ORIGINAL RESEARCH

Rare and *de novo* duplications containing *SHOX* in clubfoot

Brooke Sadler ,[1] Gabe Haller,[2] Lilian Antunes,[1] Momchil Nikolov,[1] Ina Amarillo,[3] Bradley Coe,[4,5] Matthew B. Dobbs,[2] Christina A. Gurnett[1]

ABSTRACT

Introduction Congenital clubfoot is a common birth defect that affects at least 0.1% of all births. Nearly 25% cases are familial and the remaining are sporadic in inheritance. Copy number variants (CNVs) involving transcriptional regulators of limb development, including *PITX1* and *TBX4*, have previously been shown to cause familial clubfoot, but much of the heritability remains unexplained.

Methods Exome sequence data from 816 unrelated clubfoot cases and 2645 in-house controls were analysed using coverage data to identify rare CNVs. The precise size and location of duplications were then determined using high-density Affymetrix Cytoscan chromosomal microarray (CMA). Segregation in families and *de novo* status were determined using qantitative PCR.

Results Chromosome Xp22.33 duplications involving *SHOX* were identified in 1.1% of cases (9/816) compared with 0.07% of in-house controls (2/2645) (p=7.98×10⁻⁵, OR=14.57) and 0.27% (38/13592) of Atherosclerosis Risk in Communities/the Wellcome Trust Case Control Consortium 2 controls (p=0.001, OR=3.97). CMA validation confirmed an overlapping 180.28 kb duplicated region that included *SHOX* exons as well as downstream non-coding regions. In four of six sporadic cases where DNA was available for unaffected parents, the duplication was *de novo*. The probability of four *de novo* mutations in *SHOX* by chance in a cohort of 450 sporadic clubfoot cases is 5.4×10⁻¹⁰.

Conclusions Microduplications of the pseudoautosomal chromosome Xp22.33 region (PAR1) containing *SHOX* and downstream enhancer elements occur in ~1% of patients with clubfoot. *SHOX* and regulatory regions have previously been implicated in skeletal dysplasia as well as idiopathic short stature, but have not yet been reported in clubfoot. *SHOX* duplications likely contribute to clubfoot pathogenesis by altering early limb development.

BACKGROUND

Congenital clubfoot, also called *talipes equinovarus*, is a common and serious birth defect that affects an estimated 1 of every 1000 live births.[1] Clubfoot is characterised by structural defects of several tissues of the foot and lower leg, which leads to abnormal positioning of foot and ankle joints.[2] If left untreated, it can result in severe disability and deformity.[3] Approximately 80% of clubfoot cases are isolated birth defects.[4] The remaining 20% of cases are due to associated malformations, or known genetic syndromes, such as distal arthrogryposis and myleomeningocele.[5] There is also strong evidence for a genetic basis for isolated clubfoot. There is a family history of clubfoot in 25% of all isolated cases.[6] Data from twin studies show a higher concordance in monozygotic (33%) than dizygotic (3%) twins,[4] and more recent data estimates heritability of isolated clubfoot at ~30%.[7]

Although few causative genes are known, progress has been made on understanding the genetics of isolated clubfoot. Thus far, the strongest genetic evidence for clubfoot is the *PITX1-TBX4-HOXC* pathway, the proper function of which is required for normal hindlimb development.[8–10] Variation in this pathway has been linked to clubfoot phenotypes through a segregating dominant mutation in *PITX1*,[11] inherited recurrent chromosome 17q23.1q23.2 microduplications containing *TBX4*,[12–14] *PITX1* haploinsufficiency,[15] and copy number variants (CNVs) and point mutations in 5′ *HOXC* genes including *HOXC12*.[16] However, these account for only ~1% of familial clubfoot and are rarely identified in non-familial cases.

Here, we report that approximately 1% of paediatric patients with clubfoot has microduplications of *SHOX*, a transcription factor involved in early limb development. These duplications include the *SHOX* gene as well as downstream regulatory regions. Four of these microduplications are *de novo* strengthening their role in clubfoot pathogenesis.

MATERIALS AND METHODS
Cohort description

All patients with clubfoot were recruited at St Louis Children's Hospital or Shriners Hospital St Louis. Only those with a clinical diagnosis of clubfoot were included. The institutional review board approved this study and all patients and/or parents provided informed consent. We excluded patients with myelomeningocele, arthrogryposis or known genetic disorders. DNA was obtained from either blood or spit from probands, additional affected family members, and parents wherever possible. We include in the present study only those who self-reported their ethnicity as European-American. When family history data are available, it was obtained from the parents of the affected child.

Exome sequencing

Exome sequence data for 816 unrelated clubfoot probands were generated at the McDonnell Genome Institute using IDT xGen Exome Panel V1 capture on Illumina HiSeq 4000 paired-end reads.

These cases included 450 sporadic cases with no family history of clubfoot in either first-degree, second-degree or third-degree relatives. In-house control data consisted of 2645 samples, including 1197 with adolescent idiopathic scoliosis, 334 with Chiari 1 Malformation and 433 with male infertility, which were were captured and sequenced using the identical IDT platform as the clubfoot cases. An additional 637 controls with amyotrophic lateral sclerosis were captured using Agilent reagents and sequenced on HiSeq 1500/2000.

