




Article

Whole Exome Sequencing of 23 Multigeneration Idiopathic Scoliosis Families Reveals Enrichments in Cytoskeletal Variants, Suggests Highly Polygenic Disease

Elizabeth A. Terhune ¹, Cambria I. Wethey ¹, Melissa T. Cuevas ¹, Anna M. Monley ^{1,2} , Erin E. Baschal ¹ , Morgan R. Bland ¹, Robin Baschal ^{1,2}, G. Devon Trahan ³, Matthew R. G. Taylor ⁴ , Kenneth L. Jones ^{3,5} and Nancy Hadley Miller ^{1,2,*}

¹ Department of Orthopedics, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA; elizabeth.a.terhune@cuanschutz.edu (E.A.T.); cambria.wethey@cuanschutz.edu (C.I.W.); Melissa.cuevas@cuanschutz.edu (M.T.C.); anna.monley@cuanschutz.edu (A.M.M.); erin.baschal@cuanschutz.edu (E.E.B.); morganrbland@gmail.com (M.R.B.); robin.baschal@cuanschutz.edu (R.B.)

² Musculoskeletal Research Center, Children's Hospital Colorado, Aurora, CO 80045, USA

³ Department of Pediatrics, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA; devon.trahan@cuanschutz.edu (G.D.T.); ken.jones@ouhsc.edu (K.L.J.)

⁴ Department of Medicine, Adult Medical Genetics Program, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA; matthew.taylor@cuanschutz.edu

⁵ Department of Cell Biology, University of Oklahoma Health Science Center, Oklahoma City, OK 73104, USA

* Correspondence: Nancy.hadley-miller@cuanschutz.edu



Citation: Terhune, E.A.; Wethey, C.I.; Cuevas, M.T.; Monley, A.M.; Baschal, E.E.; Bland, M.R.; Baschal, R.; Trahan, G.D.; Taylor, M.R.G.; Jones, K.L.; et al. Whole Exome Sequencing of 23 Multigeneration Idiopathic Scoliosis Families Reveals Enrichments in Cytoskeletal Variants, Suggests Highly Polygenic Disease. *Genes* **2021**, *12*, 922. <https://doi.org/10.3390/genes12060922>

Academic Editor: Cathy L. Raggio

Received: 18 May 2021

Accepted: 10 June 2021

Published: 16 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Adolescent idiopathic scoliosis (AIS) is a lateral spinal curvature $>10^\circ$ with rotation that affects 2–3% of healthy children across populations. AIS is known to have a significant genetic component, and despite a handful of risk loci identified in unrelated individuals by GWAS and next-generation sequencing methods, the underlying etiology of the condition remains largely unknown. In this study, we performed exome sequencing of affected individuals within 23 multigenerational families, with the hypothesis that the occurrence of rare, low frequency, disease-causing variants will co-occur in distantly related, affected individuals. Bioinformatic filtering of uncommon, potentially damaging variants shared by all sequenced family members revealed 1448 variants in 1160 genes across the 23 families, with 132 genes shared by two or more families. Ten genes were shared by >4 families, and no genes were shared by all. Gene enrichment analysis showed an enrichment of variants in cytoskeletal and extracellular matrix related processes. These data support a model that AIS is a highly polygenic disease, with few variant-containing genes shared between affected individuals across different family lineages. This work presents a novel resource for further exploration in familial AIS genetic research.

Keywords: idiopathic scoliosis; exome sequencing; spine; polygenic; variants; musculoskeletal disease; cytoskeleton; extracellular matrix

1. Introduction

Adolescent idiopathic scoliosis (AIS) is a structural lateral spinal curvature $\geq 10^\circ$ that affects 2–3% of healthy children [1], with females at the greatest risk for severe progression [2–4]. In individuals with severe progressive curvatures, life-long problems of cosmetic deformity, respiratory compromise, back pain, and degenerative disease often arise and in many cases require surgical intervention, placing a significant economic burden upon our healthcare system [5]. AIS is known to be highly heritable, however, our knowledge of its etiology is severely limited, both in terms of the individuals at risk for curve initiation and those likely to experience curve progression. Understanding key risk variants or genetic pathways leading to AIS holds the potential to improve patient care by way of targeted clinical treatments or prognostics for detecting AIS risk or risk of curvature progression.

Multiple studies support the genetic foundation of AIS, with sibling recurrence risks reported to be near 18%, and heritability estimates of approximately 87.5% [6–8]. To understand the genetics of AIS, traditional approaches including genome wide association studies (GWAS), exome sequencing, and familial linkage studies have been applied to multiple populations. GWAS of unrelated individuals with AIS have identified potential common risk alleles, notably those in or near *LBX1* [9–23] and *GPR126/ADGRG6* [23–27], which are the most well-replicated genetic findings to date across populations. Next-generation sequencing studies (i.e., whole exome sequencing) both within families and unrelated individuals with AIS have identified *rare* variants in extracellular matrix genes that may contribute to the AIS phenotype [28–30]. Recent genetic and functional studies have led to varied hypotheses of AIS etiology, including dysfunction within neuroinflammatory pathways [31,32], the cartilage matrisome [33], cilia, the cytoskeleton [34–38], or the vestibular system [39–43]. However, the inability thus far to relate specific genetic variants to the biology of AIS and the wide variation in which AIS presents are indicative of the complex heterogeneity of this disorder.

Disease susceptibility variants for complex diseases may collectively be common in the general population, but specific variants may be rare within families or affected individuals. Family studies present unique advantages over case-control studies, as they may reveal rare disease-associated variants enriched within the family that are amenable to targeted sequencing. Sequencing studies of multiple large families can reveal multiple variants within a gene or molecular pathway that contribute to the disease phenotype [44,45].

Within our laboratory, a pilot exome sequencing study of five AIS families identified an enrichment of damaging uncommon variants in cilia, extracellular matrix, and cytoskeletal genes, although no specific variants or genes were found to be present across all families [46].

This led us to propose the hypothesis that the development of AIS may be due to damaging variants within a specific set of pathways or molecular classes, rather than being driven by just a few select ‘AIS genes’. In this study, we expand upon our previous work and present exome sequencing of affected individuals from 23 AIS families, interpreted using gene enrichment analyses to identify overrepresented functional categories. We then investigate specific genes containing variants in multiple families via genotyping confirmation in additional affected and unaffected members of the family to assess how closely these variants track with the AIS phenotype.

2. Materials and Methods

An overview of the methodology for this study—including subject enrollment, sample collection, sequencing, and bioinformatic filtering strategies—are provided in Figure 1.

2.1. Subjects

Study subjects were enrolled as previously described [29,38,46]. Inclusion in the sequencing pool required a standing anteroposterior spinal radiograph showing $\geq 10^\circ$ curvature by the Cobb method with pedicle rotation, and no evidence of congenital deformity or other co-existing genetic disorders [47–49].

Blood samples were collected from all study subjects as described previously [38]. DNA was then extracted from whole blood using standard phenol chloroform protocols or the QIA-GEN Genra PureGene Blood Kit. DNA quality was verified by Qubit (Invitrogen, Waltham, MA, USA), agarose gel electrophoresis and Nanodrop (Thermo Scientific, Waltham, MA, USA).

2.2. Family Selection for Exome Sequencing

Twenty-three large families were selected through a tiered process wherein each pedigree was reviewed and evaluated by five project experts. Selection was based on the number of affected individuals in the family, the severity of their scoliosis curvatures, and the estimated genetic relationship between enrolled, affected individuals. Families with other musculoskeletal conditions or AIS from multiple sides of the family were excluded.

Three to five affected individuals were selected for sequencing per family based on the degree of hypothesized genetic distance between them, availability of high-quality DNA, and severity of spinal curvature.

