



## REVIEW ARTICLE OPEN

## The cartilage matrixome in adolescent idiopathic scoliosis

Carol A. Wise<sup>1,2,3,4</sup>, Diane Sepich<sup>5</sup>, Aki Ushiki<sup>6,7</sup>, Anas M. Khanshour<sup>1</sup>, Yared H. Kidane<sup>1</sup>, Nadja Makki<sup>8</sup>, Christina A. Gurnett<sup>9,10,11</sup>, Ryan S. Gray<sup>12</sup>, Jonathan J. Rios<sup>1,2,3,4</sup>, Nadav Ahituv<sup>6,7</sup> and Lila Solnica-Krezel<sup>5</sup>

The human spinal column is a dynamic, segmented, bony, and cartilaginous structure that protects the neurologic system and simultaneously provides balance and flexibility. Children with developmental disorders that affect the patterning or shape of the spine can be at risk of neurologic and other physiologic dysfunctions. The most common developmental disorder of the spine is scoliosis, a lateral deformity in the shape of the spinal column. Scoliosis may be part of the clinical spectrum that is observed in many developmental disorders, but typically presents as an isolated symptom in otherwise healthy adolescent children. Adolescent idiopathic scoliosis (AIS) has defied understanding in part due to its genetic complexity. Breakthroughs have come from recent genome-wide association studies (GWAS) and next generation sequencing (NGS) of human AIS cohorts, as well as investigations of animal models. These studies have identified genetic associations with determinants of cartilage biogenesis and development of the intervertebral disc (IVD). Current evidence suggests that a fraction of AIS cases may arise from variation in factors involved in the structural integrity and homeostasis of the cartilaginous extracellular matrix (ECM). Here, we review the development of the spine and spinal cartilages, the composition of the cartilage ECM, the so-called “matrixome” and its functions, and the players involved in the genetic architecture of AIS. We also propose a molecular model by which the cartilage matrixome of the IVD contributes to AIS susceptibility.

*Bone Research* (2020)8:13

; <https://doi.org/10.1038/s41413-020-0089-0>

## INTRODUCTION

The human spine serves multiple structural and physiologic roles. An essential function of the spinal column in all vertebrates is to integrate the brain and nervous system with the axial skeleton and simultaneously protect the spinal cord. For humans, the symmetry and position of the spine is key for bipedalism and for maintaining the center of gravity over the pelvis.<sup>1</sup> The spine, with its 33 vertebrae, 23 intervertebral discs (IVDs), ligaments, and facet joints, withstands biomechanical forces while providing flexibility in three dimensions (Fig. 1). The stability of this multicomponent system over a lifetime is remarkable, although eventually the wear and tear from environmental stresses, plus genetic contributions, can evoke degenerative changes primarily in the IVD of the lumbar spine.<sup>2</sup> Other common pathologies of the adult spine include tumors, fractures, lumbar spondylosis, stenosis, and segmental instability or deformity. Scoliosis is a lateral spinal deformity that is common and exhibits increasing prevalence with age and degeneration of the spine.<sup>3</sup> In children, rare developmental disorders of embryogenesis due to malformed vertebrae, muscle weakness, or neurologic impairment can affect the integrity of the spine.<sup>1</sup> However, the most frequent spinal disorder presenting in children is idiopathic scoliosis, which unlike adult scoliosis occurs in the absence of other associated pathologies or

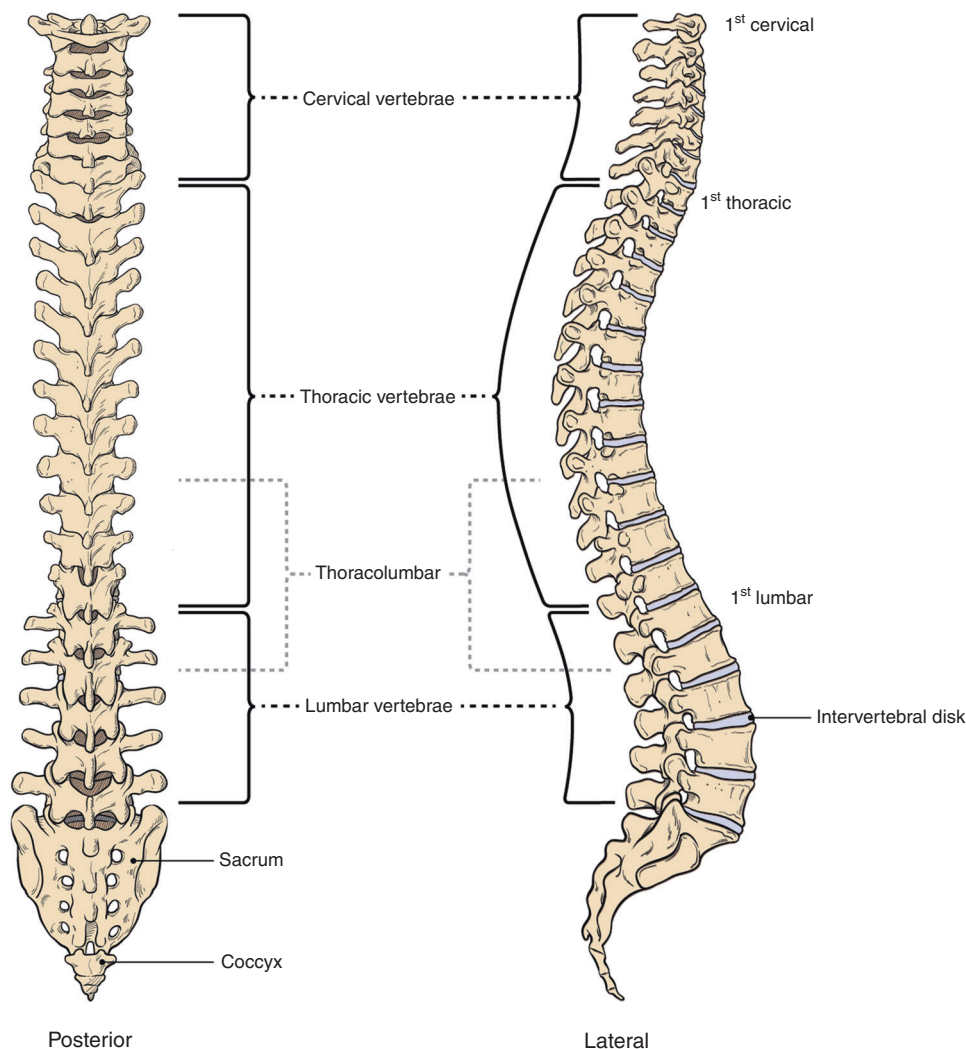
degenerative changes. Idiopathic scoliosis, defined by lateral deformity greater than 10 degrees on coronal radiographs, usually presents during the adolescent growth spurt. Adolescent idiopathic scoliosis (AIS) is the most common pediatric musculoskeletal disorder, occurring in ~3% of school age children, >29 million children worldwide<sup>4</sup> (Fig. 2). Untreated, progressive AIS is associated with restrictive lung disease, pain, and severe deformity later in life.<sup>5</sup> Although curve patterns vary, right-sided thoracic deformities are typical for reasons that are not understood. A child whose scoliosis continues to worsen before the peak of growth velocity is at greatest risk of serious deformity, but after skeletal maturity the risk of “curve progression” decreases sharply.<sup>6</sup> Orthotics, i.e., a back bracing, may be prescribed to help control progression until skeletal maturity, and not to necessarily correct existing deformity.<sup>7,8</sup> Bracing is considered to have “failed” if the curve continues to progress to ~50 degrees before closure of growth plates as measured from iliac crest or hand X-rays.<sup>9</sup> In this scenario, the standard of care is surgical correction, with the goal of reducing the deformity, fusing the involved vertebral bodies to prevent continued progression, and achieving balance of the head, shoulders, and trunk over the pelvis. Typical surgery involves implanting metal instrumentation to provide the forces to straighten and derotate the deformity.<sup>1</sup>

<sup>1</sup>Center for Pediatric Bone Biology and Translational Research, Texas Scottish Rite Hospital for Children, 2222 Welborn St., Dallas, TX 75219, USA; <sup>2</sup>McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, TX 75235, USA; <sup>3</sup>Departments of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX 75235, USA; <sup>4</sup>Orthopaedic Surgery, University of Texas Southwestern Medical Center, Dallas, TX 75235, USA; <sup>5</sup>Department of Developmental Biology, Washington University School of Medicine, St. Louis, MO 63110, USA; <sup>6</sup>Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA 94158, USA; <sup>7</sup>Institute for Human Genetics, University of California San Francisco, San Francisco, CA 94158, USA; <sup>8</sup>Department of Anatomy and Cell Biology, University of Florida, College of Medicine, Gainesville, FL 32610, USA; <sup>9</sup>Departments of Neurology, Washington University School of Medicine, St. Louis, MO 63110, USA; <sup>10</sup>Pediatrics, Washington University School of Medicine, St. Louis, MO 63110, USA; <sup>11</sup>Orthopaedic Surgery, Washington University School of Medicine, St. Louis, MO 63110, USA and <sup>12</sup>Department of Pediatrics, Dell Pediatric Research Institute, University of Texas at Austin Dell Medical School, Austin, TX 78723, USA

Correspondence: Carol A. Wise (carol.wise@tsrh.org)