Exome alignment and BAM file processing

The sequencing reads (ie, fastq files) from all clubfoot cases and control samples were aligned to the human genome reference (GRCh37) using BWA-MEM (V.0.7.15). The resulting BAM files were sorted and duplicates were marked using Picard Mark-Duplicates (V.2.9.0). Re-alignment intervals for each BAM file were determined using GATK (V.3.5) RealignerTargetCreator using a list of known indel sites (Mills and 1 kg indels data from the GATK resource bundle ftp://ftp.broadinstitute.org/bundle/b37/), local re-alignment and base quality recalibration were then performed by GATK (V3.5) IndelRealigner and BaseRecalibrator, respectively.

Exome CNV analysis

CNV analysis was performed using the software package FishingCNV that is designed to identify rare CNVs from exome sequencing data without the need for a paired control. We previously used this programme to successfully identify CNVs in an adolescent idiopathic scoliosis cohort.[17] This programme uses an algorithm to prioritise rare variants, and compares coverage depth in a test sample against the background distribution, as well as principal components analysis to remove batch effects. The programme extracts coverage information from GATK processed BAM files,[18] then normalises read depth information at each exon[19] into Reads Per Kilobase per Million mapped reads (RPKM) files. Following our previous protocol, to identify only the most confident CNV calls, we included only CNVs with an average RPKM of at least 10, the mean log-ratio coverage between background and sample of -1.2 to -0.8 (consistent with a deletion) or 0.4 to 0.8 (consistent with a duplication) and Holm–Bonferroni adjusted p value less than 10^{-3}.

Sex chromosome aneuploidy calculations

We estimated sex chromosome aneuploidy by comparing the total read depth on chromosomes X and Y to the average total read depth for autosomes. Read depth was calculated for all samples using Depth of Coverage from the Genome Analysis Tool Kit V.3.7 (GATKV.3.7). For females, the average X:Autosome ratio was ~1 average and Y:Autosome ratio was <0.5. For males, the average X:autosome ratio was <0.5 and the average Y:autosome ratio was ~1.5.

External control samples

Control samples included the Atherosclerosis Risk in Communities (ARIC) and the Wellcome Trust Case Control Consortium 2 (WTCCC2) NBS and 58C cohorts, for a total of 13 692 samples. ARIC/WTCCC2 controls were profiled on the Affymextrix SNP6 array and processed as per Coe *et al* (2014).[20] Manual inspection was performed to exclude low-quality CNVs. Due to inherent noise in array data in the pseudoautosomal regions (PARs), an additional check filter was performed on the samples based on their overall noise level as measured by SD of all genomic probes to look for evidence of lost sensitivity in noisier samples. This

method detects CNVs from complete chromosomal deletions or duplications down to those 10 Kb in size.

CNV validation using Affymetrix Cytoscan

The Affymetrix Cytoscan HD chromosomal microarray (CMA) consists of more than 750 000 SNPs and 1.9 million non-polymorphic probes, and results in 5–10 Kb resolution. Chromosome Analysis Suite was used to view and summarise chromosomal aberrations across the exomes. This array enables the performance of high-resolution, genome-wide CNV analysis and provides genotyping information for the detection of copy neutral loss/absence of heterozygosity, and was used to validate and determine the size of the *SHOX* containing CNVs detected in clubfoot probands. In some cases, CNVs were broken into multiple contiguous pieces by Cytoscan. We only considered pieces as part of the same CNV if they were <20 000 bp apart.

Quantitative PCR of unaffected parents to determine *de novo* status

For validation of CNVs using quantitative PCR (qPCR), we designed three primer pairs to exonic sequence within the CNV. Primers were tested for quality and efficiency using PCR and real-time qPCR standard curves, respectively. QC was performed on the tested DNA using the Qubit Fluorimeter (Invitrogen). qPCR is performed on a 96 well high-sided low-profile PCR plates (Thermo Scientific) on the Applied Biosystems StepOne Real-Time PCR System. Positive and negative controls with known gains or losses as previously determined by Affymetrix Cytoscan in the gene of interest were included in each run. The *RPL27* housekeeping gene was used as a reference.

De novo enrichment analysis and statistical analysis

Fisher's exact tests were performed in R V.3.6.1 on the number of *SHOX*-containing CNVs in cases versus controls. The R package 'denovolyzeR' was used to assess the enrichment of *de novo* coding variants on a per gene basis using the estimated *de novo* mutation rate per gene derived previously from a large exome-sequenced control cohort.[21] We estimated the proportion of *de novo* variants contributing to clubfoot based on the observed number of protein-altering *de novo* mutations observed in our cohort of clubfoot sporadic uniplex trios (n=450) compared with expected ($=Nx((M_1–M_2)/M_1)$, where N is the total number of trios and M_1 and M_2 are the observed and expected count of protein-altering *de novo* mutations per trio, respectively).

