GENERAL ARTICLE

TBX6 compound inheritance leads to congenital vertebral malformations in humans and mice

Nan Yang^{1,2,3,†}, Nan Wu^{4,5,6,7,†}, Ling Zhang^{1,†}, Yanxue Zhao^{4,5}, Jiaqi Liu^{4,5}, Xiangyu Liang⁸, Xiaojun Ren¹, Weiyu Li¹, Weisheng Chen^{4,5}, Shuangshuang Dong¹, Sen Zhao^{4,5}, Jiachen Lin^{4,5}, Hang Xiang¹, Huadan Xue⁹, Lu Chen¹, Hao Sun⁹, Jianguo Zhang^{4,5,6}, Jiangang Shi¹⁰, Shuyang Zhang^{5,11}, Daru Lu¹, Xiaohui Wu¹², Li Jin¹, Jiandong Ding⁸, Guixing Qiu^{4,5,6}, Zhihong Wu^{4,5,13,†}, James R. Lupski^{7,14,15,‡} and Feng Zhang^{1,2,3,†,*}

¹Obstetrics and Gynecology Hospital, NHC Key Laboratory of Reproduction Regulation (Shanghai Institute of Planned Parenthood Research), School of Life Sciences, Fudan University, Shanghai 200011, China, ²Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing 211116, China, ³Shanghai Key Laboratory of Female Reproductive Endocrine Related Diseases, Shanghai 200011, China, ⁴Department of Orthopedic Surgery, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100730, China, ⁵Beijing Key Laboratory for Genetic Research of Skeletal Deformity, Beijing 100730, China, ⁶Medical Research Center of Orthopedics, Chinese Academy of Medical Sciences, Beijing 100730, China, ⁷Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA, ⁸State Key Laboratory of Molecular Engineering of Polymers, Department of Macromolecular Science, Fudan University, Shanghai 200438, China, ⁹Department of Radiology, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100730, China, ¹⁰Second Department of Spine Surgery, Changzheng Hospital, The Second Military Medical University, Shanghai 200003, China, ¹¹Department of Cardiology, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100730, China, ¹²Shanghai Kidney Development and Pediatric Kidney Disease Research Center, Institute of Developmental Biology and Molecular Medicine, Fudan University, Shanghai 200433, China, ¹³Department of Central Laboratory, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100730, China, ¹⁴Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030, USA and ¹⁵Texas Children's Hospital, Houston, TX 77030, USA

*To whom correspondence should be addressed at: Obstetrics and Gynecology Hospital, Fudan University, Shanghai 200011, China. Tel: +86 2131246783; Email: zhangfeng@fudan.edu.cn

[†]These authors contributed equally to this work.

[‡]These authors are equal senior authors.

Received: August 6, 2018. Revised: September 29, 2018. Accepted: October 5, 2018

© The Author(s) 2018. Published by Oxford University Press. All rights reserved.

For Permissions, please email: journals.permissions@oup.com

Abstract

Congenital vertebral malformations (CVMs) are associated with human *TBX6* compound inheritance that combines a rare null allele and a common hypomorphic allele at the *TBX6* locus. Our previous *in vitro* evidence suggested that this compound inheritance resulted in a *TBX6* gene dosage of less than haploinsufficiency (i.e. <50%) as a potential mechanism of *TBX6*-associated CVMs. To further investigate this pathogenetic model, we ascertained and collected 108 Chinese CVM cases and found that 10 (9.3%) of them carried *TBX6* null mutations in combination with common hypomorphic variants at the second *TBX6* allele. For *in vivo* functional verification and genetic analysis of *TBX6* compound inheritance, we generated both null and hypomorphic mutations in mouse *Tbx6* using the CRISPR-Cas9 method. These *Tbx6* mutants are not identical to the patient variants at the DNA sequence level, but instead functionally mimic disease-associated *TBX6* variants. Intriguingly, as anticipated by the compound inheritance model, a high penetrance of CVM phenotype was only observed in the mice with combined null and hypomorphic alleles of *Tbx6*. These findings are consistent with our experimental observations in humans and supported the dosage effect of *TBX6* in CVM etiology. In conclusion, our findings in the newly collected human CVM subjects and *Tbx6* mouse models consistently support the contention that *TBX6* compound inheritance causes CVMs, potentially via a gene dosage-dependent mechanism. Furthermore, mouse *Tbx6* mutants mimicking human CVM-associated variants will be useful models for further mechanistic investigations of CVM pathogenesis in the cases associated with *TBX6*.

