Human Molecular Genetics, 2020, Vol. 29, No. 22 3717–3728

doi: 10.1093/hmg/ddaa236
Advance Access Publication Date: 26 October 2020
Association Studies Article

ASSOCIATION STUDIES ARTICLE

Genetic association and characterization of *FSTL5* in isolated clubfoot

Anas M. Khanshour[1], Yared H. Kidane[1], Julia Kozlitina[2,†], Reuel Cornelia[1], Alexandra Rafipay[3], Vanessa De Mello[3], Mitchell Weston[4], Nandina Paria[1], Aysha Khalid[1], Jacqueline T. Hecht[5], Matthew B. Dobbs[6], B. Stephens Richards[7,8], Neil Vargesson[3], F. Kent Hamra[9], Megan Wilson[4], Carol Wise[1,2,8,10], Christina A. Gurnett[11] and Jonathan J. Rios[1,2,8,10,*]

[1]Center for Pediatric Bone Biology and Translational Research, Scottish Rite for Children, Dallas, TX 75219, USA, [2]McDermott Center for Human Growth and Development, UT Southwestern Medical Center, Dallas, TX 75390, USA, [3]School of Medicine, Medical Sciences & Nutrition, Institute of Medical Sciences, University of Aberdeen, Aberdeen, AB25 2ZD, Scotland, UK, [4]Department of Anatomy, University of Otago, Dunedin 9016, New Zealand, [5]Department of Pediatrics, McGovern Medical School, University of Texas Health, Houston, TX 77030, USA, [6]Paley Orthopedic and Spine Institute, West Palm Beach, FL 33407, USA, [7]Department of Orthopaedics, Scottish Rite for Children, Dallas, TX 75219, USA, [8]Department of Orthopaedic Surgery, UT Southwestern Medical Center, Dallas, TX 75390, USA, [9]Department of Obstetrics and Gynecology, Cecil H. & Ida Green Center for Reproductive Biology Sciences, UT Southwestern Medical Center, Dallas, TX 75390, USA, [10]Department of Pediatrics, UT Southwestern Medical Center, Dallas, TX 75390, USA and [11]Department of Neurology, School of Medicine, Washington University, St. Louis, MO 63130, USA

*To whom correspondence should be addressed. Tel: +1 2146487658; Fax: +1 2145597872; Email: Jonathan.Rios@tsrh.org

Abstract

Talipes equinovarus (clubfoot, TEV) is a congenital rotational foot deformity occurring in 1 per 1000 births with increased prevalence in males compared with females. The genetic etiology of isolated clubfoot (iTEV) remains unclear. Using a genome-wide association study, we identified a locus within *FSTL5*, encoding *follistatin-like* 5, significantly associated with iTEV. *FSTL5* is an uncharacterized gene whose potential role in embryonic and postnatal development was previously unstudied. Utilizing multiple model systems, we found that *Fstl5* was expressed during later stages of embryonic hindlimb development, and, in mice, expression was restricted to the condensing cartilage anlage destined to form the limb skeleton. In the postnatal growth plate, *Fstl5* was specifically expressed in prehypertrophic chondrocytes. As *Fstl5* knockout rats displayed no gross malformations, we engineered a conditional transgenic mouse line (*Fstl5*^{LSL}) to overexpress *Fstl5* in skeletal osteochondroprogenitors. We observed that hindlimbs were slightly shorter and that bone mineral density was reduced in adult male, but not female, *Prrx1*-cre;*Fstl5*^{LSL} mice compared with control. No overt clubfoot-like deformity was observed in *Prrx1*-cre;*Fstl5*^{LSL} mice, suggesting FSTL5 may function in other cell types to contribute to iTEV pathogenesis. Interrogating published mouse embryonic single-cell expression data showed that *Fstl5* was expressed in cell lineage

[†]Julia Kozlitina, http://orcid.org/0000-0001-7720-2290
Received: August 10, 2020. **Revised:** September 28, 2020. **Accepted:** October 14, 2020

subclusters whose transcriptomes were associated with neural system development. Moreover, our results suggest that lineage-specific expression of the *Fstl* genes correlates with their divergent roles as modulators of transforming growth factor beta and bone morphogenetic protein signaling. Results from this study associate *FSTL5* with iTEV and suggest a potential sexually dimorphic role for *Fstl5 in vivo*.

Introduction

Talipes equinovarus (clubfoot, TEV) is a congenital rotational foot deformity occurring in ~1 per 1000 live births and affecting either or both limbs (1) (Fig. 1A). Infants are often treated with non-operative methods to gradually manipulate the foot to a corrected position, though relapse of the deformity may occur and require surgical intervention (2,3). Syndromic TEV coincides with other diagnoses, such as cerebral palsy or spina bifida, and may be caused by single-gene mutations or large chromosomal abnormalities (1,4). Most often, however, clubfoot is isolated (iTEV), presenting as a distal limb malformation without significant co-occurring sequelae. The prenatal developmental etiology of iTEV remains poorly understood, though it is generally accepted that bone, muscle, nerve and tendon may all contribute to its development.

Mendelian inheritance of iTEV has been described (5–7). Linkage analysis and resequencing identified a novel dominant-negative mutation in human paired-like homeodomain 1 (*PITX1*) that segregated with clubfoot in a multigenerational family (8). Subsequent analyses identified rare copy number variations in chromosomal regions including the *PITX1* or human T-box transcription factor 4 genes, among others, in clubfoot (9–12), though these accounted for only a small fraction of clubfoot cases. Although *PITX1* haploinsufficiency was implicated in clubfoot, misexpression of *PITX1* from deletion or structural rearrangement of the 5′ regulatory region was shown to cause Liebenberg syndrome, which is characterized by partial arm-to-leg transformation (13). Therefore, alterations in these master regulators of limb development, through genetic mutation, may lead to pleiotropic limb malformations.

Genetically engineered mouse models have shown that haploinsufficiency or misexpression of *Pitx1* results in phenotypes resembling clubfoot and Liebenberg syndrome, respectively (9,13). More recently, a spontaneous mutation in *Limk1* was identified in the '*pma*' mouse, which developed a fully penetrant recessive clubfoot-like deformity (14). In homozygous *pma/pma* mice, increased embryonic hindlimb expression and activation of Limk1 were associated with inactivation of Cofilin, leading to defects in hindlimb muscle innervation. Results from these and other mouse models (15) suggest multiple developmental origins of clubfoot.

iTEV is considered a complex genetic disorder likely impacted by a spectrum of associated genetic loci, and the prevalence varies greatly between ethnic populations and between genders (1,5,16). Here, we report results from a genome-wide associated study (GWAS) of iTEV, which identified an associated locus within *FSTL5*. We show *Fstl5* is expressed in the embryonic hindlimb and the postnatal growth plate, and conditional overexpression of *Fstl5* in osteochondroprogenitors resulted in sexually dimorphic differences in skeletal development in mice. Unlike other *Fstl* genes, our results suggest that *Fstl5* likely does not modulate transforming growth factor beta (TGF-ß) and bone morphogenetic protein (BMP) signaling. Finally, using results from publicly available mouse embryonic single-cell expression profiling, we provide evidence for a neurogenic role of *Fstl5* in clubfoot.

Results

FSTL5 is associated with iTEV

To identify novel loci associated with iTEV, we performed a GWAS with 399 Caucasian iTEV subjects (previously described; 17) and 7820 ethnicity-matched controls from the population-based Atherosclerosis Risk in Communities study (ARIC; dbGAP: phs000280.v3.p1) (18) (Supplementary Table 1, Supplementary Material, Fig. 1A). Following imputation and quality control filtering, 7 794 536 autosomal variants with overall minor allele frequency >0.02 and imputation quality (Rsq) > 0.3 were tested for association with iTEV using logistic regression, adjusting for gender and 10 principal components. A single locus on chromosome 4 exceeded genome-wide significance (Fig. 1B, Supplementary Material, Fig. 1B and C). The top-associated variant [rs76973778, Rsq = 0.87, $P = 1.92 \times 10^{-8}$; OR = 0.35 (0.26–0.60)] was located within intron 3 of the *Follistatin-like* 5 (*FSTL5*) gene (Fig. 1C). Because clubfoot is twice as frequent in males compared with females (6), we investigated the association of *FSTL5* with iTEV in males and females separately. The *FSTL5* association was more strongly evident in males ($P = 4.41 \times 10^{-8}$, OR = 0.26 [0.19–0.50]) compared with females ($P = 0.02$, OR = 0.51 [0.33–0.92]), which was attributed to reduced statistical power in analyses restricted to females (Fig. 1D and E; Supplementary Material, Table 2). To confirm the association with *FSTL5*, rs76811724, which was in high linkage-disequilibrium with rs76973778 ($R^2 = 0.94$; $D' = 1$), was genotyped in an independent ethnicity-matched cohort (see Methods, Supplementary Material, Table 1). The rs76811724 variant showed a highly significant and directionally concordant association with iTEV in the discovery cohort (Rsq = 0.89, $P = 2.06 \times 10^{-7}$; OR = 0.39 [0.26–0.58]) and the replication cohort ($P = 2.70 \times 10^{-3}$, OR = 0.62 [0.44–0.87]). After combining both the discovery and replication cohorts, the association of rs76811724 exceeded genome-wide significance ($P = 2.99 \times 10^{-9}$; OR = 0.49 [0.38–0.64]) (Fig. 1F, Supplementary Material, Table 3), providing further evidence for an association of *FSTL5* with iTEV.