Comparative genomic analysis

The human genomic sequence was retrieved from the UCSC genome browser (www.ucsc.edu/genome). The orthologous sequences from vertebrate species including dog, opossum, chicken, frog and multiple fish species were obtained from the Evolutionarily Conserved Region (ECR) Browser (https://ecrbrowser.dcode.org).[22] As per previous studies, the settings of the ECR browser were adjusted to identify sequences with a minimal length of 100 bp and a 70% homology between species.[22 23]

RESULTS

SHOX duplications are enriched in patients with clubfoot

SHOX-containing duplications were identified in nine out of 816 (1.1%) clubfoot probands and were the only CNVs significantly enriched in cases (table 1). Identical analysis of our in-house control samples revealed only two duplications out of 2645 controls (0.07%) (p=7.98×10^{-5}, OR=14.57). This

Table 1 Frequency of X chromosome duplications and deletions in cases and controls

CNV region	CNV type	Chromosomal coordinates	Cases (n=816)	Controls (n=2645)	P value
47, XXY	Duplication	Entire chromosome	5	0	7.4×10^{-4}
Xp22.33	Duplication	2 00 854–2 729 493	9	2	7.9×10^{-5}
Xp22.33	Duplication	1 401 596–3 631 294	11	31	0.4
Xp11.4	Duplication	37 701 034–38 009 147	1	2	0.6
Xp11.23	Duplication	48 265 385–48 270 814	1	3	0.7
Xq22.2-q23	Duplication	102 973 984–109 310 624	2	1	0.2
Xq26.3	Duplication	134 978 401–135 586 622	2	1	0.2
Xq28	Duplication	152 241 445–155 235 144	9	14	0.07
Xp11.23	Deletion	41 307 142–46 508 232	2	0	0.5
Xq28	Deletion	155 003 533–155 235 853	1	1	0.4

enrichment was also significant when compared with 13 692 ARIC/WTCCC2 controls, where 38 individuals were found to have high-confidence *SHOX* duplications (p=0.001, OR=3.97). All nine chromosome Xp22.33 duplications detected using FishingCNV on exome data were validated using Affymetrix Cytoscan arrays and the size of online supplementary tables S1 and S2. The duplications were more precisely determined, and varied from 235 kb to 2.528 Mb online supplementary tables S1 and S2. No microdeletions of *SHOX* were detected in the clubfoot probands.

Aneuploidy in the sample

We evaluated 816 clubfoot probands for sex chromosome aneuploidy and sex chromosome CNVs using exome read depth data. Five out of 564 male probands (0.9%) were discovered to have 47, XXY or Klinefelter Syndrome based on X:Autosome ratios consistent with female sex (range: 0.85–1.1) but also had Y:Autosome ratios consistent with male sex (range: 1.4–1.7) and were reported as biological male. We removed these aneuploid samples prior to further CNV analysis. There were no clubfoot probands with 45, X (Turner Syndrome) or 47, XXX. Our in-house controls and ARIC/WTCCC2 controls could not be used for comparison due to the exclusion of patients with aneuploidy from both datasets prior to generation of genetic data.

Duplication of *SHOX* and downstream regions containing known evolutionary conserved enhancers

All nine clubfoot probands share an overlapping 180.28 kb duplication of *SHOX* and portions of the 200 kb downstream position effect (PE) region, containing known ECRs and enhancers (figure 1). An affected great uncle of 7 470 001 (individual 7 470 004) with clubfoot was also evaluated and found to have a small 1.41 kb chromosome Xp22.33 duplication on Affymetrix Cytoscan array, although his duplication was much smaller than the proband. When he is included, the overlapping duplicated region shared by all 10 clubfoot cases is narrowed down to a 1.41 kb portion of the downstream PE region (figure 1). This 1.41 kb region contains a 131 bp ECR composed of transposable elements and simple repeats.

Clinical phenotypes associated with chromosome Xp22.33 duplications

Clinical evaluations revealed that seven out of nine probands with the duplication were men (77.7%), whereas the entire clubfoot cohort evaluated was 68% men. Height data were available for seven out of nine clubfoot probands. Six patients were <50th percentile of height for their age, and five of the seven were <10th percentile (table 2). Three probands were
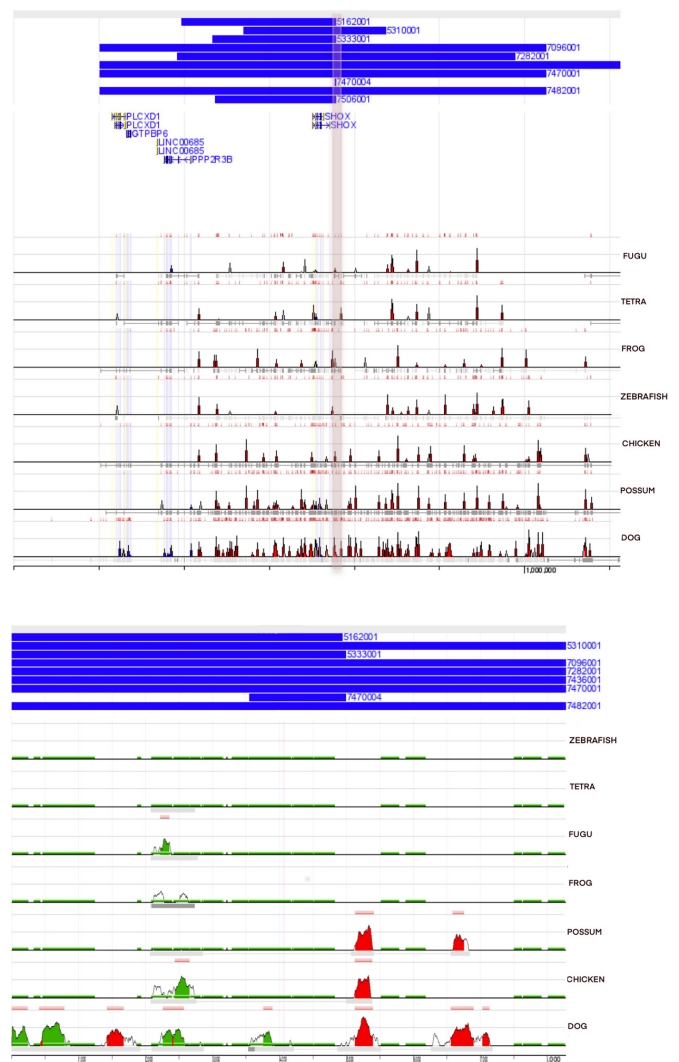