Received: 29 October 2019 Accepted: 27 January 2020

Published online: 09 March 2020



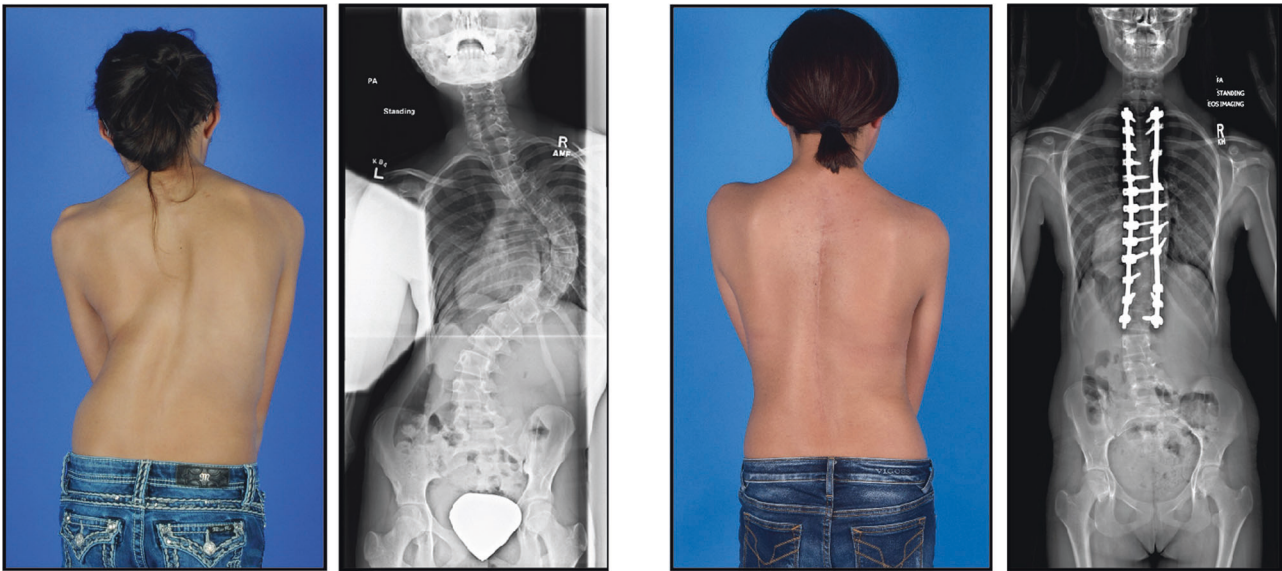
**Fig. 1** The human spine viewed from the back (posterior) and side (lateral)

For centuries, physicians have speculated on the causes of AIS.<sup>10</sup> In what tissues does it originate, and why the curious onset during the adolescent growth spurt? Case reports and epidemiologic studies dating back more than 45 years have consistently demonstrated significant but complex heritability in AIS.<sup>11–13</sup> The assumption from these models is that AIS, in a typical patient, arises from a cumulative effect of several to many genetic risk factors. Certain patients, however, may harbor more deleterious genetic factors, for example, those with a family history of the disease, that increase their risk of developing a scoliosis phenotype. Another curious feature of AIS is its striking sexual dimorphism, with girls having a more than fivefold greater risk than boys of progressive deformity requiring treatment.<sup>14</sup> These observations suggest that the genetic factors underlying AIS may differ between the sexes. In this model, it is hypothesized that the total mutational burden required to “activate” progressive AIS in males will be greater than that in females, a phenomenon known as the Carter effect.<sup>15,16</sup> It follows from this concept that males with progressive AIS will harbor more high-risk, deleterious mutations. As with other complex genetic disorders, we expect many heritable mutations with varying effect sizes and frequencies will comprise the total mutational burden in AIS. Variants that are rare in a population tend to be more penetrant and confer a greater disease risk in an individual, whereas common variants are expected to be less penetrant and to confer relatively smaller effects.<sup>17</sup>

The underlying genetic, molecular, and cellular causes of AIS have remained enigmatic for centuries. However, modern genomic and systems biology approaches in human patient cohorts are beginning to elucidate the genetic changes underlying AIS risk. Simultaneously, precise and powerful forward and reverse genetic approaches in vertebrate model systems are enabling researchers to rapidly identify new AIS candidate genes that may be screened in clinical cohorts, as well as to directly test the phenotypic and mechanistic consequences of mutations discovered in human patients. From these efforts the field of AIS research has seen a convergence on factors that control the development of the IVD, particularly on molecules that maintain or comprise the extracellular matrix (ECM). For the rest of this review, we focus on the biologic factors that control IVD development and their genetic contributions to AIS susceptibility. We also discuss the implications of these discoveries for new hypothesis-driven research and therapeutic interventions.

#### **BIOLOGIC PATHWAYS OF SPINE DEVELOPMENT**

The diverse functions of the spine in strength, flexibility, balance, and protection are enabled by a striking segmented organization along the vertical axis that is the product of an exquisitely controlled molecular network called the “segmentation clock”. The details of the segmentation clock have been described in a



**Fig. 2** Adolescent idiopathic scoliosis in a female patient. Photograph and standing radiograph (left) reveal a typical right-sided thoracic curve and postural imbalance that is largely corrected with surgical implants (right)

number of excellent reviews to which we refer the reader.<sup>18,19</sup> Briefly, in early development, presomitic mesodermal cells proliferate and move from the tail bud at a specific rhythmic pace, creating somites that emerge at the anterior tip and progress along the embryonic axis (Fig. 3). The somites then develop into a dorsal epithelial dermomyotome that differentiates into skin and muscle, and a ventral sclerotome that forms the vertebral bodies, ribs, tendons, and cartilages of the spine, i.e., the vertebral growth plates and IVDs. In humans, a pair of somites forms every 6–12 h between days 20 and 32 of embryonic development.<sup>20</sup> This finely timed process is controlled by three major signaling pathways—Fibroblast Growth Factor (FGF), Notch, and Wnt/ $\beta$ -catenin—that set up waves of gene expression through “circuits” controlling the timing of oscillations, spatial expression, and synchronicity of somitogenesis.<sup>18,21</sup> The subsequent formation of the sclerotome, i.e., future spine and associated structures, requires that a subset of mesodermal cells in the ventral somite region undergo an epithelial-to-mesenchymal transition and then migrate toward the midline to surround the notochord (axial mesoderm). This two-step transition is regulated by Wnt and Fgf signaling pathways.<sup>22</sup> The vertebrate spine also derives from the axial mesoderm in addition to the paraxial mesoderm (somites) (Fig. 3).

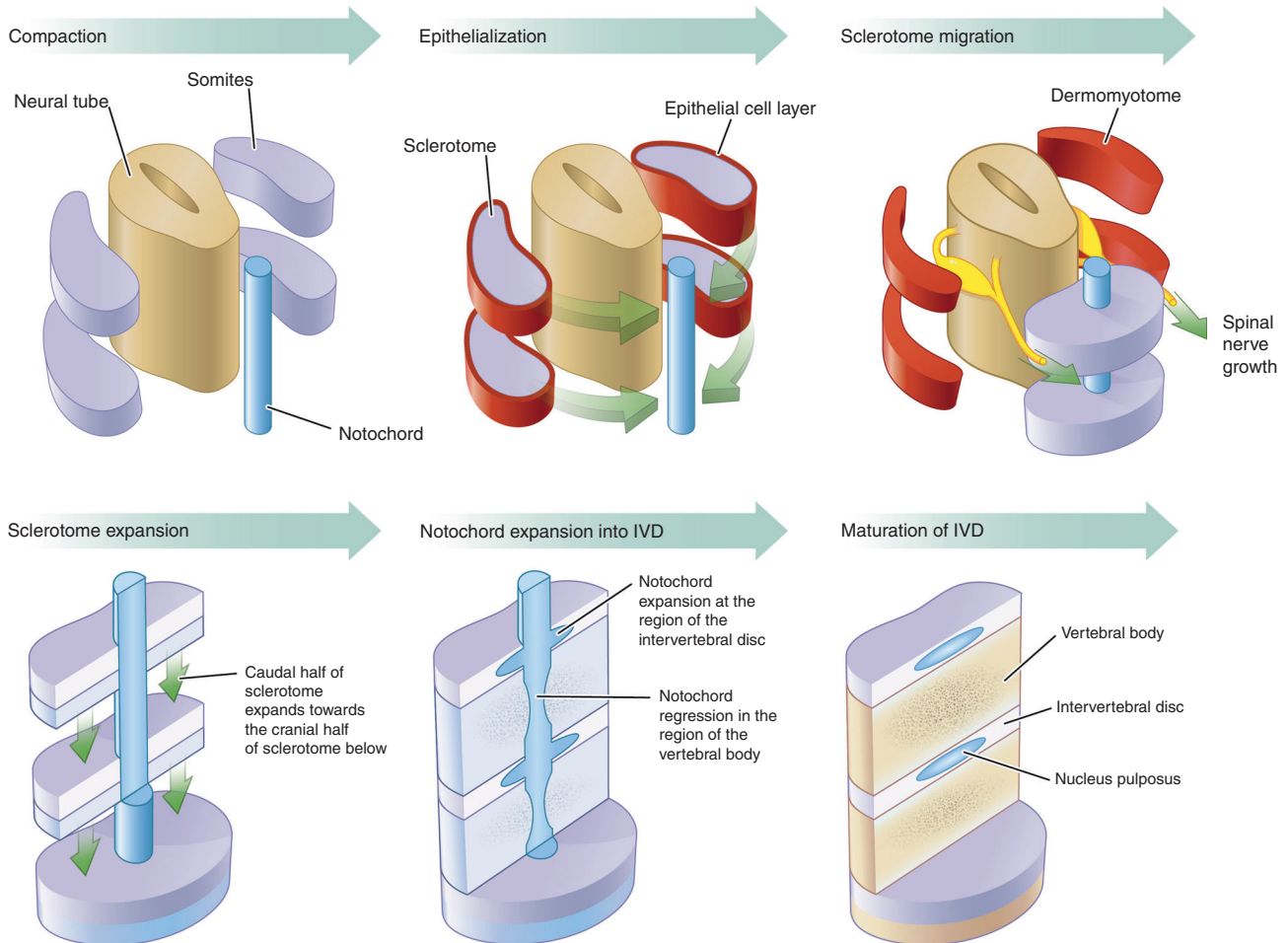
#### DEVELOPMENTAL REGULATION OF IVD BIOGENESIS