Embryonic expression of *Fstl5*

Little is known regarding the expression of *Fstl5* in the hindlimb, and the top-associated rs76973778 was not positioned within an annotated regulatory element (Supplementary Material, Fig. 2) and was not previously annotated as an expression quantitative trait locus in human tissues. Therefore to assess whether Fstl5 may regulate hindlimb development *in vivo*, we first examined *Fstl5* expression in multiple diverse species. In the developing embryonic bat limb, an emerging species to study conserved regulatory mechanisms of limb development (19), expression profiling showed *Fstl5* was expressed throughout multiple developmental timepoints, albeit at lower levels

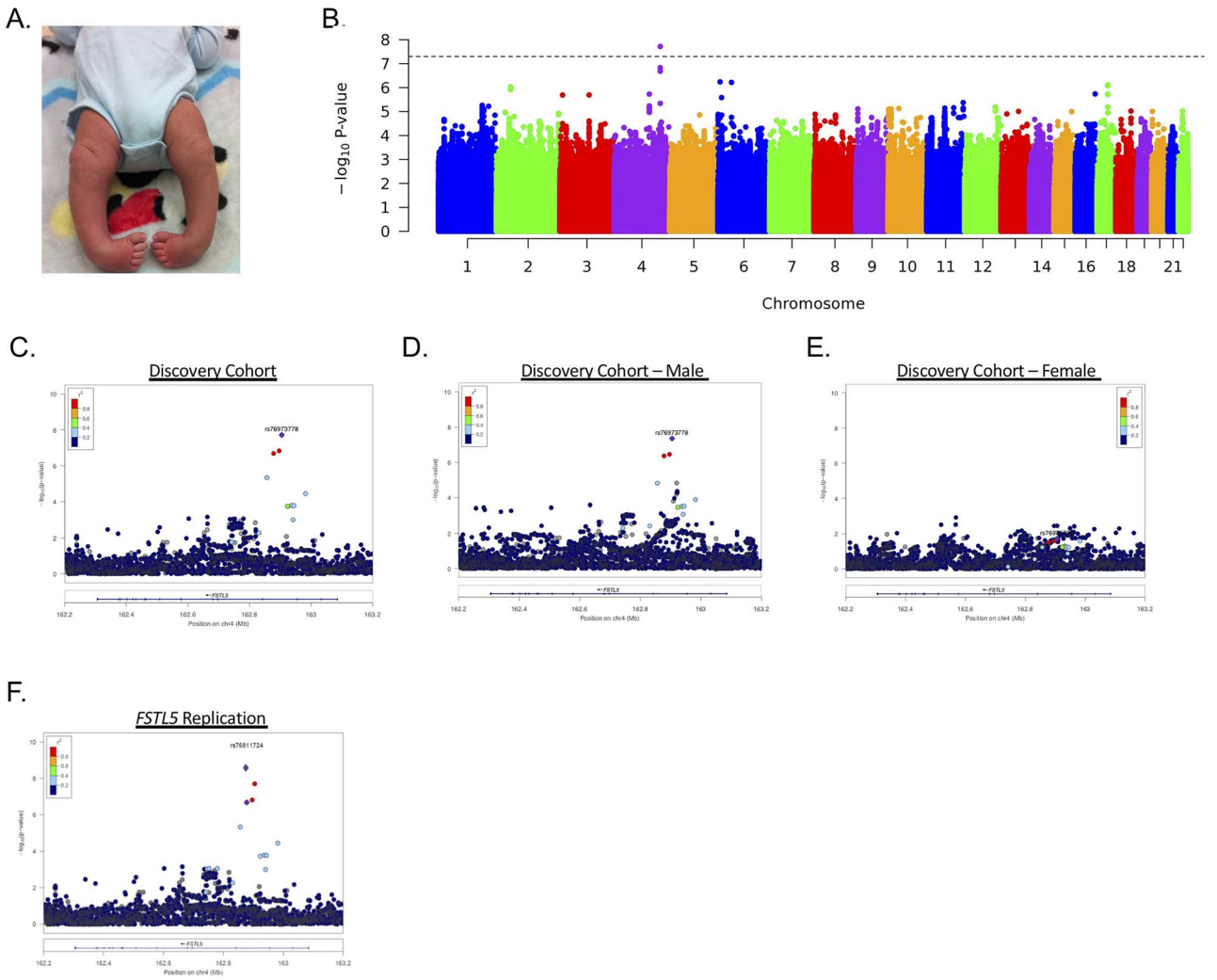

Figure 1. *Fstl5* is associated with isolated clubfoot. **(A)** Clinical image of an infant with bilateral clubfoot. **(B)** Manhattan plot of genome-wide association results. Dashed line represents genome-wide significance (5×10^{-8}). Only the *Fstl5* locus exceeded genome-wide significance. **(C–E)** LocusZoom plot showing association of rs76973778 (purple circle) within the *FSTL5* locus in (C) the entire cohort, (D) males only, and (E) females only. **(F)** Variant rs76811724 (purple circle) was genotyped for replication. This variant exceeded genome-wide significance following meta-analysis combining the original and replication cohorts (purple diamond).

compared with *Fstl1* (Fig. 2A). In developing chick embryos, *Fstl5* expression was restricted to the brain early, then was later evident in the developing limbs (Fig. 2B).

In mice, RNA-seq analysis of embryonic limb buds detected increasing *Fstl5* expression from E12.5 to E13.5 stages of development (Fig. 2C). We then evaluated expression of multiple *Fstl* genes in mouse E15.5 forelimb and hindlimb by reverse transcription quantitative polymerase chain reaction (RT-qPCR). *Fstl1* had the highest expression; however, only *Fstl5* was differentially expressed between forelimb and hindlimb, with significantly higher expression in embryonic mouse hindlimb (Fig. 2D, Supplementary Material, Fig. 3). In all assays, *Fstl5* was expressed at lower levels than *Fstl1*, which is strongly expressed throughout the entire limb (20). Therefore, we hypothesized *Fstl5* expression may either be expressed at low levels throughout the entire developing hindlimb or show a spatially restricted expression pattern, leading to overall low levels of expression. *Fstl5* mRNA expression was detected by *in situ* hybridization in multiple tissues of the developing E15.5 mouse, including the limb, forebrain and heart (Fig. 2E). Closer examination of the mouse hindlimbs showed *Fstl5* was strongly and specifically expressed within the condensing cartilage anlage, the primary source for chondrocytes during endochondral ossification (Fig. 2F). Together, these results suggest a potential role for *Fstl5* during later stages of embryonic limb development, following specification and outgrowth.

Fstl5 is expressed in the postnatal growth plate

To evaluate postnatal expression of *Fstl5 in vivo*, we utilized Sprague–Dawley rats harboring a beta-galactosidase insertion ($Fstl5^{\beta geo}$) within intron 3 of *Fstl5* that resulted in a premature truncation and significantly reduced expression (Fig. 3A and B). Heterozygous $Fstl5^{+/\beta geo}$ and knockout $Fstl5^{\beta geo/\beta geo}$ rats were born at expected Mendelian ($P = 0.38$) and male:female (Binomial $P = 0.82$) ratios with no obvious skeletal or other developmental abnormalities (Supplementary Material, Tables 4 and 5). Using the endogenous $Fstl5^{\beta geo}$ reporter, we found β-galactosidase expression was restricted to the prehypertrophic chondrocytes