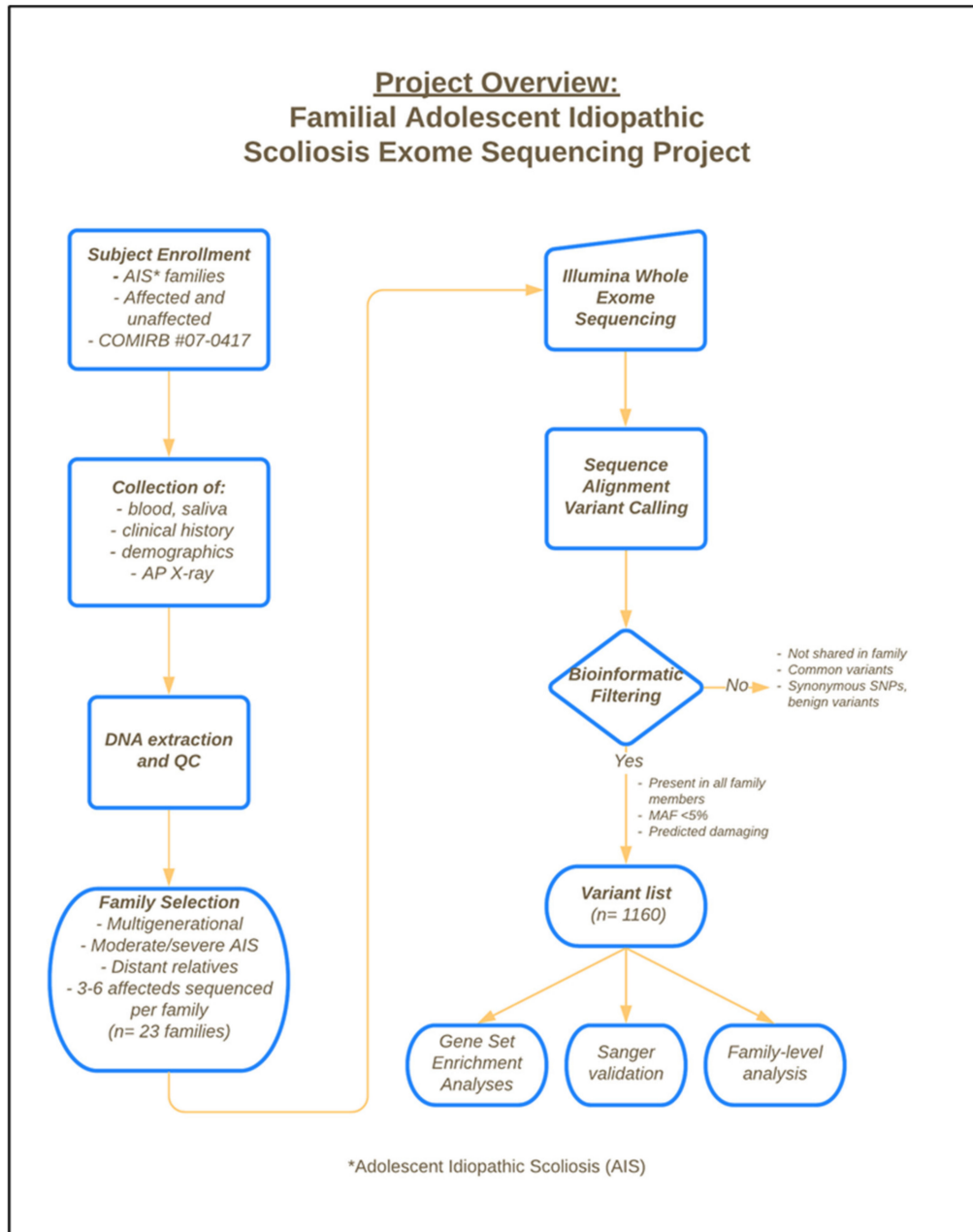


Figure 1. Project overview of subject enrollment, sample collection and extraction, exome sequencing, and bioinformatic filtering strategy.

Pedigrees for these families and a summary of clinical information for all sequenced individuals are provided in Supplementary Files S1 and S2. On the pedigrees, degree of spinal curvature as measured by Cobb angle is indicated in numbers below affected individuals (i.e., 20D). Individuals with known double curves have both listed (i.e., 26/30D), as are any triple curvatures. All individuals with listed curvatures were examined radiographically, and any found to be unaffected on physical exam are labeled as “confirmed negative”. Proband for each family is indicated with an arrow.

2.3. Whole Exome Sequencing

Exome capture was completed using 1 µg of genomic DNA from 86 individuals across 23 families using the Agilent SureSelect Human V5 (51 Mb) exon capture kit. Samples were sequenced with a 2 × 100 bp run on an Illumina HiSeq 2500 at the Otogenetics Corporation facility in Atlanta, GA, with a minimum average coverage of 50X guaranteed per sample.

2.4. Bioinformatic Filtering

Whole exome reads were aligned to GRCh38 and variants were identified with FreeBayes as previously described [38,46]. Candidate variants were filtered by SnpEff (version 4.1g) [50] along with custom scripts to retain only non-synonymous SNPs, coding indels, and variants affecting splice sites. Known artifacts and variants whose frequency was greater than 0.05 in the ExAC database (r0.3) [51] were also stripped. If the variant was annotated in the dbNSFP database (version 3.0) [52,53], it was retained only if at least one of the included variant prediction algorithms (SIFT, Polyphen2, LRT, MutationTaster) scored it as “damaging”, signifying that the resulting change to the encoded protein had a predicted functional consequence. Variants that were not shared by all sequenced members of the family were not retained. Variants with a Minor Allele Frequency (MAF) < 0.05 that remained after the above filters were applied were retained for further analysis in separate gene sets. This MAF threshold was intentionally set higher than typical rare variant thresholds to account for the prevalence of the disease (2–3% of the general population), working under the hypothesis that low frequency variants may contribute to the high prevalence of this disease.

2.5. Genotyping

Variants appearing in multiple families that were present in the GO functional categories for “cytoskeleton” or “extracellular matrix” (or related terms) were prioritized for genotyping. Additional enrolled affected and unaffected family members were sequenced at the variant site by the Sanger method to establish whether the variant segregated with the AIS phenotype.

PCR was conducted in 20 µL reactions containing 10 µL Premix D (Epicentre Biotechnologies, Madison, WI, USA), 0.2 µL Taq Polymerase (Sigma, St. Louis, MO, USA), 60 ng genomic DNA, and 10 µM Forward and Reverse Primers. PCR reactions were run on a SimpliAmp Thermocycler (Fisher Scientific, Waltham, MA, USA) with a touchdown PCR protocol [38]. Primer sequences were obtained from Integrated DNA Technologies and are provided in Supplementary Files S2. Sanger sequencing was performed by Quintara Biosciences and chromatograms were analyzed using the CodonCode Aligner v9.0 (CodonCode Corporation, Centerville, MA, USA, <https://www.codoncode.com/index.htm> (accessed on 16 June 2021)).

Pedigrees for each sequenced family are provided in Supplementary Files S1 and were created using PedigreeXP software (PC Pal, <https://www.pedigreeexp.com> (accessed on 16 June 2021)).

2.6. Gene Set Overrepresentation Analyses

Gene set overrepresentation analyses (GOA) were performed on MAF < 0.05 gene lists with duplicate genes removed ($n = 1160$) from the combined list from each family. Both DAVID and EnrichR websites were used, as described below.

DAVID: DAVID (Database for Annotation, Visualization and Integrated Discovery) v6.8 was used to identify the significant GO terms and clusters in each gene list (<https://david.ncifcrf.gov>, accessed on 20 April 2021) [54,55]. The input gene list of 1160 genes resulted in 1146 DAVID IDs. Functional annotation clustering was used on our dataset with default settings and the GOTERM_All and GOTERM_CC_Direct annotation categories.

Enrichr: The same input gene lists as used for DAVID were used in Enrichr, a gene enrichment software developed by the Ma’ayan laboratory [56,57] (<https://maayanlab.cloud/Enrichr/>, accessed on 27 April 2021). Additionally, family-specific gene lists were

separately inputted into EnrichR. We report results from the 2018 GO term Cellular Component and KEGG pathways. Volcano plots and charts were generated with Appyter (<https://appyters.maayanlab.cloud/#/>, accessed on 27 April 2021).

3. Results

To identify rare and low frequency variants associated with familial AIS, we performed whole exome sequencing on 23 multigenerational IS families (3–5 individuals per family, 86 individuals in total). Pedigrees for all families are provided in Supplementary Files S1, and clinical information for sequenced individuals is provided in Supplementary Files S2. Whole exome sequencing was performed using DNA extracted from whole blood, as described in the Methods. Illumina HiSeq reads were mapped to the human reference genome (hg39), and a minimum of 50X average coverage was obtained for each sample.

We then filtered the list of variants, requiring that each retained variant was present in all sequenced members of the family, was predicted to be damaging, and had an ExAC minor allele frequency (MAF) of <0.05. This MAF filter thus included both low-frequency (MAF 1–5%) and rare variants (MAF < 1%). These filters resulted in a total of 1448 variants in 1160 genes across the 23 families, with 11 to 128 variants identified in each family (median = 51 variants). Figure 2 provides a summary of variant information identified across all families. Nonsynonymous single nucleotide polymorphisms (SNPs) constituted the majority of filtered variants (88%, $n = 1281$), followed by non-frameshift deletions (4%, $n = 52$), non-frameshift insertions (2%, $n = 34$), frameshift deletions (2%, $n = 25$), stop gains (2%, $n = 22$), frameshift insertions (2%, $n = 22$), and splice sites (1%, $n = 12$). About half (48%) of variants had a minor allele frequency (MAF) of 0.01–0.05, 22% had an MAF 0.001–0.01, 18% had an MAF < 0.001, and 12% were predicted to be novel. Chromosomes with the largest number of variants were chromosome 1 ($n = 150$) and chromosome 19 ($n = 119$). Full details of each variant, sorted by family, are provided in Supplementary Files S3.

We next searched for genes containing variants in multiple families, with the hypothesis that shared genes would be more likely to contribute to the IS phenotype. Most variant-containing genes ($n = 1028$ of 1160 total, 88.6%) were specific to only one family. 132 genes were shared by at least 2 families, 38 were shared between 3+ families, and 10 were shared by 4+ families. No genes were shared among all families. Table 1 provides a list of genes found within multiple families. Of this list, four genes had a previously observed association with either idiopathic, degenerative, or infantile scoliosis (*DLL3* [58,59], *AHNAK* [60], *TTN* [61], *ANKRD11* [62]).

Although few genes were shared across families, we hypothesized that AIS families may share an enrichment of variants in specific functional categories. To identify categories of genes which were overrepresented with damaging variants within the AIS families, we performed gene set overrepresentation analysis (GOA) of our resulting filtered gene lists. We entered the MAF < 0.05 gene lists from individual families and combined data from all families into DAVID to identify overrepresented gene ontologies (GO Terms) (Table 2). The most overrepresented GO Cell Component categories across all families using DAVID were “microtubule” ($p = 1.62 \times 10^4$, 1.98 fold enrichment), “slit diaphragm” ($p = 8.88 \times 10^4$, 10.66 fold enrichment), and “intermediate filament” ($p = 8.88 \times 10^4$, 2.57 fold enrichment). Categories within extracellular matrix (ECM) GO terms, including “proteinaceous extracellular matrix” and “collagen trimer” also showed mild enrichments (Table 2) and the top enriched KEGG term was “ECM receptor activation” (Figure 2E). A full list of all GO functional categories is provided in Supplementary Files S2. We also used GOA to analyze an overrepresentation of cytogenic bands, as specific chromosomal regions have been linked with IS development by linkage analyses [63–71]. The most overrepresented band was 16p13.3 ($p < 0.001$), as described in Supplementary Files S2. This region was previously identified as significant in familial linkage analyses of a large AIS familial cohort. A neighboring region, distal chromosome 16p11.2 duplication, has been recently identified as a significant risk factor for severe AIS.