In early mouse embryonic development, the sclerotome migrates out from the somites to surround the notochord in a pattern of condensed and less condensed segments corresponding to future IVD and vertebral bodies, respectively, along the anterior/posterior axis. The paired box 1 (*Pax1*) transcription factor is a well-described marker of sclerotome development which is induced via factors secreted from the notochord including: Sonic Hedgehog (*Shh*) and through noggin-dependent inhibition of bone morphogenetic protein (BMP) signaling.<sup>23</sup> During embryonic spine development, the notochord regresses from regions of developing vertebral bodies and expands in regions of the future IVD (Fig. 3). The mechanism of this physical transformation of the notochord remains unresolved but may be related to the mechanical force exerted by the developing vertebral bodies.<sup>23–25</sup> Eventually, the notochord tissue gives rise to the nucleus pulposus (NP), the gel-like inner core of the IVD, and *Pax1* expressing cells contribute to the region of the outer annulus fibrosus (AF) IVD.<sup>26</sup> The spontaneous mutant mouse strains “*undulated*” and “*scoli*” display

varying spinal phenotypes resulting from mutations in the *Pax1* locus.<sup>27–30</sup> Two undulated strains (*un*, *Un<sup>ex</sup>*) exhibit mild kink tail phenotypes due to hypomorphic *Pax1* mutations. By comparison, the homozygous *Un<sup>s</sup>/Un<sup>s</sup>* strain harbors a 120 kb deletion encompassing *Pax1* and completely lacks lower thoracic and lumbar vertebral units. Of note, *Un<sup>s</sup>/Un<sup>s</sup>* displays a more severe phenotype than a homozygous targeted knockout of *Pax1*, suggesting that noncoding sequences outside of *Pax1* may drive expression of additional loci related to spine development. Indeed, analysis of *Pax1* enhancers using transgenic assays in mouse identified a strong extra-genic enhancer “Xe1” which drives expression in the sclerotome and whose loss presumably contributes to the *Un<sup>s</sup>/Un<sup>s</sup>* phenotype.<sup>31</sup> More comprehensive analysis of enhancers, using transcriptomic and chromatin immunoprecipitation sequencing (ChIP-seq) studies of *Pax1*-positive flow-sorted E12.5 vertebral column cells, found enrichment of signals at loci related to ECM, cartilage development, and collagen fibrillogenesis biological pathways. Moreover, *Pax1*-negative cells display a general downregulation of ECM genes, e.g., *Col11A1*, suggesting *Pax1* is a critical transcriptional regulator of ECM matrix biogenesis in the outer AF. Proteoglycan accumulation (e.g., aggrecan) is also significantly reduced in animals lacking *Pax1*.<sup>32,33</sup> It is clear from these and other studies that *Pax1*, together with its paralog *Pax9*, act synergistically to drive development of IVD of the entire vertebral spine.<sup>34</sup>

*Sox9* is a master transcriptional regulator of chondrogenesis.<sup>33,35–38</sup> *Sox5/6* is expressed in the notochord and the sclerotome, where they share partially redundant roles and cooperate with *Sox9* to drive the expression of chondrogenesis genes in the growth plate.<sup>39–44</sup> *Sox5* and *6* are clearly important in IVD development as well. Both are expressed in the notochord and the sclerotome, and the notochordal sheath fails to form in *Sox5<sup>-/-</sup>;Sox6<sup>-/-</sup>* mice. The consequence of this is failure of the NP to form. In addition to severe chondrodysplasia, mice lacking *Sox6* alone develop a kinked tail that occurs in segments where the NP is small and misplaced,<sup>43,45</sup> while *Sox5<sup>-/-</sup>;Sox6<sup>-/-</sup>* double mutant mice display a virtual absence of cartilage and chondrogenic tissues of the spine.<sup>46</sup> Other striking features of *Sox5<sup>-/-</sup>;Sox6<sup>-/-</sup>* double mutant mice are the downregulation of genes encoding major ECM components including collagen II, aggrecan, and cartilage link protein, and a general delay in differentiation of the inner AF (IAF).<sup>32,45,47</sup> *Sox9* controls *Sox5/6* expression and is required for the normal patterning and development of the spine,



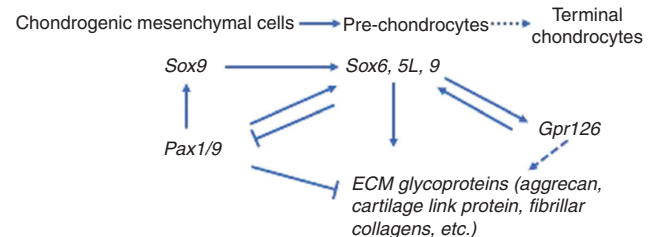


**Fig. 3** Cartoon depiction of the stages of somitogenesis and maturation of the intervertebral disc. Somites (lavender) become epithelialized to form the dermomyotome (red) and sclerotome. Cells of the sclerotome migrate from pairs of paraxial somites and condense around the notochord in a characteristic pattern of more or less condensed regions. These alternating regions (blue and lavender in the figure) give rise to the vertebral bodies and AF, respectively. The NP arises by a different developmental path by swelling in the region of the IVD and pinching off of the notochord in the vertebral bodies

failing to generate obvious patterned vertebral bodies or IVD tissues when ablated in embryonic mice.<sup>48</sup> *Sox9* is continuously required for homeostasis of the spine in adult mice, where its ablation in skeletally mature mice leads to kyphoscoliosis, disc compression, and degeneration of the IVD.<sup>37</sup> Interestingly, *Pax1* transcripts are upregulated when *Sox5* and *6* together are removed genetically from the IVD anlagen, suggesting regulatory feedback between *Sox5/6* and *Pax1* signaling, which may help to define the regulatory programs driving development of the inner and outer AF. Indeed, *Pax1*, as well as its paralog *Pax9*, also assists in the robust gene expression of *Sox5*. Thus, *Pax1/9* and *Sox9* function in concert with BMP and transforming growth factor (TGF) signaling and are critical to initiate expression of chondrogenic genes during early IVD differentiation. Later, the expression of *Pax1* is restricted to the outer annulus by the *Sox5/Sox6* negative feedback mechanism<sup>26,49</sup> (Fig. 4).

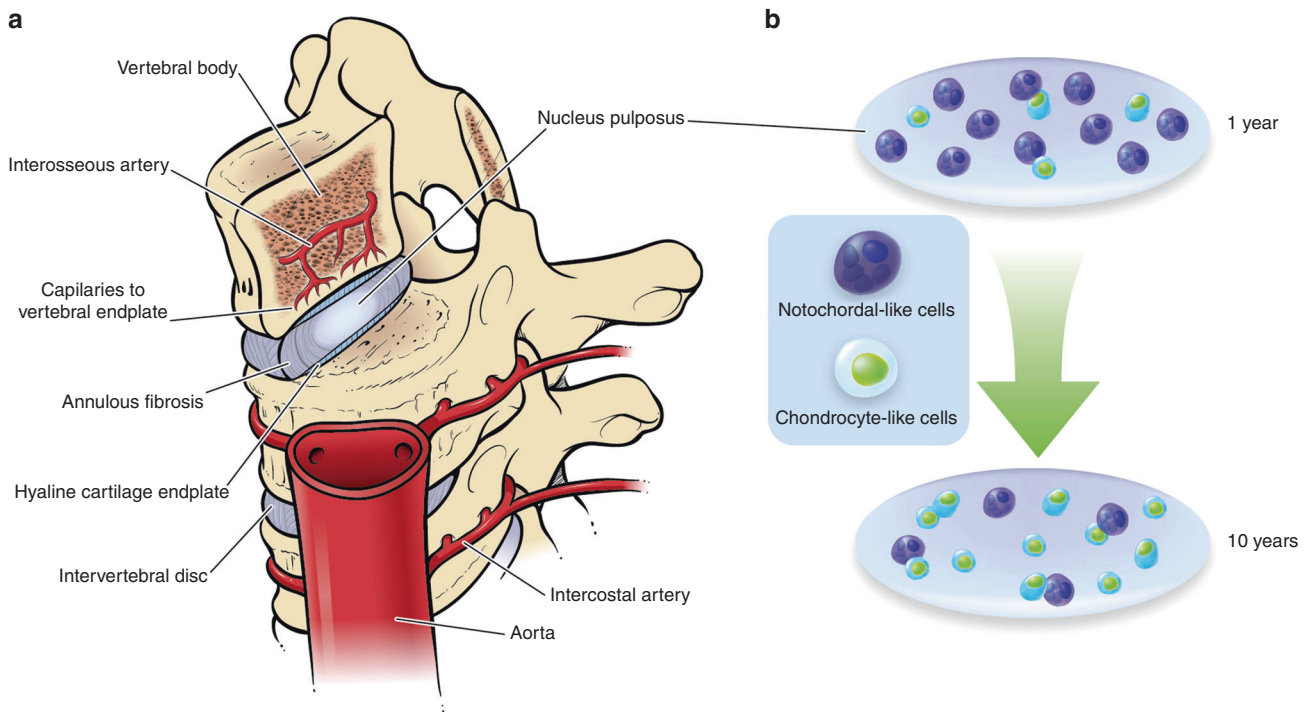
**BEARING THE LOAD—THE POSTNATAL INTERVERTEBRAL DISC**

In humans, the spine grows most rapidly in the first 5 years of life, with an average T1 to S1 segment length increase of 10 cm during this time (2 cm per year). There is a slower T1 to S1 growth from age 5 to 10 years of ~5 cm until adolescence (1 cm per year). From age 10 years to adulthood, T1 to S1 grows an additional 10 cm; this includes the adolescent growth spurt (2 cm per year).<sup>50</sup> There



**Fig. 4** Schematic of the *Pax/Sox*-mediated extracellular matrix pathway in cartilages and IVD. Genes that have been associated with AIS are shown in bold. *Pax1* expression precedes expression of the *Sox* trio in IVD development. The *Sox* trio directly activates the transcription of many ECM genes. *Gpr126* and *Sox9* positively regulate each other in IVD and chondrocyte cells. In the presence of *Sox9*, *Pax1/9* may inhibit expression of aggrecan and possibly other ECM genes in a context-dependent manner

are 23 IVDs in the fully formed human spine that collectively contribute about one third of its height.<sup>51</sup> The vertebrate IVD is composed of fibrocartilaginous tissues encapsulating a gel-like NP tissue, and consequently is considered a cartilaginous tissue joint that connects the vertebral bodies (Fig. 5a). Cells within the IVD are described as “chondrocyte-like” and “fibroblast-like”. Within the IVD, the NP is surrounded by a lamellar structure aptly named



