HHS Public Access

Author manuscript

Genet Med. Author manuscript; available in PMC 2020 July 01.

Published in final edited form as:

Genet Med. 2019 July ; 21(7): 1548–1558. doi:10.1038/s41436-018-0377-x.

TBX6-associated congenital scoliosis (TACS) as a clinically distinguishable subtype of congenital scoliosis: further evidence supporting the compound inheritance and *TBX6* gene dosage model

A full list of authors and affiliations appears at the end of the article.

Abstract

Purpose—To characterize clinically measurable endophenotypes, implicating the *TBX6* compound inheritance model.

Methods—Patients with congenital scoliosis (CS) from China (N=345, Cohort 1), Japan (N=142, Cohort 2), and the USA (N=10, Cohort 3) were studied. Clinically measurable endophenotypes were compared according to the *TBX6* genotypes. A mouse model for *Tbx6* compound inheritance [N=52] was investigated by micro-CT. A clinical diagnostic algorithm (TACScore) was developed to assist in clinical recognition of *TBX6*-associated CS (TACS).

Results—In Cohort 1, TACS patients [N=33] were significantly younger at onset than the remaining CS patients (P=0.02), presented with one or more hemivertebrae/butterfly vertebrae (P=4.9×10^{-8}), and exhibited vertebral malformations involving the lower part of the spine (T8–S5, P=4.4×10^{-3}); observations confirmed in two replication cohorts. Simple rib anomalies were prevalent in TACS patients (P=3.1×10^{-7}), while intraspinal anomalies were uncommon (P=7.0×10^{-7}). A clinically usable TACScore was developed with an AUC of 0.9 (P=1.6×10^{-15}). A *Tbx6*$^{-/mh}$ (mild-hypomorphic) mouse model supported that gene-dosage effect of *TBX6* underlies the TACS phenotype.

Conclusion—TACS is a clinically distinguishable entity with consistent clinically measurable endophenotypes. The type and distribution of vertebral column abnormalities in *TBX6/Tbx6* compound inheritance implicate subtle perturbations in gene dosage as a cause of spine developmental birth defects responsible for about 10% of CS.

We have deposited all the data involving this study on http://www.discostudy.org/data/DISCO2018-1.html, and the access to the data is available upon request (http://discostudy.org/contactus.html).

*To whom correspondence should be addressed: Nan Wu, MD, (dr.wunan@pumch.cn), and Guixing Qiu, MD, (qiuguixingpumch@126.com); both from the Department of Orthopedic Surgery, Beijing Key Laboratory for Genetic Research of Skeletal Deformity, Medical Research Center of Orthopedics, all at Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, No. 1 Shuaifuyuan, Beijing 100730, China. Dr. Nan Wu is also affiliated with the Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA.

†Dr. J. Liu, Dr. N. Wu, Dr. N. Yang, Dr. K. Takeda, and Dr. W. Chen contributed equally to this article.

‡These authors share senior authorship.

Conflict of interest: J.R.L has stock ownership in 23andMe, 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).

Supplementary Material

Supplementary Material is available at GIM online.

Keywords

Congenital Scoliosis (CS); 16p11.2/*TBX6*; compound inheritance model; genotype-phenotype correlation; gene dosage

INTRODUCTION

Recent progress in elucidating the genetic contributions to disease and molecular etiology of clinical phenotypes presents opportunities to further subclassify human disease traits such as scoliosis according to their underlying genetic etiologies. Such genetic/genomic subclassification can result in a virtuous cycle from bench to bedside to better understand the biological perturbations underlying disease traits and pathophysiological bases of disease.[1,2] To further increase the specificity of molecular diagnosis and the clinical application of genetic testing and clinical genomics, precise correlations between genotype and phenotype, allelic series, and the clinical consequences of combinations of biallelic variants at a locus need to be established.[3,4] The recent report of a compound inheritance model in congenital scoliosis (CS) provides a genetic biomarker with which to potentially clinically subclassify scoliosis and define a specific clinical disease entity (TACS, *TBX6*-Associated CS) which is associated with a particular combination of variant alleles – biallelic variants consisting of a Loss-of-Function (LoF) *TBX6* lesion and a common risk hypomorphic allele *in trans*– as an underlying genetic etiology and further explore the hypothesized *TBX6* gene dosage effect.[5]

Congenital scoliosis is a form of spinal curvature that can be caused by vertebral malformations potentially resulting from defects of formation, defects of segmentation, or a combination of the two.[6] The prevalence of CS is approximately 0.5–1 per 1,000 live births.[7] As a major contributor to childhood and adolescent disability, CS affects patients' lives and activities of daily living both physically and psychologically.[8] CS can arise from developmental spine defects that result from perturbations in somitogenesis.[7] Successful somitogenesis requires the spatial and temporal regulation of a complicated gene-interaction network in which *TBX6* plays an essential role. Our previous work elucidated the *TBX6* compound inheritance model accounts for approximately 7.9%−10.6% of sporadic CS in the Chinese population.[5] Similar analyses of independent Japanese and European CS cohorts, in multiple patients from different world populations and genetic backgrounds, supported the proposed *TBX6* compound inheritance model.[9,10]

We previously showed the CS phenotype could result from compound inheritance[5]: consisting of null variant alleles at the *TBX6* locus, either a deletion copy number variant (CNV) of the 16p11.2 region[11,12] or LoF induced by single nucleotide variants (SNVs) and indels in *TBX6*, which maps within the 16p11.2 region.[13] With the delineation of the hypothesis of the compound inheritance and *TBX6* gene dosage model,[5] we were motivated to more precisely characterize the genotype-phenotype relationships of the specific genotypic combination of alleles at the *TBX6* locus observed in CS cases with *TBX6* variant alleles [i.e. the compound inheritance model] and the clinically observed endophenotypes characterizing the CS. In addition, we sought to explore the compound inheritance and gene dosage hypothesis in both humans and mice. Moreover, further investigations are required to

test the potential utility of molecular diagnosis and clinical genomics in the ‘precision medicine directed’ clinical evaluation and management of CS.

MATERIALS AND METHODS

Participants

We recruited two independent cohorts and a multicenter case series from China, Japan, and the United States. The discovery set (Cohort 1) consists of unrelated sporadic CS patients of Chinese Han descent in the DISCO (Deciphering disorders Involving Scoliosis and COmorbidities, http://discostudy.org/) study from Peking Union Medical College Hospital (PUMCH) between October 2010 and January 2016, in which clinical diagnoses of CS were confirmed by radiological imaging.[5] The first replication cohort (Cohort 2) comprised Japanese patients with CS recruited from Japan, in compliance with the selection criteria as reported previously.[10] The second replication case series (Cohort 3) comprised patients with 16p11.2 deletion/*TBX6* LoF variants collected from the United States. Genomic DNA was extracted from the peripheral blood or saliva samples. Informed consent was obtained from each participant or the corresponding guardian. The study was approved by the institutional review boards of PUMCH, RIKEN, Baylor College of Medicine, and the other participating hospitals.

Genetic Analyses

A genome-wide copy number variant analysis was performed in 20 sporadic CS patients in Cohort 1.[5] Quantitative polymerase chain reaction (qPCR) analysis was conducted to screen for 16p11.2 deletions in the remaining 325 patients from Cohort 1 and in the 142 patients from Cohort 2. An orthogonal experimental approach to measure copy number alteration, customized comparative genomic hybridization (CGH) microarrays or digital droplet PCR (ddPCR), was used to independently confirm the deletion CNV candidates. The 16p11.2 deletion case series (Cohort 3) was identified by chromosomal microarray analysis (CMA) (Table S1).[12,14,15]

The entire *TBX6* gene and its upstream region were amplified and analyzed by Sanger sequencing in all patients from Cohort 1[5] and 121 patients from Cohort 2[10]. *TBX6* variants were also detected by exome sequencing (ES) and verified by Sanger sequencing in the remaining 21 patients from Cohort 2 and one patient (BH8084) from Cohort 3 (Table S1). Based on our previous findings,[5] we analyzed the hypomorphic allele in patients with 16p11.2 deletions and *TBX6* LoF variants.