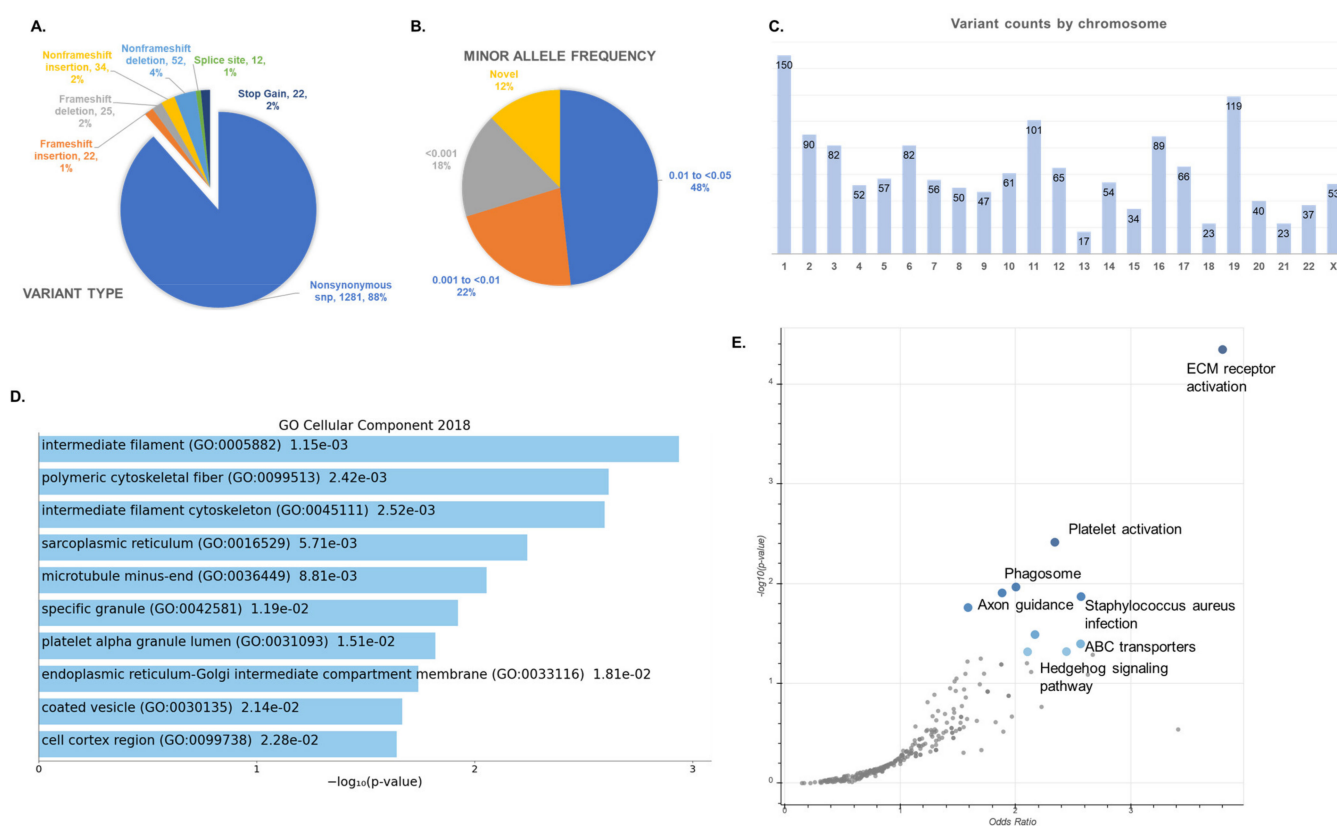


Figure 2. Summary information for total gene list ($n = 1160$) of familial AIS-associated variants across families with a minor allele frequency (MAF) < 0.05 passing bioinformatic filtering, as described in the Methods. (A) Variant type with n variants and % of total provided. (B) Minor allele frequency of all variants using ExAC. (C) Variant counts by chromosome. (D) Top GO Cellular Component, 2018 terms using EnrichR. See Table 2 for top GO terms using DAVID. (E) Volcano plot of top enriched KEGG terms over expected.

Table 1. Genes containing variants in multiple idiopathic scoliosis (IS) families. Genes with the same variant identified across families are indicated with an *. Citations are given for genes previously associated with scoliosis (including idiopathic, degenerative, or infantile). Full variant results by family are provided in Supplementary Files S3.

N Families Containing Variant	Gene Names
2 families	<i>TMEM52, GCFC2, RTP5, SLC10A6, SAMD9L, PTPRD, ASCL1, GALC, CD34, ALPPL2, FAM189B, NTRK1, CHIT1, DCBLD2, COL6A5, ATG9B, NUTM2F, NME3, GFAP, KRTAP10-5, ZNF644, FLG, OR6P1, CSMD1, UQCRB, LMNTD2, THAP11, MLLT1, OR14A2, VWA3B, CFC1B, MSH3, AKAP3, GNB1L, RAB36, FGD1, THOC3, TNXB, AHNAK2, XYLT1, APOL3, CACFD1, RPL3L, PIGT, CTNNA3, LRIT1, TYMP, ASXL2, GPAT2, AARD, MGA, LGALS9C, C18orf65, AHCY, CSMD2, VPS41, BHLHE22, CCDC68, SYNJ1, HMGCS2, SYN2, DNASE1L3, CASP12, FSIP2, TET2, FAM153B, FAM153A, PFAS, PDZD2, KLHL32, ANKRD18A, MCM8, NPHP4, USP32, COL21A1, SSTR5, CEACAM21, FAM26F, OR4A5, KRT81, BDNF, SLC1A7, SPTA1, TNK2, PTPRE, CYB5R2, MTUS2, FANCA, KATNAL2, OTC, GPIHBP1, CLEC18B, SLC22A31, SETDB1</i>
3 families	<i>GPRIN1, SKIV2L, PRR23D1, TBC1D26, KIR3DL1, TTN [61], KRTAP4-3, DLL3 [58,59], OBSCN, KCNQ5, CCDC168, PRR25, ANKRD11 [62], HRC, GPR179, USP26, FCGBP, KCNN3, CACNA1H, ANKRD30B, KIAA1875, MRC1, ACOT4, KIAA0556, ERCC6L, DYSE, CEP170, AHNAK [60]</i>
4 families	<i>POM121, PKD1L2, TMPRSS13, KRT2, ZNF717, PDE4DIP</i>
≥ 5 families	<i>AL589743.1, TPRX1, FAM47A, WIPF3</i>

Table 2. Overrepresented Gene Ontology, Cellular Component terms using the combined Minor Allele Frequency (MAF) < 0.05 gene list from all families (input $n = 1160$). The genes under the terms “cytoplasm” and “plasma membrane” are limited to 50 for brevity. The full list of enriched GO Terms (all GO annotation sets) is provided in Supplementary Files S2.

Term	Count	%	<i>p</i> Value	Genes	Fold Enrichment
microtubule	36	3.14	1.62×10^{-4}	<i>INVS, DNAH1, DNAH7, TUBAL3, DCTN1, DNAH6, IQGAP1, CAMSAP2, GOLGA2, TEK1, DVL1, KIF13B, TEK2, CEP170, KIF21B, FSD1, KIF1A, PCNT, DYNC2H1, DNAH14, MTUS2, KCNAB2, SHROOM1, HAU5, EML2, DLG1, SYNJ1, KATNAL2, INCENP, KIF26A, KIFC1, FEZ1, EHHADH, TTLL11, EIF3A, GAS8</i>	1.98
slit diaphragm	5	0.44	6.761×10^{-4}	<i>TRPC6, KIRREL2, NPHS1, MAGI2, IQGAP1</i>	10.66
intermediate filament	17	1.48	8.88×10^{-4}	<i>FLG, DSP, KRTAP13-4, KRT2, KRTAP26-1, KRTAP27-1, KRT79, KRT10, GFAP, KRT28, KRT37, PKP2, KRTAP6-1, SYNC, NES, KRT6A, PRPH</i>	2.57
Z disc	17	1.48	1.42×10^{-3}	<i>PPP1R12A, SYNPO2, AHNAK2, ATP2B4, SLC4A1, NEB, ANK3, ADRA1A, IGFN1, RYR3, MYPN, TTN, OBSCN, HRC, PDE4B, SYNC, CRYAB</i>	2.46
spindle	17	1.48	1.85×10^{-3}	<i>INVS, DIDO1, SPAG8, NUMA1, DCTN1, DCTN3, HEPACAM2, HAU5, EML2, INCENP, KIFC1, NUP85, CLTCL1, ANKRD53, CEP170, E4F1, KBTBD8</i>	2.40
cytoplasm	347	30.28	3.04×10^{-3}	<i>*RPL5, ANKLE1, MTRR, HDAC10, RGSL1, ABCA12, BACH2, WDR87, ENDOV, TBK1, CYP2D7, MPRIP, PPP4R2, C7ORF31, FAM65C, PLCE1, TRIM26, RTTN, KIF21B, ADGB, SDS, BSX, KRT2, MTUS2, AFAP1, SCRIB, LCE1E, MAPK8IP2, EML2, PPP1R3G, INPP4B, KATNAL2, TEPI, FEZ1, ZSCAN26, WDPCP, TTLL11, FLYWCH1, ALPK2, PFN1, RIN1, TRIB2, KPRP, SPTBN5, PARPBP, DHX8, PCDH15, C1ORF198, AGAP1, XPC</i>	1.13
lateral plasma membrane	10	0.87	3.37×10^{-3}	<i>CEACAM1, DLG1, KCNBI, MTCL1, ABCC6, PTPRO, DVL1, ANK3, IQGAP1, NKD2</i>	3.22
cytoskeleton	36	3.14	3.73×10^{-3}	<i>TENM1, DRC7, MTCL1, EVPL, PPL, GPHN, FGD1, SYNE1, KIAA0556, CNN2, AKAP12, EPB41L5, SGCD, EPB41L2, TNKS1BP1, PLEK2, NPHP4, CTNNA3, TRIM67, DSP, VASP, FARP2, AFAP1, FRMD4A, KCNAB2, PTPN13, ARHGAP24, APBB1IP, UBXN11, FRMD7, TRIP10, FILIP1, NF2, RIN1, PFN1, TRIB2</i>	1.66
neuron projection	25	2.18	6.37×10^{-3}	<i>TENM1, GPI, TENM2, TENM3, AHCY, TENM4, PTPRO, IQGAP1, STON2, PARK2, DVL1, CADM1, DTNBP1, ATP2B4, ANK3, PTPN13, SSTR5, SYNJ1, FRMD7, RAPGEF2, FAS, ATP13A2, NF2, PFN1, CPEB2</i>	1.80
centrosome	38	3.32	1.08×10^{-2}	<i>ERCC6L, NUMA1, DCTN1, VPS4B, DCTN3, HEPACAM2, CDC14A, RPGR, TTC28, CAMSAP2, PPP4R2, TMEM67, C7ORF31, CHEK1, PDE4B, CEP170, NPHP4, RTTN, NEK3, CCDC141, PCNT, ZNF322, DCAF13, LRRCC1, PPP1R12A, CEP135, DYX1C1, CCDC116, IFT140, CEP131, KLHL21, MTUS2, PDE4DIP, HAU5, ALS2, CROCC, FEZ1, ALMS1</i>	1.52