Introduction

Congenital vertebral malformations (CVMs) have an estimated prevalence of ~1–2 in 1000 live births (1,2). Vertebral malformations include wedge vertebra, hemivertebra, block vertebra and butterfly vertebra (3). Human subjects with CVMs also commonly manifested anomalies including rib fusion, malalignment or missing ribs (4). CVMs can cause spinal curvature and scoliosis and result in three-dimensional thoracic deformity or even thoracic insufficiency syndrome, which seriously reduces the quality of life (5).

Genetic factors play an important role in disease pathogenesis. Robust correlation between genotype and phenotype is vital to understanding human biology, to the veracity of molecular diagnosis and for genetic counseling regarding recurrence risk (6). Mendelian disorders with complete penetrance can be well explained by monogenic inheritance with classic dominant (monoallelic variants) or recessive (biallelic variants) modes for trait segregation (7). However, it is always challenging to illuminate molecular underpinnings of incomplete penetrance of diseases based on the genetic variant information at a single locus. Therefore, digenic and polygenic inheritances were postulated to explain incomplete penetrance (8,9). Meanwhile, epigenetic and environmental modifications were also frequently involved in diseases by interaction with genetic factors (10,11). However, genetic modifiers at the same locus were underrecognized in phenotypic manifestation, especially the common regulatory variants in noncoding regions. These variants could be missed by routinely used technologies for genetic diagnosis such as whole-exome sequencing and chromosomal microarray analysis (6,12–14).

Previous clinical experience has shown that heterozygous null mutations of the *TBX6* gene in human chromosome 16p11.2 may not be sufficient to cause the manifestation of clinically obvious CVMs, whereas we intriguingly identified a high-frequency (~44% in Chinese and 33% in Caucasians; i.e. common variants), noncoding and mild-hypomorphic allele of *TBX6* as a key genetic modifier in human CVMs (Supplementary Material, Fig. S1) (15). The compound inheritance of the common *TBX6* mild-hypomorphic allele and any rare *TBX6* null variant (including *TBX6*/16p11.2 deletion or *TBX6* nonsense/frame-shift variant) led to CVMs with high penetrance across

human populations (15–17), potentially via a lowered gene dosage that was measurably less than haploinsufficiency. Here, we conducted a bedside-to-bench study to investigate the compound inheritance of *TBX6* on CVMs *in vivo*. Both *Tbx6* null and mild-hypomorphic mutants were generated in mouse models to functionally mimic the human disease-associated *TBX6* variants and further test the compound inheritance hypothesis.

Results

TBX6 compound inheritance leading to CVMs in a new Chinese cohort

As we previously reported, human CVM phenotypes can be caused by *TBX6* compound inheritance with a common hypomorphic allele combined with a null allele (15). For further verification of this pathogenetic model, the *TBX6* locus was carefully investigated in 108 newly collected Chinese CVMs cases. Intriguingly, we found that 10 (9.3%) of them carried *TBX6* heterozygous null variants, including 9 *TBX6*/16p11.2 deletions and 1 *TBX6* stop-gain mutation (c.933C > A, p.C311*) (Table 1). Notably, all of these 10 cases also carried the common mild-hypomorphic variants at the second *TBX6* allele (Table 1). These findings further supported the pathogenetic model of *TBX6* compound inheritance in human CVMs. The CVMs phenotypes of these 10 cases are shown in Figure 1.

Heterozygous null mutation of *Tbx6* not sufficient to cause CVMs in mice

To further investigate *Tbx6* compound inheritance in mice, we employed the CRISPR-Cas9 technology to induce *Tbx6* frame-shift mutations in mice (FVB/NJ strain) for a *Tbx6* null mutant (18). A *Tbx6* mutation with 1-bp insertion in exon 2 (NM_011538.2) before the key T-box domain was obtained, resulting in a frame-shift allele damaging the key T-box domain (Supplementary Material, Fig. S2). Notably, no homozygote of this *Tbx6* frame-shift mutation was detected in 98 offsprings [29 *Tbx6*^{+/+} (wild type) and 69 *Tbx6*^{+/-} mice] from heterozygote crosses. This observation of distortion from the expected

Table 1. Vertebral malformations of 10 *TBX6*-associated subjects in our new Chinese CVM cohort

Human subject	Sex	Age (year) ^a	First null allele	<i>TBX6</i> gene Second mild-hypomorphic allele ^b	Vertebral malformations
XR139	M	2	16p11.2/ <i>TBX6</i> deletion	Yes	Wedge vertebra and hemivertebra
XR330	M	4	16p11.2/ <i>TBX6</i> deletion	Yes	Hemivertebra
XR468	M	3	16p11.2/ <i>TBX6</i> deletion	Yes	Butterfly vertebrae and hemivertebra
XR480	F	11	16p11.2/ <i>TBX6</i> deletion	Yes	Hemivertebra and segmental defect
XR522	M	8	16p11.2/ <i>TBX6</i> deletion	Yes	Hemivertebra
XR529	M	14	16p11.2/ <i>TBX6</i> deletion	Yes	Hemivertebra
XR605	F	5	16p11.2/ <i>TBX6</i> deletion	Yes	Hemivertebra and an additional lumbar vertebra
XR623	M	8	16p11.2/ <i>TBX6</i> deletion	Yes	Hemivertebra
XR625	M	7	c.933C > A (p.C311*)	Yes	Hemivertebra and an additional lumbar vertebra
XR636	F	9	16p11.2/ <i>TBX6</i> deletion	Yes	Butterfly vertebra