**Fig. 5** The human IVD. The IVD is largely avascular, receiving nutrition through the cartilaginous endplate. Callout depicts replacement/conversion of notochordal-like cells in the nucleus pulposus with chondrocyte-like cells in the first 10 years of life

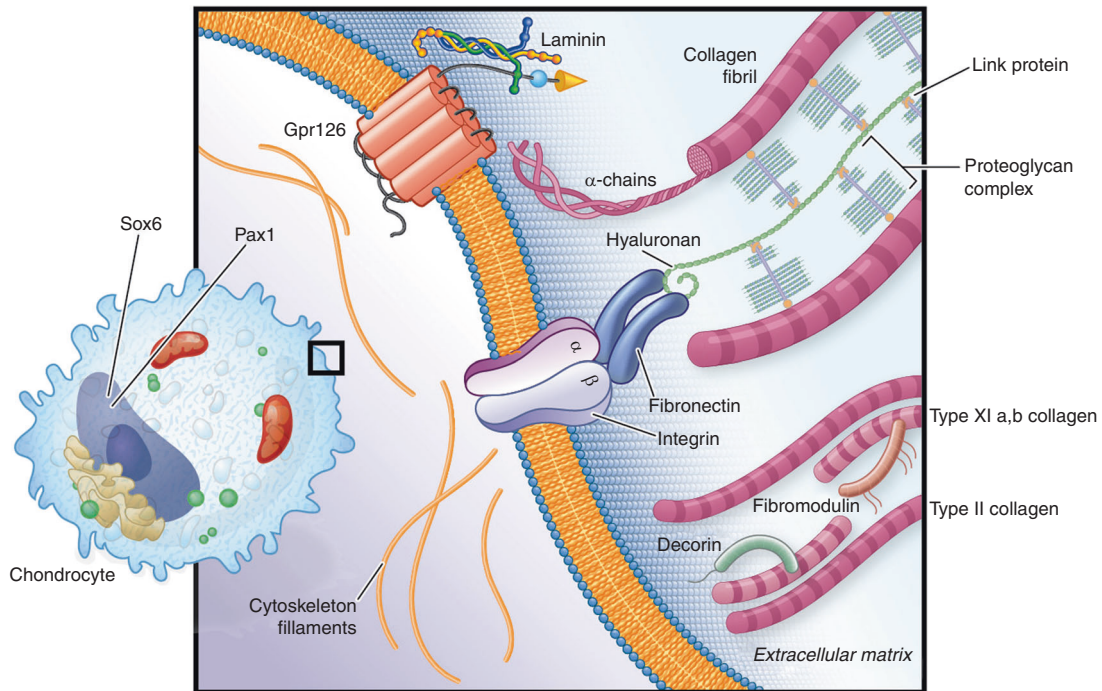
the AF. The ring-shaped AF has two distinct parts, an inner AF made of chondrocytic cells nearest the NP and expressing type II collagen, and an outer AF made of fibroblast-like cells expressing mostly type I collagen. In the outer AF, concentric lamellae of collagen fibers are organized in alternating angles that are oblique to the spinal axis, a so-called “angle-ply” structure. This arrangement conveys mobility and flexibility in multiple planes of motion.<sup>23,52</sup> The IVD is capped by the cartilage endplate (CEP) that connects to cartilage fibers extending from the AF.

The CEP has the critical role of supplying oxygen and nutrients to cells within the IVD, which is in fact the largest avascular structure in the body (Fig. 5a). While the cells of the outer AF are fed by capillaries in the surrounding soft tissues, the remaining cells of the AF and NP receive their nutrients from capillaries that penetrate the vertebral bodies and terminate in loops within the CEP. Nutrients and metabolites then move to and from cells of the disc through the CEP and dense ECM, mostly by diffusion. The IVD is relatively acellular, with about 9 000 “fibroblast-like” cells per mm<sup>3</sup> in the AF and only 4 000 “chondrocyte-like” cells per mm<sup>3</sup> in the NP.<sup>52</sup> In the latter case, notochordal-like cells that begin as large vacuolated structures are slowly replaced by distinctly smaller and less metabolic cells in postnatal development, up to ~10 years of age in humans (Fig. 5b). At this point, notochordal-like cells are virtually absent and replaced by cells with a chondrocyte-like morphology.<sup>23</sup> The transition of cells to a chondrocyte-like phenotype coincides with a change in ECM composition and firming of the NP due to collagen fibril accumulation. The precise transition states of the cells within the NP, and underlying mechanism driving their phenotypes, are still unclear. Cell-tracing experiments using specific Cre drivers such as *Shh-Cre* and *Noto-Cre* demonstrate that cells of notochordal origin are detected in adult NP cells of the mouse.<sup>24</sup> However, not all cells are labeled in these experiments, pointing to heterogeneity that could be due to spatio-temporal differences in NP cell content or even differing origins. Clearly a key role of the NP is to orchestrate the production of healthy ECM.<sup>53</sup> Experimental evidence suggests that the large notochordal-like

“precursor” cells produce proteoglycans faster than small chondrocyte-like cells. Moreover, the distribution and assembly of proteoglycans differs between the two cell types.<sup>54</sup> Consequently, it is reasonable to conjecture that the proportions of cell types within the NP may determine ECM production and thereby influence the strength and integrity of the IVD.

#### COMPOSITION AND ROLE OF THE INTERVERTEBRAL DISC MATRISOME

In the intact disc, osmotic swelling within the NP is resisted radially by the collagen fibers of the AF and axially by the endplates and vertebrae. The AF then effectively absorbs compressive and torsional loads.<sup>23,52</sup> Thus, the three components of the disc are “greater than the sum of the parts”, serving to support and maintain each other as they collectively bear loading of the spine while conferring mobility. Deficiencies in any one part of the disc can impact the whole and, in the mature organ, lead to irreversible damage. The ECM is a critical functional compartment of the IVD, providing mechanical strength, regulating nutritional flow as noted, and conveying physiologic signals back to the cells<sup>55</sup> (Fig. 6). The composition of the acellular ECM, the so-called “matrisome” has been described in a number of excellent reviews to which we direct the reader.<sup>56–58</sup> Here we describe the ECM in brief, with particular reference to the cartilaginous matrisome of the IVD. Predominant components of the cartilage matrisome are fibrillar collagens, chondroitin sulfate proteoglycans, and hyaluronic acid. These large molecules form the basic structural network of the matrisome that allows other factors and cells to interact.<sup>56,58</sup> The IVD produces multiple collagens including types I, II, III, V, VI, IX, XI, XII, and XIV. Types I and II collagens are predominant in the AF and NP, respectively as noted. The minor, less abundant collagen types V and XI form hybrid fibrils with types I and II collagens, respectively, where they are thought to help maintain an appropriately narrow diameter to the collagen fibrils in each compartment. Proteomic studies of the bovine NP suggest that various heterotrimeric combinations may be represented in the



**Fig. 6** Cartoon depiction of a cartilage cell and surrounding matrix. Fibrillar collagens are shown in pink, while sulfated proteoglycan (e.g. aggrecan, perlecan) is shown as a complex with link protein and hyaluronan. The adhesion G protein receptor GPR126 is depicted interacting with a laminin ligand, although its physiologic interaction in cartilage has not been described

type V/XI collagen pool of the disc. What determines this composition is not clear, but may devolve from preferences of the different C-propeptide domains of the various procollagen chains, their relative expression levels in the endoplasmic reticulum where the helices are formed, and/or differences in cell phenotype within the NP both spatially and temporally, for example chondrocytes versus notochordal cells.<sup>59</sup> While collagen is the primary structural component of the disc ECM, elastin fibers are also key for structural support and resistance to deformation.<sup>51,57</sup>

The proteoglycan content is highest in the NP and declines outward toward the AF. In the NP, chondroitin sulfate-linked proteoglycans interact with collagens to retain hydration and form a gel-like matrix that resists compression and provides height to the disc. The predominant proteoglycan of the IVD matrisome is aggrecan, present in both NP and AF.<sup>23</sup> Aggrecan is aptly named as it aggregates along chains of hyaluronan, a high molecular weight polysaccharide, through a small “link” protein called cartilage link protein.<sup>60</sup> The overall negative charge of this complex attracts water and leads to osmotic swelling within the loose meshwork of collagen fibrils in the NP<sup>23,51,61</sup> (Fig. 6). Aggrecan is composed of a core protein linked to both chondroitin sulfate and keratan sulfate chains. Chondroitin sulfate is the major side chain at birth, but over time the proportion of keratan sulfate substitution increases. The shift in side chain composition of aggrecan seems to mirror the transition of notochordal-like cells to chondrocyte-like cells and may reflect differing aggrecan production by the two cell types and less access to oxygen that is needed for chondroitin sulfation intermediates.<sup>52,53</sup> Hyaluronan and link protein also change in abundance over time in the disc. While hyaluronan increases, link protein abundance decreases with age and its proteolytic processing increases. There is evidence that a peptide called “N-link” released by this process actually stimulates ECM (glycosaminoglycan and collagen) production via BMP2 receptor SMAD1/5 signaling.<sup>62,63</sup>

Other proteoglycans in the disc function as accessory proteins that may, for example, participate in collagen assembly and help

to regulate the appropriate width and length of fibrils. For example, small leucine-rich repeat proteins (SLRPs) are proteoglycans such as versican, decorin, biglycan, fibromodulin, lumican, and perlecan that contain characteristic domains including carboxy-terminal leucine-rich repeats (Fig. 6). These proteins regulate collagen fibril growth rates, size, content, and width.<sup>56</sup> Other proteins effectively bridge the structural components of the ECM, re-enforcing and strengthening the network and also connecting it to cells and to other soluble factors. The multi-domain structures of these connectors facilitate direct binding to cell surface receptors, other ECM connectors, and soluble growth factors. Example glycoproteins in this class include the fibronectins and laminins. Like collagens, laminins are triple-helical proteins made up of  $\alpha$ ,  $\beta$ , and  $\gamma$  chains that combine via a central coiled-coil domain. Multifunctional domains of the laminins can interact with cell surface receptors such as integrins and also other components of the ECM, thereby mediating cell–cell interactions. Fibronectin, a large secreted glycoprotein, dimerizes through C-terminal disulfide bonds where it binds to collagens, gelatins, integrins, and other ECM molecules. In particular, fibronectin binding to integrin receptors promotes their clustering in the cell membrane, which in turn leads to further assembly of fibronectin into fibrils around the cells.<sup>56–58</sup> In the adult disc, fibronectin proteolysis into fragments is correlated with the grade of disc degeneration, and has been shown to stimulate catabolic events that lead to the same.<sup>52</sup>