Phenotype Evaluation

We conducted detailed phenotypic analyses of the spine, ribs and intraspinal anomalies in Cohort 1 and Cohort 2. The age of onset indicated the age at which an individual first came to medical attention, exhibiting or presenting with features such as asymmetric shoulder height or shoulder blade prominence. Morphologically, the vertebral malformations were classified as defect of formation, defect of segmentation and mixed malformations.[16] Variations in the number of vertebral bodies were also described. The incidence of rib defects, along with the rib structure and the number of alterations were also analyzed. The

rib defect subtypes were classified as simple or complex.[17] Intraspinal anomalies were defined as any defect involving the spinal cord, such as a tethered cord or syringomyelia.

Generation and Evaluation of the *Tbx6* Gene Edited Mice

Zygotes from FVB mice were edited by CRISPR/Cas9[18] to generate a *Tbx6* frameshift allele and in independent experiments a *Tbx6* hypomorphic allele.[19] Matings of strains with different *Tbx6* genotypes were performed to derive the *Tbx6* locus biallelic variant mouse model with a *Tbx6* LoF variant in exon 2 *in trans* with the hypomorphic haplotype in the *T* binding site in the promoter of *Tbx6* on the second allele ($Tbx6^{\text{wild-type (wt)/-}}$ X $Tbx6^{\text{wt/mild-hypomorphic(mh)}}$ to derive $Tbx6^{-/\text{mh}}$).[19] Multiple adult animals [N = 92, 35–45 days] were evaluated by micro-CT (Bruker, Belgium) to assess the phenotypes; this included 52 with compound inheritance, $Tbx6^{-/\text{mh}}$. No randomization was used and no blinding was done in the animal study.

Development of the TACS Predictive Model

We developed a multivariate model of the risk score (TACScore) to clinically predict TACS. The TACScore was derived from Cohort 1 (training dataset) by logistic regression[20], in which all reliable variables that were significantly associated ($P<0.05$) with TACS were entered into a multivariate model to discover the optimal predictors for distinguishing TACS from all the enrolled subjects with CS using binary logistic regression analysis. A points system was developed for making these complex statistical models useful to clinicians and to evaluate the clinical utility of the TACScore, in which the point of each predictor was assigned according to the product of the corresponding β coefficient and value of the predictors with point totals corresponding to the risk estimate.[20] The receiver operating characteristic (ROC) was used to assess the effectiveness. Cutoff points for the TACScore were determined with the Youden index (J=maxc[sensitivity+specificity-1][21]) to define the maximum potential effectiveness of the predictive models. Then the TACScore was validated independently in Cohort 2 (testing dataset).

Statistical Analysis

The Mann-Whitney U test was used to analyze the ages, the numbers of vertebral and rib malformations. The sex association and prevalence of different types of vertebral malformations were compared using the Pearson χ^2 test or the Fisher exact test. The odds ratio (OR) with 95% confidence interval (CI) was used to assess the influences of *TBX6* LoF variants on vertebral, rib and intraspinal malformations. Statistical analyses were performed with SPSS version 15.0 (SPSS, USA). $P<0.05$ was considered statistically significant. The variance was similar between the groups that are statistically compared. (Methodological details are provided in the Supplementary Materials.)

RESULTS

We have enrolled 345 Chinese patients with sporadic CS (156 males and 189 females, 12y [7–15] as the median age at enrollment [the interquartile range, IQR], 3y [1–9] at onset, Figure S1), including 237 previously reported[5] and 108 new CS patients (Figure 1). For Cohort 2, 142 Japanese patients with CS were recruited in Japan (79 previously reported[10]

and 63 new patients, Figure 1). Ten patients with evidence for compound inheritance, 16p11.2 deletion/*TBX6* LoF variants, who were systematically evaluated for vertebral phenotypes were enrolled in Cohort 3 (Figure 1 and Table 1). Collectively, twenty-six patients with proximal 16p11.2 deletions, five patients with frameshift variants (c.1250_1251insT, c.266_267insC, c.704_705insG, c.1169_1170insC, and c.1179_1180delAG) and two patients with stop-gain variants (c.844C>T [p.R282*], c.933C>A [p.C311*]) in the *TBX6* gene were identified in Cohort 1 (Table 1, Figure 1, Figure S2–4, and Table S2). Similarly, ten 16p11.2/*TBX6* deletions, two *TBX6* frameshift variants (c.156delG and c.935_936insGA), one nonsense variant (c.699G>A [p.W233*]), one splice-site variant (c.119–1G>A) and one novel missense variant (c.333G>T [p.M111I]), which had been confirmed as a LoF variant,[10] were identified in Cohort 2 (Figure 1, Table 1, Figure S4B and Table S2). Further haplotype analysis showed that all patients with *TBX6* LoF variants had the 'T-C-A' risk haplotype (the co-occurrence of three common SNPs, namely, rs2289292, rs3809624 and rs3809627[5,22]) *in trans* on the opposite allele in Cohort 1 and Cohort 2 (Table 1). Thus, the percentage of new CS patients explained by the *TBX6* compound inheritance model (10/108 [9.3%] in Cohort 1 and 15/142 [10.6%] in Cohort 2) was consistent with our previous findings (23/237 [9.7%]).[5] In Cohort 3, ten patients with *TBX6* deleterious variants (nine 16p11.2 deletions and one frameshift variant [c.469_470insCGGC, p.R157fs], Figure S4C, Table 1) were identified. Furthermore, we identified thirteen patients with fourteen *TBX6* missense variants and one patient with an in-frame insertion variant from Cohort 1; these 14 subjects were excluded from the in-depth systematic phenotypic and clinical endophenotypic data analyses described below as the functional effects of these variants were uncertain.

Distinct Endophenotype of Patients with TACS

We divided patients in Cohort 1 into two groups based on the *TBX6* genotype. The 33 CS patients containing *TBX6* LoF variants and the *in trans* 'T-C-A' risk haplotype were classified as those with compound inheritance at *TBX6* and their form of CS as TACS, and the remaining 298 patients were classified as non-TACS (Figure 1).

The TACS patients were significantly younger than the non-TACS at the age of onset (TACS, 2y [1–3]; non-TACS, 3y [1–9]; P=0.02 by the Mann-Whitney U test; Table 2). A higher proportion of male than female was found in the TACS group in Cohort 1 (TACS, 21/33 [63.6%]; non-TACS, 128/298 [43.0%]; P=0.03 by the Pearson χ^2 test; Table 2). Overall, the vertebral anomalies were less complex in the TACS group as defined by the number of vertebral malformations (TACS, 2 [1–2]; non-TACS, 4 [2–6]; $P=2.8\times10^{-9}$; Table 2 and Figure 2). Remarkably, all spinal deformities in the TACS group originated from the defect of vertebral formations, in which 29 (87.9%) had a simple type ($P=9.2\times10^{-14}$, Table 2 and Figure 2). Specifically, all TACS patients exhibited hemivertebrae or butterfly vertebrae (TACS, 33/33 [100%]; non-TACS, 171/298 [57.4%]; $P=4.9\times10^{-8}$; OR 1.7; 95% CI 1.6–1.9; Table 2 and Figure S6A). More segmented hemivertebrae/butterfly vertebrae (TACS, 1 [1–1.5]; non-TACS, 0 [0–1]; $P=1.2\times10^{-12}$; Table 2 and Figure 2) and fewer block vertebrae (TACS, 0 [0–0]; non-TACS, 0 [2–4]; $P=8.8\times10^{-10}$; Table 2 and Figure 2) were noted in the TACS group.

Of note, the vertebral segment location and spine level distribution regarding the type of abnormal vertebrae in the TACS group were distinct from that of the non-TACS group (Figure 2). In the TACS group, the lower part of the spine (T8–S5) was more frequently involved (TACS, 33/33 [100%]; non-TACS, 247/298 [82.9%]; P=4.4×10^{-3}; Table 2 and Figure 2). While the incidence of rib anomalies was comparable between the two groups, more simplified rib anomalies (e.g. a localized fusion of two ribs or an absence of one rib) versus complex abnormalities (e.g. multiple extensive rib fusions or adjacent large chest wall defects) were more frequent in the TACS patients (TACS, 25/33 [75.8%]; non-TACS, 88/298 [29.5%]; P=3.1×10^{-7}; Table 2 and Figure S6B). Furthermore, fewer intraspinal malformations (e.g. tethered cords or syringomyelia) were observed in the TACS group (TACS, 1/33 [3.0%]; non-TACS, 131/298 [44.0%]; P=7.0×10^{-7}; Table 2).