^aAge means latest time of vertebral evaluation.^bThe mild-hypomorphic haplotype of human *TBX6* was defined as three SNPs including rs2289292, rs3809624 and rs3809627.

F, female; M, male.

Figure 1. CVM phenotypes of the human subjects associated with *TBX6* compound inheritance. The subject numbers are shown above each radiograph. The arrows indicate the vertebral malformations.

1:2:1 Mendelian ratio is consistent with the contention that homozygous null mutations of mouse *Tbx6* are embryonically lethal (19). Remarkably, the mice carrying *Tbx6* heterozygous null mutation did not manifest any evidence of obvious CVM phenotypes, which was consistent with the observations in human subjects.

Hypomorphic mutation leading to mild reduction of *Tbx6* expression in vivo

The *in vitro* assays investigating the common hypomorphic allele of human *TBX6* suggested that both single-nucleotide polymor-

phisms (SNPs) rs3809624 and rs3809627 in the 5' noncoding regulatory region of human *TBX6* were potentially responsible for a 'mild-hypomorphic' effect that reduced the *TBX6* expression level to ~70% of the wild-type allele (Supplementary Material, Fig. S1) (15). Therefore, a mouse *Tbx6* regulatory mutant with the correspondingly downregulated expression level was required to functionally mimic the human *TBX6* mild-hypomorphic allele. The 5' noncoding region of *TBX6* is only partially conserved between human and mouse. Therefore, perfectly 'knocking-in' the human *TBX6* mild-hypomorphic allele in mice is not readily applicable. Instead, four mutations (designated R1 to R4) at the potential regulatory region of *Tbx6* were obtained for further functional screening for *Tbx6* mild-hypomorphic alleles using

the luciferase reporter assays (Supplementary Material, Fig. S3B). We found that R1 to R3 mutants significantly reduced *Tbx6* gene expression (Supplementary Material, Fig. S3C). Intriguingly, the R1 mutation was shown to have a similar mild-hypomorphic effect with that of the common *TBX6* mild-hypomorphic allele in humans (Supplementary Material, Fig. S1) (15). Therefore, the mouse R1 mutation was empirically considered as a *Tbx6* mild-hypomorphic (*Tbx6^{mh}*) mutation to functionally mimic the known human *TBX6* mild-hypomorphic allele and used to recapitulate the compound inheritance model (Fig. 2A–C). Furthermore, the mild-hypomorphic effect of *Tbx6^{mh}* was also confirmed by *in vivo* assay of whole-mount RNA in situ hybridization in E9.5 mouse embryos (Fig. 2D). Notably, even though the DNA sequence of *Tbx6^{mh}* was not identical to that of the common hypomorphic allele in the patients, the *Tbx6^{mh/mh}*

mice showed neither CVMs nor other developmental defects, which was consistent with the observations that human *TBX6* mild-hypomorphic homozygotes, are common in healthy populations (Supplementary Material, Fig. S1) (15).

Compound inheritance of *Tbx6* leading to mouse CVM phenotypes

The FVB/NJ strain of mice with five genotypes (*Tbx6^{+/+}*, *Tbx6^{+/mh}*, *Tbx6^{mh/mh}*, *Tbx6^{+/-}* and *Tbx6^{mh/-}*) were constructed. In each group, embryos were harvested at E14.5 and stained by Alcian blue as described previously (20). No obvious CVM phenotypes were observed in the embryos of *Tbx6^{+/+}*, *Tbx6^{+/mh}*, *Tbx6^{mh/mh}* or *Tbx6^{+/-}*. Remarkably, CVMs phenotypes were observed in

Figure 2. Generation of *Tbx6* mild-hypomorphic (*mh*) mutation in mice. (A) Eight base pairs (chr7:126781240-126781242 and chr7:126781244-126781248, GRCm38/mm10) in the upstream regulatory region of mouse *Tbx6* were deleted in the mild-hypomorphic mutant allele (*Tbx6^{mh}*). The mutant sequence of *Tbx6^{mh}* was confirmed by Sanger sequencing. (B) The conserved upstream regulatory region of *Tbx6* was shown. The sequence of single-guide RNA target for genome editing was highlighted. (C) The mild-hypomorphic effect of *Tbx6^{mh}* on gene expression was investigated using the luciferase reporter assay in mouse-derived P19CL6 cells (27,28). The graph shows mean values of normalized luciferase activity and their standard errors. +, wild type; ****, $P < 0.0001$. (D) The mild-hypomorphic effect of *Tbx6^{mh}* was also confirmed by *in vivo* assay of whole-mount RNA in situ hybridization in E9.5 mouse embryos. Caudal is to the left and rostral to the right. The scale bar represents 250 μ m.