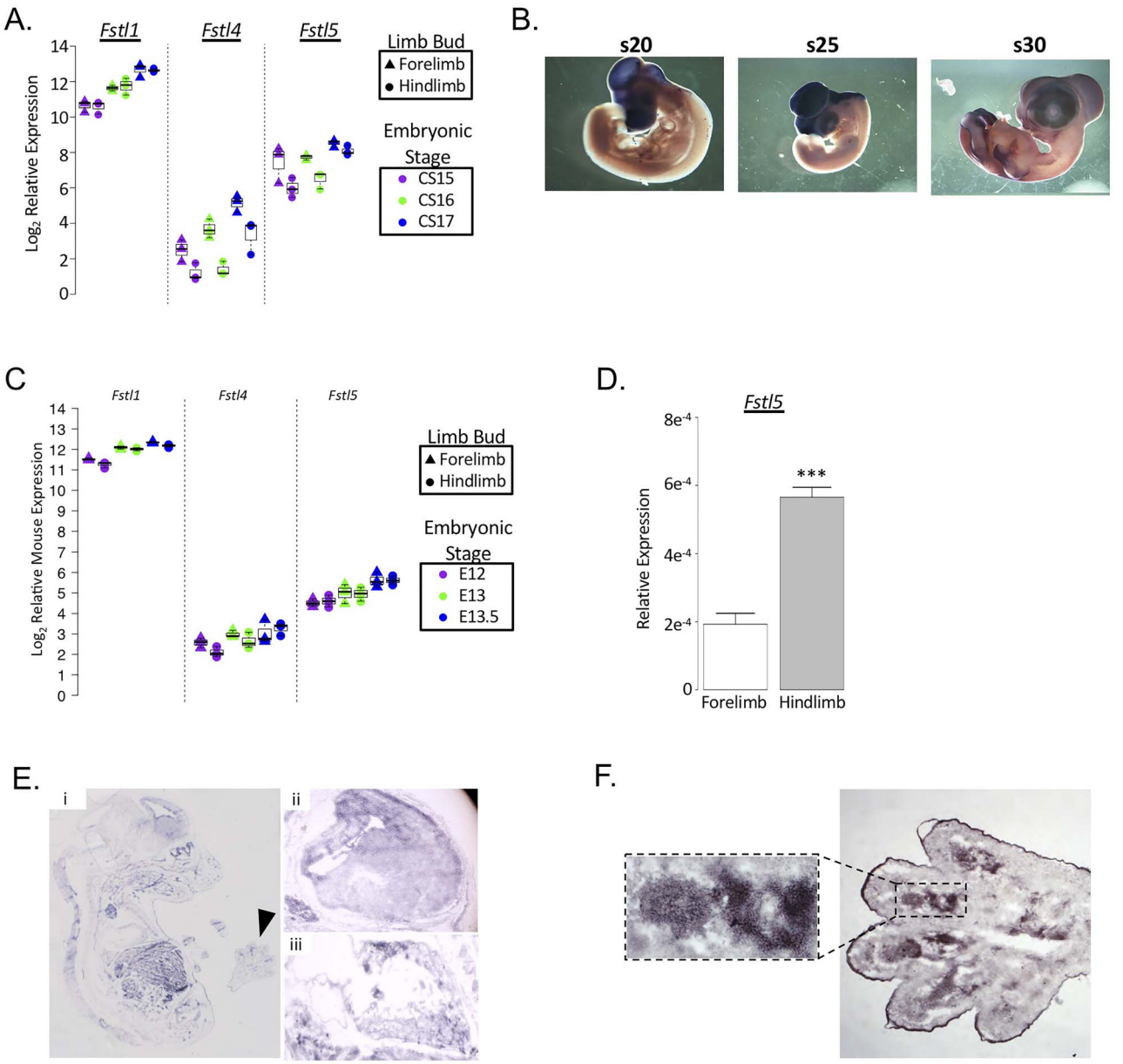

Figure 2. Evolutionarily conserved expression of *Fstl5* in the developing limb. (**A**) Expression of bat *Fstl1*, *Fstl4* and *Fstl5* in developing forelimb and hindlimb at embryonic stages CS15, CS16 and CS17 quantified using RNA-seq. Timepoints correspond to similarly staged mouse limb development shown in (**C**). (**B**) Whole-mount *in situ* hybridization showing embryonic expression of chick *Fstl5*. Expression becomes evident during later stages of limb development. (C) Expression of mouse *Fstl1*, *Fstl4* and *Fstl5* in developing forelimb and hindlimb at embryonic stages E12, E13 and E13.5 quantified using RNA-seq. Timepoints correspond to similarly staged bat limb development shown in (A). (**D**) Differential expression of mouse *Fstl5* in E15.5 forelimb and hindlimb measured by RT-qPCR ($n = 3$ independent biological replicates). Data represented as mean ± standard error. Statistically significant differences were determined using two-sided *t*-test. ***, $P < 0.001$. (**E**) Mouse E15.5 *in situ* hybridization showing *Fstl5* expression throughout the (i) entire embryo, (ii) forebrain, (iii) heart, and in the hindlimb (*arrowhead*). (**F**) *In situ* hybridization of mouse embryonic E15.5 distal hindlimb showing localized expression of *Fstl5* to the condensing cartilage mesenchyme.

of the femur growth plate and, to a lesser extent, in hypertrophic chondrocytes in 2-week-old rats (Fig. 3C). During endochondral ossification, the embryonic cartilage condensate gives rise to chondrocytes whose proliferation and maturation in the growth plate regulate subsequent bone formation (21,22). Thus, the restricted expression of *Fstl5* in the embryonic cartilage condensate and postnatal prehypertrophic chondrocytes as well as the lack of gross skeletal malformations in $Fstl5^{\beta geo/\beta geo}$ rats suggest a non-essential role for *Fstl5* in chondrocyte maturation (23).

Sexually dimorphic skeletal development in *Prrx1*-cre;$Fstl5^{LSL}$ mice

We sought to test the impact of conditional *Fstl5* overexpression in the osteochondral cell lineage. Conditional overexpression of *Fstl5* was tested in mice for two reasons. First, loss of *Fstl5* expression did not result in any observable skeletal phenotype in $Fstl5^{\beta geo/\beta geo}$ rats. Second, the orthologous position of the clubfoot-associated rs76973778 allele was within a short segment deleted from the mouse genome, preventing generation