Tbx6 Gene Edited Mouse Model for Compound Inheritance

We engineered the *Tbx6* LoF variant (*Tbx6*$^{-}$) by introducing a 1-bp insertion in exon 2. The *Tbx6*^{mh} allele was generated as a 'functional equivalent' to the human mild hypomorphic allele (Figure S7).[19] Gene expression *in vitro* was down-regulated by the *Tbx6*^{mh} mutant to approximately 65% of the wild-type gene, which was close to the 70% dosage level of the human *TBX6* mild-hypomorphic allele (Figure S7C).[19] These engineered *Tbx6* alleles were used to construct strains with different genotypic combinations.

As predicted, and consistent with literature observations, no homozygotes for the LoF variant were identified in liveborn animals and expected Mendelian ratios of particular genotypic combinations were distorted in liveborns;[19] consistent with embryonic lethality of the *Tbx6*$^{-/-}$ null animals.[23] We acquired through genetic matings and phenotypically assessed mice with five specific genotypes: *Tbx6*$^{wt/wt}$ (N=10; wt, wild-type), *Tbx6*$^{wt/-}$ (N=10), *Tbx6*$^{wt/mh}$ (N=10), *Tbx6*$^{mh/mh}$ (N=10), and *Tbx6*$^{-/mh}$ (N=52). Consistently, only the *Tbx6*$^{-/mh}$ mice exhibited vertebral malformations (48/52 [92.3%], P=9.3×10^{-9} [OR, 13.0; 95% CI, 5.1–33.3], Figure S8 and Table S3). Intriguingly, as observed in human TACS patients, all *Tbx6*$^{-/mh}$ mice had vertebral malformations involving the lower part of the spine (Figure 2 and Table S3). Defects of vertebral column formation were present in most of the *Tbx6*$^{-/mh}$ mice (36/52 [69.2%], P=4.9×10^{-5} [OR, 3.3; 95% CI, 2.2–4.9], Figure 2 & S8 and Table S3). The recapitulation of the type, extent, and distribution of vertebral malformations in the engineered compound inheritance model in mice (Figure 2) further supports the compound inheritance and gene dosage model for TACS and implicates biological perturbations in vertebral column malformations in this type of CS.

Worldwide Multicenter Replications

For the first replication cohort (Cohort 2), we recruited 142 CS patients from Japan and identified fifteen TACS patients (Figure 1, Table 1 and Table S2). As expected, all TACS patients had one or more hemivertebrae/butterfly vertebrae (TACS in Cohort 2, 15/15 [100%]; non-TACS, 99/127 [78.0%]; P=0.04 [OR, 1.3; 95% CI, 1.2–1.4]). All TACS patients in Cohort 2 exhibited malformed vertebrae at the lower part of the spine (T8–S5), which is consistent with the patterns observed in Cohort 1 (Table 1, Table S2 and Figure 2). This pattern, of vertebral malformations and distribution of defects in the vertebral column, was observed only in the mice with compound inheritance and gene dosage perturbations

below that of haploinsufficiency (i.e. *Tbx6*$^{wt/-}$), but not equivalent to *Tbx6*$^{-/-}$ combination of alleles.

Furthermore, we collected another replication case series (Cohort 3) containing 10 TACS patients in the United State. Intriguingly, hemivertebrae/butterfly vertebrae involving the lower half of the spine were again observed. The replication of our observations for this distinct genotypic combination in two additional independent cohorts from distinct world populations provide substantial evidence in support of the compound inheritance and *TBX6* gene dosage model and speaks to the universality of TACS and the compound inheritance model in human and medical genetics.

TACScore to Clinically Predict TACS

We developed a model to predict TACS from the phenotypic data, and clinically measurable endophenotypes, obtained from Cohort 1 by logistic regression.[20] The final multivariate risk model, the TACScore, integrated variables including (1) segmented hemivertebrae/butterfly vertebrae involving the lower half of the spine (T8–S5), (2) the number of vertebral malformations, (3) the presence of intraspinal defects and (4) the type of rib defect (Figure 3A and Table S4). The calculated score of each variant parameter was linearly correlated with the corresponding risk in the regression model and defined from −8 to 4 (Figure 3B). The cutoff point was selected as ≥3 to achieve the highest Youden index (J=84.8%), with sensitivity, specificity, and accuracy levels higher than 90% (Table S5). The area under the curve (AUC) of the ROC curve was 0.9 (P=1.6×10^{-15}; 95% CI, 0.9–1.0; Figure 3C). Importantly, we validated the TACScore in Cohort 2 with an AUC of 0.8 (P=1.5×10^{-4}; 95% CI, 0.7–0.9; Figure 3C and Table S6). Therefore, we introduced a clinical assessment pipeline for CS to efficiently identify high-risk patients for TACS and potential compound inheritance at the *TBX6* locus (Figure 3D).

DISCUSSION

In this study, we expanded our knowledge of the contribution of the *TBX6* compound inheritance to CS by providing international patient cohort studies and animal model-based evidence that: i) 9.6% (33/345) and 10.6% (15/142) of the patients in Cohort 1 and Cohort 2, respectively, could be parsimoniously explained by the *TBX6* compound inheritance gene dosage model (Table 1 and Figure 1); ii) all patients in the three cohorts had the risk haplotype *in trans* with a *TBX6* LoF variant, with the exception of one patient [BCM01] for whom haplotype information was not available and one patient [BH8084] for whom allelic phasing of the identified variants was not available (Table 1); and iii) only the *Tbx6*$^{-/mh}$ genotype presented with vertebral malformations in the gene-edited mice (Figure 2 & S8 and Table S3). Moreover, we elucidated a novel genetically and phenotypically defined disease entity, which carries a *TBX6* LoF variant *in trans* with the hypomorphic allele and manifests a distinguishable constellation of clinical features and CS endophenotypes. Furthermore, we developed the TACScore system and proposed a clinical practice algorithm (Figure 3) for evaluating CS and predicting TACS. Finally, we show qualitatively the same kind (hemivertebrae/butterfly vertebrae) and quantitatively the same distribution (lower

spine) of vertebral malformations with the compound inheritance combination of variant alleles and *TBX6*/*Tbx6* gene dosage model in human and mice.

TBX6 is a member of the T-box family.[24] Variants in *Tbx6* lead to phenotypes of vertebral and rib defects in mice.[23,25] *Tbx6* is important for cell fate in the paraxial mesoderm structure[26] and formation of the segmental boundary[27], which is the precursor of the vertebral column.[28] Therefore, the genetic and developmental malfunction in *Tbx6* perturbation of gene dosage model may parsimoniously explains the hemivertebrae/butterfly vertebrae presenting in all TACS patients (Table 1 and Figure 2). Interestingly, $Tbx6^{-/-}$ produces an embryonic lethal phenotype with expanded tailbud and lack of posterior somites,[23] suggesting that *Tbx6* functions to maintain the specification of the posterior paraxial mesoderm.[7] In zebrafish, *tbx6* is required in the posterior of the embryo during the posterior part of primary neurulation and all of the secondary neurulation.[29,30] Therefore, this developmental process optimally explains the observation that all TACS patients and $Tbx6^{-/mh}$ gene-edited mice show vertebral malformations anatomically localized at the lower part of the spine.