Table 2. Cont.

Term	Count	%	<i>p</i> Value	Genes	Fold Enrichment
axon	23	2.01	1.11×10^2	<i>NTRK1, EPHA5, CNTNAP2, TENM3, KCNB1, PTPRO, DAB2IP, DTNBP1, IQGAP1, IGSF9, ROBO1, SPTA1, ALCAM, ALS2, FEZ1, DVL1, KIF13B, CHRNA10, KIF21B, NEK3, LDLRAP1, CRYAB, VPS16</i>	1.77
spectrin	4	0.35	1.29×10^2	<i>SPTA1, SPTBN5, EPB41L2, SPTBN2</i>	7.58
sarcoplasmic reticulum	7	0.61	1.30×10^2	<i>ATP2A3, ITPR1, CLEC18B, MRV11, ANK3, XDH, RYR3</i>	3.51
plasma membrane	272	23.73	1.48×10^2	<i>*RGSL1, SLC4A1, ABCA12, ABRA, SLC4A5, PIEZO1, PLCE1, GPR179, DYNC2H1, EPHA5, IL15RA, UNC5A, OR1J4, KCNK13, CACNA2D2, SCRIB, ANK3, SYTL5, PHKA2, HSPG2, SYTL3, KEL, FEZ1, OR6C6, KCNQ5, WDPCP, RIN1, EPHA1, OR5AK2, CFB, TAS2R42, SLC22A1, PCDH15, MGST2, C2CD4A, OR10T2, XPC, PCDH12, IQGAP1, MST1R, PPL, IGSF9, OR52L1, ART1, EPB41L5, NCSTN, EPB41L2, PLCG2, KCNN3, APOB</i>	1.13
proteinaceous extracellular matrix	25	2.18	2.55×10^2	<i>FBN2, TNXB, LAMC3, ADAMTS12, ADAMTS10, ADAMTS14, GPC1, SLIT2, MUC4, IMPG1, TECTB, AMBN, FN1, MMP10, MMP16, CILP, COL4A3, COL4A6, COL8A2, COL21A1, COL9A3, COL6A5, COL9A2, MATN3, MATN2</i>	1.59
adherens junction	8	0.70	2.56×10^2	<i>EPHA5, EPB41L5, CEACAM1, TNK2, PKP2, FMN1, CTNNA3, NF2</i>	2.73
sarcolemma	11	0.96	2.65×10^2	<i>DLG1, SGCD, AHNAK, KCNB1, AHNAK2, DYSE, DTNBP1, ANK3, SYNC, SLC8B1, RYR3</i>	2.21
presynapse	9	0.79	3.25×10^2	<i>SYT3, SYNJ1, DVL1, SYT15, SCRIB, STON2, SYTL5, SYTL3, PARK2</i>	2.40
apical plasma membrane	26	2.27	3.54×10^2	<i>MTCL1, PTPRO, SIPA1L3, SLC4A5, PARD6B, DSTYK, SLCO2B1, CD34, DUOX2, GPIHBP1, SPTBN2, SLC22A11, MUC17, SLC10A2, ABCC6, AKR1A1, FN1, NRG1, ABCA7, CEACAM1, CDHR2, RAPGEF2, CHRFA7A, WDPCP, SLC26A4, KCNK1</i>	1.52
collagen trimer	11	0.96	4.25×10^2	<i>MSR1, COL27A1, COL7A1, COL4A3, COL8A2, SAAL1, COL4A6, COL21A1, COL9A2, COL6A5, OTOL1</i>	2.04

Next, we performed GOA on each family individually, to determine whether the cytoskeletal functional categories seen in the combined gene list were driven by a single family or many families. Overall, a cytoskeletal GO term was detected among the significantly enriched terms in every family except Family H, J, and O (Supplementary Files S2). Every family possessed at least one variant in a cytoskeletal gene.

We next selected variants of interest for genotyping to determine whether the variant segregated with the AIS phenotype within each relevant family. Specific variants were genotyped if they were present in functional categories of interest (cytoskeletal or extracellular matrix GO terms), and if additional affected and unaffected members of the family were enrolled in our study. Ten genes (*ANKRD11, COL21A1, COL6A5, FGD1, NPHP4, OBSCN, TNXB, CTNNA3, NTRK1, and PDE4DIP*) containing multiple variants across families were genotyped. Of these, variants within *TNXB, CTNNA3, NTRK1, and PDE4DIP* showed segregation of the variant with the IS phenotype in additional affected

and unaffected family members (Supplementary Files S2). *TNXB* (tenascin XB) localizes to the major histocompatibility region on chromosome 6 and encodes an ECM glycoprotein and has been associated to classic Ehlers-Danlos syndrome (EDS), a connective tissue disorder with scoliosis as one characteristic of the phenotype [72]. The *TNXB* variants within our dataset (Family G: *TNXB*:NM_019105:exon24:c.C8192G:p.P2731R; Family S: *TNXB*:NM_019105:exon12:c.G4444A: p.V1482M) do not appear in the EDS variant database to date [73–75]. *CTNNA3* (Catenin α 3) encodes a protein within the vinculin family of cell-cell adhesion proteins, and has been linked with certain types of cardiomyopathy [76]. *NTRK1* (Neurotrophic Receptor Tyrosine Kinase 1) encodes a protein that binds neurotrophin peptide and participates within neuronal cell differentiation and specification of neuronal cell subtypes [77]. *PDE4DIP* encodes myomegalin, which anchors phosphodiesterase 4D to the Golgi/centrosome regions. An isoform of myomegalin was recently shown to form a complex with AKAP9 and CDK5RAP to link the pericentrosomal complex to the microtubule-nucleating complex [78]. The remaining variants showed no segregation or only partial segregation of the variant with the AIS phenotype.

4. Discussion

In this study, we report an enrichment of predicted damaging variants in cytoskeletal and extracellular matrix (ECM) functional categories within adolescent idiopathic scoliosis (AIS) families. Additionally, this cohort showed minimal overlap in specific genes across families. Our results support and add to the growing evidence that AIS is a highly polygenic disorder [7,8,28,79,80] in which multiple variants of variable effect size, potentially in combination with epigenetic and environmental phenomenon [81], contribute to the disease phenotype.

The ECM is a dynamic molecular network of proteoglycans, glycoproteins, minerals, and related proteins that plays a critical role within musculoskeletal tissues [82]. The ECM provides critical structural networks and scaffolding to tissues, and contributes to cellular signaling, growth, and repair [83]. Studies of genes encoding the “matrisome”, the proteins composing and relating to the ECM, have identified dozens of causal ECM mutations for connective tissue and musculoskeletal disorders [83]. A polygenic burden of rare variants in musculoskeletal collagen genes was linked to AIS risk in a large case-control cohort [28], and our group previously showed mild variant enrichments in our pilot exome sequencing study of five AIS families [46]. Candidate ECM genes including *HSPG2* [29] and *FBN2* [30] have also been linked to AIS in specific cohorts. Several collagen genes (*COL8A2*, *COL4A3*, *COL6A5*, *COL27A1*, *COL7A1*, *COL21A1*, *COL9A2*, *COL9A3*, *COL4A6*) and perlecan (*HSPG2*) appeared in our families (Supplementary Files S3). Candidate studies of ECM genes in AIS [63,84–86] were launched, in part, because scoliosis is a common phenotype of monogenic connective tissue disorders including Marfan and Ehlers-Danlos syndromes (EDS) [72,74]. *TNXB* (tenascin XB) variants, which have been associated with classic EDS, were found in two families in our current cohort as well as one family in our previous study [46]. The *TNXB* variant within Family G (*TNXB*:NM_019105:exon24:c.C8192G:p.P2731R) appeared in the Leiden Open Variation Database for EDS but was predicted as “benign” (<https://databases.lovd.nl/shared/variants/0000313458#00021614>, accessed on 4 May 2021). Functional data will be required to determine the pathogenicity of these variants. In a recent review, Wise et al. proposed that genetic variants affecting the cartilage matrisome, specifically of the intravertebral disc, may be involved in a subset of AIS cases [33]. Our results provide support to polygenic variants in ECM genes as a contributing factor to AIS in some family lineages.