#### ORGANIZATION OF THE IVD MATRISOME

Within the disc, the AF is a collagen-rich fibrillar ECM constructed to withstand tensile and torsional forces transmitted outward from the gelatinous NP. Microfibrils are an important structure distinct from collagen and elastin fibrils.<sup>51,57</sup> The primary structural component of microfibrils in the IVD are the large, structurally distinctive fibrillin glycoproteins (FBN1, 2, 3). Associated with the fibrillins are a number of ancillary proteins, such as microfibril-associated glycoprotein 1, that provide specialized functions. In



**Table 1.** Susceptibility loci in AIS. Table of significant associations and OMIM-related phenotypes

Genome region	Marker	Candidate gene(s)	OMIM disorder	Reference
10q24.31	rs11190870	<i>LBX1</i>		59
6q24.1	rs6570507	<i>GPR126</i>	Lethal congenital contracture syndrome	61
9p22.2	rs3904778	<i>BNC2</i>		76
20p11.22	rs6137473	<i>PAX1</i>	Otofacial cervical syndrome 2	42
9q34.2	rs687621	<i>ABO</i>		68
11p15.2	rs1455114	<i>SOX6</i>	Craniofacial dyssynostosis with short stature	68
16q23.3	rs4513093	<i>CDH13</i>		68
18q21.33	rs4940576	<i>BCL2</i>		62
2q36.1	rs13398147	<i>PAX3/EPHA4</i>	Craniofacial-deafness-hand syndrome	62
1p36.32	rs241215	<i>AJAP1</i>		62
17q24.3	rs12946942	<i>SOX9, KCNJ2</i>	Campomelic dysplasia/familial atrial fibrillation	60
11q24.1	rs35333564	<i>MIR4300HG</i>		68
4q21.23	rs141903557	<i>LOC101928978</i>		111
1q21.2	rs11205303	<i>MTMR11</i>		96
1q42.13	rs12029076	<i>ARF1</i>		96
22q11.21	rs1978060	<i>TBX1</i>	Velocardiofacial syndrome	96
12p12.3	rs2467146	<i>LINC02378/MIR3974</i>		96
8p23.2	rs11787412	<i>CSMD1</i>		96
9p13.3	rs188915802	<i>KIF24</i>		96
6q14.1	rs658839	<i>BCKDHB/FAM46A</i>	Osteogenesis imperfecta type XVIII	96
7p15.1	rs160335	<i>CREB5</i>		96
6q22.1	rs482012	<i>NT5DC1</i>		96
7p22.3	rs11341092	<i>LOC101927021/UNCX</i>		96
1q32.2	rs17011903	<i>PLXNA2</i>	Beckwith–Wiedemann syndrome	96
7p21.2	rs397948882	<i>AGMO/MEOX2</i>		96
16q12.2	rs12149832	<i>FTO</i>	Growth retardation, developmental delay, and facial dysmorphism	96

Regions in the human genome providing statistically significant evidence of harboring an AIS risk factor are denoted by their OMIM locus number (column one) where applicable. The markers shown are those that provided most significant evidence in the region

the OAF and IAF microfibrils co-localize with elastin fibers. While in the OAF microfibrils closely align with collagen lamellae of the ECM, in the IAF fibrillin-positive microfibrils are also seen in an elastin-negative region around cells. In the NP microfibrils are found almost exclusively in the pericellular region, away from elastin fibers in the ECM.<sup>51</sup> The distribution of microfibrils may reflect differing roles in the OAF, which must absorb the brunt of compression and torsional forces, compared with the IAF and NP that are subjected to relatively minor, isotropic forces. It is also now appreciated that the microfibrillar niche is a command center for regulating cellular behavior. This is apparently achieved through direct interactions with cell surface receptors as well as by serving as a reservoir for various growth factors, particularly in the TGF family.<sup>51,64,65</sup>

#### PATHWAYS OF IVD DEVELOPMENT AND SUSCEPTIBILITY TO ADOLESCENT IDIOPATHIC SCOLIOSIS

The majority of genetic risk in AIS has been defined thus far by hypothesis-free genome-wide association studies (GWAS) that map the location of genetic risk factors in the genome relative to fixed markers, usually single nucleotide polymorphisms (SNPs).<sup>66</sup> By definition, SNPs are common in the population and consequently can be informative in statistical tests such as logistic regression that measure frequency differences between affected cases and matched controls. Because of the many tests performed in a typical GWAS (so-called “multiple testing”), a significance level of  $P < 5 \times 10^{-8}$  is generally required for any association to be

declared significant (GWAS-level significance). Replication of GWAS-significant loci is necessary to independently validate candidate SNPs and provide realistic estimates of effect size and fraction of risk conferred across populations. It is important to note that while GWAS are proven to be powerful for mapping disease loci, further fine-mapping approaches are needed to define the underlying causal mutations. Multiple AIS risk loci that were originally identified in GWAS have been further validated by replication and meta-analysis studies, as given in Table 1. Importantly, while these studies have mapped disease loci near or within likely candidate genes, the majority have not pinpointed causal sequence changes driving genetic associations. Nonetheless, as shown below, the number of validated loci around genes that are known to determine early spine development is striking. Further, pathway enrichment analysis of a recent multi-ethnic GWAS meta-analysis provided statistically significant evidence that biologic processes of cartilage development and/or maintenance are impacted in AIS.<sup>67</sup>

An association distal to the *PAX1* gene was originally discovered by a GWAS in 3102 individuals of European descent (North American) and replicated in independent North American, Japanese, and East Asian female cohorts.<sup>68</sup> Further investigation revealed that the association was driven almost entirely by females (combined rs6137473  $P = 2.15 \times 10^{-10}$ , OR = 1.30) but not males (combined rs6137473  $P = 0.71$ , OR = 1.08), identifying it as an apparently sex-specific AIS locus. Resequencing highly conserved sequence elements within the associated interval in AIS subjects refined the signal to a putative enhancer near the *Xe1*

enhancer described earlier. The new enhancer, called *PEC7*, was shown to drive reporter gene expression in zebrafish somites, an activity that was disrupted by AIS-associated SNPs.<sup>68</sup> These data suggest that alterations in the spatial and/or temporal control of *PAX1* expression can increase risk of spinal deformity (AIS). The association signal within an intron of the *SOX6* gene implicates its transcriptional regulation in AIS. *SOX6* is expressed in several tissues including testis, muscle, and spinal cord in addition to chondrocytes.<sup>45</sup> However, it is tempting to speculate that AIS-associated variation in the *SOX6* enhancer could specifically diminish its expression in cartilaginous tissues and thus dysregulate *SOX5/6/9*, *PAX1*, and possibly even *GPR126* (see below) in these tissues. *GPR126* (OMIM 612243, alias *ADGRG6*, or *VIGR*) is a member of the adhesion G protein-coupled receptor family with multiple cellular functions ranging from cell adhesion to migration, to name a few.<sup>69</sup> In humans, *GPR126* is highly expressed in cartilages and the IVD. In mouse embryo, *Gpr126* is highly expressed in the proliferating cartilage of the spine, suggesting its role in spinal development could be exerted through effects on chondrogenesis.<sup>70</sup> As outlined below, further experiments in conditional mutants have supported this hypothesis, and in particular, have shown that *Gpr126* is a positive regulator of cartilage differentiation and possibly its homeostasis. In the chondrogenic ATDC5 cell line, *Gpr126* mRNA expression increases with early chondrogenic differentiation.<sup>71</sup> *Gpr126* overexpression in this cell line increases expression of the cartilage marker genes *Col2a1* (encoding type II collagen) and *Acan* (encoding aggrecan), while *Gpr126* knockdown decreases their expression.<sup>72</sup> *Gpr126* and *Sox9* are also positive regulators of each other, as knockdown or knockout of either reduces expression of the other in ATDC5 cells, chondrocytic N1511 cells, or adult IVD.<sup>37,72</sup> In zebrafish, *Gpr126* is required for Schwann cell myelination, an effect that could be overcome by elevating cAMP levels with forskolin, suggesting a classic signaling mechanism through G proteins.<sup>73</sup> *Gpr126* receptor signaling can be self-activated through autoproteolytic processing.<sup>74</sup> *Gpr126* additionally mediates certain activities such as Schwann cell development through interaction with laminin-211 ligand.<sup>75</sup> However, the mechanisms by which *Gpr126* signals in chondrocyte cells are yet to be characterized.

Whether genetic variation of other associated loci contributes to mechanisms of cartilage biogenesis awaits investigation. For example, *CDH13* encodes the cadherin 13, or “T-cadherin” protein, a unique member of the cadherin family that lacks transmembrane and cytoplasmic domains, attaching to the plasma membrane via a glycosylphosphatidylinositol anchor.<sup>76</sup> During development T-cadherin is highly expressed throughout the early developing sclerotome, with later expression restricted to the caudal sclerotome and migrating neural crest cells.<sup>77</sup> Strong *CDH13* expression is also reported in a chondrocytic cell line and rib cartilage, but its role in these tissues is unknown.<sup>67</sup> ABO encodes glycosyltransferase isoforms that catalyze the transfer of carbohydrates to the H antigen, forming the antigenic structures of the ABO blood group [MIM:616093]. *ABO* is expressed at varying levels in a variety of tissues and is reportedly undetected in cartilage. *BNC2* (OMIM 608669) encoding Basonuclin 2, is a highly conserved transcription factor belonging to the *C2H2* group of zinc finger proteins. In humans, *BNC2* is produced in a variety of tissues including bone and cartilage. The strongest and most reproduced association with AIS in multiple studies is with SNP rs11190870 that is located near the end of the *LBX1* gene. *LBX1* is encoded head-to-head with another gene *FLJ41350* (also called *LOC399806* or *LBX1-AS1* (for *LBX1* antisense RNA 1)). *LBX1* (OMIM 604255) encodes the ladybird late homeobox transcription factor. *Lbx1* is specifically expressed early in embryogenesis in the developing dermamyotome and is suggested to control the expression of genes that guide lateral migration of muscle precursors and maintain their migratory potential.<sup>78</sup> *Lbx1* is also critical for specification of dorsal interneuron fates and

somatosensory function.<sup>79</sup> Thus, *LBX1* may be implicated in myogenic and neurogenic etiology of human AIS.