TACS does not follow 'conventional' rare variant Mendelian inheritance expectations likely reflecting the embryonic lethality of *Tbx6/TBX6* homozygous null alleles. Instead, TACS is caused by biallelic variants at a locus, consistent with autosomal recessive trait manifestation, but the presence of a single rare LoF variant *in trans* with one common hypomorphic allele; a gene dosage that is less than haploinsufficiency but not zero as with homozygous null alleles. This genetic model (compound inheritance and TACS) may provide an explanation for the scenarios when monoallelic variants and phenotypes are perceived to be following a dominant inheritance pattern but for a disease trait with incomplete penetrance or the disease pattern observed is pseudodominance as found when the carrier state occurs at a high frequency in the population. Because the 'T-C-A' haplotype is common worldwide (44% among Asians and 33% among Europeans, but <1% among Africans[31]), the haplotype is likely to be parsed and filtered out by the current genomic analytical pipelines. Therefore, for some single-locus genetic models, the functional effects of individual variant alleles, as well as the combination of alleles should be considered when interpreting the genetic data, and specific phenotype observed, for the 'unsolved' cases. Moreover, the compound inheritance model may be particularly relevant to recurrent rearrangement CNV[32] loci with relatively high mutational frequency rates[33] wherein deletion CNV can result in haploinsufficiency for many gene loci (e.g. the 0.6-Mb 16p11.2 deletion CNV contains 27 annotated genes).[5,15] It is important to know the functional effects of individual variant alleles, but also the functional consequences of combinations of variant alleles at a locus.

A significantly small proportion of the TACS patients in Cohort 1 were female (Table 2, 12/33 [36.4%] in the TACS patients and 170/298 [57.0%] in the non-TACS patients in Cohort 1), which might be due to the interference of the 16p11.2 deletion and *TBX6* variants in the formation of the female reproductive system.[34] Notably, pleiotropic effects of 16p11.2 deletion CNV include other traits related to obesity[35], and cognitive phenotypes including intellectual disability and psychiatric disorders[15,36], birth defects such as CAKUT (congenital anomalies of the kidney and urinary tract)[37], and may even be involved in other

undiscovered phenotypes or diseases. Thus, TACS patients are recommended to be evaluated systematically with long-term follow-up (Figure 3D).

More importantly, from an orthopedic surgeon's clinical perspective, hemivertebrae/butterfly vertebrae are the most remarkable features of TACS (Figure 2). Surgical intervention and treatment at an early age results in a satisfactory prognosis for the patient and ameliorates scoliosis.[38] For example, patient XH101 was diagnosed as having CS with a L2 hemivertebra at 3 years of age, and molecular testing indicated that she had a frameshift variant in *TBX6* with the *in trans* 'T-C-A' risk haplotype. The TACS diagnosis was made by integrating the molecular testing and clinical phenotypes. We performed L2 hemivertebra resection with short segment internal fixation. We removed the instrument at 3-years post-operation when she had a balanced spine. The last follow-up, which was at 9 years post-operation, showed an almost completely normal growth of the spine and normal curvature of the spinal column (Figure S9).

To realize clinical applications from bench to bedside, we developed a TACScore prediction algorithm as an objective adjuvant clinical measure to guide clinical management for patients with CS. The TACScore was validated in two independent cohorts and further integrated into the proposed guidelines for evaluating the risk of TACS (Figure 3). Considering the possibility of phenotypic expansion,[4] the prediction algorithm should include multiple variables associated with TACS comprehensively (i.e. endophenotypes) instead of relying on one particular phenotype (e.g. scoliosis). This clinical phenotypic assessment provides efficient information to physicians and families regarding whether the patient is suitable for 16p11.2/*TBX6* variant detection or a genome-wide test. Clinical ES[39] or genome-wide analysis are reserved for cases with low-risk TACScores or high-risk TACScores that are negative for 16p11.2/*TBX6* pathogenic variants. Furthermore, the cutoff point of the TACScore could be adjusted according to medical practice experience and continued future observations as well as the local medical resources available to achieve the maximum health economic benefits.

Conclusions

We have defined a new subtype of CS; namely TACS, containing a *TBX6* LoF variant compound with a hypomorphic allele. TACS is a clinical entity defined by consistent clinically measurable endophenotypes: i.e., younger age at onset, hemivertebrae/butterfly vertebrae involving the lower part of the spine, simple rib anomaly, and fewer vertebrae and intraspinal defects. TACScore can guide clinical management and genetic and clinical genomic testing. Human and mouse studies further confirm the compound inheritance and gene dosage model, and provide insights into potential biological consequences for spine development of perturbations in *TBX6/Tbx6* gene dosage and expression. Such genetic models may be important to other birth defects.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Jiaqi Liu, MD[1,2,3,†], Nan Wu, MD[1,2,4,5,†,*], Deciphering Disorders Involving Scoliosis and COmorbidities (DISCO) study, Nan Yang, PhD[6,7,8,†], Kazuki Takeda, MD[9,10,†], Weisheng Chen, MD[1,11,†], Weiyu Li, BSc[6,7], Renqian Du, DDS, PhD[5], Sen Liu, MD[1,2,4], Yangzhong Zhou, MD[2,12], Ling Zhang, PhD[6,7], Zhenlei Liu, MD[2,13], Yuzhi Zuo, MD[1,2,4], Sen Zhao, BS[1,2], Robert Blank, MD, PhD[14], Davut Pehlivan, MD[5], Shuangshuang Dong, BSc[6,7], Jianguo Zhang, MD[1,2,4], Jianxiong Shen, MD[1,2,4], Nuo Si, PhD[15,16], Yipeng Wang, MD[1], Gang Liu, MD[1,2,4], Shugang Li, MD[1], Yanxue Zhao, BS[1,2], Hong Zhao, MD[1], Yixin Chen, MD[1,2], Yu Zhao, MD[1], Xiaofei Song, BS[5], Jianhua Hu, MD[1], Mao Lin, MD[1,2,11], Ye Tian, MD[1], Bo Yuan, PhD[5], Keyi Yu, MD[1], Yuchen Niu, MS[2,17], Bin Yu, MD[1], Xiaoxin Li, MS[2,17], Jia Chen, MD[1,2], Zihui Yan, BS[1,2,11], Qiankun Zhu, MD[1,2], Xiaolu Meng, MS[15,16], Xiaoli Chen, PhD[18], Jianzhong Su, PhD[19], Xiuli Zhao, PhD[15,16], Xiaoyue Wang, PhD[16], Yue Ming, MD[20], Xiao Li, MD[21], Cathleen L. Raggio, MD[22], Baozhong Zhang, MD[1], Xisheng Weng, MD[1,2,4], Shuyang Zhang, MD, PhD[2,23], Xue Zhang, MD, PhD[2,15,16], Kota Watanabe, MD, PhD[10], Japan Early Onset Scoliosis Research Group[10], Li Jin, PhD[6], Yiping Shen, PhD[24,25], Nara L. Sobreira, MD, PhD[26], Jennifer E. Posey, MD, PhD[5], Philip Giampietro, MD, PhD[27], Baylor-Hopkins Center for Mendelian Genomics[26], Pengfei Liu, PhD[5,28], Zhihong Wu, MD, PhD[2,4,17,‡], Shiro Ikegawa, MD, PhD[9,‡], James R. Lupski, MD, PhD, DSc (hon)[5,29,30,‡], Feng Zhang, PhD[6,7,8,‡], Guixing Qiu, MD[1,2,4,*,‡]

Affiliations

[1]Department of Orthopedic Surgery, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China.

[2]Beijing Key Laboratory for Genetic Research of Skeletal Deformity, Beijing, China.

[3]Department of Breast Surgical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.

[4]Medical Research Center of Orthopedics, Chinese Academy of Medical Sciences, Beijing, China.

[5]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA.

[6]Obstetrics and Gynecology Hospital, State Key Laboratory of Genetic Engineering at School of Life Sciences, Institute of Metabolism and Integrative Biology, Fudan University, Shanghai, China.

[7]NHC Key Laboratory of Reproduction Regulation, Shanghai Institute of Planned Parenthood Research, Fudan University, Shanghai, China.

[8]Shanghai Key Laboratory of Female Reproductive Endocrine Related Diseases, Shanghai, China.

[9]Laboratory of Bone and Joint Diseases, Center for Integrative Medical Sciences, RIKEN, Tokyo, Japan.

[10]Department of Orthopedic Surgery, Keio University School of Medicine, Tokyo, Japan.

[11]Graduate School of Peking Union Medical College, Beijing, China.

[12]Department of Internal Medicine, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China.

[13]Department of Neurosurgery, Xuanwu Hospital, Capital Medical University, Beijing, China.

[14]Department of Medicine, Medical College of Wisconsin, Milwaukee, USA.

