/hg19 (1).

Karyotyping and Fluorescence in Situ Hybridization

Karyotyping was performed on cultured EBV-LCLs cells in the diagnostic setting following previously described standard protocols (2), and the International System for Human Cytogenetic Nomenclature (ISCN, 2020) was used to describe chromosomal abnormalities. Chromosome slides were made according to routine procedures. Metaphases were analysed after GTG-banding and FISH experiments were performed on chromosome slides using the probe RP11-668P3 specific for 15q26.3 (Empire Genomics, Williamsville, NY, USA), and probes CEP 15 and CEP X (Vysis, Abbott, Abbott Park, IL, USA) for centromeres of chromosome 15 and chromosome X respectively.

Ultra-high molecular weight DNA extraction

Ultra-high molecular weight (UHMW) DNA was isolated from 1-1.5 million cultured EBV-LCLs, with the SP Blood & Cell Culture DNA Isolation Kit following the manufacturers' instructions (Bionano Genomics®, San Diego, CA, USA) as previously described. In brief, genomic DNA was released by treating cells with LBB lysis buffer, and bound to a nanobind disk, followed by washing steps and elution in the provided elution buffer (3).

Bionano optical genome mapping (Saphyr system)

The DLS (Direct Label and Stain) DNA Labeling Kit (Bionano Genomics®, San Diego, CA, USA) was used to label the UHMW DNA molecules. 750 ng of genomic DNA was labelled with Direct Label Enzyme (DLE-1) and DL-Green fluorophores. DL-Green fluorophores excess was washed out, followed by an overnight DNA backbone counterstaining. The labelled UHMW genomic DNA was loaded on a Saphyr chip® for linearization and visualisation with the Saphyr system (Bionano Genomics, San Diego USA). The Bionano Solve software version v3.6.1 executed the *de novo* assembly and Variant Annotation Pipelines. Output was analysed by a CNV and a SV pipeline, enabling detection of unbalanced aberrations based on differences in normalized molecule coverage, and detection of structural variants based on comparison of labelling patterns between the sample genome map and a reference genome respectively. Interpretation of results was performed using Bionano Access software v1.6.1. To filter on quality of results, confidence values were set as follows: for insertion/deletion=0, inversion=0.01, duplications= -1, translocation=0 and CNV=0. For SV calls, an optical mapping dataset of 204 human control samples (provided by Bionano Genomics) was used to filter out common variants (3).

Whole genome sequencing (WGS)

WGS was outsourced to BGI (BGI, Hongkong, China) on a BGISeq500 sequencing platform using a paired-end module of 2x 100 base pairs with a minimal median coverage of 30-fold per genome.

BWA V.0.78 was used for read mapping to the GRCh37/hg19 reference genome build and bam quality control was performed with Qualimap V.2.2.1. To ensure data quality, several quality metrics were checked, including insert size, percentage mapped reads, percentage duplicated mapped reads, coverage, percentage of bases with more than 20-fold coverage and error rate. The resulting alignment files were subjected to several variant calling pipelines. Variant calling of single nucleotide variants and small indels (SNVs/indels) was carried out using xAtlas V.0.1 and variants were subsequently annotated with an inhouse developed annotation pipeline, that uses the Variant Effect Predictor (VEP V.91) and Gencode V.34lift37 basic gene annotations. Additionally, information on population allele frequency was added from GnomAD V.2.1.1 and from an inhouse database. For genetic variants in genes associated with a known disease, inhouse gene panel information was added. Other included annotations were CADD score V.1.6, SpliceAI, OMIM or KEGG pathways. Detection of Runs of Homozygosity was performed using Plink V.1.07 applying the following parameters: homozyg-window-het=3, homozyg-snp=50 and homozyg-kb=300. Known pathogenic short tandem repeats (STRs) were analysed with Expansion Hunter V.3.1.2. using default settings. CNVs were identified with two CNV calling algorithms, being Control-FREEC (4) and Canvas Copy Number Variant Caller (5), which both use read depth to detect copy number changes. SV calling was performed using the Manta Structural Variant Caller V.1.1.0 (6), that uses a paired-end and split read evidence approach to identify SVs. CNVs and SVs were annotated with an inhouse developed pipeline, based on ANNOVAR (7) and Gencode V.34lift37 basic gene annotations. Additional information on population allele frequency was added from GnomAD V.2.1, 1000G V.8 and GoNL SV database. SNVs/indels, CNVs and SVs were prioritized from WGS by applying a customized inhouse pipeline. Additionally, SNVs/indels were assessed by a phenotype-based variant prioritization using the Exomiser software package (version 13.1.0 with 2202 databases) with a default presets for both exome and genome analyses (8, 9). Prioritized variants were visually inspected in IGV (10).