The cytoskeleton comprises actin, microtubule, and intermediate filament networks and is responsible for a multitude of functions including molecular transport, cellular stability, molecular signaling, cell migration, and cell division [87]. The cytoskeleton and ECM are intricately connected within musculoskeletal tissues and play important roles in mechanotransduction, tissue stability and response to biomechanical loading [88,89]. Every family within our dataset had at least one cytoskeletal genetic variant, and 20/23

families showed an enriched cytoskeletal GO term (Supplementary Files S2). Variants in the centriolar protein *POC5* have been associated with scoliosis through human and animal studies [90–92]. The cytoskeletal kinesin *kif6* was shown to be necessary for proper spine development in zebrafish [34], and mutations in *kif6* also appeared in a zebrafish ENU genetic screen for scoliosis [93].

Defects in cilia, microtubule-based projections critical for cell signaling and fluid flow, have been linked to scoliosis in animal models [31,34,35,37,94,95]. A variant in the ciliary kinesin *KIF7* was found within one AIS family and specific *KIF7* mutations produced scoliosis in zebrafish [38]. Additionally, an overrepresentation of cilia variants was observed in our pilot exome sequencing cohort [46]. However, functional roles of the cytoskeleton and ECM in the development of IS have not yet been demonstrated in humans.

This study contains several limitations. Our familial cohort is relatively small and our ability to understand the impact of rare or low frequency variants in relation to the expression of a complex genetic disease, such as AIS, is limited [96]. Low frequency variants frequently lie outside the scope of large statistical association studies, and therefore may contribute to the “missing heritability” that accompanies many complex traits [97]. The majority of our discovered variants are heterozygous, and we do not know without functional testing whether these variants have a true dominant-negative effect on the resulting protein. To confirm our observed functional pathway enrichments, validate the presence of specific genetic variants, and obtain adequate statistical power, a larger cohort of AIS families must be sequenced. Additional functional studies of specific variants, with particular focus on those most likely to be damaging (i.e., stopgain, frameshift mutations) will be required to demonstrate causality [93].

Collectively, our results provide support to the hypothesis of ECM and cytoskeletal involvement in AIS etiology through what is, to our knowledge, the largest sequencing study of AIS families to date. These results suggest that there are many specific genes that can collectively increase disease risk, although there may be affected pathways that are shared across families. We hypothesize that (1) individual AIS families harbor low frequency mutations in different functional categories, resulting in different subsets of AIS, and/or (2) individuals with AIS require mutations in multiple gene categories (i.e., the cytoskeleton and ECM), resulting in mild dysfunction across several molecular pathways, to cause disease. Specifically, these results suggest that mild mutations in cytoskeletal or ECM genes may play a role in AIS etiology. Ultimately, this work will assist in the ability to predict the onset of adolescent idiopathic scoliosis and the risk of progressive disease and, thus, lead to the development of more personalized treatments for individuals with AIS.

5. Conclusions

This work presents a novel set of candidate variants found in affected individuals from 23 AIS families. In agreement with our previous WES study of five AIS families, we observed an overrepresentation of variants in cytoskeletal and ECM functional categories, with few specific genes shared across families. Overall, this work paints a picture of AIS genetic etiology as highly polygenic and specific to individual family lineages. An analytical approach that integrates data from family-based sequencing with genetic association studies, an understanding of study population variation, population stratification and genetic heterogeneity, and advances in clinical phenotyping will enhance our ability to define the genetic complexity of this disorder.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/genes12060922/s1>: File S1: Family Pedigrees. Degree of spinal curvature is indicated with a D (degrees). Individuals marked “confirmed positive” were confirmed as affected with AIS but were unable to provide us with an X-ray for measurement. Exome sequenced individuals are indicated. File S2: Table S1: Clinical information from all exome sequenced individuals; Table S2: Enriched cytobands from MAF < 0.05 gene list; Table S3: Genotyping of genes in multiple families and in functional categories of interest; Table S4: Full Gene Ontology (GO) Term enrichment results

($p < 0.01$) for all functional categories using DAVID; File S3: Detail of all filtered variants, sorted by family. All filtered variants had a minor allele frequency (MAF) < 0.05 , were present in all sequenced members of the family, and were predicted to be damaging to the resulting protein by at least one algorithm. Details of the exome sequencing strategy and bioinformatic filtering process are provided in the Methods.

Author Contributions: Conceptualization: N.H.M., K.L.J., M.R.G.T., E.E.B. and E.A.T.; Methodology: C.I.W., M.T.C., A.M.M., G.D.T., M.R.B., E.E.B. and R.B.; Software: G.D.T. and K.L.J.; Validation: C.I.W.; Formal analysis: K.L.J., G.D.T. and E.A.T.; Investigation: N.H.M.; Resources: N.H.M.; Data curation: N.H.M. and E.A.T.; Writing—original draft preparation: E.A.T.; Writing—review and editing: N.H.M. and A.M.M.; Visualization: E.A.T.; Supervision: N.H.M., E.E.B. and E.A.T.; Project administration: N.H.M.; Funding acquisition: N.H.M. and E.E.B. All authors have read and agreed to the published version of the manuscript.

Funding: Funding for NHM's laboratory and this study was provided by NIH/NIAMS R01AR068292. Additional funds were provided by Children's Hospital Colorado.

Institutional Review Board Statement: Written informed consent was obtained from study subjects who were enrolled in accordance with protocols approved by the Johns Hopkins School of Medicine Institutional Review Board and the University of Colorado Anschutz Medical Campus Institutional Review Board (Colorado Multiple Institutional Review Board, Studies #06-1161 and #07-0417). All procedures involving human participants were performed in accordance with the ethical standards of these institutional review boards, the 1964 Declaration of Helsinki and its later amendments, or comparable ethical standards.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The authors affirm that all data necessary for confirming the conclusions of this article are represented fully within the article and its supplementary files, including the complete lists of filtered variants for each family.

Acknowledgments: Patient data was collected and managed using REDCap (Research Electronic Data Capture) tools hosted at the University of Colorado Anschutz Medical Campus. REDCap is a secure, web-based application designed to support data capture for research studies, providing: (1) an intuitive interface for validated data entry; (2) audit trails for tracking data manipulation and export procedures; (3) automated export procedures for seamless data downloads to common statistical packages; and (4) procedures for importing data from external sources. We thank Lee Niswander and Bruce Appel for their helpful discussions regarding the overall design of this study, as well as assistance selecting families for sequencing. Above all, we thank the study subjects and their families for contributing their time, samples, and information to this research effort.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Asher, M.A.; Burton, D.C. Adolescent idiopathic scoliosis: Natural history and long term treatment effects. *Scoliosis* **2006**, *1*, 2. [[CrossRef](#)] [[PubMed](#)]
2. Kane, W.J.; Moe, J.H. A scoliosis-prevalence survey in Minnesota. *Clin. Orthop. Relat. Res.* **1970**, *69*, 216–218. [[CrossRef](#)]
3. Weinstein, S.L. Adolescent idiopathic scoliosis: Prevalence and natural history. In *The Pediatric Spine: Principles and Practice*; Weinstein, S.L., Ed.; Raven Press: New York, NY, USA, 1994; pp. 463–478.
4. Weinstein, S.L.; Zavala, D.C.; Ponseti, I.V. Idiopathic scoliosis: Long-term follow-up and prognosis in untreated patients. *J. Bone Jt. Surg. Am. Vol.* **1981**, *63*, 702–712. [[CrossRef](#)]
5. Bozzio, A.E.; Hu, X.; Lieberman, I.H. Cost and Clinical Outcome of Adolescent Idiopathic Scoliosis Surgeries-Experience From a Nonprofit Community Hospital. *Int. J. Spine Surg.* **2019**, *13*, 474–478. [[CrossRef](#)]
6. Tang, N.L.; Yeung, H.Y.; Lee, K.M.; Hung, V.W.; Cheung, C.S.; Ng, B.K.; Kwok, R.; Guo, X.; Qin, L.; Cheng, J.C. A relook into the association of the estrogen receptor [alpha] gene (PvuII, XbaI) and adolescent idiopathic scoliosis: A study of 540 Chinese cases. *Spine* **2006**, *31*, 2463–2468. [[CrossRef](#)] [[PubMed](#)]
7. Khanshour, A.M.; Wise, C.A. The Genetic Architecture of Adolescent Idiopathic Scoliosis. In *Pathogenesis of Idiopathic Scoliosis*; Machida, M., Weinstein, S.L., Dubousset, J., Eds.; Springer: Tokyo, Japan, 2018; pp. 51–74. [[CrossRef](#)]
8. Ward, K.; Ogilvie, J.; Argyle, V.; Nelson, L.; Meade, M.; Braun, J.; Chettier, R. Polygenic inheritance of adolescent idiopathic scoliosis: A study of extended families in Utah. *Am. J. Med. Genet A* **2010**, *152A*, 1178–1188. [[CrossRef](#)] [[PubMed](#)]