[15]The McKusick–Zhang Center for Genetic Medicine, Institute of Basic Medical Sciences, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China.

[16]The State Key Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China.

[17]Department of Central Laboratory, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China.

[18]Department of Medical Genetics, Capital Institute of Pediatrics, Beijing, China.

[19]College of Biomedical Engineering, The Eye Hospital, Wenzhou Medical University, Wenzhou, Zhejiang, China.

[20]PET-CT Center, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.

[21]Department of Radiology, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China.

[22]Department of Orthopedic Surgery, Hospital for Special Surgery, NY, USA.

[23]Department of Cardiology, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China.

[24]Division of Genetics and Genomics, Boston Children’s Hospital, Boston, Massachusetts, USA

[25]Harvard Medical School, Boston, Massachusetts, USA.

[26]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

[27]Department of Pediatrics, Drexel University College of Medicine, Philadelphia, Pennsylvania, USA.

[28]Baylor Genetics Laboratory, Houston, Texas, USA.

[29]Departments of Pediatrics, Baylor College of Medicine, Houston, Texas, USA.

[30]Texas Children's Hospital, Houston, Texas, USA.

Acknowledgments:

We thank all the individuals, families, and physicians involved in the study for their participation. We thank the nurses from the department of orthopedic surgery of Peking Union Medical College Hospital for assistance with patient enrollment. We thank the members of Japan Early Onset Scoliosis Research Group: Drs. Noriaki Kawakam, Toshiaki Kotani, Hideki Sudo, Ikuho Yonezawa, Koki Uno, Hiroshi Taneichi, Kei Watanabe, Shohei Minami, Hideki Shigematsu, Ryo Sugawara, Yuki Taniguchi, and Nao Ootomo. We thank Dr. Sally Dunwoodie and Dr. Gavin Chapman from Victor Chang Cardiac Research Institute for their collaborations. We appreciate the support of Dr. Ellen Wald at the University of Wisconsin-Madison.

This research was funded in part by the National Natural Science Foundation of China (81822030 and 81501852 to N.W., 81472045 and 81772301 to G.Q., 31625015, 31571297 and 31771396 to F.Z., 81472046 and 81772299 to Z.W., 31521003 to L.J., 81672123 to J.Z., 7162029 to X.C.), Beijing Natural Science Foundation (7172175 to N.W., 7162029 to X.C.), Beijing Nova Program (Z161100004916123 to N.W.), Beijing Nova Program Interdisciplinary Collaborative Project (xxjc201717 to N.W.), 2016 Milstein Medical Asian American Partnership Foundation Fellowship Award in Translational Medicine (to N.W.), The Central Level Public Interest Program for Scientific Research Institute (2016ZX310177 to N.W.), PUMC Youth Fund & the Fundamental Research Funds for the Central Universities (3332016006 to N.W.), CAMS Initiative Fund for Medical Sciences (2016-I2M-3–003 to G.Q. and N.W., 2016-I2M-2–006 and 2017-I2M-2–001 to Z.W.), the Distinguished Youth Foundation of Peking Union Medical College Hospital (JQ201506 to N.W.), the 2016 PUMCH Science Fund for Junior Faculty (PUMCH-2016–1.1 to N.W.), and the National Key Research and Development Program of China (No. 2016YFC0901501 to S.Z). Also supported by Japan Agency for Medical Research and Development (AMED, No. 17ek0109280h0001 and 17824969 to S.I.) and Japan Orthopedics and Traumatology Research Foundation (No. 358 to K.T.), and the US National Institutes of Health, National Institute of Neurological Disorders and Stroke (NINDS R01 NS058529 and R35 NS105078 to J.R.L), National Human Genome Research Institute/National Heart, Lung, and Blood Institute (NHGRI/NHLBI UM1 HG006542 to J.R.L), and the National Human Genome Research Institute (NHGRI K08 HG008986 to J.E.P).

References:

1. Manolio TA, Fowler DM, Starita LM, et al. Bedside back to bench: building bridges between basic and clinical genomic research. Cell. 2017;169(1):6–12. [PubMed: 28340351]
2. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. JAMA. 2014;312(18):1870–1879. [PubMed: 25326635]
3. Maver A, Lovrecic L, Volk M, et al. Phenotype-driven gene target definition in clinical genome-wide sequencing data interpretation. Genet Med. 2016;18(11):1102–1110. [PubMed: 27031083]
4. Karaca E, Posey JE, Coban Akdemir Z, et al. Phenotypic expansion illuminates multilocus pathogenic variation. Genet Med. 2018 10.1038/gim.201833.
5. Wu N, Ming X, Xiao J, et al. *TBX6* null variants and a common hypomorphic allele in congenital scoliosis. N Engl J Med. 2015;372(4):341–350. [PubMed: 25564734]
6. Hedequist D, Emans J. Congenital scoliosis. J Am Acad Orthop Surg. 2004;12(4):266–275. [PubMed: 15473678]
7. Sparrow DB, Chapman G, Smith AJ, et al. A mechanism for gene-environment interaction in the etiology of congenital scoliosis. Cell. 2012;149(2):295–306. [PubMed: 22484060]
8. Shen J, Wang Z, Liu J, Xue X, Qiu G. Abnormalities associated with congenital scoliosis: a retrospective study of 226 Chinese surgical cases. Spine (Phila Pa 1976). 2013;38(10):814–818. [PubMed: 23197014]
9. Lefebvre M, Duffourd Y, Jouan T, et al. Autosomal recessive variations of *TBX6*, from congenital scoliosis to spondylocostal dysostosis. Clin Genet. 2017;91(6):908–912. [PubMed: 27861764]
10. Takeda K, Kou I, Kawakami N, et al. Compound heterozygosity for null mutations and a common hypomorphic risk haplotype in *TBX6* causes congenital scoliosis. Human Mutation. 2017;38(3): 317–323. [PubMed: 28054739]

11. Shen Y, Chen X, Wang L, et al. Intra-family phenotypic heterogeneity of 16p11.2 deletion carriers in a three-generation Chinese family. Am J Med Genet B Neuropsychiatr Genet. 2011;156(2):225–232. [PubMed: 21302351]
12. Al-Kateb H, Khanna G, Filges I, et al. Scoliosis and vertebral anomalies: additional abnormal phenotypes associated with chromosome 16p11.2 rearrangement. Am J Med Genet A. 2014;164A(5):1118–1126. [PubMed: 24458548]
13. Sparrow DB, McInerney-Leo A, Gucev ZS, et al. Autosomal dominant spondylocostal dysostosis is caused by mutation in *TBX6*. Hum Mol Genet. 2013;22(8):1625–1631. [PubMed: 23335591]
14. Shen Y, Irons M, Miller DT, et al. Development of a focused oligonucleotide-array comparative genomic hybridization chip for clinical diagnosis of genomic imbalance. Clin Chem. 2007;53(12): 2051–2059. [PubMed: 17901113]
15. Shinawi M, Liu P, Kang SH, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. J Med Genet. 2010;47(5):332–341. [PubMed: 19914906]
16. McMaster MJ, Ohtsuka K. The natural history of congenital scoliosis. A study of two hundred and fifty-one patients. J Bone Joint Surg Am. 1982;64(8):1128–1147. [PubMed: 7130226]
17. Tsirikos AI, McMaster MJ. Congenital anomalies of the ribs and chest wall associated with congenital deformities of the spine. J Bone Joint Surg Am. 2005;87(11):2523–2536. [PubMed: 16264130]
18. Platt RJ, Chen S, Zhou Y, et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell. 2014;159(2):440–455. [PubMed: 25263330]
19. Yang N, Wu N, Zhang L, et al. *TBX6* compound inheritance leads to congenital vertebral malformations in humans and mice. Hum Mol Genet. 2018 10.1093/hmg/ddy358.
20. Sullivan LM, Massaro JM, D'Agostino RB, Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med. 2004;23(10):1631–1660. [PubMed: 15122742]
21. Reiser B Measuring the effectiveness of diagnostic markers in the presence of measurement error through the use of ROC curves. Stat Med. 2000;19(16):2115–2129. [PubMed: 10931515]
22. Liu J, Zhou Y, Liu S, et al. 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. 2018;137(6–7):553–567. [PubMed: 30019117]
23. Chapman DL, Papaioannou VE. Three neural tubes in mouse embryos with mutations in the T-box gene *Tbx6*. Nature. 1998;391(6668):695–697. [PubMed: 9490412]
24. Papapetrou C, Putt W, Fox M, Edwards YH. The human *TBX6* gene: cloning and assignment to chromosome 16p11.2. Genomics. 1999;55(2):238–241. [PubMed: 9933572]
25. White PH, Farkas DR, McFadden EE, Chapman DL. Defective somite patterning in mouse embryos with reduced levels of *Tbx6*. Development. 2003;130(8):1681–1690. [PubMed: 12620991]
26. Chapman DL, Agulnik I, Hancock S, Silver LM, Papaioannou VE. *Tbx6*, a mouse T-Box gene implicated in paraxial mesoderm formation at gastrulation. Dev Biol. 1996;180(2):534–542. [PubMed: 8954725]
27. Zhao W, Ajima R, Ninomiya Y, Saga Y. Segmental border is defined by *Ripply2*-mediated *Tbx6* repression independent of *Mesp2*. Dev Biol. 2015;400(1):105–117. [PubMed: 25641698]
28. Dahmann C, Oates AC, Brand M. Boundary formation and maintenance in tissue development. Nat Rev Genet. 2011;12(1):43–55. [PubMed: 21164524]
29. Copp AJ, Greene ND, Murdoch JN. The genetic basis of mammalian neurulation. Nat Rev Genet. 2003;4(10):784–793. [PubMed: 13679871]
30. Nikaido M, Kawakami A, Sawada A, Furutani-Seiki M, Takeda H, Araki K. *Tbx24*, encoding a T-box protein, is mutated in the zebrafish somite-segmentation mutant fused somites. Nat Genet. 2002;31(2):195–199. [PubMed: 12021786]
31. The 1000 Genomes Project Consortium, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56–65. [PubMed: 23128226]