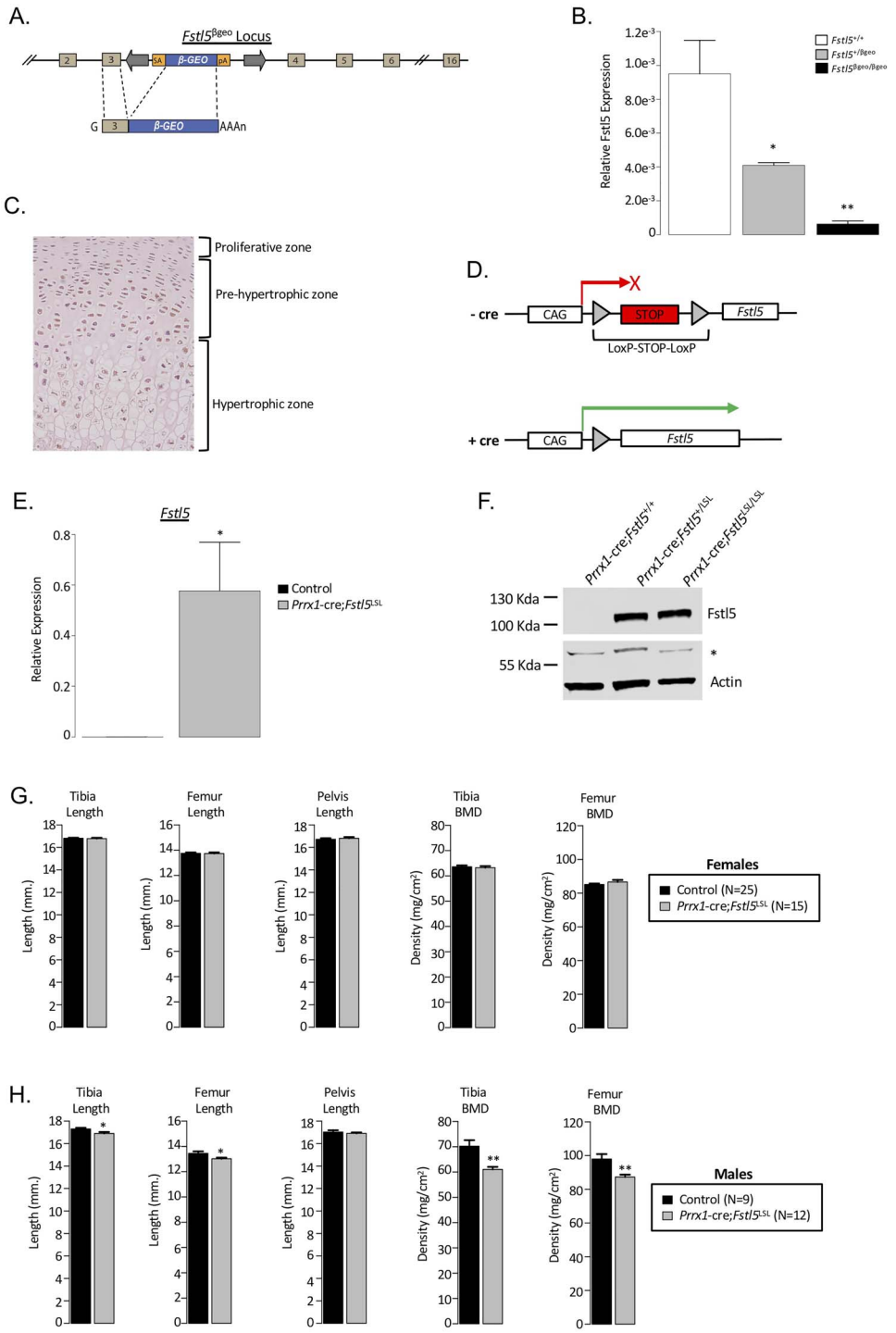

Figure 3. *In vivo* characterization of *Fstl5*. (**A**) Schematic diagram of *Sleeping Beauty* insertion of β-galactosidase within intron 3 of rat *Fstl5*, predicted to result in premature truncation. (**B**) Expression of *Fstl5* in brain tissue from control ($Fstl5^{+/+}$), heterozygous ($Fstl5^{+/\beta geo}$) and homozygous ($Fstl5^{\beta geo/\beta geo}$) rats ($n = 3$–5 animals per genotype). Statistically significant differences were determined by a one-sided *t*-test. Data represented as mean +/− standard error. (**C**) Immunohistochemistry showing β-galactosidase expression in the prehypertrophic chondrocytes of the femur growth plate of 2-week-old $Fstl5^{\beta geo}$ rats. Expression diminishes in the hypertrophic chondrocytes of the growth plate. Similar results were seen in tibia and metatarsal growth plates (data not shown). (**D**) Schematic diagram of the $Fstl5^{LSL}$ conditional overexpression cassette within the mouse *ROSA* locus. (**E** and **F**) Confirmation of cre-induced overexpression of *Fstl5* in BSCs cultured from *Prrx1*-cre;$Fstl5^{LSL}$ mice compared to control by (E) qPCR and (F) western blot. For western blotting, beta-actin (Actin) is shown as loading control. Asterisk indicates background band from Fstl5 antibody. (**G, H**) Skeletal phenotypes of 3-month-old (G) female and (H) male *Prrx1*-cre;$Fstl5^{LSL}$ mice (gray) compared with age- and gender-matched littermate controls (black). Data represented as mean ± standard error. Statistically significant differences were determined using two-sided *t*-tests. ∗, $P < 0.05$; ∗∗, $P < 0.01$.

of allele-specific knock-in mice (Supplementary Material, Fig. 2). Therefore, we engineered *Fstl5*^{LSL} mice with the *Fstl5* coding sequence under the control of a CAG promoter flanked by LoxP sites (Fig. 3D). The *Fstl5*^{LSL} expression cassette was introduced into the ROSA locus using CRISPR/Cas9, and resultant founders were out-crossed and subsequently inter-crossed with *Prrx1*-cre mice to generate heterozygous *Prrx1*-cre;*Fstl5*$^{+/LSL}$ breeders. Heterozygous mice were crossed to produce control (without *Prrx1*-cre or *Fstl5*$^{+/+}$), heterozygous, and homozygous (*Prrx1*-cre;*Fstl5*$^{LSL/LSL}$) mice. To confirm Cre-dependent overexpression of *Fstl5*, bone marrow from 5-month-old mice were flushed, and plastic-adherent bone stromal cells (BSCs) were expanded in culture and tested for Cre-dependent *Fstl5* expression. *Fstl5* was not detected in control BSCs; however, Cre-dependent *Fstl5* was robustly expressed in BSCs from *Prrx1*-cre;*Fstl5*^{LSL} mice (Fig. 3E and F; Supplementary Material, Fig. 4). *Prrx1*-cre;*Fstl5*^{LSL} mice showed no overt malformations and were visually indistinguishable from littermate control mice. We evaluated skeletal development using X-ray and dual-energy X-ray absorptiometry imaging. No differences in tibia and femur length or bone mineral density (BMD) were evident in 3-month-old female *Prrx1-Fstl5*^{LSL} mice compared to control (Fig. 3G). In contrast, tibia and femur lengths were slightly, though significantly, reduced in 3-month-old male *Prrx1*-cre;*Fstl5*^{LSL} mice compared to control (Fig. 3H). Additionally, both tibia and femur BMD were significantly reduced in male *Prrx1*-cre;*Fstl5*^{LSL} mice compared to control (Fig. 3H).

Fstl5-expressing cell lineages are enriched in neuropathic gene signatures

As no overt limb malformations were evident in *Fstl5*$^{\beta geo}$ knockout rats or in *Prrx1*-cre;*Fstl5*^{LSL} conditional overexpression mice, we sought putative functional pathways through which Fstl5 may influence hindlimb morphogenesis and development. For this, we evaluated *Fstl5* expression using available mouse embryonic single-cell transcriptomes (24) (Fig. 4A). Consistent with our *in situ* expression results (Fig. 2E), the highest levels of *Fstl5* expression were evident in neuronal cell lineages, with lower expression in skeletal lineages, such as '*Limb Mesenchyme*', '*Osteoblast*', '*Chondrocyte Progenitor*' and '*Chondrocyte/Osteoblast*' (Supplementary Material, Fig. 5). For each skeletal lineage, expression of *Fstl5* was significantly enriched in a single subcluster of cells (Fig. 4A). As the transcriptional landscape of lineage subclusters may help to broadly inform future cell fates (25), we extracted all genes with significantly enriched expression within *Fstl5*-expressing subclusters and performed pathway analyses. *Fstl5*-expressing subclusters were significantly enriched for genes involved in neural development and synaptic transmission (Fig. 4B). Neural system genes were significantly enriched in the *Fstl5*-expressing subclusters of the *Chondrocyte progenitor* ($P = 6.60e^{-8}$), *Chondrocyte and Osteoblast* ($P = 3.46e^{-10}$) and *Osteoblast* ($P = 1.71e^{-11}$) cell lineages. These results implicate a potential role for *Fstl5* in the neuropathogenesis of iTEV.

Lineage-specific gene expression suggests functional differences in Fstl proteins

Fstl proteins were originally named for having follistatin-like domains, suggesting potential roles in regulating TGF-ß and BMP signaling, as has been described for Fstl1 (26–28). However, other Fstl family members are less well characterized. Prior phylogenetic protein analyses identified sequence similarities between Fstl4 and Fstl5, which clustered most closely with Fstl1, whereas Fstl3 and Follistatin (Fst) were both unique (20,29,30). We queried expression of *Fst* and *Fstl* genes across the 38 mouse embryonic cell lineages defined by single-cell sequencing of mouse embryos (24). Surprisingly, hierarchical clustering distinguished patterns of expression between *Fstl* genes that resembled phylogenetic analyses based on protein structure (20,30) (Fig. 4C). Both *Fstl4* and *Fstl5* clustered together with the highest expression in neural cell lineages. Consistent with skeletal abnormalities in *Fstl1* knockout mice (27), *Fstl1* expression was highest in mesenchymal and skeletal cell lineages. *Fstl3* showed the highest expression in the cardiac cell lineage, among others, and mice lacking *Fstl3* were previously shown to develop cardiac abnormalities resulting in hypertension (31). These results suggest lineage-specific embryonic expression of *Fstl* genes in mice correlates with proposed functional differences among the Fstl family of proteins.

Discussion

We identified a locus within a previously uncharacterized gene, *FSTL5*, associated with isolated clubfoot. Analysis of *Fstl5* in multiple vertebrate embryos identified conserved expression in the embryonic hindlimb mesenchyme and neural tissues. Conditional overexpression of *Fstl5* in the embryonic limb mesenchyme (using *Prrx1*-cre) resulted in reduced BMD in male, but not female, mice, though no gross malformations were evident. Finally, we show that evolutionary divergence in the essential roles of Fstl proteins resembles differences in their lineage-specific expression in developing mouse embryos (24).

Follistatin-like 5—the protein encoded by *FSTL5*—is a member of the secreted protein acidic rich in cysteines family of proteins (30). Follistatin is a secreted TGF-ß and BMP antagonist, and because Fstl1 harbors a domain with modest sequence homology to Follistatin, Fstl proteins were hypothesized to also regulate TGF-ß and BMP signaling (20,32,33). Consistent with this, *Fstl1* knockout mice developed lung and skeletal defects and died shortly after birth from respiratory insufficiency (26,27). Skeletal abnormalities included defects in axial skeleton patterning, long bone dysplasia and a hindfoot rotational deformity caused by malposition of the distal fibula (27). Moreover, *Fstl1* was shown to promote chondrogenic lineage differentiation of mesenchymal stromal cells *ex vivo* (34). However, functional similarities between Fstl1 and other Fstl proteins, such as Fstl5, were unclear.

Although the function of FSTL5 was largely unstudied, its mouse homolog (Fstl5) shared highest sequence similarity to Fstl4, another uncharacterized protein, and Fstl1 (20). *Fstl4* expression was localized to neural tissues, and though neurologic differences were observed in *Fstl4* knockout mice, a lack of gross developmental deformities suggested Fstl4 likely does not regulate BMP signaling (35,36). Like *Fstl4*, *Fstl5* was expressed more highly in neural tissues, and rats lacking *Fstl5* showed no gross developmental malformations. Our results suggest that Fstl5, like Fstl4, does not modulate TGF-ß and BMP signaling and is not essential for embryonic skeletal development, unlike Fst and Fstl1 (26–28,37). Moreover, lineage-specific expression of *Fstl* genes in mouse embryos is supported by the functional differences between Fstl proteins and differences in developmental phenotypes among genetically engineered mouse models. Expression levels of both *Fstl4* and *Fstl5* were highest in neural tissues, whereas *Fstl3* and *Fstl1* were expressed highest in cardiac and osteochondroprogenitor lineages, respectively.