### GENETICS OF AIS CURVE PROGRESSION

Identifying patients with a high likelihood of curve progression is a top priority of AIS clinical management. Consequently, a comprehensive search for severity-correlated SNPs was conducted in a GWAS by limiting cases to those with spinal deformity above 40 degrees after skeletal maturity or surgical correction. Through a two-stage association study using a total of 12 000 Japanese subjects, a novel locus, rs12946942 on chromosome 17q24.3, showed significant association and was replicated in a Han Chinese population.<sup>80</sup> rs12946942 and the LD block containing it are in a gene desert between *SOX9* and *KCNJ2*. The block contains six enhancer elements, one of which forms a long-range chromatin loop to *SOX9* in a prostate cancer cell line, and SNPs within the enhancer appear to direct allele-specific gene expression. It is posited that variants in the region may confer significant risk of progressive AIS by controlling scoliosis-related spatio-temporal *SOX9* expression.

A later study in the Japanese cohort identified association with SNPs in MIR4300HG, the gene encoding MIR4300 microRNA. Correlated SNPs were mapped within a putative enhancer in a conserved topologically associated domain (TAD). So-called TADs are large, self-interacting genomic regions that are implicated in gene regulation.<sup>81</sup> Although MIR4300 expression has been noted in bone, cartilage, IVD, and spinal cord among other tissues, its function and potential role in AIS is not yet known.<sup>82</sup>

### THE IVD MATRISOME AND RARE GENETIC VARIATION

Heritability studies (calculated as the square of the correlation coefficient,  $r^2$ ) estimate that <7% of the total genetic variance underlying AIS has been discovered.<sup>67</sup> More recently, DNA sequence-based approaches have been applied to both AIS case/control cohorts and to individual AIS families, with the goal of identifying rare disease-causing variants that would be missed by standard GWAS and that could account for “missing” heritability.<sup>83</sup> The discovery of rare disease alleles has been bolstered by large-scale publicly available reference sets such as the Genome Aggregated Database (gNOMAD) that provide summary data for sequenced alleles.<sup>84</sup> The search space for rare genetic risk factors is often limited to the exons of the genome, the so-called “exome” for practical reasons, e.g., cost, computation time, and biologic interpretability. As the costs of sequencing decrease, the field is moving beyond exome sequencing (ES) toward genome sequencing (GS). Simultaneously, the tools for identifying and interpreting noncoding rare disease variants are rapidly improving.<sup>85</sup> Single rare variants may be tested for enrichment in cases versus controls or co-segregation with disease in families. Alternatively, a potentially more informative approach is to statistically measure the burden of rare alleles per gene (i.e., coding sequences) in cases versus controls. The first ES burden study of 91 AIS cases and 337 controls showed *FBN1* and *FBN2*, encoding fibrillins-1 and -2, respectively, as the most significantly associated genes. Dominantly inherited mutations in *FBN1* are known to cause Marfan syndrome (OMIM 154700). Marfan syndrome patients typically present with distinctive skeletal features including increased height, disproportionately long limbs and digits, anterior chest deformity, mild-to-moderate joint laxity, scoliosis, dural ectasia (widening of the dural sac around the spinal cord), and a narrow, highly arched palate with crowding of the teeth. Cardiovascular features can be life-threatening and include mitral valve prolapse, mitral regurgitation, dilatation of the aortic root, and aortic regurgitation.<sup>86</sup> Mutations in *FBN2* are associated with the related disorder Congenital Contractural Arachnodactyly (CCA) (OMIM 121050). The mutations



identified in AIS tended to be associated with tall stature and larger spinal curves, but without the cardiac manifestations or other complications that are typical of MFS or CCA,<sup>87</sup> suggesting they exert hypomorphic effects on the functions of these proteins.

Other sequence-based studies have highlighted rare mutations in ECM genes. In a pathway-based analysis of severe (i.e. >40 degree curves) AIS, restricting the analysis to novel (not found in dbSNP at that time) coding variants revealed association with musculoskeletal collagen genes, particularly *COL11A2*.<sup>88</sup> Dominant inheritance of mutations in *COL11A1*, *COL11A2*, and *COL2A1* cause Stickler or Marshall syndromes (OMIM 604841, 184840, 154780). The Stickler/Marshall syndromes are hereditary conditions characterized by distinctive, flattened facial features, abnormalities of the vitreous, conductive hearing loss, scoliosis or kyphosis, and joint problems including joint hypermobility and early onset arthritis. Many Stickler patients exhibit Pierre Robin sequence, a combination of features (cleft palate, glossoptosis, micrognathia) that can lead to feeding problems and difficulty breathing.<sup>89,90</sup> On average, significantly more novel variants in ECM genes were observed in the AIS cohort compared with controls, suggesting that the polygenic accumulation of novel variants within potentially many ECM genes contributes to AIS risk. More recently, an exome-wide association study that was restricted to severe AIS identified a variant in the *SLC39A8* gene. *SLC39A8* is a ubiquitously expressed divalent metal cation transporter that, not surprisingly, has been genetically linked to disorders of numerous organ systems in humans including blood pressure, body mass index (BMI), circulating manganese levels, and circulating cholesterol levels.<sup>91</sup> The AIS-associated variant altered alanine at position 391 to threonine. This single amino acid change was also correlated with greater spinal curvature, decreased height, increased BMI, and lower plasma manganese in AIS cases. Mutations in the zebrafish homolog caused vertebral anomalies and mild spinal curvatures that were noted to be reminiscent of zebrafish knockouts of the gene encoding alpha-1 chain of type VIII collagen described below.<sup>92</sup> This observation suggests the possibility of a functional interaction between *Slc39a8* and the extracellular matrix in zebrafish, and possibly other vertebrates including humans. Taken together, rare variant-based studies are interpreted to suggest that rare, hypomorphic variants in genes responsible for Mendelian connective tissue disorders may be contributing factors in AIS.

Families with multiple affected members provide the opportunity to map and identify disease variants by their co-segregation with AIS. In one study, filtering for rare variants (minor allele frequency <5% in the Exome Aggregation Consortium (EXAC)) identified by ES in five multigenerational families followed by gene ontology term enrichment analyses produced top results for actin- and microtubule-based projections and ECM.<sup>93</sup> In a separate hypothesis-driven study, the *HSPG2* gene encoding the SLRP perlecan was selected as a candidate per ES of an extended AIS family. One particular *HSPG2* variant, Asn786Ser, was marginally overrepresented in a larger case-control cohort.<sup>94</sup> Recessively inherited mutations in the *HSPG2* gene cause Schwartz–Jampel syndrome type I (OMIM 255800), a rare disorder characterized by joint contractures, myotonic myopathy (muscle stiffness), dystrophy of epiphyseal cartilages, eye abnormalities, short stature, distinct facial features and propensity to malignant hyperthermia.<sup>95</sup> In a separate study, linkage analysis and ES of a Swedish pedigree with an autosomal dominant inheritance of AIS identified a rare, nonsynonymous variant (V2287I) in the *CELSR2* gene encoding cadherin EGF LAG seven-pass G-type receptor 2. Furthermore, analysis of tagging SNPs revealed significant association ( $P=0.0001$ ) to this variant in a set of 1739 Swedish-Danish AIS cases and 1812 controls but not with case-control cohorts from Japan and the US. It is interesting to note that *CELSR2* is a planar cell polarity cadherin, like *GPR126*, that encodes an adhesion-class G protein receptor. Mice that lack *Celsr2* develop progressive

hydrocephalus and display dysgenesis of ependymal cilia. Further family-based studies hold promise for discovering higher-impact AIS disease alleles, although reduced penetrance of the phenotype is an acknowledged challenge. Continued genome sequencing in both populations and families is warranted to fully define the polygenic architecture of the disease.

## MODELS OF AIS PATHOGENESIS

Taken together, evidence from human genetic studies and animal models supports a role for vertebral cartilages, particularly the cartilage matrisome, in AIS pathogenesis. This model focuses attention on the IVD due to its role in withstanding and dissipating the forces of spinal growth. Radiographic studies of the spine in patients with progressive scoliosis are consistent with this model. In AIS patients that were followed longitudinally, wedging of the involved discs preceded wedging of the involved vertebrae until the end of the growth spurt, at which point vertebral wedging progressed somewhat more.<sup>96</sup> Furthermore, MRI studies of AIS patients found that progressive scoliosis was associated with displacement of the NP toward the convexity of the curve.<sup>96</sup> Whether the relatively subtle hypomorphic genetic changes in components of the IVD ECM identified with AIS so far are necessary or sufficient to cause AIS is not clear at present. Consequently, the ability to affect spatio-temporal expression of specific genes with precision, and to create knock-ins of disease-associated mutations in vertebrate models will be key for addressing these questions.

The specific tissues in which AIS originates has been difficult to discern, as bone and soft tissue components of the spine in human patients appear to develop and function normally, at least superficially. Likewise, although AIS manifests around the adolescent growth spurt, it has been unclear whether its origins trace back to deficiencies in early development. Research addressing these questions has been severely limited by a lack of appropriate animal models that recapitulate a true AIS phenotype. This issue is likely a reflection of several factors, including reduced penetrance and the relatively late onset that may be missed in typical forward genetic screens. Moreover, the genetic factors identified in humans thus far point to hypomorphic mutations and variations in regulatory elements that may manifest in tissue- and temporal-specific ways. Fortunately, breakthroughs in this regard have come from use of the Cre/loxP system to create tissue-specific and inducible knockouts of AIS candidate genes. For example, removing *Gpr126* using a Col2Cre driver produced thoracic scoliosis beginning at age P20 and becoming more prevalent with age. From this study it was posited that AIS may arise from deficiencies in a specific osteoprogenitor cell type expressing type II collagen.<sup>97</sup> The study also suggested that scoliosis may have its beginnings in early development. While cell proliferation was unaffected in rib and spinal cartilage (IVD) cells, a slight increase in apoptotic cells was noted in both locations at P1 before the expected onset of spinal curvature. This and additional experiments defined a critical role for *Gpr126* on cartilage development, including effects on the morphogenesis of the annulus fibrosis, chondrocyte survival, and the expression of *Gal3st4* encoding galactose-3-O-sulfotransferase 4.<sup>98</sup> These results also support the hypothesis that the dysregulation, and possibly loss, of specific cell populations early in postnatal development leads to a “weakened” IVD, perhaps through deficiencies in the ECM, that is vulnerable to the increasing axial and torsional forces exerted during rapid growth.