Validation and segregation with (nested) PCR, Agarose Gel Electrophoresis and Sanger sequencing

The variant was validated by breakpoint-spanning PCRs and evaluation by Agarose Gel Electrophoresis in all individuals with DNA available. For amplification of the left breakpoint, long-range PCR (LR-PCR) followed by a nested conventional PCR was performed. For the right breakpoint, a conventional PCR was applied for amplification, also including a control PCR reaction for *B3GALNT2.* Primers were designed using Primer3web v4.1.0 software. LR-PCR and (nested) conventional PCRs were performed according to the manufacturer's protocols using the Q5 High-Fidelity 2X Master Mix (New England BioLabs inc.) and the Amplitaq Gold 360 Master Mix (Thermo Fisher Scientific) respectively. For segregation analysis, the agarose gels of both breakpoints were inspected for the presence or absence of a PCR product of the expected size. For individual III-7, the PCR products were enzymatically cleaned with Exonuclease I and FastAP, followed by Sanger sequencing using a routine diagnostic workflow (11). Finally, Sanger sequencing traces were analysed using the Chromas Lite v2.1.1 software package (Technelysium) to verify the exact sequence at both breakpoints that was seen in WGS data.

qPCR for LRRC28, SOX3 and FGF13 expression

For the quantitative real-time PCR (qPCR) experiments, RNA was isolated from cultured EBV-LCLs from four affected individuals and six controls following the NuleoSpin RNA isolation protocol (Machery-Nagel). RNA was converted to cDNA with the iScript cDNA Synthesis Kit (BioRad). qPCR primers were designed with Primer3 v4.1.0 software. qPCR was performed according to standard protocol with GoTaq 2x master mix (Promega). *GUSB* was included in the qPCR experiments as reference gene. For all primers, standard curves were made with a cDNA control sample in a dilution series (20x, 80x, 320x, 1280x, 5120x dilutions in MilliQ). For all qPCR experiments a blank sample was taken along for each primer pair, and all samples were tested in duplicate. Relative expression of the genes of interest (*LRRC28, SOX3* and *FGF13)* was calculated by normalizing Ct values for these genes with Ct values of *GUSB* and normalizing individual delta Ct scores to the mean of delta Ct scores of

controls. Relative expression of the genes of interest in affected individuals was compared to controls with a paired T-test. P-values below 0.05 were considered statistically significant.

Primers for PCR and qPCR

Goal of primer pair	**Forward primer**	**Reverse primer**	**Product size (bp)**	**Notes**
LR-PCR spanning left breakpoint	GATCGTCTGTGATGGTTAGGTG	GTGATGTCAGCAAGTGGGATAC	7,824	Forward primer specific to the locus on chrX; reverse primer specific to the locus on chr15
Conventional (nested) PCR spanning left breakpoint	GCTAGTAAGGGCACATAGAGC	TGCCTCAATGTTCTTCAGGG	836	Forward primer specific to the locus on chrX; reverse primer aligning to the locus on chr15, but supplementary alignment to different loci
Conventional PCR spanning right breakpoint	TACTATAGAGAGCACCACCACAC	CAGGAGCCACGCACATAATG	426	Forward primer specific to the locus on chr15; reverse primer specific to the locus on chrX
PCR control using exon 3 of *B3GALNT2*	AAATGGGCATGAGGAAACG	AAGCTTAGCAACTTTTACTCAACATC	238	
qPCR on exon 5 of *LRRC28* (gene of interest)	CCATTGGGTCTCTTGTAAAACTC	TCGAAGATGACGTAAAGCTCTC	104	
qPCR on *SOX3* (gene of interest)	TGGAGAACTGCAACGCCTACGC	GATCACGGCAGAAATCACCAACTC	204	Primers as in (12)
qPCR on *FGF13* (gene of interest)	CAGCCGACAAGGCTACCAC	GTTCCGAGGTGTACAAGTATCC	185	Primers as in (12)
qPCR on *GUSB* (gene for normalization)	CTGTACACGACACCCACCAC	TACAGATAGGCAGGGCGTTC	245	

X-inactivation studies

X-inactivation studies were performed in the diagnostic workflow using DNA derived from blood of all female SV carriers as previously described (13).

In silico analysis of regulatory functions

Regulatory functions of the regions affected by the SV were assessed using the UCSC genome browser (14), visualizing datasets from the GeneHancer database (15) and the Open Regulatory Annotation database (ORegAnno; http://www.oreganno.org/dump/) (16). To assess TFBS in the SV,

we downloaded the complete dataset from ORegAnno (ORegAnno_Combined_2014.09.15.tsv), that also includes data from PAZAR and JASPAR databases (16-18). Additionally, candidate cis-Regulatory Elements that are active in the human developing limb were examined using SCREEN: Search Candidate cis-Regulatory Elements by ENCODE (Registry of cCREs V3) (19) after liftover of the relevant sequences to GRCh38/hg38 (chrX:140398896-140436824 and chr15:99255115-99420157).