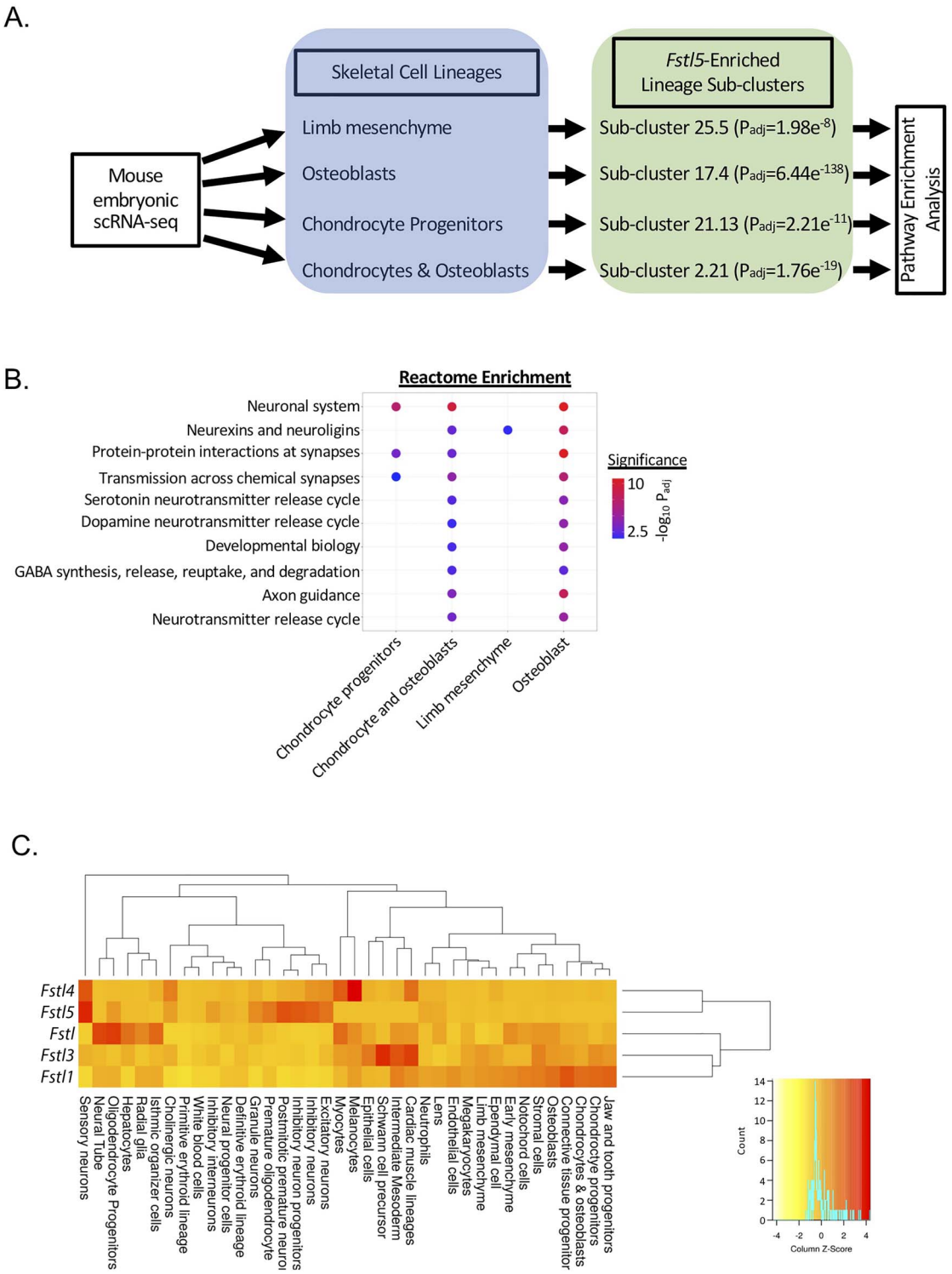

Figure 4. Analyses of mouse embryonic single-cell gene expression. (**A**) Schematic of *Fstl5* expression analyses using publicly available results from mouse embryonic scRNAseq (ref. 24). (**B**) Enrichment of Reactome pathways for *Fstl5*-expressing subclusters. (**C**) Bi-clustering of *Fstl* genes and their expression across the various cell lineages.

Here, the identification of human *FSTL5* associated with iTEV by GWAS, together with comparative analyses of *Fstl5* across diverse genetic models (rodents, birds and bats), point to a neurogenic mechanism and potential sexually dimorphic role for *FSTL5* in iTEV pathophysiology. A neuropathologic mechanism for clubfoot development is further supported by the recent characterization of a mouse model of peroneal muscular atrophy (14). Additional experiments are now required to understand what role Fstl5 plays in the nervous system and how alterations in Fstl5 expression/function may contribute to limb deformity. The

conditional transgenic mouse model, *Fstl5*^{LSL}, reported here now opens the door to more in depth analyses of how *Fstl5* expression in distinct cell lineages can impact limb development. As conditional overexpression in skeletal progenitor cells did not result in significant limb deformity, the impact on limb development following overexpression in neuro-lineage cells remains to be tested. Finally, analysis of *FSTL5* expression in human nervous system tissues is warranted; however, the correlation to clubfoot pathogenesis may be limited since infants with isolated clubfoot mostly undergo non-operative treatments. Our results implicate a neurogenetic pathology to clubfoot pathogenesis.

Materials and Methods

Cohort descriptions

For the discovery cohort analyses, non-Hispanic white clubfoot subjects were recruited from St Louis Children's Hospital and St Louis Shriners Hospital. The study protocol was approved by the Institutional Review Board, and all subjects and/or parents gave informed consent. Patients were diagnosed at infancy with TEV (clubfoot) based on the physical examination findings by a single orthopedic surgeon. Exclusion criteria included additional congenital anomalies, developmental delay or known underlying etiologies such as arthrogryposis, myelomeningocele or myopathy. Control subjects consisted of an ethnicity-matched subset of the ARIC (18). The ARIC cohort is a prospective community-based recruitment study focused primarily on atherosclerosis and cardiovascular outcomes. Ethnicity was confirmed for both cohorts by principal component analysis compared with HapMap populations (Supplementary Material, Fig. 1A).

For replication analyses, non-Hispanic white clubfoot subjects were recruited from Scottish Rite for Children (Dallas, TX), Shriners Hospital for Children (Houston, TX), the St. Louis Children's Hospital (St. Louis, MO) and St. Louis Shriners Hospital (St. Louis, MO). All subjects provided written informed consent approved by their respective Institutional Review Boards. Inclusion and exclusion criteria were the same as for the discovery cohort. Ethnicity-matched control subjects used for replication were recruited, in part, from Scottish Rite for Children (Dallas, TX). Additional control subjects were included from the Dallas Heart Study (38)—a multiethnic population-based probability sampling of Dallas County (Dallas, TX)—and from a randomized trial with work-place recruitment from the Dallas-Fort Worth metroplex (39). Recruitment was unrelated to clubfoot. The allele frequency of the rs76811724 variant among control subjects of the replication cohort was similar to the allele frequency among non-Hispanic whites from the gnomAD consortium (frequency = 0.08).

Genome-wide imputation and association

Genome-wide genotyping of iTEV subjects was performed using the Affymetrix 6.0 microarray as previously described (17). Control subjects available from the population-based Atherosclerosis Risk in Communities were genotyped on the Affymetrix 6.0 microarray, and genotype data were obtained from dbGAP [dbGAP: phs000280]. Quality-control was performed to exclude individuals with evidence of chromosomal abnormalities, gender inconsistencies and excessive genotype missingness. Case and control datasets were merged and genotypes harmonized using Plink v1.9 (40). Duplicate and related samples were identified using a subset of linkage disequilibrium (LD)-pruned variants, and only unrelated individuals were included in the imputation. LD-pruned variants were used for principal component analysis including HapMap Phase 3 samples for ethnicity verification (Supplementary Material, Fig. 1A). Prior to imputation, individual single nucleotide polymorphisms (SNPs) were excluded based on missingness rate (>5%), minor allele frequency (<0.01), deviation from Hardy–Weinberg equilibrium (HWE $P < 1 \times 10^{-4}$ in controls), non-variant monomorphic SNPs, and SNPs with a differential missingness rate between case and control cohorts. After filtering, 549 589 autosomal SNPs were available for imputation.