Figure 3. In vivo assays using mouse embryos support the compound inheritance model, consisting of a *Tbx6* null allele (–) and a mild-hypomorphic allele (*mh*), for the dosage-dependent manifestation of CVMs. (A) Dorsal views of E14.5 mouse embryos stained by Alcian blue. One embryo without obvious vertebral malformations is shown. (B) One embryo with vertebral malformations is shown. Red dots and arrows indicate vertebral malformations and rib fusions, respectively. The scale bar represents 1 mm. (C) No obvious vertebral malformations were observed in the E14.5 mouse embryos of *Tbx6*^{+/+}, *Tbx6*^{+/mh}, *Tbx6*^{mh/mh} or *Tbx6*^{+/-}, whereas the *Tbx6*^{mh/-} embryos manifested vertebral malformations with a high penetrance. Sample sizes are shown below columns. *****, $P < 0.0001$.

37 (93%) of 40 embryos of *Tbx6*^{mh/-} ($P < 1.4 \times 10^{-22}$ when compared with the wild type; Fisher's exact test). This observation in mice is strong genetic evidence in support of the *TBX6* compound inheritance model in human CVMs (Fig. 3A–C and Supplementary Material, Fig. S4).

Since phenotyping mild CVMs in stained embryos is technically challenging, the CVM penetrance might be underestimated in our embryo study. Therefore, we also investigated adult mice (aged from 35 to 45 days) using micro-computed tomography (CT) scanning. Consistently, no CVM was observed in the adult mice of *Tbx6*^{+/+}, *Tbx6*^{+/mh}, *Tbx6*^{mh/mh} or *Tbx6*^{+/-} (10 adult mice examined for each genotype). Intriguingly, all the 10 adult *Tbx6*^{mh/-} mice showed CVMs in their micro-CT images (Fig. 4A–C, Supplementary Material, Fig. S5 and Movie S1) ($P < 1.1 \times 10^{-5}$ when compared with the wild type; Fisher's exact test). Our in vivo observations in the *Tbx6* mouse models strongly support the conclusion that *Tbx6* heterozygous null mutation is not sufficient to cause overt CVMs. Instead, the biallelic combination of *Tbx6* null and mild-hypomorphic mutants can result in CVMs potentially via a further adverse effect on *Tbx6* gene expression and a decrement below that for haploinsufficiency (i.e. <50%; note that the *Tbx6*^{+/-} mice have no CVM), but not completely null (*Tbx6*^{-/-} is embryonic lethal).

Discussion

We previously reported that the combination of a null allele and a mild-hypomorphic allele of human *TBX6* was associated with CVMs (15). Among the newly collected 108 Chinese subjects with CVMs, 10 (9.3%) cases are *TBX6*-associated and all of them fit the compound inheritance model. This incident was similar with that (~10%) across human populations (15,16).

In this study, we also functionally validated the *Tbx6* compound inheritance model in mice. By employing the CRISPR-Cas9 technology to edit the coding region and 5' noncoding regulatory

region of mouse *Tbx6*, we obtained a null mutant (*Tbx6*⁻) and a mild-hypomorphic mutant (*Tbx6*^{mh}), respectively. After mating and vertebral phenotyping, we found that only the offsprings with the *Tbx6*^{mh/-} genotype manifested CVMs, which strongly supports the compound inheritance model of *TBX6* in human CVMs (15). Also, these mouse models that functionally mimic disease-associated *TBX6* alleles can be used for investigating pathogenetic mechanisms of CVMs in the cases associated with *TBX6* compound inheritance.

The clinical genomics community currently relies on the 'rare variant' hypothesis and has identified an abundance of rare but patient-enriched variants (21,22). This has proven an efficient approach to the identification of disease-causing genes, molecular diagnosis in clinical genomics, functional annotation of human genes and a molecular entry point into elucidating disease biology. But in some circumstances, as documented for the *TBX6*-associated CVM phenotype in this study, rare variants alone cannot achieve a robust genotype–phenotype correlation that is required for prenatal genetic diagnosis. Some previous studies proposed the biallelic compound inheritance of a rare null allele and a low-frequency regulatory allele for syndromic conditions. For example, a rare null mutation of *RBM8A* combined with one low-frequency regulatory SNP on the other allele caused human thrombocytopenia with absent radii (MIM 274000) (12). Furthermore, a rare loss-of-function mutation of *TXNL4A* and a low-frequency 34 bp deletion in the core promoter region on the other allele resulted in Burn-McKeown syndrome (MIM 608572) (23).