Fish (teleosts) are emerging as powerful tools for understanding developmental mechanisms of spinal deformities. Scoliotic curves have been observed in various fish species, including early studies in “curveback” guppy that demonstrated a heritable phenotype reminiscent of human idiopathic scoliosis.<sup>99</sup> It has been conjectured that the cranial to caudal spinal loads that fish experience while swimming may recapitulate, to some extent, the same

directional forces on the human spine due to walking upright.<sup>100</sup> Zebrafish (*Danio rerio*) in particular is a workhorse laboratory fish that has become a staple in the armamentarium of tools to dissect human developmental disorders. Attractive features of the zebrafish include its relatively rapid generation time of 6–12 weeks, external and near-transparent development, and genetic and genomic tractability. Early models of spinal deformity in zebrafish were mapped to mutations in the *Col8a1a* gene encoding alpha-1 chain of collagen type VIII. These mutants exhibited vertebral anomalies, reminiscent of human “congenital” scoliosis, that were correlated with early defects of notochord development.<sup>101</sup> Similar morphologic changes were later noted in *Slc39a8* knockout zebrafish as described above, raising the question of a possible pathogenic overlap between congenital scoliosis that involves frank vertebral anomalies, and idiopathic scoliosis in which the vertebrae appear to be morphologically normal. Evidence in support of this came from another gene-targeting study in zebrafish that produced the first genetically defined developmental model of IS. In that study, genetic removal of an atypical protein tyrosine kinase called *Ptk7* evoked a rotational spinal deformity beginning at late larval stages and progressing until sexual maturity. This model also recapitulates human IS in that there was a strong bias toward severe curves in females. Further genetic manipulations that removed maternal copies of *Ptk7* in early development led to loss-of-function mutants with vertebral anomalies reminiscent of congenital scoliosis, suggesting that the same genetic mutations could result in idiopathic versus congenital scoliosis phenotypes in a context-dependent manner.<sup>102</sup> This study of *ptk7* mutants implicated abnormal spinal fluid flow. A more recent study further implicated neuroinflammation as an underlying cause of spine curve formation in *ptk7* mutants.<sup>103</sup> A correlation between inflammation and onset of scoliosis was also supported by independent studies in which the transcription factor *Stat3* was removed from maternal and zygotic zebrafish. During later development, *stat3* mutant zebrafish exhibit stunted growth, scoliosis, excessive inflammation, and fail to thrive. In humans, dominant-negative mutations in *STAT3* cause Hyper IgE syndrome marked by dysregulation of immunity as well as late-onset scoliosis reminiscent of AIS.<sup>104</sup>

The ability to create knock-ins of specific human mutations, whether in exons or in disease-associated enhancers such as *Xe1/PEC7*, using genome editing tools will be critical for hypothesis-testing in vertebrate systems. It remains to be seen whether single variants in individual genes will be sufficient to create AIS-like phenotypes in mouse. More likely, in keeping with the complex heritability of AIS, it may be combinations of genetic changes affecting multiple components of the cartilage matrisome that will be required to evoke progressive deformity. The introduction of multiple genetic changes into individual animals by breeding strategies or sequential genome editing may be required to model AIS appropriately.

### THERAPEUTIC STRATEGIES FOR AIS

Natural-history studies have shown AIS measuring less than 30 degrees at the cessation of growth rarely increases in adulthood. Fortunately, the majority of patients are in this category but nevertheless require careful monitoring with radiography to monitor the curve. Curves greater than 50 degrees, however, will slowly progress throughout adulthood at a rate of 0.75–1.00 degrees per year. Therefore, the AIS patient with an immature skeleton and scoliosis of greater than about 20 degrees is at risk for progression and may be treated by bracing.<sup>14</sup> Bracing addresses both biologic and biomechanical susceptibility in AIS by relieving asymmetric loading in the spine and thereby allowing symmetric growth. Although bracing has shown efficacy in preventing progression in clinical trials,<sup>8</sup> it is typically prescribed to be worn 12–16 h each day over a course of years, and compliance is notoriously poor.<sup>8,105</sup>

Surgical correction is required when bracing fails to halt progression, not only to address the appearance of the deformity but also to prevent future pulmonary complications. Lung volume normally increases dramatically during the juvenile and adolescent years, approximately doubling from the age of 10 years to skeletal maturity. Progressive AIS is associated with diminished lung volumes, particularly for curves  $\geq 70$  degrees. Patients with curves reaching 100 degrees are at significant risk of symptomatic restrictive pulmonary disease. Given all of these issues, the first goal of surgical approaches to AIS is to halt progression beyond an acceptable limit of  $\sim 50$  degrees Cobb angle, while a secondary goal is to achieve some correction of the existing deformity.<sup>1</sup>

The fact that bracing can be successful suggests that future biologic therapies aimed at preventing progression should target the same “window of opportunity”. As outlined in this review, genomic studies of human AIS point to the biogenesis and/or maintenance of the ECM of the IVD in disease susceptibility. Future therapies consequently could be aimed at stimulating proper ECM production in cells of the IVD in patients with mild AIS, with the goal of providing better biomechanical and/or biologic resistance to asymmetric growth forces. Approaches to IVD-targeted therapy could perhaps build upon the data armamentarium developed by the intervertebral disc degeneration (IVDD) research community, as well as on other successful gene therapy efforts.<sup>106–109</sup> Therapies proposed to stimulate ECM production in IVDD include local delivery of growth factors or viral-based gene therapy, or adoptive transfer of adult mesenchymal stem cells. In this regard, the avascular nature of the IVD and its limited systemic connection could prove advantageous for direct in vivo therapeutic delivery, similar to the eye. Direct injection of therapeutic viral vectors has shown such promise for correcting several inherited eye disorders and it is now being pursued in clinical trials.<sup>108</sup> In IVDD, increased proteoglycan synthesis has been observed after direct injection of adenovirus, adeno-associated virus, or lentivirus expressing *Tgf- $\beta$* , latent membrane protein 1 (*Lmp1*), *Sox9*, and other factors in rabbit models.<sup>106</sup> Cell-based therapies would aim to replace or regenerate resident IVD cells, likely using endogenous progenitor cells or MSCs to increase ECM production. However, harvesting endogenous cells would require surgery, and the cells would be predicted to harbor the same genetic susceptibilities as resident IVD cells. The number of resident NP cells in the IVD is expected to be at its peak in individuals younger than 20 years of age.<sup>107</sup> Although the cellular status in AIS patients is not well understood, delivering additional progenitors into the IVD environment might not prove beneficial. Modern genome editing technologies provide another promising line of gene therapy. Specifically, clustered regularly interspaced short palindromic repeat (CRISPR)-associated nucleases can be used to cleave specific sites, with subsequent removal or replacement of targeted DNA sequences.<sup>108</sup> The benefits of genome editing are its simplicity and precision in removing or correcting an offending mutation. This is particularly relevant where simply stimulating endogenous gene expression would not apply. For example, genome editing could be envisioned to correct aberrant amino changes in type XI collagen in the AIS IVD. Another advantage of CRISPR-based genome editing is the efficiency of creating multiple genetic changes within each targeted cell, which may be necessary for more complex disorders such as AIS. It should be noted that while genome editing holds promise, its clinical application to any acquired or inherited human disorder is in its infancy until potential side effects (off-target changes and possible leakage to germline) are fully understood and controlled. Considering AIS, biologic therapies are particularly attractive given the “window of opportunity” in children with mild AIS who are apparently otherwise healthy.

Prevention of AIS altogether has been a futuristic goal but not inconceivable given the emerging understanding of underlying mechanisms. For some patients, improving mechanical function in the IVD by boosting glycoprotein content and hydration early in

life could be envisioned as a preventative strategy, perhaps by small molecule or peptide therapies, or by better nutritional delivery to the disc.<sup>110</sup> In this scenario, the late onset and ostensibly healthy metabolism in AIS could prove advantageous, analogous to preventing osteoporosis by early calcium and vitamin D intake. It is notable however that in the IVD, the ECM is key in governing the rate of molecular transport through the long pathway from the cartilaginous endplate to the relatively sparse NP and/or AF cells. Aggrecan in particular intercalates with the collagen framework and thereby largely determines solute distribution and mobility.<sup>110</sup> Clearly, appropriate *in vitro* and *in vivo* model systems and more fundamental understanding of the genetic, cellular, and developmental mechanisms of AIS are needed to advance any biologic therapeutic strategies.

## SUMMARY

The embryologic development of the spine is an exquisitely controlled process both spatially and temporally. Deleterious mutations in single genes controlling somitogenesis can lead to rare Mendelian disorders affecting spinal integrity. Idiopathic scoliosis, usually occurring at adolescence, is the most common developmental disorder of the spine in children. AIS is genetically complex, but emerging data both from GWAS and next-generation sequencing of human AIS cohorts have highlighted a role for determinants of cartilage biogenesis and development of the IVD, particularly pathways controlling extracellular matrix, or matrisome production. DNA sequence variants have been identified that potentially alter the transcriptional regulation of matrisome genes, or that create hypomorphic changes in protein-coding sequences. Combinations of these variations may effectively create tissue-specific alterations, with major effects on the composition and/or organization of the IVD matrisome in cells of both NP and AF. Cellular maturation in the NP occurs in the first decade of life in humans and also may be a contributor to AIS. In either scenario, impairments in the disc may cause asymmetric responses to the strong torsional forces along the spine during adolescent growth, leading the spine to twist away from these forces. In contrast to the disc in adult IVD degeneration, the IVD in AIS does not exhibit overt signs of degeneration. Given this and the fact that current bracing therapies during the adolescent growth spurt can be effective suggests that there is a window of opportunity for providing biologic therapies to control or prevent AIS and subsequent curve progression. The IVD of AIS patients may be amenable to direct *in vivo* delivery of molecular or nutritional therapies aimed at preserving its structural integrity. These factors encourage optimism and justify continued research to understand underlying disease mechanisms and to develop robust model systems for *in vivo* testing.