Supplemental results

X-inactivation studies

We performed X-inactivation studies by quantifying methylation of the human androgen receptor locus including all ten confirmed female carriers (II-4, II-7, II-10, II-11, II-13, II-14, III-6, III-12, IV-2 and IV-5). We found that the majority of these individuals exhibited random X-inactivation (21-79%; individual II-4, II-7, II-11, III-12, IV-2 and IV-5), three individuals showed skewed X-inactivation (11-20% and 80-89%; individual II-10, II-14 and III-6) and only in individual II-13 extreme skewing of X-inactivation was seen (≤10% and ≥90%; Figure S6). Whereas these results do not explain why III-6 is the only female expressing split-hand, we can conclude that increased X-chromosome inactivation is not the protective mechanism for non-penetrance, as is illustrated by the female carriers with random X-inactivation without any limb abnormalities. Although our experimental set-up did not prove which X-chromosome was inactivated (e.g. the one with or without the SV) and interpretation of X-inactivation studies has intrinsic challenges (20-23), our results could partially be explained if individual III-6 would have skewed inactivation of the X-chromosome without the SV. This would make the X-chromosome containing the SV more active, resulting in individual III-6 having a mild phenotype, whereas in theory the other female carriers might show (extreme) skewing of the X-chromosome containing the SV. This would explain why they do not exhibit any limb abnormalities, although the subclinical foot abnormalities and random X-inactivation in individual II-7 do not fit with this hypothesis. Alternatively, it is not unthinkable that – like individual II-10 – individuals II-13 and II-

14 are subclinically or mildly affected and/or that the available information obtained per family history is incomplete, which would also largely explain our observations.

Web resources

https://omim.org/

http://www.oreganno.org/dump/

https://screen.wenglab.org/

Supplemental references

1. de Ligt J, Boone PM, Pfundt R, Vissers LE, de Leeuw N, Shaw C, et al. Platform comparison of detecting copy number variants with microarrays and whole-exome sequencing. Genom Data. 2014;2:144-6.
2. Bates SE. Classical cytogenetics: karyotyping techniques. Methods Mol Biol. 2011;767:177-90.
3. Mantere T, Neveling K, Pebrel-Richard C, Benoist M, van der Zande G, Kater-Baats E, et al. Optical genome mapping enables constitutional chromosomal aberration detection. Am J Hum Genet. 2021.
4. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schleiermacher G, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics. 2012;28(3):423-5.
5. Roller E, Ivakhno S, Lee S, Royce T, Tanner S. Canvas: versatile and scalable detection of copy number variants. Bioinformatics. 2016;32(15):2375-7.
6. Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics. 2016;32(8):1220-2.
7. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
8. Jacobsen JOB, Kelly C, Cipriani V, Robinson PN, Smedley D. Evaluation of phenotype-driven gene prioritization methods for Mendelian diseases. Brief Bioinform. 2022;23(5).
9. Smedley D, Jacobsen JO, Jäger M, Köhler S, Holtgrewe M, Schubach M, et al. Next-generation diagnostics and disease-gene discovery with the Exomiser. Nat Protoc. 2015;10(12):2004-15.
10. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24-6.
11. de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, et al. Diagnostic exome sequencing in persons with severe intellectual disability. N Engl J Med. 2012;367(20):1921-9.
12. Haines B, Hughes J, Corbett M, Shaw M, Innes J, Patel L, et al. Interchromosomal insertional translocation at Xq26.3 alters SOX3 expression in an individual with XX male sex reversal. J Clin Endocrinol Metab. 2015;100(5):E815-20.

13. Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW. Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. Am J Hum Genet. 1992;51(6):1229-39.
14. Lee BT, Barber GP, Benet-Pagès A, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2022 update. Nucleic Acids Res. 2022;50(D1):D1115-d22.
15. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database (Oxford). 2017;2017.
16. Lesurf R, Cotto KC, Wang G, Griffith M, Kasaian K, Jones SJ, et al. ORegAnno 3.0: a community-driven resource for curated regulatory annotation. Nucleic Acids Res. 2016;44(D1):D126-32.
17. Portales-Casamar E, Arenillas D, Lim J, Swanson MI, Jiang S, McCallum A, et al. The PAZAR database of gene regulatory information coupled to the ORCA toolkit for the study of regulatory sequences. Nucleic Acids Res. 2009;37(Database issue):D54-60.
18. Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ, et al. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res. 2014;42(Database issue):D142-7.
19. Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, Adrian J, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature. 2020;583(7818):699-710.
20. Orstavik KH. X chromosome inactivation in clinical practice. Hum Genet. 2009;126(3):363-73.
21. Sharp A, Robinson D, Jacobs P. Age- and tissue-specific variation of X chromosome inactivation ratios in normal women. Hum Genet. 2000;107(4):343-9.
22. Shvetsova E, Sofronova A, Monajemi R, Gagalova K, Draisma HHM, White SJ, et al. Skewed X-inactivation is common in the general female population. Eur J Hum Genet. 2019;27(3):455-65.
23. Amos-Landgraf JM, Cottle A, Plenge RM, Friez M, Schwartz CE, Longshore J, et al. X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. Am J Hum Genet. 2006;79(3):493-9.