9. Takahashi, Y.; Kou, I.; Takahashi, A.; Johnson, T.A.; Kono, K.; Kawakami, N.; Uno, K.; Ito, M.; Minami, S.; Yanagida, H.; et al. A genome-wide association study identifies common variants near LBX1 associated with adolescent idiopathic scoliosis. *Nat. Genet.* **2011**, *43*, 1237–1240. [[CrossRef](#)] [[PubMed](#)]
10. Fan, Y.H.; Song, Y.Q.; Chan, D.; Takahashi, Y.; Ikegawa, S.; Matsumoto, M.; Kou, I.; Cheah, K.S.; Sham, P.; Cheung, K.M.; et al. SNP rs11190870 near LBX1 is associated with adolescent idiopathic scoliosis in southern Chinese. *J. Hum. Genet.* **2012**, *57*, 244–246. [[CrossRef](#)] [[PubMed](#)]
11. Jiang, H.; Qiu, X.; Dai, J.; Yan, H.; Zhu, Z.; Qian, B.; Qiu, Y. Association of rs11190870 near LBX1 with adolescent idiopathic scoliosis susceptibility in a Han Chinese population. *Eur. Spine J. Off. Publ. Eur. Spine Soc. Eur. Spinal Deform. Soc. Eur. Sect. Cerv. Spine Res. Soc.* **2013**, *22*, 282–286. [[CrossRef](#)]
12. Gao, W.; Peng, Y.; Liang, G.; Liang, A.; Ye, W.; Zhang, L.; Sharma, S.; Su, P.; Huang, D. Association between common variants near LBX1 and adolescent idiopathic scoliosis replicated in the Chinese Han population. *PLoS ONE* **2013**, *8*, e53234. [[CrossRef](#)]
13. Londono, D.; Kou, I.; Johnson, T.A.; Sharma, S.; Ogura, Y.; Tsunoda, T.; Takahashi, A.; Matsumoto, M.; Herring, J.A.; Lam, T.P.; et al. A meta-analysis identifies adolescent idiopathic scoliosis association with LBX1 locus in multiple ethnic groups. *J. Med. Genet.* **2014**, *51*, 401–406. [[CrossRef](#)]
14. Chen, S.; Zhao, L.; Roffey, D.M.; Phan, P.; Wai, E.K. Association of rs11190870 near LBX1 with adolescent idiopathic scoliosis in East Asians: A systematic review and meta-analysis. *Spine J. Off. J. N. Am. Spine Soc.* **2014**, *14*, 2968–2975. [[CrossRef](#)]
15. Liang, J.; Xing, D.; Li, Z.; Chua, S.; Li, S. Association Between rs11190870 Polymorphism Near LBX1 and Susceptibility to Adolescent Idiopathic Scoliosis in East Asian Population: A Genetic Meta-Analysis. *Spine* **2014**, *39*, 862–869. [[CrossRef](#)]
16. Zhu, Z.; Tang, N.L.; Xu, L.; Qin, X.; Mao, S.; Song, Y.; Liu, L.; Li, F.; Liu, P.; Yi, L.; et al. Genome-wide association study identifies new susceptibility loci for adolescent idiopathic scoliosis in Chinese girls. *Nat. Commun.* **2015**, *6*, 8355. [[CrossRef](#)]
17. Chettier, R.; Nelson, L.; Ogilvie, J.W.; Albertsen, H.M.; Ward, K. Haplotypes at LBX1 have distinct inheritance patterns with opposite effects in adolescent idiopathic scoliosis. *PLoS ONE* **2015**, *10*, e0117708. [[CrossRef](#)]
18. Grauers, A.; Wang, J.; Einarsdottir, E.; Simony, A.; Danielsson, A.; Akesson, K.; Ohlin, A.; Halldin, K.; Grabowski, P.; Tenne, M.; et al. Candidate gene analysis and exome sequencing confirm LBX1 as a susceptibility gene for idiopathic scoliosis. *Spine J. Off. J. N. Am. Spine Soc.* **2015**, *15*, 2239–2246. [[CrossRef](#)]
19. Liu, S.; Wu, N.; Zuo, Y.; Zhou, Y.; Liu, J.; Liu, Z.; Chen, W.; Liu, G.; Chen, Y.; Chen, J.; et al. Genetic Polymorphism of LBX1 Is Associated With Adolescent Idiopathic Scoliosis in Northern Chinese Han Population. *Spine* **2017**, *42*, 1125–1129. [[CrossRef](#)] [[PubMed](#)]
20. Nada, D.; Julien, C.; Samuels, M.E.; Moreau, A. A Replication Study for Association of LBX1 Locus With Adolescent Idiopathic Scoliosis in French-Canadian Population. *Spine* **2018**, *43*, 172–178. [[CrossRef](#)]
21. Li, Y.L.; Gao, S.J.; Xu, H.; Liu, Y.; Li, H.L.; Chen, X.Y.; Ning, G.Z.; Feng, S.Q. The association of rs11190870 near LBX1 with the susceptibility and severity of AIS, a meta-analysis. *Int. J. Surg.* **2018**, *54*, 193–200. [[CrossRef](#)] [[PubMed](#)]
22. Man, G.C.; Tang, N.L.; Chan, T.F.; Lam, T.P.; Li, J.W.; Ng, B.K.; Zhu, Z.; Qiu, Y.; Cheng, J.C. Replication Study for the Association of GWAS-associated Loci with Adolescent Idiopathic Scoliosis Susceptibility and Curve Progression in a Chinese Population. *Spine* **2018**. [[CrossRef](#)]
23. Kou, I.; Takahashi, Y.; Johnson, T.A.; Takahashi, A.; Guo, L.; Dai, J.; Qiu, X.; Sharma, S.; Takimoto, A.; Ogura, Y.; et al. Genetic variants in GPR126 are associated with adolescent idiopathic scoliosis. *Nat. Genet.* **2013**, *45*, 676–679. [[CrossRef](#)] [[PubMed](#)]
24. Xu, J.F.; Yang, G.H.; Pan, X.H.; Zhang, S.J.; Zhao, C.; Qiu, B.S.; Gu, H.F.; Hong, J.F.; Cao, L.; Chen, Y.; et al. Association of GPR126 gene polymorphism with adolescent idiopathic scoliosis in Chinese populations. *Genomics* **2015**, *105*, 101–107. [[CrossRef](#)] [[PubMed](#)]
25. Qin, X.; Xu, L.; Xia, C.; Zhu, W.; Sun, W.; Liu, Z.; Qiu, Y.; Zhu, Z. Genetic Variant of GPR126 Gene is Functionally Associated With Adolescent Idiopathic Scoliosis in Chinese Population. *Spine* **2017**, *42*, E1098–E1103. [[CrossRef](#)] [[PubMed](#)]
26. Kou, I.; Watanabe, K.; Takahashi, Y.; Momozawa, Y.; Khanshour, A.; Grauers, A.; Zhou, H.; Liu, G.; Fan, Y.H.; Takeda, K.; et al. A multi-ethnic meta-analysis confirms the association of rs6570507 with adolescent idiopathic scoliosis. *Sci. Rep.* **2018**, *8*, 11575. [[CrossRef](#)]
27. Haller, G.; Alvarado, D.; McCall, K.; Yang, P.; Cruchaga, C.; Harms, M.; Goate, A.; Willing, M.; Morcuende, J.A.; Baschal, E.; et al. A polygenic burden of rare variants across extracellular matrix genes among individuals with adolescent idiopathic scoliosis. *Hum. Mol. Genet.* **2016**, *25*, 202–209. [[CrossRef](#)]
28. Baschal, E.E.; Wethey, C.I.; Swindle, K.; Baschal, R.M.; Gowan, K.; Tang, N.L.; Alvarado, D.M.; Haller, G.E.; Dobbs, M.B.; Taylor, M.R.; et al. Exome sequencing identifies a rare HSPG2 variant associated with familial idiopathic scoliosis. *G3* **2014**, *5*, 167–174. [[CrossRef](#)]
29. Buchan, J.G.; Alvarado, D.M.; Haller, G.E.; Cruchaga, C.; Harms, M.B.; Zhang, T.; Willing, M.C.; Grange, D.K.; Braverman, A.C.; Miller, N.H.; et al. Rare variants in FBN1 and FBN2 are associated with severe adolescent idiopathic scoliosis. *Hum. Mol. Genet.* **2014**, *23*, 5271–5282. [[CrossRef](#)]
30. Rose, C.D.; Pompili, D.; Henke, K.; Van Gennip, J.L.M.; Meyer-Miner, A.; Rana, R.; Gobron, S.; Harris, M.P.; Nitz, M.; Ciruna, B. SCO-Spondin Defects and Neuroinflammation Are Conserved Mechanisms Driving Spinal Deformity across Genetic Models of Idiopathic Scoliosis. *Curr. Biol.* **2020**. [[CrossRef](#)]
31. Van Gennip, J.L.M.; Boswell, C.W.; Ciruna, B. Neuroinflammatory signals drive spinal curve formation in zebrafish models of idiopathic scoliosis. *Sci. Adv.* **2018**, *4*, eaav1781. [[CrossRef](#)]
32. Wise, C.A.; Sepich, D.; Ushiki, A.; Khanshour, A.M.; Kidane, Y.H.; Makki, N.; Gurnett, C.A.; Gray, R.S.; Rios, J.J.; Ahituv, N.; et al. The cartilage matrisome in adolescent idiopathic scoliosis. *Bone Res.* **2020**, *8*, 13. [[CrossRef](#)]