Figure 1 X-chromosome duplications involving *SHOX* in patients with clubfoot. (Top) The location of the duplicated region as determined by AffyCytoscan array data from all nine *SHOX* duplication carriers and the affected great uncle (7470004) with the smallest duplication magnified in the bottom portion. The smallest overlapping region to all patients with clubfoot is in the pink rectangle. Conserved regulatory elements across seven species are shown (evolutionary conserved region browser data). (bottom) Zoomed in view of the region shared among nine duplication carriers and individual 7 470 004 that shows the smallest overlapping 1.41 Kb duplication that contains a single evolutionarily conserved region.

Table 2 Demographic and phenotypic information on *SHOX* duplication carriers

ID	Current age	Sex	Type of clubfoot	CDC height percentile	Duplication size (Kb)	De novo	Additional history
5162001	13.8	F	LTEV	48	301.166	Unknown	Adolescent idiopathic scoliosis
5310001	14.2	M	BTEV	Unknown	277.152	Unknown	
5333001	14.2	M	BTEV	Unknown	241.258	Y	Developmental delay, chr 16p13.11 duplication
7436001	7.9	M	BTEV	56	2528.216	Y	
7470001	1.9	M	BTEV	4.7	847.043	Y	Bilateral camptodactyly, great uncle with clubfoot
7282001	3.0	M	BTEV	3.4	659.028	N	
7506001	2.4	M	BTEV	8.4	235.578	N	Amniotic band syndrome, bilateral symbrachydactyl, sixth nerve palsy
7482001	2.1	F	BTEV	0.1	847.043	Y	Torticollis, chr 2q37.3 deletion, developmental delay
7096001	16.8	M	BTEV	3.7	847.043	unknown	Developmental delay
7470004	45.3	M	BTEV	67	1.415	unknown	Great uncle of 7470001

BTEV, bilateral talipes equinovarus; LTEV, left talipes equinovarus.

found on follow-up to have mild global developmental delay, and two of these probands were also found to have additional known recurrent CNVs associated with developmental delay phenotypes along with the SHOX duplication. Eight of the nine probands had bilateral clubfoot. Two probands presented with hand abnormalities, including symbrachydactly and campto-dactly. One proband had adolescent idiopathic scoliosis. Only one proband (7470001) had a family history of clubfoot, and his affected great uncle also had a chromosome Xp22.33 microdeletion, although much smaller than the proband.

Four *de novo SHOX* duplications among those with available DNA

Of the nine probands with *SHOX* duplications, DNA was available for both unaffected parents of six probands. DNA of one set of parents was analysed with Affymetrix Cytoscan and the other five sets were analysed with qPCR. Four out of six *SHOX* duplications were not present in the unaffected parents but were confirmed in the proband, resulting in a *de novo* rate of 66% (4/6) for chromosome Xp22.33 duplications. When we consider that only 450 out of the 816 clubfoot probands evaluated were considered sporadic cases because they did not have a family history of clubfoot, the probability of detecting four *de novo* mutations in the *SHOX* gene by chance in a cohort of this size (n=450) is 5.4×10^{-10}.

DISCUSSION

We report evidence supporting a new role for *SHOX* gene duplications in clubfoot pathogenesis. First, we have shown that approximately 1% of clubfoot probands have a microduplication of *SHOX* and that there is a significant enrichment of *de novo* duplications. Short stature homeobox (*SHOX*) is located in the PAR1 on the distal end of the short arms of the X and Y chromosomes, and encodes a homeodomain transcription factor expressed in the developing forelimb and lower leg, as well as the first and second pharyngeal arches.[24] *SHOX* escapes X-inactivation, hence XX and XY individuals both have two copies of *SHOX*.[25] *SHOX*-related short stature phenotypes, including growth failure and limb deformities, are observed in patients with Turner syndrome who have only one copy of *SHOX*.[26] Point mutations or deletions resulting in haploinsufficiency of *SHOX* are also associated with idiopathic growth retardation,[27] and idiopathic short stature (ISS) and are

causative of Leri–Weill dyschondrosteosis (LWD), a dominant skeletal dysplasia presenting with mesomelic limb shortening and curving of the radius known as Madelung deformity.[28 29] Homozygous loss of *SHOX* or compound heterozygous *SHOX* defects result in the more severe phenotype known as Langer mesomelic dysplasia (LMD).[23 30–32] Mesomelic involvement of the lower legs observed in patients with LWD and LMD with *SHOX* defects is consistent with the predominant anatomic abnormalities previously described in patients with idiopathic clubfoot, in which skeletal muscle hypoplasia is most evident in the lower legs but not the calf.[33]