Here, we recapitulate the compound inheritance model in mice. The compound inheritance model of *TBX6* in humans showed that common noncoding variant alleles could be important genetic modifiers and contributors to variability of expression and disease manifestations (i.e. penetrance). Such common variant alleles may be routinely filtered or removed from consideration of causality by analytical algorithms when conducting clinical genome and exome sequencing for rare diseases and

Figure 4. Micro-CT results of the adult mice. (A) Dorsal and ventral views of adult mice were revealed by micro-CT scanning. A mouse without obvious vertebral malformations is shown. (B) A mouse with vertebral malformations is shown. White arrowhead and arrows indicate vertebral malformations and rib fusions, respectively. (C) Micro-CT results of the adult mice with genotypes of *Tbx6*^{+/+}, *Tbx6*^{+/mh}, *Tbx6*^{mh/mh}, *Tbx6*^{+/-} and *Tbx6*^{mh/-}. Sample size is shown below each column. ****, *P* < 0.0001.

Mendelian disorders (6). In some circumstances, as the *TBX6* compound inheritance model showed here, the coexistence of copy number variations (CNVs) and SNPs at the same locus could result in distorted calculations of the significance in associating SNPs to disease (24). Therefore, the contributions of common regulatory variants to congenital developmental disorders are underrecognized.

Regulatory variants, as genetic modifiers, could downregulate gene dosage mildly (i.e. mild-hypomorphic) even in the homozygous state no clinical phenotype is manifest (Supplementary Material, Fig. S1). But when a common mild-hypomorphic variant is combined with a rare damaging variant at the same locus, the compound inheritance could manifest obvious CVM phenotype via further reducing gene expression level in humans and we now recapitulate that in mice (Supplementary Material, Fig. S1). Fortunately, the Genotype-Tissue Expression (GTEx) project has identified thousands of regulatory expression quantitative trait loci (eQTL) variants across human tissues (25). Perhaps for other loci more robust genotype-phenotype correlation, and the clinical implementation of molecular diagnosis and genetic counseling, will be further enabled by accounting for eQTL variants when interpreting or predicting the potential clinical consequences of any rare pathogenic variants.

Materials and Methods

The CVMs cohort

The newly collected Han Chinese cases with CVMs were enrolled at Peking Union Medical College Hospital (PUMCH). The clinical diagnoses of CVMs were confirmed by radiological imaging (15). This work was approved by the institutional review boards at Fudan University and PUMCH. Written informed consents were obtained from the participants (those who were ≥18 years of age at the time of enrollment) or their guardians (for participants who were <18 years of age).

Quantitative Polymerase Chain Reaction (PCR)

Quantitative PCR analysis was conducted to screen for the *TBX6*/16p11.2 deletion in human subjects. The primers and protocol are shown in Supplementary Material, Tables S1 and S2. The experiments were conducted using the SYBR Green Realtime PCR Master Mix (TOYOBO, Osaka, Japan) and the ABI Prism 7900HT Sequence Detection System. Three replicates were conducted for each assay. One sample with *TBX6*/16p11.2 deletion and one non-deletion sample were selected as the

positive and negative controls, respectively. Both of them were previously confirmed by gender-matched array-based comparative genomic hybridization (CGH). Other experimental details were described in our previous work (15). The SDS 2.4.1 software (ABI, Foster City, CA) was used for data analysis.

Array-based CGH

CNV analysis was conducted using Agilent oligonucleotide-based CGH microarrays. Some experimental details were previously described (26). The genomic DNA extracted from peripheral blood leukocytes of each subject and the gender-matched reference DNA (Promega, Mannheim, Germany) were respectively fragmented using *AluI* and *RsaI* enzyme digestion. DNA labeling was conducted using the Agilent SureTag DNA Labeling Kit. Different fluorescence dyes were used for DNA labeling of each subject (Cy5-dUTP) and the reference DNA (Cy3-dUTP). Each labeled subject DNA was hybridized together with the labeled reference DNA onto Agilent human CGH microarrays. DNA processing, microarray handling and data analyses were conducted by following the Agilent oligonucleotide CGH protocol (version 6.0).

TBX6 gene sequencing and haplotyping in CVMs subjects

The entire *TBX6* gene and its ~1 kb upstream region were amplified using long-range PCR for Sanger sequencing. The sequences were mapped to human reference genome (GRCh37/hg19) to identify the sequence mutations. Haplotyping of common *TBX6* variants was conducted using clone sequencing. The primers for amplification and sequencing are shown in [Supplementary Material, Tables S3 and S4](#).