32. Lupski JR. Structural variation mutagenesis of the human genome: Impact on disease and evolution. Environ Mol Mutagen. 2015;56(5):419–436. [PubMed: 25892534]
33. Lupski JR. Genomic rearrangements and sporadic disease. Nat Genet. 2007;39(7 Suppl):S43–47. [PubMed: 17597781]
34. Sandbacka M, Laivuori H, Freitas E, et al. *TBX6, LHX1* and copy number variations in the complex genetics of Mullerian aplasia. Orphanet J Rare Dis. 2013;8:125. [PubMed: 23954021]
35. Bachmann-Gagescu R, Mefford HC, Cowan C, et al. Recurrent 200-kb deletions of 16p11.2 that include the *SH2B1* gene are associated with developmental delay and obesity. Genet Med. 2010;12(10):641–647. [PubMed: 20808231]
36. Weiss LA, Shen Y, Korn JM, et al. Association between microdeletion and microduplication at 16p11.2 and autism. N Engl J Med. 2008;358(7):667–675. [PubMed: 18184952]
37. Verbitsky M, Westland R, Perez A, et al. The copy number variation landscape of congenital anomalies of the kidney and urinary tract. Nat Genet. In press.
38. Guo J, Zhang J, Wang S, et al. Surgical outcomes and complications of posterior hemivertebra resection in children younger than 5 years old. J Orthop Surg Res. 2016;11(1):48. [PubMed: 27113726]
39. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med. 2013;369(16):1502–1511. [PubMed: 24088041]

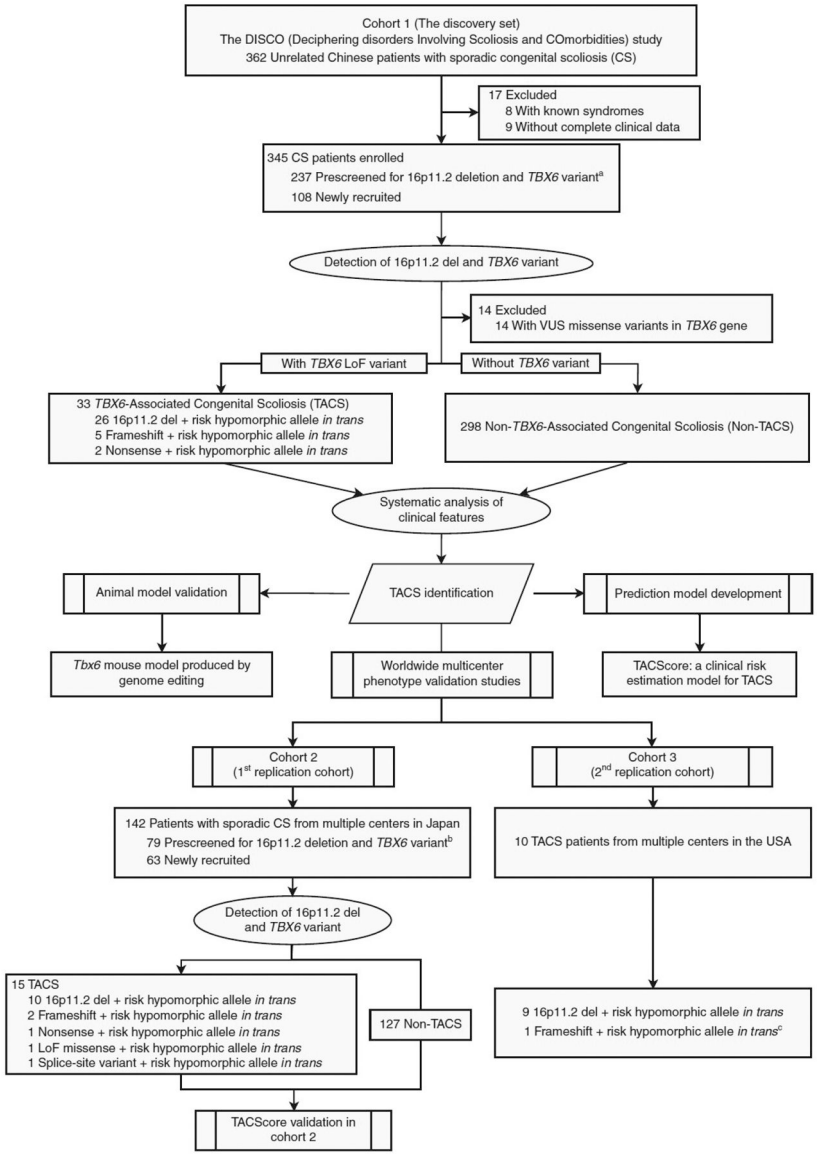

Figure 1. Workflow for *TBX6*-Associated Congenital Scoliosis (TACS) Study Participants in the Multicenter Cohorts and the Animal Model Genotype-Phenotype Analyses.

The discovery set (Cohort 1) consisted of 345 unrelated Chinese patients with sporadic CS. The status of the *TBX6* variants was screened, and the clinical characteristics were reviewed and compared according to the *TBX6* genotypes, including vertebral, rib, and intraspinal anomalies. To recapitulate and investigate the phenotypic consequences of the compound inheritance and *TBX6*/*Tbx6* gene dosage genetic model, a mouse strain with a specific combination of *Tbx6* alleles was constructed. *Tbx6* alleles were individually engineered and a mouse strain for the compound inheritance model was constructed by mating, introducing a truncated allele *in trans* with a mild-hypomorphic (mh) allele; $Tbx6^{-/mh}$. Furthermore, the *TBX6*-associated CS score (TACScore) was developed to guide and increase the efficiency of diagnosing TACS from clinically observed endophenotypes.

Abbreviations: CS, congenital scoliosis; VUS, variants with unknown significance; LoF, loss-of-function; TACS, *TBX6*-associated CS; Non-TACS, non-*TBX6*-associated CS.

[a] Cohort 1 comprised the 237 sporadic CS patients enrolled between October 2010 and June 2014 in our previous study.[5]

[b] Cohort 2 comprised 79 Japanese CS patients from 94 previously described CS patients,[10] after excluding 15 patients without complete clinical data.

[c] No information regarding whether the risk haplotype was *in trans* or *in cis* with the *TBX6* frameshift variant.