33. Buchan, J.G.; Gray, R.S.; Gansner, J.M.; Alvarado, D.M.; Burgert, L.; Gitlin, J.D.; Gurnett, C.A.; Goldsmith, M.I. Kinesin family member 6 (kif6) is necessary for spine development in zebrafish. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **2014**, *243*, 1646–1657. [[CrossRef](#)] [[PubMed](#)]
34. Grimes, D.T.; Boswell, C.W.; Morante, N.F.; Henkelman, R.M.; Burdine, R.D.; Ciruna, B. Zebrafish models of idiopathic scoliosis link cerebrospinal fluid flow defects to spine curvature. *Science* **2016**, *352*, 1341–1344. [[CrossRef](#)] [[PubMed](#)]
35. Hayes, M.; Gao, X.; Yu, L.X.; Paria, N.; Henkelman, R.M.; Wise, C.A.; Ciruna, B. ptk7 mutant zebrafish models of congenital and idiopathic scoliosis implicate dysregulated Wnt signalling in disease. *Nat. Commun.* **2014**, *5*, 4777. [[CrossRef](#)] [[PubMed](#)]
36. Konjikusic, M.J.; Yeetong, P.; Boswell, C.W.; Lee, C.; Roberson, E.C.; Ittiwut, R.; Suphapeetiporn, K.; Ciruna, B.; Gurnett, C.A.; Wallingford, J.B.; et al. Mutations in Kinesin family member 6 reveal specific role in ependymal cell ciliogenesis and human neurological development. *PLoS Genet.* **2018**, *14*, e1007817. [[CrossRef](#)] [[PubMed](#)]
37. Terhune, E.A.; Cuevas, M.T.; Monley, A.M.; Wethey, C.I.; Chen, X.; Cattel, M.V.; Bayrak, M.N.; Bland, M.R.; Sutphin, B.; Devon Trahan, G.; et al. Mutations in KIF7 Implicated in Idiopathic Scoliosis in Humans and Axial Curvatures in Zebrafish. *Hum. Mutat* **2020**. [[CrossRef](#)]
38. Carry, P.M.; Duke, V.R.; Brazell, C.J.; Stence, N.; Scholes, M.; Rousie, D.L.; Hadley Miller, N. Lateral semi-circular canal asymmetry in females with idiopathic scoliosis. *PLoS ONE* **2020**, *15*, e0232417. [[CrossRef](#)]
39. Hawasli, A.H.; Hullar, T.E.; Dorward, I.G. Idiopathic scoliosis and the vestibular system. *Eur. Spine J. Off. Publ. Eur. Spine Soc. Eur. Spinal Deform. Soc. Eur. Sect. Cerv. Spine Res. Soc.* **2015**, *24*, 227–233. [[CrossRef](#)]
40. Pialasse, J.P.; Descarreaux, M.; Mercier, P.; Blouin, J.; Simoneau, M. The Vestibular-Evoked Postural Response of Adolescents with Idiopathic Scoliosis Is Altered. *PLoS ONE* **2015**, *10*, e0143124. [[CrossRef](#)]
41. Haumont, T.; Gauchard, G.C.; Lascombes, P.; Perrin, P.P. Postural instability in early-stage idiopathic scoliosis in adolescent girls. *Spine* **2011**, *36*, E847–E854. [[CrossRef](#)]
42. Simoneau, M.; Lamothe, V.; Hutin, E.; Mercier, P.; Teasdale, N.; Blouin, J. Evidence for cognitive vestibular integration impairment in idiopathic scoliosis patients. *BMC Neurosci.* **2009**, *10*, 102. [[CrossRef](#)]
43. Mathieson, I.; McVean, G. Differential confounding of rare and common variants in spatially structured populations. *Nat. Genet.* **2012**, *44*, 243–246. [[CrossRef](#)]
44. McClellan, J.; King, M.C. Genetic heterogeneity in human disease. *Cell* **2010**, *141*, 210–217. [[CrossRef](#)]
45. Baschal, E.E.; Terhune, E.A.; Wethey, C.I.; Baschal, R.M.; Robinson, K.D.; Cuevas, M.T.; Pradhan, S.; Sutphin, B.S.; Taylor, M.R.G.; Gowan, K.; et al. Idiopathic Scoliosis Families Highlight Actin-Based and Microtubule-Based Cellular Projections and Extracellular Matrix in Disease Etiology. *G3* **2018**, *8*, 2663–2672. [[CrossRef](#)]
46. Shands, A.R.; Eisberg, H.B. The incidence of scoliosis in the state of Delaware; a study of 50,000 minifilms of the chest made during a survey for tuberculosis. *J. Bone Jt. Surg. Am. Vol.* **1955**, *37*, 1243–1249. [[CrossRef](#)]
47. Kane, W.J. Scoliosis prevalence: A call for a statement of terms. *Clin. Orthop. Relat. Res.* **1977**, *126*, 43–46. [[CrossRef](#)]
48. Armstrong, G.W.; Livermore, N.B., 3rd; Suzuki, N.; Armstrong, J.G. Nonstandard vertebral rotation in scoliosis screening patients. Its prevalence and relation to the clinical deformity. *Spine (Phila Pa 1976)* **1982**, *7*, 50–54. [[CrossRef](#)]
49. Cingolani, P.; Platts, A.; Wang le, L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [[CrossRef](#)] [[PubMed](#)]
50. Lek, M.; Karczewski, K.J.; Minikel, E.V.; Samocha, K.E.; Banks, E.; Fennell, T.; O'Donnell-Luria, A.H.; Ware, J.S.; Hill, A.J.; Cummings, B.B.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285–291. [[CrossRef](#)] [[PubMed](#)]
51. Liu, X.; Jian, X.; Boerwinkle, E. dbNSFP: A lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum. Mutat* **2011**, *32*, 894–899. [[CrossRef](#)]
52. Liu, X.; Wu, C.; Li, C.; Boerwinkle, E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum. Mutat* **2016**, *37*, 235–241. [[CrossRef](#)]
53. Huang da, W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc* **2009**, *4*, 44–57. [[CrossRef](#)] [[PubMed](#)]
54. Huang da, W.; Sherman, B.T.; Lempicki, R.A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **2009**, *37*, 1–13. [[CrossRef](#)] [[PubMed](#)]
55. Chen, E.Y.; Tan, C.M.; Kou, Y.; Duan, Q.; Wang, Z.; Meirelles, G.V.; Clark, N.R.; Ma'ayan, A. Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **2013**, *14*, 128. [[CrossRef](#)]
56. Kuleshov, M.V.; Jones, M.R.; Rouillard, A.D.; Fernandez, N.F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S.L.; Jagodnik, K.M.; Lachmann, A.; et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**, *44*, W90–W97. [[CrossRef](#)] [[PubMed](#)]
57. Giampietro, P.F.; Raggio, C.L.; Reynolds, C.; Ghebranious, N.; Burmester, J.K.; Glurich, I.; Rasmussen, K.; McPherson, E.; Pauli, R.M.; Shukla, S.K.; et al. DLL3 as a candidate gene for vertebral malformations. *Am. J. Med. Genet. A* **2006**, *140*, 2447–2453. [[CrossRef](#)]
58. Loomes, K.M.; Stevens, S.A.; O'Brien, M.L.; Gonzalez, D.M.; Ryan, M.J.; Segalov, M.; Dormans, N.J.; Mimoto, M.S.; Gibson, J.D.; Sewell, W.; et al. Dll3 and Notch1 genetic interactions model axial segmental and craniofacial malformations of human birth defects. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* **2007**, *236*, 2943–2951. [[CrossRef](#)]