Mouse strain and animal husbandry

Animal studies were approved by the institutional review board at Fudan University and was carried out in accordance with the recommendation in the Guide for the Care and Use of Laboratory Animals of the US National Institutes of Health. The mice were bred in specific pathogen-free barrier facility. FVB/NJ was the strain background used in all mouse experiments.

Genome editing in mice using the CRISPR-Cas9 method

We employed the CRISPR-Cas9 technology to edit the mouse genomes as reported (18). The targets of guide RNAs in the mouse genome are shown in [Supplementary Material, Table S5](#). Guide RNA and Cas9 mRNA were pooled (10 ng/μl for each RNA) and injected into the zygotes from FVB/NJ strain mice (200 cells for each editing). The founder (F0) mice were crossed with wild-type FVB/NJ mice to obtain offspring. Mice were genotyped using genomic DNAs from toe clips by PCR and Sanger sequencing. The PCR primers used for amplification and sequencing are shown in [Supplementary Material, Table S6](#).

Cell culture and luciferase reporter assay

The P19CL6 cells were cultured and induced to differentiate into cardiomyocytes as described previously (27). The cells were harvested every 24 h after treatment with dimethyl sulfox-

ide (28). To construct reporter plasmids, we amplified the DNA fragments that included the potential regulatory elements of mouse *Tbx6* ([Supplementary Material, Fig. S3B](#)). The primers for amplification are shown in [Supplementary Material, Table S7](#). The veracity of all constructs was confirmed by Sanger sequencing. The fragments obtained from *Tbx6* wild type and potential regulatory alleles were respectively constructed into the PGL3-Basic vector (Promega) enabling fusion to the reporter gene. The luciferase reporter assay was conducted as described in our previous study (15).

Whole-mount RNA in situ hybridization

Female and male mice were mated in the evening. Vagina plug was checked in the next morning and embryos in female mice with plug were recorded as E0.5. Whole-mount RNA in situ hybridization was performed according to Saga et al. (29). Embryos were harvested at E9.5 and genotyped by PCR and Sanger sequencing using genomic DNAs extracted from yolk sac. Then the embryos were washed three times in phosphate buffer saline containing 0.1% Tween 20 and fixed them in 4% paraformaldehyde overnight at 4°C. Probes for *Tbx6*, covering complementary DNA 414-1210, was obtained by reverse transcription PCR. Hybridization was carried out using standard procedures with probe concentrations of 1000 ng/ml. Bound digoxigenin-11-UTP labeled riboprobes were detected with alkaline phosphatase conjugated anti-digoxigenin antibodies (Roche, Mannheim, Germany). All the images were recorded by Axio Zoom.V16 (Zeiss, Oberkochen, Germany).

Embryonic somite staining by Alcian blue

Embryos were harvested at E14.5 and genotyped using genomic DNAs from yolk sac by PCR and Sanger sequencing. Embryos were stained with Alcian blue 8GX as described previously (20). All embryos were stained for 5 days at room temperature. After staining, embryos were kept in 2:1 benzyl benzoate/benzyl alcohol. All the images were recorded by Axio Zoom.V16 (Zeiss).

Micro-CT imaging of adult mouse spinal column and ribs

Each adult mouse was anesthetized by intraperitoneal injection of 2% pentobarbitalum natricum. Tomography imaging data sets were obtained using a high-resolution micro-CT scanner (SkyScan 1176, Bruker). The scanning parameters were set as follows: spatial resolution of 35 μm pixel size, X-ray tube voltage of 50 kVp, X-ray tube current of 497 mA, 0.5 mm Al filter, 2° rotation step and rotated 360° around the vertical axis. These imaging data sets were reconstructed using the GPURecon Server.

Statistical analysis

The Fisher's exact test was used to investigate different prevalence of CVM phenotypes between the wild-type and *Tbx6* mutant mice. The unpaired t-test was used for statistical analysis of luciferase reporter assays. A two-sided $P < 0.05$ was considered to indicate statistical significance.

Supplementary Material

[Supplementary Material](#) is available at HMG online.

Electronic Database Information

GTEx, <http://www.gtexportal.org/>; OMIM, <http://www.omim.org/>; UCSC Genome Browser, <http://genome.ucsc.edu/>.

Acknowledgements

We would like to thank Ting Ni, Hongbo Nie, Yi Wang, Yanhua Wu and Xueyan Yang for technical advice and support; Shixue Liu for experimental assistance; and Jianxiong Shen, Yipeng Wang, Yu Zhao, Hong Zhao, Ye Tian, Shugang Li, Qiyi Li and Jianhua Hu for their generous help in patient collection.