Imputation was performed with MACH v1.0 (41) and Minimac (42) on the merged case-control cohort following pre-phasing with the 1000 Genomes Phase 3 reference panel. Imputation accuracy (Rsq) was investigated for different ranges of allele frequency, and a marked decline in the percent of high-quality imputed variants (with Rsq >0.3) was evident for those with minor allele frequency <2% (Supplementary Material, Table 6, Supplementary Material, Fig. 1B). As a majority of imputed variants were rare or monomorphic and thus poorly imputed, only variants with frequency >2%, Rsq > 0.3, and without significant HWE deviation were included for association analysis. Following postimputation filtering, imputed dosages of 7 794 536 variants were included for logistic regression analysis with *Mach2dat*, which included gender and 10 principal components as covariates. Association results were plotted using R and LocusZoom (43).

Replication genotyping was performed using allele-discrimination assays (ThermoFisher). The top-associated variant (rs76973778) was genotyped using a commercially available assay; however, the commercial assay was unable to accurately discriminate between genotypes, which was attributed to adjacent variants affecting the design specificity. Therefore, a custom allele-discrimination assay was designed to genotype the highly correlated rs76811724 (Supplementary Material, Table 7). The replication cohort consisted of ethnicity-matched subjects with isolated clubfoot recruited at St. Louis Children's Hospital, Shriners Hospital-St. Louis, Scottish Rite for Children and Shriners Hospital for Children-Houston. All subjects provided written informed consent approved by the respective Institutional Review Boards. Genotypes were merged with the discovery cohort, and association analysis performed using logistic regression analysis including gender as a covariate.

Production of *Fstl5*$^{\beta geo}$ rats

Generation of *Fstl5*$^{\beta geo}$ rats were previously described (44,45). Briefly, *Sleeping Beauty β-Geo* trap transposons (46) were used to select mutant rat spermatogonial libraries *in vitro* (44). Spermatogonia comprising a selected library were then transplanted into rat testes for production of mutant spermatozoa (47). Recipient males were bred with wild-type females to produce a random panel of donor spermatozoa-derived mutant rat strains enriched with gene traps in protein coding genes (44). Genomic sites of transposon integration were defined in the newly generated mutant rats by splinkerette PCR (44) with sequence analysis alignment on genome build RGSC v3.4 (Rn4).

For this study, testis cells from a rat harboring a *Sleeping Beauty β-Geo* trap transposon in intron 3 of *Fstl5* (44) were thawed from cryopreservation and used to derive a donor spermatogonial stem cell line that was transplanted into recipient males, as described (48). Recipient males were bred with wild-type females to regenerate the *Fstl5* mutant rat strain. Two founders were established from cryopreservation, both harboring the *Fstl5* intron 3 *β-Geo* trap transposon. *Fstl5*

gene-specific PCR primers near *Sleeping Beauty* integration sites were used in combination with transposon-specific primers to genotype progeny (Supplementary Material, Table 8). *Fstl5* mutant rats were housed in individually ventilated, Lab Products 2100 cages in a dedicated room with atmosphere controls set to 72°F, 45–50% humidity during a 12 h light/dark cycle (i.e. Light cycle = 6:00 a.m.–6:00 p.m., Central Standard Time adjusted for daylight savings time). Rats were fed Harlan Teklad Irradiated 7912, LM-485 Mouse/Rat Diet, 5% fat Diet with a continuous supply of reverse osmosis water. All rat protocols were approved by the University of Texas Southwestern Medical Center.

Expression analysis

For analysis of *Fstl5*$^{\beta geo}$ rats, brain tissue was harvested from 3-week-old rats and homogenized in 500 µl TRIzol reagent (Invitrogen). Samples were vortexed and incubated at room temperature for 5 min and 100 µl chloroform added. Samples were vortexed and centrifuged at 12 000×*g* for 15 min at 4°C. RNA was precipitated using 250 µl isopropyl alcohol and incubated at room temperature for 10 min followed by centrifugation at 12 000×*g* for 10 min at 4°C. The RNA pellet was subsequently washed twice with 500 µl 75% ethanol followed by centrifugation at 7500×*g* for 5 min at 4°C. RNA was rehydrated in 50 µl nuclease free water and concentrations estimated using a Nanodrop spectrophotometer. *Fstl5* expression was assayed using SYBR green qPCR together with *Gapdh* for normalization. Relative expression levels were compared using one-sided *t*-tests compared with *Fstl5*$^{+/+}$ control. Rat qPCR primer sequences are listed in Supplementary Material, Table 8. Experimental protocols were approved by the University of Texas Southwestern Medical Center.

For mouse *Fstl* limb bud expression, forelimb and hindlimb buds were dissected from E15.5 embryos, RNA extracted using PureLink RNA Mini kit (Ambion), and samples pooled (1 l per pool). cDNA was generated using iScript (BioRad) and qPCR performed using SybrSelect Mastermix (ThermoFisher). Assays were run in triplicate and expression normalized to *ActinB*. Relative expression levels were compared using two-sided *t*-tests. Mouse qPCR primer sequences are listed in Supplementary Material, Table 8. Experimental procedures were approved by the Otago University Animal Ethics committee.

Immunohistochemistry

Femurs were dissected from 2-week-old *Fstl5*$^{\beta geo}$ rats and fixed in 10% formalin followed by decalcification in 14% EDTA in PBS and embedded in paraffin. Four-micron sections were cut and placed on 3-aminopropyltriethoxysaline (APES)-treated slides. Slides were dried at 37°C, deparaffinized, and digested in 0.25% Trypsin in 1 mM EDTA for 3 min at 37°C incubator. Endogenous peroxidase activity was deactivated using 3% hydrogen peroxide in Methanol for 10 min at room temperature and rinsed in PBS with 0.1% Tween 20 (PBST). Sections were blocked in 10% Normal Goat Serum (Dako) for 30 min at room temperature and incubated overnight at 4°C with Anti-Beta Galactosidase Mouse antibody (Promega, 1:500). Following, slides were washed in PBST 3 times and incubated with HRP-conjugated Goat Anti Mouse IgG Antibody, (Millipore) for 1.5 h and washed in PBST 3×. The DAB substrate–chromogen (DAKO) was applied for 10 min, counterstained in Mayer's hematoxylin for 2 min, and washed in distilled water. Sections were dehydrated in alcohol, cleared in xylene, and mounted with Cytoseal XYL (Thermo Scientific) prior to imaging with a DP73 Olympus microscope using Cell Sens software.

In situ hybridization

A 505 bp region of the mouse *Fstl5* mRNA was amplified from embryonic E15.5 hindlimb cDNA using oligonucleotide primers (Fstl5F:5′-GGAAGGCTAAGCTCTGCATATT-3′, and Fstl5R:5′-GCACTACAGAGAGTGGTTTCAG-3′). The PCR product was cloned into pGEMTEasy (Promega) and sequenced to confirm orientation. Dioxigenin (DIG)-labeled probe synthesis (sense and antisense) was carried out as published previously (49).

Embryos from time-mated mice were removed and frozen in OCT prior to sectioning (10 µm) for *in situ* hybridization. Thawed sections were air-dried and fixed with 4% paraformaldehyde (PFA)/phosphate buffered saline (PBS) for 10 min, washed with PBT (PBS with 10% Tween-20) three times before proteinase K digestion (1 µg/ml for 10 min). Following, sections were re-fixed with PFA/PBS and acetylated (625 µl acetic anhydride and 295 ml 0.1 M triethanolamine) for 10 min at room temperature. Sections were washed three times with PBT prior to hybridization with the appropriate sense or antisense probe in hybridization buffer (50% formamide, 5× saline sodium citrate buffer (SSC), 5× Denhardts, 250 µg/ml yeast RNA, 500 µg/ml herring sperm DNA). Slides were incubated overnight at 65°C in a humidified chamber. Post-hybridization washes were done at 65°C with a series of wash buffers (1xSSC/50% formamide; 2xSSC and 0.2x SSC (twice)) for 20 min each. Sections were washed twice at room temperature with maleic acid buffer with Tween 20 (MABT) before blocking with 20% heat inactivated sheep serum (HISS), 2% Roche blocking powder in MABT. To detect the DIG-probe, slides were incubated overnight at 4°C with anti-DIG-alkaline phosphatase (Roche) diluted 1:2500 in 5% HISS/MABT. Unbound antibody was removed by washing with MABT, followed by NTM buffer (100 mM NaCl, 100 mM Tris-HCl (pH 9.5), 50 mM $MgCl_2$) with nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl-phostphate (BCIP) (Roche) to detect alkaline phosphatase. Following the color reaction, slides were washed and fixed with 4% PFA/PBS before mounting for imaging. Sections were imaged on an Olympus AX70 light microscope. All mouse experimental protocols were approved by the Otago University Animal Ethics committee.

Chicken *Fstl5* was cloned from a cDNA library from day 4 chicken embryos using PCR primers and subcloned into a pBS vector. In-situ hybridization analysis was carried out using previously published methods (50,51). Experimental procedures were performed in accordance with UK Home Office Animal licensing and in accordance with University of Aberdeen ethical review committees.