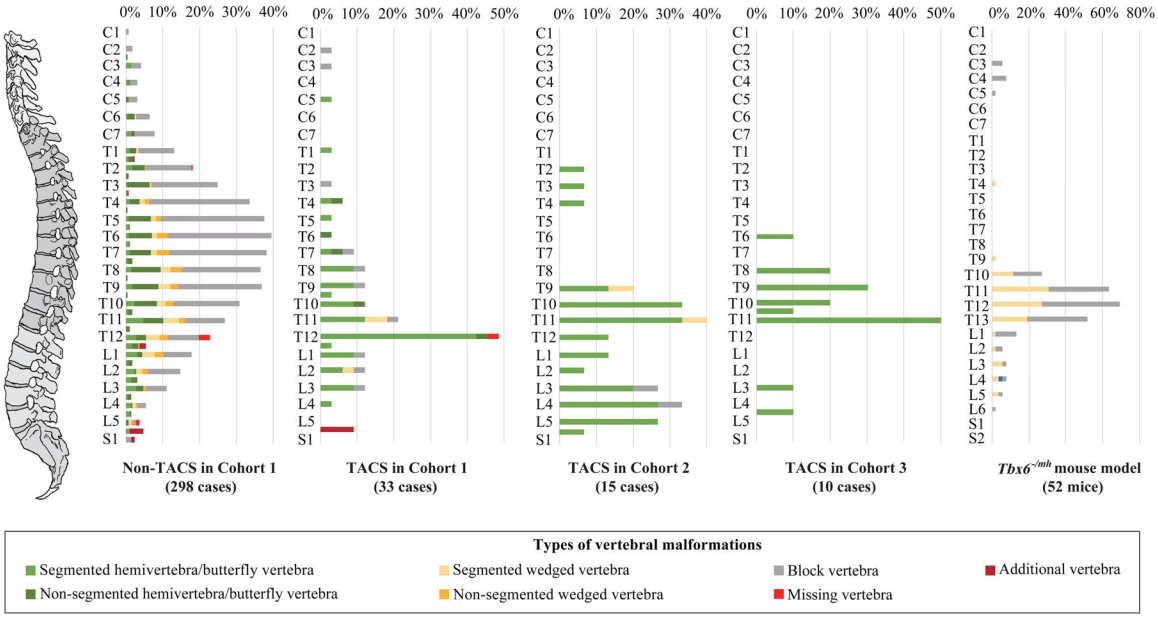

Figure 2. Comparison of the Distribution Regarding Abnormal Vertebrae in Non-TACS and TACS Patients and in *Tbx6*$^{-/mh}$ Mouse Compound Inheritance and Gene Dosage Model.

The X-axis shows the frequency of each malformation in each vertebra and Y-axis shows the vertebral distribution in the spine. The vertebral malformations in the non-TACS group (N=298) were normally distributed. There were 58 TACS patients in total in three worldwide cohorts (Cohort 1 from PUMCH in China, N=33; Cohort 2 from the multiple centers in Japan, N=15; Cohort 3 from multiple centers in the USA, N=10). The *Tbx6*$^{-/mh}$ engineered mice [N=52] exhibited a distinct phenotype, namely, a defect of formation involving the lower part of the spine. In addition, malformations in the upper and middle thoracic spine were significantly less involved (P=7.3×10^{-5} and 3.0×10^{-6}, respectively) in the TACS group than in the non-TACS group.

Abbreviation: CS, congenital scoliosis; TACS, *TBX6*-associated CS; Non-TACS, non-*TBX6*-associated CS; *mh*, mild-hypomorphic; C, cervical vertebra; T, thoracic vertebra; L, lumbar vertebra; S, sacral vertebra.

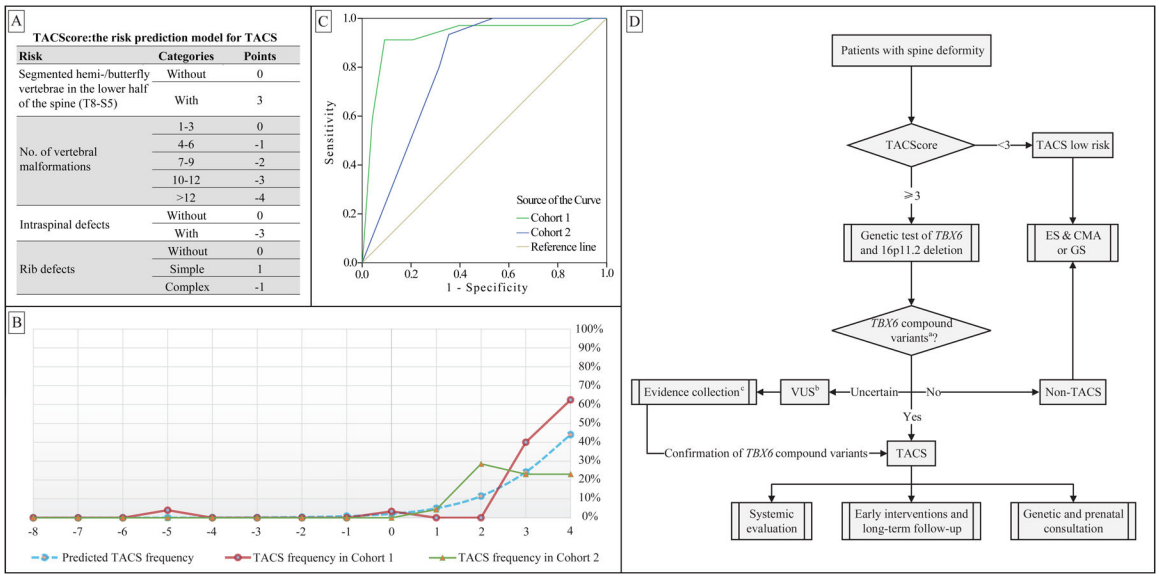

Figure 3. Developing and Validation of the Risk Prediction Model and Diagnostic Pipeline of TACS. (A) TACScore: risk predictive model for TACS.

The final multivariate risk model was developed through a binary logistic regression analysis of detailed phenotypic data from Cohort 1. **(B) The predictive efficacy of the TACScore**. The X-axis shows the spectrum of TACScore and Y-axis shows the predicted TACS frequency and the percentage of TACS patients in all CS patients with each calculated score in Cohort 1 and Cohort 2. The TACScore presented excellent predictive efficacy by comparing the predicted TACS risk to the real TACS frequency. The cutoff point was selected as 3 to achieve the highest accuracy. **(C) ROC Curve for the TACScore in Cohort 1 and Cohort 2**. AUCs were 0.9 (P=1.6×10^{-15}; 95% CI, 0.9–1.0) for the discovery cohort (Cohort 1) and 0.8 (P=1.5×10^{-4}; 95% CI, 0.7–0.9) for the validation cohort (Cohort 2). **(D) A proposed guideline for predicting and evaluating TACS**. The risk of TACS evaluated by TACScore is suggested to perform prior to genetic testing. After the detection of TACS, a systemic evaluation, with early interventions and genetic consultation were recommended. Abbreviation: CS, congenital scoliosis; TACS, *TBX6*-associated CS; ES, exome sequencing; CMA, chromosomal microarray analysis; GS, genome sequencing.

[a] *TBX6* compound variant contains a 16p11.2 deletion/*TBX6* loss-of-function variant in the compound heterozygous configuration with the risk haplotype providing a hypomorphic variant.

[b] VUS, variants with unknown significance.

[c] Evidence from large-scale case-control studies, pedigree analysis, and functional studies are needed.

Table 1.

Genotype and Phenotype Information of TACS Patients from Multiple Centers Worldwide.