Conflict of Interest statement. J.R.L. has stock ownership in 23 and Me, is a paid consultant for Regeneron Pharmaceuticals and is a co-inventor on multiple United States and European patents related to molecular diagnostics for inherited neuropathies, eye diseases and bacterial genomic fingerprinting. The Department of Molecular and Human Genetics at Baylor College of Medicine derives revenue from the chromosomal microarray analysis and clinical exome sequencing offered in the Baylor Genetics Laboratory (<http://bmgl.com>).

Funding

National Key Research and Development Program of China (2016YFC0905100, 2016YFC1100300 and 2016YFC0901501); National Natural Science Foundation of China (31625015, 31521003, 31571297, 31771396, 51533002, 81822030, 81472046 and 81772299); Shanghai Municipal Science and Technology Major Project (2017SHZDZX01); Beijing Nova Program (Z161100004916123); Beijing Natural Science Foundation (7172175); CAMS Initiative Fund for Medical Sciences (2016-I2M-3-003); and the US National Human Genome Research Institute/National Heart Lung and Blood Institute supported Baylor Hopkins Center for Mendelian Genomics (grant UM1 HG006542).

References

- Giampietro, P.F., Dunwoodie, S.L., Kusumi, K., Pourquie, O., Tassy, O., Offiah, A.C., Cornier, A.S., Alman, B.A., Blank, R.D., Raggio, C.L. et al. (2009) Progress in the understanding of the genetic etiology of vertebral segmentation disorders in humans. *Ann. N.Y. Acad. Sci.*, **1151**, 38–67.
- Alexander, P.G. and Tuan, R.S. (2010) Role of environmental factors in axial skeletal dysmorphogenesis. *Birth Defects Res. C Embryo Today*, **90**, 118–132.
- Hedequist, D. and Emans, J. (2007) Congenital scoliosis: a review and update. *J. Pediatr. Orthop.*, **27**, 106–116.
- Offiah, A., Alman, B., Cornier, A.S., Giampietro, P.F., Tassy, O., Wade, A. and Turnpenny, P.D. (2010) Pilot assessment of a radiologic classification system for segmentation defects of the vertebrae. *Am. J. Med. Genet. A*, **152A**, 1357–1371.
- Flynn, J.M., Emans, J.B., Smith, J.T., Betz, R.R., Deeney, V.F., Patel, N.M. and Campbell, R.M. (2013) VEPTR to treat nonsyndromic congenital scoliosis: a multicenter, mid-term follow-up study. *J. Pediatr. Orthop.*, **33**, 679–684.
- Biesecker, L.G. and Green, R.C. (2014) Diagnostic clinical genome and exome sequencing. *N. Engl. J. Med.*, **370**, 2418–2425.
- McKusick, V.A. (2007) Mendelian inheritance in man and its online version, OMIM. *Am. J. Hum. Genet.*, **80**, 588–604.
- Posey, J.E., Harel, T., Liu, P., Rosenfeld, J.A., James, R.A., Coban Akdemir, Z.H., Walkiewicz, M., Bi, W., Xiao, R., Ding, Y. et al. (2017) Resolution of disease phenotypes resulting from multilocus genomic variation. *N. Engl. J. Med.*, **376**, 21–31.
- Weiner, D.J., Wigdor, E.M., Ripke, S., Walters, R.K., Kosmicki, J.A., Grove, J., Samocha, K.E., Goldstein, J.I., Okbay, A., Bybjerg-Grauholm, J. et al. (2017) Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nat. Genet.*, **49**, 978–985.
- Shah, S., Bonder, M.J., Marioni, R.E., Zhu, Z., McRae, A.F., Zernakova, A., Harris, S.E., Liewald, D., Henders, A.K., Mendelson, M.M. et al. (2015) Improving phenotypic prediction by combining genetic and epigenetic associations. *Am. J. Hum. Genet.*, **97**, 75–85.
- Sparrow, D.B., Chapman, G., Smith, A.J., Mattar, M.Z., Major, J.A., O'Reilly, V.C., Saga, Y., Zackai, E.H., Dormans, J.P., Alman, B.A. et al. (2012) A mechanism for gene-environment interaction in the etiology of congenital scoliosis. *Cell*, **149**, 295–306.
- Albers, C.A., Paul, D.S., Schulze, H., Freson, K., Stephens, J.C., Smethurst, P.A., Jolley, J.D., Cvejic, A., Kostadima, M., Bertone, P. et al. (2012) Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome. *Nat. Genet.*, **44**, 435–439.
- Zhang, F. and Lupski, J.R. (2015) Non-coding genetic variants in human disease. *Hum. Mol. Genet.*, **24**, R102–R110.
- Wapner, R.J., Martin, C.L., Levy, B., Ballif, B.C., Eng, C.M., Zachary, J.M., Savage, M., Platt, L.D., Saltzman, D., Grobman, W.A. et al. (2012) Chromosomal microarray versus karyotyping for prenatal diagnosis. *N. Engl. J. Med.*, **367**, 2175–2184.
- Wu, N., Ming, X., Xiao, J., Wu, Z., Chen, X., Shinawi, M., Shen, Y., Yu, G., Liu, J., Xie, H. et al. (2015) TBX6 null variants and a common hypomorphic allele in congenital scoliosis. *N. Engl. J. Med.*, **372**, 341–350.
- Takeda, K., Kou, I., Kawakami, N., Iida, A., Nakajima, M., Ogura, Y., Imagawa, E., Miyake, N., Matsumoto, N., Yasuhiko, Y. et al. (2017) Compound heterozygosity for null mutations and a common hypomorphic risk haplotype in TBX6 causes congenital scoliosis. *Hum. Mutat.*, **38**, 317–323.
- Lefebvre, M., Duffourd, Y., Jouan, T., Poe, C., Jean-Marcais, N., Verloes, A., St-Onge, J., Riviere, J.B., Petit, F., Pierquin, G. et al. (2017) Autosomal recessive variations of TBX6, from congenital scoliosis to spondylocostal dysostosis. *Clin. Genet.*, **91**, 908–912.
- Platt, R.J., Chen, S., Zhou, Y., Yim, M.J., Swiech, L., Kempton, H.R., Dahlman, J.E., Parnas, O., Eisenhaure, T.M., Jovanovic, M. et al. (2014) CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell*, **159**, 440–455.
- Chapman, D.L. and Papaioannou, V.E. (1998) Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature*, **391**, 695–697.
- Jegalian, B.G. and De Robertis, E.M. (1992) Homeotic transformations in the mouse induced by overexpression of a human Hox3.3 transgene. *Cell*, **71**, 901–910.
- Schork, N.J., Murray, S.S., Frazer, K.A. and Topol, E.J. (2009) Common vs. rare allele hypotheses for complex diseases. *Curr. Opin. Genet. Dev.*, **19**, 212–219.
- Lupski, J.R., Belmont, J.W., Boerwinkle, E. and Gibbs, R.A. (2011) Clan genomics and the complex architecture of human disease. *Cell*, **147**, 32–43.