Generation of the *Fstl5*^{LSL} mouse line

A vector containing the CAG promoter flanked by LoxP sites upstream of the Fstl5 coding sequence was designed, produced and purified using VectorBuilder (https://en.vectorbuilder.com) and Cyagen US (Santa Clara, CA). Mice were engineered using CRISPR/Cas9 via pronuclear injection of 1-cell stage C57BL/6 N embryos. Injections included 25 ng/µl Cas9 protein (Integrated DNA Technologies), 25 ng/µl crRNA (5′-CGCCCATCTTCTAGAAAGAC-3′) complexed with tracRNA scaffold, and 10 ng/µl circular plasmid, diluted in TE buffer. Following injection, embryos were implanted into pseudo-pregnant females. Multiple founder lines were established and maintained by out-crossing to C57BL/6 J mice.

Fstl5 ^{LSL} mice were crossed to *Prrx1*-cre mice, and resulting *Prrx1*-cre;*Fstl5*$^{+/LSL}$ mice were crossed to generate control (no *Prrx1*-cre or *Fstl5*$^{+/+}$), *Prrx1*-cre;*Fstl5*$^{+/LSL}$ heterozygous, and *Prrx1*-cre;*Fstl5*$^{LSL/LSL}$ homozygous mice. *Cre* and *Fstl5*^{LSL} genotyping primer sequences are provided in Supplementary Material, Table 8. Cre-induced overexpression was validated by qPCR and western blotting using cultured BSCs from adults *Prrx1*-cre;*Fstl5*^{LSL} mice compared to control.

Mouse protocols were approved by the UT Southwestern Medical Center Institutional Animal Care and Use Committee.

Cell culture, expression analysis and western blotting

For *ex vivo* BSC culture, bone marrow was flushed from adult mice and cultured in alpha-MEM (Thermofisher Scientific) with 10% FBS (Sigma) and 1% antibiotics (penicillin/streptomycin) (Thermofisher Scientific) for 5 days at 37°C and 5% CO_2. Following, media was refreshed and non-plastic adherent cells washed away. The cells were then plated in 12-well plates and 100 mm plates for RNA and protein extraction, respectively and grown to confluence.

For RNA extraction, the cells were harvested in RLT Plus buffer with beta-marcaptoethanol (Sigma). The cells were passed through Qiashredder columns (Qiagen) prior to RNA extraction. RNeasy Plus Mini Kit (Qiagen) was used to extract RNA following manufacturer's recommendations. The samples were quantified using a Nanodrop1000 spectrophotometer (Thermofisher Scientific). cDNA synthesis was performed using the High Capacity RNA to cDNA Kit (Thermofisher Scientific). The qPCR assays were performed using SYBR Green PCR Master Mix (Thermofisher Scientific). Protein was extracted using RIPA buffer with protease inhibitor cocktail (Roche) and concentration was determined using the Pierce BCA Protein Assay (Thermofisher Scientific). Approximately 10 µg of protein was loaded on SDS-PAGE. The antibody detection was performed using rabbit anti-Fstl5 (1:2000) and rabbit anti-Actb (1:5000) (Cell Signaling Technologies) in 5% BSA in TBST buffer overnight at 4°C. After washing, goat anti-rabbit (1:10 000) IRDye secondary antibody (Licor Biosciences) was incubated in 5% non-fat dry milk in TBST buffer for 1 h at room temperature in the dark. Following washing, images were acquired using the Odyssey CLx system (Licor Biosciences).

Skeletal assessments

For live-animal rodent imaging, mice were anesthetized using 2% inhaled isoflurane. Images were acquired using a Faxitron UltrafocusDXA. Bone densities and bone lengths were measured using Faxitron software and compared with littermate controls using two-sided t-tests.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C and HHSN268201100012C). The authors thank the staff and participants of the ARIC study for their important contributions. Funding for GENEVA was provided by National Human Genome Research Institute (grant U01HG004402) (E. Boerwinkle). We thank H. Hobbs and J. Cohen for contributing control samples for replication genotyping, Nadav Ahituv for sharing RNA-seq data for both bat and mouse embryonic limb buds, Tommy Hyatt for designing the custom genotyping assay and members of the UT Southwestern Transgenic Core facility, including John Ritter, Mylinh Nguyen and Robert Hammer. Publicly available mouse embryonic expression analysis results were provided online at https://oncoscape.v3.sttrcancer.org/atlas.gs.washington.edu.mouse.rna/landing (24). The authors acknowledge the contributions and support of the Center for Excellence in Clubfoot Research at Scottish Rite for Children, including Shawne Faulks and Kristhen Atala.

Conflict of Interest Statement. The authors have no conflict of interest to declare.

Funding

Fstl5 mutant rats were produced by the NIH Mutant Rat Resource at UT Southwestern Medical Center (R24RR03232601, R24OD011108, R01HD036022 and 5R01HD053889). This study was supported by funding from the Scottish Rite for Children Research Fund (J.J.R.), Shriners Hospital for Children (J.T.H) and the National Institutes of Health award R01HD043342 (J.T.H.). A.R. and V.D.M. were supported by BBSRC EastBIO DTP PhD Studentships.

References

1. Wynne-Davies, R. (1972) Genetic and environmental factors in the etiology of talipes equinovarus. *Clin. Orthop. Relat. Res.*, **84**, 9–13.
2. Siebert, M.J., Karacz, C.M. and Richards, B.S. (2020) Successful Ponseti-treated clubfeet at age 2 years: what is the rate of surgical intervention after this? *J. Pediatr. Orthop.*, **40**, 597–603.
3. Faulks, S. and Richards, B.S. (2009) Clubfoot treatment: Ponseti and French functional methods are equally effective. *Clin. Orthop. Relat. Res.*, **467**, 1278–1282.
4. Brewer, C., Holloway, S., Zawalnyski, P., Schinzel, A. and FitzPatrick, D. (1998) A chromosomal deletion map of human malformations. *Am. J. Hum. Genet.*, **63**, 1153–1159.
5. Werler, M.M., Yazdy, M.M., Mitchell, A.A., Meyer, R.E., Druschel, C.M., Anderka, M., Kasser, J.R. and Mahan, S.T. (2013) Descriptive epidemiology of idiopathic clubfoot. *Am. J. Med. Genet. A*, **161A**, 1569–1578.
6. Lochmiller, C., Johnston, D., Scott, A., Risman, M. and Hecht, J.T. (1998) Genetic epidemiology study of idiopathic talipes equinovarus. *Am. J. Med. Genet.*, **79**, 90–96.
7. Beals, R.K. (1978) Club foot in the Maori: a genetic study of 50 kindreds. *N. Z. Med. J.*, **88**, 144–146.
8. Gurnett, C.A., Alaee, F., Kruse, L.M., Desruisseau, D.M., Hecht, J.T., Wise, C.A., Bowcock, A.M. and Dobbs, M.B. (2008) Asymmetric lower-limb malformations in individuals with homeobox PITX1 gene mutation. *Am. J. Hum. Genet.*, **83**, 616–622.
9. Alvarado, D.M., McCall, K., Aferol, H., Silva, M.J., Garbow, J.R., Spees, W.M., Patel, T., Siegel, M., Dobbs, M.B. and Gurnett, C.A. (2011) Pitx1 haploinsufficiency causes clubfoot in humans and a clubfoot-like phenotype in mice. *Hum. Mol. Genet.*, **20**, 3943–3952.
10. Lu, W., Bacino, C.A., Richards, B.S., Alvarez, C., VanderMeer, J.E., Vella, M., Ahituv, N., Sikka, N., Dietz, F.R., Blanton, S.H. and Hecht, J.T. (2012) Studies of TBX4 and chromosome