Duplications of the *SHOX* region are less often described, though they have been associated with tall stature in one study.[34] However, positive effects on height appear to require duplication of the entire *SHOX* coding and regulatory region, as partial duplications have been associated with short stature.[32 35] For instance, a girl with severe short stature and agenesis of the right tibia and fibula was found to have large duplications flanking either side of the *SHOX* gene, but a normal copy number of the gene itself.[36] In fact, *SHOX* microduplications represent a small fraction of the aetiology of ISS because *SHOX* enhancer duplications may impair gene expression as extra copies of the regulatory region can dysregulate *SHOX*.[37] Many of our clubfoot probands with SHOX duplications also had short stature and none of them were tall. While variable hand anomalies were observed in two of our clubfoot probands with the *SHOX* duplication, we do not know whether the preferential involvement of the lower leg is a consequence of selection bias due to our recruitment from clubfoot clinic or reflects enhanced susceptibility of the lower leg. It is also important to note that analysis of CNVs using exome data has some limitations, most notably that potential intronic or intergenic regions CNVs acting through regulatory effects may be missed. As in CNV analysis of whole genome data, detection is only as good as the sequence coverage, and more even coverage will always lead to cleaner results.

The non-coding region downstream of *SHOX* that is located within the 180 kb common clubfoot deletion interval is known as a position effect (PE) of *SHOX*[23] because alterations can affect gene expression even when the gene is intact.[38] Because the *SHOX* duplication common region shared among all patients with clubfoot also contains long-range enhancers of the gene,[30 39] we cannot exclude the possibility that these non-coding regions also play a role in clubfoot pathogenesis. Deletions of these

enhancers, without involvement of the *SHOX* transcript, have previously been shown to be enriched in patients with ISS and LWS and absent in controls,[23 30 31] although there is evidence of incomplete penetrance in families.[40]

Evolutionary sequence conservation has previously been used as a guide to locate potential enhancers downstream of the *SHOX* gene,[23 41 42] and is a good predictor for the identification of *cis*-regulatory elements, particularly in developmental genes.[43] Comparative genomic analysis identified eight highly conserved non-genic elements (CNEs) located in the *SHOX* downstream region, and three have previously been shown to have *cis*-regulatory activity in the developing limb bud in chicken embryos.[23] Duplications of these PEs and CNEs are not as well characterised as deletions. While we narrowed the causative clubfoot genetic region to a 180 kb interval that contains both the *SHOX* gene and a large downstream regulatory region, the interval may be as small as 1.41 kb if we include the duplication in an affected great uncle that was much smaller than that present in the proband (online supplementary tables S1 and S2). It is important to note that the variant in the great uncle and corresponding proband are not the same variant; therefore, they may not be inherited from a common ancestor, so this does not impact the number of *de novo* variants found here. A single ECR is located in the very smallest 1.41 kb region but is only conserved back to the dog lineage. However, previous research of *SHOX* enhancers has shown that some less highly evolutionarily conserved elements also have enhancer activity.[40]

In a prior analysis of autosomal CNVs, our group showed that CNVs involving transcription factors or transcriptional regulators of hindlimb development segregate with clubfoot in several families.[16] These include specific regulators of hindlimb development, including *PITX1*, *TBX4* and 5′ *HOXC* genes.[10] Therefore, it is not surprising that we find an enrichment of CNVs in yet another transcription factor, this time one that is expressed on the X-chromosome. Our identification of four *de novo* duplications of the *SHOX* gene strongly suggests a causative role in clubfoot pathogenesis. Further evidence for a role of *SHOX* in clubfoot pathogenesis is provided by studies showing that recurrent chromosome 17q23.1-q23.2 microduplications involving Tbx4, a transcriptional regulator of the closely related gene Shox2 during limb development in mice,[44] are among the most common known causes of familial clubfoot.[12]

Although the vast majority of the patients in this study had isolated clubfoot, at least when recruited for the study during infancy, three out of nine duplication carriers were noted to have mild developmental disabilities on long-term follow-up. Two of these individuals had other CNVs, in addition to the *SHOX* duplication. The chromosome 16p13.11 duplication and chromosome 2q37.3 microdeletion that we identified in these patients are both known to cause neurodevelopmental abnormalities,[45 46] but not clubfoot, and are therefore likely explain their developmental disabilities since most of our *SHOX* duplication carriers are developmentally normal. Other investigators have also noted a tendency for individuals to have multiple CNVs.[47] Recent studies have shown that there is a slightly higher prevalence of neurodevelopmental difficulties in children with isolated clubfoot which is also consistent with this observation.[48]

We also identified five individuals in our clubfoot cohort with 47,XXY karyotypes, which is approximately six times greater than the population average of 0.15% or ~1/660 males.[49] Although there are nearly 800 genes duplicated in patients with 47,XXY karyotype, we hypothesise that duplication of *SHOX* may be responsible for the clubfoot phenotype in these males. We did not identify any females with clubfoot who had an 47,XXX

karyotype, suggesting that males may be particularly susceptible to the effects of *SHOX* duplications. Overall, the over-representation of males with *SHOX* duplications in our study is consistent with the 3:1 male:female ratio for clubfoot. Interestingly, the opposite is seen in patients with LWD and *SHOX* deficiency,[50] as Madelung deformity is both more common and more severe in females. Our identification of *SHOX* duplications and enrichment of 47,XXY karyotype abnormalities in patients with clubfoot, along with our prior work identifying CNVs involving other regulators of early limb development, including *PITX1*, *TBX4* and 5′ *HOXC* genes, suggests that it may be useful to screen patients with clubfoot for CNVs, particularly when there is any concern about development.