23. Wieczorek, D., Newman, W.G., Wieland, T., Berulava, T., Kaffe, M., Falkenstein, D., Beetz, C., Graf, E., Schwarzmayr, T., Douzgou, S. et al. (2014) Compound heterozygosity of low-frequency promoter deletions and rare loss-of-function mutations in TXNL4A causes Burn-McKeown syndrome. *Am. J. Hum. Genet.*, **95**, 698–707.
24. Liu, J., Zhou, Y., Liu, S., Song, X., Yang, X.Z., Fan, Y., Chen, W., Akdemir, Z.C., Yan, Z., Zuo, Y. et al. (2018) The coexistence of copy number variations (CNVs) and single nucleotide polymorphisms (SNPs) at a locus can result in distorted calculations of the significance in associating SNPs to disease. *Hum. Genet.*, **137**, 553–567.
25. The GTEx Consortium (2013) The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.*, **45**, 580–585.
26. Boone, P.M., Bacino, C.A., Shaw, C.A., Eng, P.A., Hixson, P.M., Pursley, A.N., Kang, S.H., Yang, Y., Wiszniewska, J., Nowakowska, B.A. et al. (2010) Detection of clinically relevant exonic copy-number changes by array CGH. *Hum. Mutat.*, **31**, 1326–1342.
27. Li, L., Gong, H., Yu, H., Liu, X., Liu, Q., Yan, G., Zhang, Y., Lu, H., Zou, Y. and Yang, P. (2012) Knockdown of nucleosome assembly protein 1-like 1 promotes dimethyl sulfoxide-induced differentiation of P19CL6 cells into cardiomyocytes. *J. Cell. Biochem.*, **113**, 3788–3796.
28. Gavrilov, S., Nuhrenberg, T.G., Ashton, A.W., Peng, C.F., Moore, J.C., Konstantinidis, K., Mummery, C.L. and Kitsis, R.N. (2012) Tbx6 is a determinant of cardiac and neural cell fate decisions in multipotent P19CL6 cells. *Differentiation*, **84**, 176–184.
29. Saga, Y., Hata, Y., Kobayashi, N., Magnuson, T., Seldin, M.F. and Taketo, M.M. (1996) MesP1: a novel basic helix-loop-helix protein expressed in the nascent mesodermal cells during mouse gastrulation. *Development*, **122**, 2769–2778.