17q23.1q23.2: an uncommon cause of nonsyndromic clubfoot. *Am. J. Med. Genet. A*, **158A**, 1620–1627.
11. Alvarado, D.M., Aferol, H., McCall, K., Huang, J.B., Techy, M., Buchan, J., Cady, J., Gonzales, P.R., Dobbs, M.B. and Gurnett, C.A. (2010) Familial isolated clubfoot is associated with recurrent chromosome 17q23.1q23.2 microduplications containing TBX4. *Am. J. Hum. Genet.*, **87**, 154–160.
12. Alvarado, D.M., Buchan, J.G., Frick, S.L., Herzenberg, J.E., Dobbs, M.B. and Gurnett, C.A. (2013) Copy number analysis of 413 isolated talipes equinovarus patients suggests role for transcriptional regulators of early limb development. *Eur. J. Hum. Genet.*, **21**, 373–380.
13. Spielmann, M., Brancati, F., Krawitz, P.M., Robinson, P.N., Ibrahim, D.M., Franke, M., Hecht, J., Lohan, S., Dathe, K., Nardone, A.M. *et al.* (2012) Homeotic arm-to-leg transformation associated with genomic rearrangements at the PITX1 locus. *Am. J. Hum. Genet.*, **91**, 629–635.
14. Collinson, J.M., Lindstrom, N.O., Neves, C., Wallace, K., Meharg, C., Charles, R.H., Ross, Z.K., Fraser, A.M., Mbogo, I., Oras, K. *et al.* (2018) The developmental and genetic basis of 'clubfoot' in the peroneal muscular atrophy mutant mouse. *Development*, **145**, 1–15.
15. Helmbacher, F., Schneider-Maunoury, S., Topilko, P., Tiret, L. and Charnay, P. (2000) Targeting of the EphA4 tyrosine kinase receptor affects dorsal/ventral pathfinding of limb motor axons. *Development*, **127**, 3313–3324.
16. Dobbs, M.B. and Gurnett, C.A. (2012) Genetics of clubfoot. *J. Pediatr. Orthop. B*, **21**, 7–9.
17. Zhang, T.X., Haller, G., Lin, P., Alvarado, D.M., Hecht, J.T., Blanton, S.H., Stephens Richards, B., Rice, J.P., Dobbs, M.B. and Gurnett, C.A. (2014) Genome-wide association study identifies new disease loci for isolated clubfoot. *J. Med. Genet.*, **51**, 334–339.
18. THE ARIC INVESTIGATORS (1989) The Atherosclerosis Risk in Communities (ARIC) study: design and objectives. The ARIC investigators. *Am. J. Epidemiol.*, **129**, 687–702.
19. Eckalbar, W.L., Schlebusch, S.A., Mason, M.K., Gill, Z., Parker, A.V., Booker, B.M., Nishizaki, S., Muswamba-Nday, C., Terhune, E., Nevonen, K.A. *et al.* (2016) Transcriptomic and epigenomic characterization of the developing bat wing. *Nat. Genet.*, **48**, 528–536.
20. Adams, D., Larman, B. and Oxburgh, L. (2007) Developmental expression of mouse Follistatin-like 1 (Fstl1): dynamic regulation during organogenesis of the kidney and lung. *Gene Expr. Patterns*, **7**, 491–500.
21. Karsenty, G. (2003) The complexities of skeletal biology. *Nature*, **423**, 316–318.
22. Kronenberg, H.M. (2003) Developmental regulation of the growth plate. *Nature*, **423**, 332–336.
23. Kozhemyakina, E., Lassar, A.B. and Zelzer, E. (2015) A pathway to bone: signaling molecules and transcription factors involved in chondrocyte development and maturation. *Development*, **142**, 817–831.
24. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J. *et al.* (2019) The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, **566**, 496–502.
25. Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F.D. and Klein, A.M. (2020) Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science*, **367**, eaaw3381.
26. Geng, Y., Dong, Y., Yu, M., Zhang, L., Yan, X., Sun, J., Qiao, L., Geng, H., Nakajima, M., Furuichi, T. *et al.* (2011) Follistatin-like 1 (Fstl1) is a bone morphogenetic protein (BMP) 4 signaling antagonist in controlling mouse lung development. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 7058–7063.
27. Sylva, M., Li, V.S., Buffing, A.A., van Es, J.H., van den Born, M., van der Velden, S., Gunst, Q., Koolstra, J.H., Moorman, A.F., Clevers, H. *et al.* (2011) The BMP antagonist follistatin-like 1 is required for skeletal and lung organogenesis. *PLoS One*, **6**, e22616.
28. Matzuk, M.M., Lu, N., Vogel, H., Sellheyer, K., Roop, D.R. and Bradley, A. (1995) Multiple defects and perinatal death in mice deficient in follistatin. *Nature*, **374**, 360–363.
29. Glusman, G., Kaur, A., Hood, L. and Rowen, L. (2004) An enigmatic fourth runt domain gene in the fugu genome: ancestral gene loss versus accelerated evolution. *BMC Evol. Biol.*, **4**, 43.
30. Sylva, M., Moorman, A.F. and van den Hoff, M.J. (2013) Follistatin-like 1 in vertebrate development. *Birth Defects Res. C Embryo Today*, **99**, 61–69.
31. Mukherjee, A., Sidis, Y., Mahan, A., Raher, M.J., Xia, Y., Rosen, E.D., Bloch, K.D., Thomas, M.K. and Schneyer, A.L. (2007) FSTL3 deletion reveals roles for TGF-beta family ligands in glucose and fat homeostasis in adults. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 1348–1353.
32. Keutmann, H.T., Schneyer, A.L. and Sidis, Y. (2004) The role of follistatin domains in follistatin biological action. *Mol. Endocrinol.*, **18**, 228–240.
33. Massague, J. and Chen, Y.G. (2000) Controlling TGF-beta signaling. *Genes Dev.*, **14**, 627–644.
34. Chaly, Y., Blair, H.C., Smith, S.M., Bushnell, D.S., Marinov, A.D., Campfield, B.T. and Hirsch, R. (2015) Follistatin-like protein 1 regulates chondrocyte proliferation and chondrogenic differentiation of mesenchymal stem cells. *Ann. Rheum. Dis.*, **74**, 1467–1473.
35. Yonehara, K., Shintani, T., Suzuki, R., Sakuta, H., Takeuchi, Y., Nakamura-Yonehara, K. and Noda, M. (2008) Expression of SPIG1 reveals development of a retinal ganglion cell subtype projecting to the medial terminal nucleus in the mouse. *PLoS One*, **3**, e1533.
36. Suzuki, R., Fujikawa, A., Komatsu, Y., Kuboyama, K., Tanga, N. and Noda, M. (2018) Enhanced extinction of aversive memories in mice lacking SPARC-related protein containing immunoglobulin domains 1 (SPIG1/FSTL4). *Neurobiol. Learn. Mem.*, **152**, 61–70.
37. Welt, C., Sidis, Y., Keutmann, H. and Schneyer, A. (2002) Activins, inhibins, and follistatins: from endocrinology to signaling. A paradigm for the new millennium. *Exp. Biol. Med. (Maywood)*, **227**, 724–752.
38. Victor, R.G., Haley, R.W., Willett, D.L., Peshock, R.M., Vaeth, P.C., Leonard, D., Basit, M., Cooper, R.S., Iannacchione, V.G., Visscher, W.A. *et al.* (2004) The Dallas Heart Study: a population-based probability sample for the multidisciplinary study of ethnic differences in cardiovascular health. *Am. J. Cardiol.*, **93**, 1473–1480.
39. Lakoski, S.G., Xu, F., Vega, G.L., Grundy, S.M., Chandalia, M., Lam, C., Lowe, R.S., Stepanavage, M.E., Musliner, T.A., Cohen, J.C. *et al.* (2010) Indices of cholesterol metabolism and relative responsiveness to ezetimibe and simvastatin. *J. Clin. Endocrinol. Metab.*, **95**, 800–809.
40. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. and Lee, J.J. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, **4**, 7.
41. Li, Y., Willer, C.J., Ding, J., Scheet, P. and Abecasis, G.R. (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.*, **34**, 816–834.

42. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. and Abecasis, G.R. (2012) Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.*, **44**, 955–959.
43. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R. and Willer, C.J. (2010) LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*, **26**, 2336–2337.
44. Izsvak, Z., Frohlich, J., Grabundzija, I., Shirley, J.R., Powell, H.M., Chapman, K.M., Ivics, Z. and Hamra, F.K. (2010) Generating knockout rats by transposon mutagenesis in spermatogonial stem cells. *Nat. Methods*, **7**, 443–445.
45. Ivics, Z., Izsvak, Z., Medrano, G., Chapman, K.M. and Hamra, F.K. (2011) Sleeping beauty transposon mutagenesis in rat spermatogonial stem cells. *Nat. Protoc.*, **6**, 1521–1535.
46. Ivics, Z., Li, M.A., Mates, L., Boeke, J.D., Nagy, A., Bradley, A. and Izsvak, Z. (2009) Transposon-mediated genome manipulation in vertebrates. *Nat. Methods*, **6**, 415–422.
47. Hamra, F.K., Gatlin, J., Chapman, K.M., Grellhesl, D.M., Garcia, J.V., Hammer, R.E. and Garbers, D.L. (2002) Production of transgenic rats by lentiviral transduction of male germ-line stem cells. *Proc. Natl. Acad. Sci. U. S. A.*, **99**, 14931–14936.
48. Wu, Z., Falciatori, I., Molyneux, L.A., Richardson, T.E., Chapman, K.M. and Hamra, F.K. (2009) Spermatogonial culture medium: an effective and efficient nutrient mixture for culturing rat spermatogonial stem cells. *Biol. Reprod.*, **81**, 77–86.
49. Yang, Y. and Wilson, M.J. (2015) Lhx9 gene expression during early limb development in mice requires the FGF signalling pathway. *Gene Expr. Patterns*, **19**, 45–51.
50. Mahony, C. and Vargesson, N. (2013) Molecular analysis of regulative events in the developing chick limb. *J. Anat.*, **223**, 1–13.
51. Vargesson, N. and Laufer, E. (2009) Negative Smad expression and regulation in the developing chick limb. *PLoS One*, **4**, e5173.