Center	Subject	Gender	Age (*year*)[a]	LoF Mutation at First Allele	Risk haplotype at Second Allele[b]	Phenotype
Cohort 1 (10 newly recruited TACS patients[c])						
Peking Union Medical College Hospital	XH139	M	3	16p11.2 del	T-C-A	T12 hemi (r) and T11 wedge vertebra (r); Missing left 12th rib
	XH330	M	4	16p11.2 del	T-C-A	T10 hemi (l) and T12 hemi (r); Missing right 10th and left 12th rib
	XH468	M	3	16p11.2 del	T-C-A	T8–T11 butterfly vertebrae and T12 hemi (l); Missing right 12th rib
	XH480	F	11	16p11.2 del	T-C-A	C2–C3 blocks and T10 hemi (right, unsegmented to T9&T11); Missing left 10th rib
	XH522	M	8	16p11.2 del	T-C-A	Hemi (r) between T9–T10; Additional rib between right 9th-10th
	XH529	M	14	16p11.2 del	T-C-A	T12 hemi (l); Missing right 12th rib
	XH605	F	11	16p11.2 del	T-C-A	T12 hemi (r) and six lumbar vertebrae; Missing left 12th rib
	XH623	M	1	16p11.2 del	T-C-A	L4 hemi (r)
	XH636	F	9	16p11.2 del	T-C-A	L1 butterfly vertebra
	XH625	M	7	c.933C>A (p.C311*)	T-C-A	T12 hemi (r) and six lumbar vertebrae; Missing left 12th rib
Cohort 2 (6 newly recruited TACS patients[d])						
RIKEN and participating hospitals in Japan	A1042	F	15	16p11.2del	T-C-A	T10 hemi, L5 butterfly vertebra
	A1076	F	9	16p11.2del	T-C-A	L2 hemi, Bilateral 13th ribs
	A9022	M	7	16p11.2del	T-C-A	T10-L1 hemi
	S1275	M	14	16p11.2del	T-C-A	T4 hemi, T2–3&L4 butterfly vertebra
	S1325	F	13	16p11.2del	T-C-A	L4–5 butterfly vertebra
	A1107	F	15	c.119–1G>A	T-C-A	L3 butterfly vertebra
Cohort 3						
Baylor College of Medicine	BCM01	F	Caucasian	16p11.2del	NA	T9 hemi

Center	Subject	Gender	Age (*year*)[a]	LoF Mutation at First Allele	Risk haplotype at Second Allele[b]	Phenotype
	BCM02	F	Hispanic	16p11.2del	T-C-A	T9, T10 and T11 hemi
	BCM03	F	Caucasian	16p11.2del	T-C-A	T10 hemi
	BCM04	M	Hispanic	16p11.2del	c.853C>T[e]	T8 and T9 butterfly vertebrae, Hemi between T10 and T11
Drexel University College of Medicine	BH8084	F	Caucasian	c.469_470insCGGC p.R157fs	T-C-A[f]	Hemi between L4 and L5
Boston Children's Hospital	BS01	F	Caucasian	16p11.2 del	T-C-A	T6 and T8 butterfly vertebrae
	BS19	M	Asian	16p11.2 del	T-C-A	T11 hemi (l)
Washington University School of Medicine	PT04	M	Caucasian	16p11.2 del	T-C-A	T11 hemi L3 butterfly vertebra
	PT05	M	Caucasian	16p11.2 del	T-C-A	T11 vertebral anomaly
Children's Hospital Central California	PT08	F	Hmong	16p11.2 del	T-C-A	T11 hemi

Abbreviation: CS, congenital scoliosis; TACS, *TBX6*-associated CS; LoF, Loss-of-Function; del, deletion; hemi, hemivertebra; hemi (*l*), hemivertebra on the left side; hemi (*r*), hemivertebra on the right side; C, cervical vertebra; T, thoracic vertebra; L, lumbar vertebra; NA, not available.

[a] Age at the time of enrollment.

[b] The risk haplotype is defined by three *TBX6* SNPs (wild-type/mutant): rs2289292 (C/T) - rs3809624 (T/C) - rs3809627 (C/A).[5,22]

[c] The detailed genotypes and phenotypes of the remaining 23 TACS patients in Cohort 1 were reported previously[5] and are listed in supplementary Table S1.

[d] The detailed genotypes and phenotypes of the remaining 9 TACS patients in Cohort 2 were reported previously[10] and are listed in supplementary Table S1.

[e] The variant is novel in European ancestry, and is predicted to be deleterious (Polyphen: possibly_damaging, SIFT: deleterious).

[f] No information regarding whether the risk haplotype was *in trans* or *in cis* with the *TBX6* frameshift mutation.

Table 2.

Comparison of Clinical Characteristics between TACS and Non-TACS Patients in Cohort 1.

Clinical characteristics	TACS (n=33)	Non-TACS (n=298)	*P* value[a]	OR	95% CI
Age of onset, median (IQR), yr	2 (1–3)	3 (1–9)	0.02[b]		
Male, No. (%)	21 (63.6)	128 (43.0)	0.03[c]	2.3	1.1–4.9
Vertebral malformation					
Classification of origins, No. (%)					
Defect of formation	29 (87.9)	69 (23.2)			
Defect of segmentation	0 (0)	69 (23.2)	9.2×10^{-14} [c]		
Mixed defects	4 (12.1)	160 (53.7)			
Incidence of hemi-/butterfly vertebrae, No. (%)	33 (100)	171 (57.4)	4.9×10^{-8} [c]	1.7	1.6–1.9
Number of vertebral malformations, median (IQR)	2 (1–2)	4 (2–6)	2.8×10^{-9} [b]		
Hemi-/butterfly vertebrae	1 (1–1.5)	1 (0–2)	0.02[b]		
Segmented hemi-/butterfly vertebrae	1 (1–1.5)	0 (0–1)	1.2×10^{-12} [b]		
Non-segmented hemi-/butterfly vertebrae	0 (0–0)	0 (0–1)	8.4×10^{-4} [b]		
Block vertebrae	0 (0–0)	0 (2–4)	8.8×10^{-10} [b]		
Incidence of malformations at each location, No. (%)					
Cervical (C1–C7)	2 (6.1)	30 (10.1)	0.8[c]	0.6	0.1–2.5
Thoracic (T1–T12)	28 (84.9)	261 (87.6)	0.6[c]	0.8	0.3–2.2
Upper thoracic (T1–T4)	2 (6.1)	117 (39.3)	7.3×10^{-5} [c]	0.1	0–0.4
Middle thoracic (T5–T8)	6 (18.2)	181 (60.7)	3.0×10^{-6} [c]	0.1	0.1–0.4
Lower thoracic (T9–T12)	24 (72.7)	180 (60.4)	0.2[c]	1.8	0.8–3.9
Lumbar (L1–L5)	15 (45.5)	111(37.3)	0.4[c]	1.4	0.7–2.9
Sacral (S1–S5)	0 (0)	4 (1.3)	1.0[c]	1.0	1.0–1.0
Lower half of the spine (T8–S5)	33 (100)	247 (82.9)	4.4×10^{-3} [c]	1.2	1.1–1.3
Rib defect					
Incidence, No. (%)	26 (78.8)	194 (65.1)	0.1[c]	2.0	0.8–4.7
Number of rib defects, median (IQR)	1 (1–1.5)	2 (0–3)	0.1[b]		
Rib missing	1 (0–1)	0 (0–1)	4.9×10^{-2} [b]		
Fused rib	0 (0–0)	0 (0–2)	2.6×10^{-3} [b]		
Type[d], No. (%)					
Simple	25 (75.8)	88 (29.5)	3.1×10^{-7} [c]	7.5	3.2–17.2
Complex	1 (3.0)	106 (35.6)	4.2×10^{-5} [c]	0.06	0.01–0.4
Intraspinal defect					

Clinical characteristics	TACS (n=33)	Non-TACS (n=298)	*P* value[a]	OR	95% CI
Incidence, No. (%)	1 (3.0)	131 (44.0)	7.0×10^{-7} [c]	0.04	0.05–0.3

Abbreviation: CS, congenital scoliosis; TACS, *TBX6*-associated CS; Non-TACS, non-*TBX6*-associated CS; CI, confidence interval; OR, odds ratio; IQR, interquartile range; No., number; C, cervical vertebra; T, thoracic vertebra; L, lumbar vertebra; S, sacral vertebra.

[a] $P<0.05$ is considered significant.

[b] Mann-Whitney U test.

[c] Pearson's chi-square test or Fisher's exact test.

[d] The rib anomalies were divided into simple or complex types following the criteria described by Tsirikos and McMaster.[17]