59. Shen, N.; Chen, N.; Zhou, X.; Zhao, B.; Huang, R.; Liang, J.; Yang, X.; Chen, M.; Song, Y.; Du, Q. Alterations of the gut microbiome and plasma proteome in Chinese patients with adolescent idiopathic scoliosis. *Bone* **2019**, *120*, 364–370. [[CrossRef](#)] [[PubMed](#)]
60. Liu, X.Y.; Wang, L.; Yu, B.; Zhuang, Q.Y.; Wang, Y.P. Expression Signatures of Long Noncoding RNAs in Adolescent Idiopathic Scoliosis. *Biomed. Res. Int.* **2015**, *2015*, 276049. [[CrossRef](#)]
61. Shin, J.H.; Ha, K.Y.; Jung, S.H.; Chung, Y.J. Genetic predisposition in degenerative lumbar scoliosis due to the copy number variation. *Spine* **2011**, *36*, 1782–1793. [[CrossRef](#)]
62. Miller, N.H.; Mims, B.; Child, A.; Milewicz, D.M.; Sponseller, P.; Blanton, S.H. Genetic analysis of structural elastic fiber and collagen genes in familial adolescent idiopathic scoliosis. *J. Orthop. Res.* **1996**, *14*, 994–999. [[CrossRef](#)]
63. Wise, C.A.; Barnes, R.; Gillum, J.; Herring, J.A.; Bowcock, A.M.; Lovett, M. Localization of susceptibility to familial idiopathic scoliosis. *Spine* **2000**, *25*, 2372–2380. [[CrossRef](#)] [[PubMed](#)]
64. Justice, C.M.; Miller, N.H.; Marosy, B.; Zhang, J.; Wilson, A.F. Familial idiopathic scoliosis: Evidence of an X-linked susceptibility locus. *Spine* **2003**, *28*, 589–594. [[CrossRef](#)]
65. Miller, N.H.; Justice, C.M.; Marosy, B.; Doheny, K.F.; Pugh, E.; Zhang, J.; Dietz, H.C., 3rd; Wilson, A.F. Identification of candidate regions for familial idiopathic scoliosis. *Spine* **2005**, *30*, 1181–1187. [[CrossRef](#)]
66. Miller, N.H.; Marosy, B.; Justice, C.M.; Novak, S.M.; Tang, E.Y.; Boyce, P.; Pettengil, J.; Doheny, K.F.; Pugh, E.W.; Wilson, A.F. Linkage analysis of genetic loci for kyphoscoliosis on chromosomes 5p13, 13q13.3, and 13q32. *Am. J. Med. Genet. A* **2006**, *140*, 1059–1068. [[CrossRef](#)] [[PubMed](#)]
67. Ocaka, L.; Zhao, C.; Reed, J.A.; Ebenezer, N.D.; Brice, G.; Morley, T.; Mehta, M.; O'Dowd, J.; Weber, J.L.; Hardcastle, A.J.; et al. Assignment of two loci for autosomal dominant adolescent idiopathic scoliosis to chromosomes 9q31.2–q34.2 and 17q25.3–qtel. *J. Med. Genet.* **2008**, *45*, 87–92. [[CrossRef](#)]
68. Esposito, T.; Uccello, R.; Caliendo, R.; Di Martino, G.F.; Gironi Carnevale, U.A.; Cuomo, S.; Ronca, D.; Varriale, B. Estrogen receptor polymorphism, estrogen content and idiopathic scoliosis in human: A possible genetic linkage. *J. Steroid Biochem. Mol. Biol* **2009**, *116*, 56–60. [[CrossRef](#)]
69. Gurnett, C.A.; Alaee, F.; Bowcock, A.; Kruse, L.; Lenke, L.G.; Bridwell, K.H.; Kuklo, T.; Luhmann, S.J.; Dobbs, M.B. Genetic linkage localizes an adolescent idiopathic scoliosis and pectus excavatum gene to chromosome 18 q. *Spine* **2009**, *34*, E94–E100. [[CrossRef](#)] [[PubMed](#)]
70. Marosy, B.; Justice, C.M.; Vu, C.; Zorn, A.; Nzegwu, N.; Wilson, A.F.; Miller, N.H. Identification of susceptibility loci for scoliosis in FIS families with triple curves. *Am. J. Med. Genet. A* **2010**, *152A*, 846–855. [[CrossRef](#)]
71. Meester, J.A.N.; Verstraeten, A.; Schepers, D.; Alaerts, M.; Van Laer, L.; Loeys, B.L. Differences in manifestations of Marfan syndrome, Ehlers-Danlos syndrome, and Loeys-Dietz syndrome. *Ann. Cardiothorac Surg.* **2017**, *6*, 582–594. [[CrossRef](#)]
72. Lao, Q.; Mallappa, A.; Rueda Faucz, F.; Joyal, E.; Veeraraghavan, P.; Chen, W.; Merke, D.P. A TNXB splice donor site variant as a cause of hypermobility type Ehlers-Danlos syndrome in patients with congenital adrenal hyperplasia. *Mol. Genet. Genom. Med.* **2021**, *9*, e1556. [[CrossRef](#)]
73. Rymen, D.; Ritelli, M.; Zoppi, N.; Cinquina, V.; Giunta, C.; Rohrbach, M.; Colombi, M. Clinical and Molecular Characterization of Classical-Like Ehlers-Danlos Syndrome Due to a Novel TNXB Variant. *Genes* **2019**, *10*, 843. [[CrossRef](#)]
74. Miller, W.L.; Merke, D.P. Tenascin-X, Congenital Adrenal Hyperplasia, and the CAH-X Syndrome. *Horm Res. Paediatr* **2018**, *89*, 352–361. [[CrossRef](#)]
75. McNally, E.; MacLeod, H.; Dellefave-Castillo, L. *Arrhythmogenic Right Ventricular Cardiomyopathy*; Adam, M.P., Ardinger, H.H., Pagon, R.A., Wallace, S.E., Bean, L.J.H., Mirzaa, G., Amemiya, A., Eds.; GeneReviews®: Seattle, WA, USA, 1993.
76. Cocco, E.; Scaltriti, M.; Drilon, A. NTRK fusion-positive cancers and TRK inhibitor therapy. *Nat. Rev. Clin. Oncol.* **2018**, *15*, 731–747. [[CrossRef](#)]
77. Bouguenina, H.; Salaun, D.; Mangon, A.; Muller, L.; Baudalet, E.; Camoin, L.; Tachibana, T.; Cianferani, S.; Audebert, S.; Verdier-Pinard, P.; et al. EB1-binding-myomegalin protein complex promotes centrosomal microtubules functions. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E10687–E10696. [[CrossRef](#)]
78. Kruse, L.M.; Buchan, J.G.; Gurnett, C.A.; Dobbs, M.B. Polygenic threshold model with sex dimorphism in adolescent idiopathic scoliosis: The Carter effect. *J. Bone Jt. Surg. Am. Vol.* **2012**, *94*, 1485–1491. [[CrossRef](#)] [[PubMed](#)]
79. Grauers, A.; Einarsdottir, E.; Gerdhem, P. Genetics and pathogenesis of idiopathic scoliosis. *Scoliosis Spinal Disord.* **2016**, *11*, 45. [[CrossRef](#)] [[PubMed](#)]
80. Perez-Machado, G.; Berenguer-Pascual, E.; Bovea-Marco, M.; Rubio-Belmar, P.A.; Garcia-Lopez, E.; Garzon, M.J.; Mena-Molla, S.; Pallardo, F.V.; Bas, T.; Vina, J.R.; et al. From genetics to epigenetics to unravel the etiology of adolescent idiopathic scoliosis. *Bone* **2020**, *140*, 115563. [[CrossRef](#)]
81. Theocharis, A.D.; Skandalis, S.S.; Gialeli, C.; Karamanos, N.K. Extracellular matrix structure. *Adv. Drug Deliv. Rev.* **2016**, *97*, 4–27. [[CrossRef](#)]
82. Lamande, S.R.; Bateman, J.F. Genetic Disorders of the Extracellular Matrix. *Anat. Rec.* **2020**, *303*, 1527–1542. [[CrossRef](#)] [[PubMed](#)]
83. Wang, S.; Qiu, Y.; Ma, Z.; Xia, C.; Zhu, F.; Zhu, Z. Expression of Runx2 and type X collagen in vertebral growth plate of patients with adolescent idiopathic scoliosis. *Connect. Tissue Res.* **2010**, *51*, 188–196. [[CrossRef](#)]
84. Carr, A.J.; Ogilvie, D.J.; Wordsworth, B.P.; Priestly, L.M.; Smith, R.; Sykes, B. Segregation of structural collagen genes in adolescent idiopathic scoliosis. *Clin. Orthop. Relat. Res.* **1992**, *274*, 305–310.
85. Montanaro, L.; Parisini, P.; Gregg, T.; Di Silvestre, M.; Campoccia, D.; Rizzi, S.; Arciola, C.R. Evidence of a linkage between matrilin-1 gene (MATN1) and idiopathic scoliosis. *Scoliosis* **2006**, *1*, 21. [[CrossRef](#)]

86. Hohmann, T.; Dehghani, F. The Cytoskeleton-A Complex Interacting Meshwork. *Cells* **2019**, *8*, 362. [[CrossRef](#)]
87. Humphrey, J.D.; Dufresne, E.R.; Schwartz, M.A. Mechanotransduction and extracellular matrix homeostasis. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 802–812. [[CrossRef](#)] [[PubMed](#)]
88. Klein-Nulend, J.; Bacabac, R.G.; Bakker, A.D. Mechanical loading and how it affects bone cells: The role of the osteocyte cytoskeleton in maintaining our skeleton. *Eur. Cell Mater.* **2012**, *24*, 278–291. [[CrossRef](#)]
89. Patten, S.A.; Margaritte-Jeannin, P.; Bernard, J.C.; Alix, E.; Labalme, A.; Besson, A.; Girard, S.L.; Fendri, K.; Fraisse, N.; Biot, B.; et al. Functional variants of POC5 identified in patients with idiopathic scoliosis. *J. Clin. Investig.* **2015**, *125*, 1124–1128. [[CrossRef](#)] [[PubMed](#)]
90. Xu, L.; Sheng, F.; Xia, C.; Li, Y.; Feng, Z.; Qiu, Y.; Zhu, Z. Common Variant of POC5 Is Associated With the Susceptibility of Adolescent Idiopathic Scoliosis. *Spine* **2018**, *43*, E683–E688. [[CrossRef](#)] [[PubMed](#)]
91. Hassan, A.; Parent, S.; Mathieu, H.; Zaouter, C.; Molidperee, S.; Bagu, E.T.; Barchi, S.; Villemure, I.; Patten, S.A.; Moldovan, F. Adolescent idiopathic scoliosis associated POC5 mutation impairs cell cycle, cilia length and centrosome protein interactions. *PLoS ONE* **2019**, *14*, e0213269. [[CrossRef](#)] [[PubMed](#)]
92. Gray, R.S.; Gonzalez, R.; Ackerman, S.D.; Minowa, R.; Griest, J.F.; Bayrak, M.N.; Troutwine, B.; Canter, S.; Monk, K.R.; Sepich, D.S.; et al. Postembryonic screen for mutations affecting spine development in zebrafish. *Dev. Biol.* **2021**, *471*, 18–33. [[CrossRef](#)] [[PubMed](#)]
93. Sternberg, J.R.; Prendergast, A.E.; Brosse, L.; Cantaut-Belarif, Y.; Thouvenin, O.; Orts-Del’Imagine, A.; Castillo, L.; Djenoune, L.; Kurisu, S.; McDearmid, J.R.; et al. Pkd2l1 is required for mechanoreception in cerebrospinal fluid-contacting neurons and maintenance of spine curvature. *Nat. Commun.* **2018**, *9*, 3804. [[CrossRef](#)] [[PubMed](#)]
94. Zhang, X.; Jia, S.; Chen, Z.; Chong, Y.L.; Xie, H.; Feng, D.; Wu, X.; Song, D.Z.; Roy, S.; Zhao, C. Cilia-driven cerebrospinal fluid flow directs expression of urotensin neuropeptides to straighten the vertebrate body axis. *Nat. Genet.* **2018**, *50*, 1666–1673. [[CrossRef](#)] [[PubMed](#)]
95. Bomba, L.; Walter, K.; Soranzo, N. The impact of rare and low-frequency genetic variants in common disease. *Genome Biol.* **2017**, *18*, 77. [[CrossRef](#)]
96. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorff, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A.; et al. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753. [[CrossRef](#)] [[PubMed](#)]
97. Harris, P.A.; Taylor, R.; Thielke, R.; Payne, J.; Gonzalez, N.; Conde, J.G. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inf.* **2009**, *42*, 377–381. [[CrossRef](#)] [[PubMed](#)]