Our results show that *SHOX* duplications are present in ~1% of all patients with clubfoot. This study will have immediate and direct impact on patient care because these genetic abnormalities can easily be detected with clinically available CMA studies. In addition, our results provide important insight into the pathogenetic mechanisms of clubfoot now that we have shown a strong link to abnormalities in multiple early transcriptional regulators of limb development. Furthermore, the enrichment of *SHOX* duplications in males may partially explain the increased prevalence of clubfoot among males. Future research employing high-resolution methods to detect even smaller X chromosome CNVs may help us to understand the clubfoot pathogenesis in the remaining unexplained cases, and may also provide insight into the role of non-coding regulatory regions in human development.

ORCID iD

Brooke Sadler http://orcid.org/0000-0002-3268-3809

REFERENCES

1 Wynne-Davies R. Genetic and environmental factors in the etiology of talipes equinovarus. *Clin Orthop Relat Res* 1972;84:9–13.

2 Basit S, Khoshhal KI. Genetics of clubfoot; recent progress and future perspectives. *Eur J Med Genet* 2018;61:107–13.

3 O'Shea RM, Sabatini CS. What is new in idiopathic clubfoot? *Curr Rev Musculoskelet Med* 2016;9:470–7.

4 Wynne-Davies R. Family studies and the cause of congenital club foot. talipes equinovarus, talipes CALCANEO-VALGUS and metatarsus varus. *J Bone Joint Surg Br* 1964;46:445–63.

5 Gurnett CA, Boehm S, Connolly A, Reimschisel T, Dobbs MB. Impact of congenital talipes equinovarus etiology on treatment outcomes. *Dev Med Child Neurol* 2008;50:498–502.

6 Lochmiller C, Johnston D, Scott A, Risman M, Hecht JT. Genetic epidemiology study of idiopathic talipes equinovarus. *Am J Med Genet* 1998;79:90–6.

7 Engell V, Nielsen J, Damborg F, Kyvik KO, Thomsen K, Pedersen NW, Andersen M, Overgaard S. Heritability of clubfoot: a twin study. *J Child Orthop* 2014;8:37–41.

8 Hasson P, DeLaurier A, Bennett M, Grigorieva E, Naiche LA, Papaioannou VE, Mohun TJ, Logan MPO. Tbx4 and Tbx5 acting in connective tissue are required for limb muscle and tendon patterning. *Dev Cell* 2010;18:148–56.

9 Logan M, Tabin CJ. Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity. *Science* 1999;283:1736–9.

10 Alvarado DM, McCall K, Hecht JT, Dobbs MB, Gurnett CA. Deletions of 5' HOXC genes are associated with lower extremity malformations, including clubfoot and vertical talus. *J Med Genet* 2016;53:250–5.

11 Gurnett CA, Alaee F, Kruse LM, Desruisseau DM, Hecht JT, Wise CA, Bowcock AM, Dobbs MB. Asymmetric lower-limb malformations in individuals with homeobox PITX1 gene mutation. *Am J Hum Genet* 2008;83:616–22.

12 Alvarado DM, Aferol H, McCall K, Huang JB, Techy M, Buchan J, Cady J, Gonzales PR, Dobbs MB, Gurnett CA. Familial isolated clubfoot is associated with recurrent chromosome 17q23.1q23.2 microduplications containing TBX4. *Am J Hum Genet* 2010;87:154–60.

13 Lu W, Bacino CA, Richards BS, Alvarez C, VanderMeer JE, Vella M, Ahituv N, Sikka N, Dietz FR, Blanton SH, Hecht JT. Studies of TBX4 and chromosome 17q23.1q23.2: an uncommon cause of nonsyndromic clubfoot. *Am J Med Genet A* 2012;158A:1620–7.

14 Peterson JF, Ghaloul-Gonzalez L, Madan-Khetarpal S, Hartman J, Surti U, Rajkovic A, Yatsenko SA. Familial microduplication of 17q23.1–q23.2 involving TBX4 is associated with congenital clubfoot and reduced penetrance in females. *Am J Med Genet A* 2014;164A:364-9–9.

15 Alvarado DM, McCall K, Aferol H, Silva MJ, Garbow JR, Spees WM, Patel T, Siegel M, Dobbs MB, Gurnett CA. Pitx1 haploinsufficiency causes clubfoot in humans and a clubfoot-like phenotype in mice. *Hum Mol Genet* 2011;20:3943–52.

16 Alvarado DM, Buchan JG, Frick SL, Herzenberg JE, Dobbs MB, Gurnett CA. Copy number analysis of 413 isolated talipes equinovarus patients suggests role for transcriptional regulators of early limb development. *Eur J Hum Genet* 2013;21:373–80.

17 Sadler B, Haller G, Antunes L, Bledsoe X, Morcuende J, Giampietro P, Raggio C, Miller N, Kidane Y, Wise CA, Amarillo I, Walton N, Seeley M, Johnson D, Jenkins C, Jenkins T, Oetjens M, Tong RS, Druley TE, Dobbs MB, Gurnett CA. Distal chromosome 16p11.2 duplications containing *SH2B1* in patients with scoliosis. *J Med Genet* 2019;56:427–33.

18 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.

19 Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat Methods* 2008;5:621–8.

20 Coe BP, Witherspoon K, Rosenfeld JA, van Bon BWM, Vulto-van Silfhout AT, Bosco P, Friend KL, Baker C, Buono S, Vissers LELM, Schuurs-Hoeijmakers JH, Hoischen A, Pfundt R, Krumm N, Carvill GL, Li D, Amaral D, Brown N, Lockhart PJ, Scheffer IE, Alberti A, Shaw M, Pettinato R, Tervo R, de Leeuw N, Reijnders MRF, Torchia BS, Peeters H, O'Roak BJ, Fichera M, Hehir-Kwa JY, Shendure J, Mefford HC, Haan E, Gécz J, de Vries BBA, Romano C, Eichler EE. Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet* 2014;46:1063–71.

21 Ware JS, Samocha KE, Homsy J, Daly MJ. Interpreting de novo variation in human disease using denovolyzeR. *Curr Protoc Hum Genet* 2015;87:7.25.1–15.

22 Ovcharenko I, Stubbs L, Loots GG. Interpreting mammalian evolution using Fugu genome comparisons. *Genomics* 2004;84:890–5.

23 Sabherwal N, Bangs F, Röth R, Weiss B, Jantz K, Tiecke E, Hinkel GK, Spaich C, Hauffa BP, van der Kamp H, Kapeller J, Tickle C, Rappold G. Long-Range conserved non-coding SHOX sequences regulate expression in developing chicken limb and are associated with short stature phenotypes in human patients. *Hum Mol Genet* 2007;16:210-22–22.

24 Hoyer-Kuhn H, Franklin J, Jones C, Blum WF, Schoenau E. Growth response to growth hormone treatment in patients with SHOX deficiency can be predicted by the Cologne prediction model. *J Pediatr Endocrinol Metab* 2018;31:25–31.

25 Lee JT, Bartolomei MS. X-Inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell* 2013;152:1308–23.

26 Clement-Jones M, Schiller S, Rao E, Blaschke RJ, Zuniga A, Zeller R, Robson SC, Binder G, Glass I, Strachan T, Lindsay S, Rappold GA. The short stature homeobox gene SHOX is involved in skeletal abnormalities in Turner syndrome. *Hum Mol Genet* 2000;9:695–702.

27 Rao E, Weiss B, Fukami M, Rump A, Niesler B, Mertz A, Muroya K, Binder G, Kirsch S, Winkelmann M, Nordsiek G, Heinrich U, Breuning MH, Ranke MB, Rosenthal A, Ogata T, Rappold GA. Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nat Genet* 1997;16:54–63.

28 Schneider KU, Sabherwal N, Jantz K, Röth R, Muncke N, Blum WF, Cutler GB, Rappold G. Identification of a major recombination hotspot in patients with short stature and SHOX deficiency. *Am J Hum Genet* 2005;77:89–96.

29 Shears DJ, Vassal HJ, Goodman FR, Palmer RW, Reardon W, Superti-Furga A, Scambler PJ, Winter RM. Mutation and deletion of the pseudoautosomal gene SHOX cause Leri-Weill dyschondrosteosis. *Nat Genet* 1998;19:70–3.

30 Benito-Sanz S, Thomas NS, Huber C, Huber C, Gorbenko del Blanco D, Del Blanco DG, Aza-Carmona M, Crolla JA, Maloney V, Rappold G, Argente J, Argente J, Campos-Barros A, Cormier-Daire V, Cormier-Daire V, Heath KE. A novel class of pseudoautosomal region 1 deletions downstream of SHOX is associated with Leri-Weill dyschondrosteosis. *Am J Hum Genet* 2005;77:533–44.

31 Chen J, Wildhardt G, Zhong Z, Röth R, Weiss B, Steinberger D, Decker J, Blum WF, Rappold G. Enhancer deletions of the SHOX gene as a frequent cause of short stature: the essential role of a 250 KB downstream regulatory domain. *J Med Genet* 2009;46:834–9.

32 Hirschfeldova K, Solc R. Comparison of SHOX and associated elements duplications distribution between patients (Léri-Weill dyschondrosteosis/idiopathic short stature) and population sample. *Gene* 2017;627:164–8.

33 Moon DK, Gurnett CA, Aferol H, Siegel MJ, Commean PK, Dobbs MB. Soft-Tissue abnormalities associated with treatment-resistant and Treatment-Responsive clubfoot: findings of MRI analysis. *J Bone Joint Surg Am* 2014;96:1249–56.

34 Upners EN, Jensen RB, Rajpert-De Meyts E, Dunø M, Aksglaede L, Juul A. Short stature homeobox-containing gene duplications in 3.7% of girls with tall stature and normal karyotypes. *Acta Paediatr* 2017;106:1651–7.

35 Benito-Sanz S, Belinchon-Martínez A, Aza-Carmona M, de la Torre C, Huber C, González-Casado I, Ross JL, Thomas NS, Zinn AR, Cormier-Daire V, Heath KE. Identification of 15 novel partial SHOX deletions and 13 partial duplications, and a review of the literature reveals intron 3 to be a hotspot region. *J Hum Genet* 2017;62:229–34.

36 Monzani A, Babu D, Mellone S, Genoni G, Fanelli A, Prodam F, Bellone S, Giordano M. Co-Occurrence of genomic imbalances on Xp22.1 in the SHOX region and 15q25.2 in a girl with short stature, precocious puberty, urogenital malformations and bone anomalies. *BMC Med Genomics* 2019;12:5.

37 Fukami M, Naiki Y, Muroya K, Hamajima T, Soneda S, Horikawa R, Jinno T, Katsumi M, Nakamura A, Asakura Y, Adachi M, Ogata T, Kanzaki S, Ssg J, Japanese SHOX study group. Rare pseudoautosomal copy-number variations involving SHOX and/or its flanking regions in individuals with and without short stature. *J Hum Genet* 2015;60:553–6.

38 Kleinjan DA, Lettice LA. Long-Range gene control and genetic disease. *Adv Genet* 2008;61:339–88.

39 Bunyan DJ, Taylor E-J, Maloney VK, Blyth M. Homozygosity for a novel deletion downstream of the SHOX gene provides evidence for an additional long range regulatory region with a mild phenotypic effect. *Am J Med Genet A* 2014;164A:2764–8.

40 Skuplik I, Benito-Sanz S, Rosin JM, Bobick BE, Heath KE, Cobb J. Identification of a limb enhancer that is removed by pathogenic deletions downstream of the SHOX gene. *Sci Rep* 2018;8:14292.

41 Benito-Sanz S, Royo JL, Barroso E, Paumard-Hernández B, Barreda-Bonis AC, Liu P, Gracía R, Lupski JR, Campos-Barros Ángel, Gómez-Skarmeta JL, Heath KE. Identification of the first recurrent PAR1 deletion in Léri-Weill dyschondrosteosis and idiopathic short stature reveals the presence of a novel SHOX enhancer. *J Med Genet* 2012;49:442–50.

42 Fukami M, Kato F, Tajima T, Yokoya S, Ogata T. Transactivation function of an ~800-bp evolutionarily conserved sequence at the SHOX 3' region: implication for the downstream enhancer. *Am J Hum Genet* 2006;78:167–70.

43 Woolfe A, Goodson M, Goode DK, Snell P, McEwen GK, Vavouri T, Smith SF, North P, Callaway H, Kelly K, Walter K, Abnizova I, Gilks W, Edwards YJK, Cooke JE, Elgar G.

Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol* 2005;3:e7.

44 Glaser A, Arora R, Hoffmann S, Li L, Gretz N, Papaioannou VE, Rappold GA. Tbx4 interacts with the short stature homeobox gene Shox2 in limb development. *Dev Dyn* 2014;243:629–39.

45 Ramalingam A, Zhou X-G, Fiedler SD, Brawner SJ, Joyce JM, Liu H-Y, Yu S. 16P13.11 duplication is a risk factor for a wide spectrum of neuropsychiatric disorders. *J Hum Genet* 2011;56:541–4.

46 Mehraein Y, Pfob M, Steinlein O, Aichinger E, Eggert M, Bubendorff V, Mannhart A, Müller S. 2q37.3 deletion syndrome: two cases with highly distinctive facial phenotype, discordant association with schizophrenic psychosis, and shared deletion breakpoint region on 2q37.3. *Cytogenet Genome Res* 2015;146:33–8.

47 Girirajan S, Rosenfeld JA, Cooper GM, Antonacci F, Siswara P, Itsara A, Vives L, Walsh T, McCarthy SE, Baker C, Mefford HC, Kidd JM, Browning SR, Browning BL, Dickel DE, Levy DL, Ballif BC, Platky K, Farber DM, Gowans GC, Wetherbee JJ, Asamoah A, Weaver DD, Mark PR, Dickerson J, Garg BP, Ellingwood SA, Smith R, Banks VC, Smith W, McDonald MT, Hoo JJ, French BN, Hudson C, Johnson JP, Ozmore JR, Moeschler JB, Surti U, Escobar LF, El-Khechen D, Gorski JL, Kussmann J, Salbert B, Lacassie Y, Biser A, McDonald-McGinn DM, Zackai EH, Deardorff MA, Shaikh TH, Haan E, Friend KL, Fichera M, Romano C, Gécz J, DeLisi LE, Sebat J, King M-C, Shaffer LG, Eichler EE. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet* 2010;42:203–9.

48 Lööf E. Additional challenges in children with idiopathic clubfoot: is it just the foot? *J Child Orthop* 2019;13:245–51.

49 Groth KA, Skakkebæk A, Høst C, Gravholt CH, Bojesen A. Clinical review: Klinefelter syndrome--a clinical update. *J Clin Endocrinol Metab* 2013;98:20–30.

50 Ross JL, Scott C, Marttila P, Kowal K, Nass A, Papenhausen P, Abboudi J, Osterman L, Kushner H, Carter P, Ezaki M, Elder F, Wei F, Chen H, Zinn AR. Phenotypes associated with SHOX deficiency. *J Clin Endocrinol Metab* 2001;86:5674–80.