## ACKNOWLEDGEMENTS

We thank Stuart Almond for help with figures, and Brad Niese, PhD for helpful discussions about disc biomechanics. This work is in part supported by the NIH (NICHD P01 HD084387 to C.A.W., N.A. and L.S.K.; and NIAMS R01AR067715 to C.G.) and the Texas Scottish Rite Hospital Research Fund (to C.A.W. and J.J.R.).

## ADDITIONAL INFORMATION

**Competing interests:** The authors declare no competing interests.

## REFERENCES

- Herring J. (ed.) *Tachdjian's Pediatric Orthopaedics*. 5th edn. (WB Saunders, Philadelphia, 2013).
- Martirosyan, N. L. et al. Genetic alterations in intervertebral disc disease. *Front Surg.* **3**, 59 (2016).
- Waldrop, R. et al. The burden of spinal disorders in the elderly. *Neurosurgery* **77** (Suppl 4), S46–S50 (2015).
- Rogala, E. J., Drummond, D. S. & Gurr, J. Scoliosis: incidence and natural history. A prospective epidemiological study. *J. Bone Joint Surg. Am.* **60**, 173–176 (1978).
- Herring, J. A. *Tachdjian's pediatric orthopaedics* (WB Saunders, Philadelphia, PA, 2008).
- Little, D. G., Song, K. M., Katz, D. & Herring, J. A. Relationship of peak height velocity to other maturity indicators in idiopathic scoliosis in girls. *J. Bone Joint Surg. Am.* **82**, 685–693 (2000).
- Katz, D. E., Herring, J. A., Browne, R. H., Kelly, D. M. & Birch, J. G. Brace wear control of curve progression in adolescent idiopathic scoliosis. *J. Bone Joint Surg. Am.* **92**, 1343–1352 (2010).
- Weinstein, S. L., Dolan, L. A., Wright, J. G. & Dobbs, M. B. Effects of bracing in adolescents with idiopathic scoliosis. *N. Engl. J. Med.* **369**, 1515–1521 (2013).
- Sanders, J. O. et al. Maturity assessment and curve progression in girls with idiopathic scoliosis. *J. Bone Joint Surg. Am.* **89**, 64–73 (2007).
- Wise, C. A. The genetic architecture of idiopathic scoliosis. in *Molecular Genetics of Pediatric Orthopaedic Disorders*. (eds Carol, A. & Wise, J. J. R.) 71–90 (Springer, New York, 2014).
- Riseborough, E. J. & Wynne-Davies, R. A genetic survey of idiopathic scoliosis in Boston, Massachusetts. *J. Bone Joint Surg. Am.* **55**, 974–982 (1973).
- Connor, J. M. et al. Genetic aspects of early childhood scoliosis. *Am. J. Med. Genet.* **27**, 419–424 (1987).
- Tang, N. L. et al. Genetic epidemiology and heritability of AIS: A study of 415 Chinese female patients. *J. Orthop. Res.* **30**, 1464–1469 (2012).
- Hresko, M. T. Clinical practice. Idiopathic scoliosis in adolescents. *N. Engl. J. Med.* **368**, 834–841 (2013).
- Carter, C. O. & Evans, K. A. Inheritance of congenital pyloric stenosis. *J. Med. Genet.* **6**, 233–254 (1969).
- 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 Joint Surg. Am.* **94**, 1485–1491 (2012).
- Antonarakis, S. E., Chakravarti, A., Cohen, J. C. & Hardy, J. Mendelian disorders and multifactorial traits: the big divide or one for all? *Nat. Rev. Genet.* **11**, 380–384 (2010).
- Pourquie, O. Vertebrate segmentation: from cyclic gene networks to scoliosis. *Cell* **145**, 650–663 (2011).
- Dunwoodie SaSD. Genetic and environmental interaction in malformation of the vertebral column. in *Molecular genetics of pediatric orthopaedic disorders*. (eds Carol, A. & Wise, J. J. R.) 131–151 (Springer, New York, 2015).
- Turnpenny, P. D. Classification and etiologic dissection of vertebral segmentation anomalies. in *Molecular genetics of pediatric orthopaedic disorders*. (eds Carol, A. & Wise, J. J. R.) 105–130 (Springer, New York, 2015).
- Liao, B. K. & Oates, A. C. Delta-Notch signalling in segmentation. *Arthropod Struct. Dev.* **46**, 429–447 (2017).
- Goto, H., Kimmey, S. C., Row, R. H., Matus, D. Q. & Martin, B. L. FGF and canonical Wnt signaling cooperate to induce paraxial mesoderm from tailbud neuromesodermal progenitors through regulation of a two-step epithelial to mesenchymal transition. *Development* **144**, 1412–1424 (2017).
- Smith, L. J., Nerurkar, N. L., Choi, K. S., Harfe, B. D. & Elliott, D. M. Degeneration and regeneration of the intervertebral disc: lessons from development. *Dis. Model Mech.* **4**, 31–41 (2011).
- Chan, W. C., Au, T. Y., Tam, V., Cheah, K. S. & Chan, D. Coming together is a beginning: the making of an intervertebral disc. *Birth Defects Res. C. Embryo Today* **102**, 83–100 (2014).
- Aszodi, A., Chan, D., Hunziker, E., Bateman, J. F. & Fassler, R. Collagen II is essential for the removal of the notochord and the formation of intervertebral discs. *J. Cell Biol.* **143**, 1399–1412 (1998).
- Sivakamasundari, V. et al. A developmental transcriptomic analysis of Pax1 and Pax9 in embryonic intervertebral disc development. *Biol. Open.* **6**, 187–199 (2017).
- Balling, R., Deutsch, U. & Gruss, P. undulated, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of Pax 1. *Cell* **55**, 531–535 (1988).
- Wallin, J. et al. The role of Pax-1 in axial skeleton development. *Development* **120**, 1109–1121 (1994).
- Wilm, B., Dahl, E., Peters, H., Balling, R. & Imai, K. Targeted disruption of Pax1 defines its null phenotype and proves haploinsufficiency. *Proc. Natl Acad. Sci. USA.* **95**, 8692–8697 (1998).
- Adham, I. M. et al. The scoliosis (sco) mouse: a new allele of Pax1. *Cytogenet Genome Res.* **111**, 16–26 (2005).
- Kokubu, C. et al. A transposon-based chromosomal engineering method to survey a large cis-regulatory landscape in mice. *Nat. Genet.* **41**, 946–952 (2009).
- Takimoto, A. et al. Differential transactivation of the upstream aggrecan enhancer regulated by PAX1/9 depends on SOX9-driven transactivation. *Sci. Rep.* **9**, 4605 (2019).



33. Lefebvre, V., Angelozzi, M. & Haseeb, A. SOX9 in cartilage development and disease. *Curr. Opin. Cell Biol.* **61**, 39–47 (2019).
34. Rodrigo, I., Hill, R. E., Balling, R., Munsterberg, A. & Imai, K. Pax1 and Pax9 activate Bapx1 to induce chondrogenic differentiation in the sclerotome. *Development* **130**, 473–482 (2003).
35. Panda, D. K., Miao, D., Lefebvre, V., Hendy, G. N. & Goltzman, D. The transcription factor SOX9 regulates cell cycle and differentiation genes in chondrocytic CFK2 cells. *J. Biol. Chem.* **276**, 41229–41236 (2001).
36. Bi, W. et al. Haploinsufficiency of Sox9 results in defective cartilage primordia and premature skeletal mineralization. *Proc. Natl Acad. Sci. USA.* **98**, 6698–6703 (2001).
37. Henry, S. P., Liang, S., Akdemir, K. C. & de Crombrughe, B. The postnatal role of Sox9 in cartilage. *J. Bone Min. Res.* **27**, 2511–2525 (2012).
38. Bi, W., Deng, J. M., Zhang, Z., Behringer, R. R. & de Crombrughe, B. Sox9 is required for cartilage formation. *Nat. Genet.* **22**, 85–89 (1999).
39. Liu, C. F. & Lefebvre, V. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. *Nucleic Acids Res.* **43**, 8183–8203 (2015).
40. Han, Y. & Lefebvre, V. L-Sox5 and Sox6 drive expression of the aggrecan gene in cartilage by securing binding of Sox9 to a far-upstream enhancer. *Mol. Cell Biol.* **28**, 4999–5013 (2008).
41. Lefebvre, V., Li, P. & de Crombrughe, B. A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. *EMBO J.* **17**, 5718–5733 (1998).
42. Lefebvre, V., Behringer, R. R. & de Crombrughe, B. L-Sox5, Sox6 and Sox9 control essential steps of the chondrocyte differentiation pathway. *Osteoarthr. Cartil.* **9**(Suppl A), S69–S75 (2001).
43. Smits, P. & Lefebvre, V. Sox5 and Sox6 are required for notochord extracellular matrix sheath formation, notochord cell survival and development of the nucleus pulposus of intervertebral discs. *Development* **130**, 1135–1148 (2003).
44. Smits, P. et al. The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev. Cell.* **1**, 277–290 (2001).
45. Lefebvre, V. Toward understanding the functions of the two highly related Sox5 and Sox6 genes. *J. Bone Min. Metab.* **20**, 121–130 (2002).
46. Smits, P., Dy, P., Mitra, S. & Lefebvre, V. Sox5 and Sox6 are needed to develop and maintain source, columnar, and hypertrophic chondrocytes in the cartilage growth plate. *J. Cell Biol.* **164**, 747–758 (2004).
47. Ikeda, T. et al. The combination of SOX5, SOX6, and SOX9 (the SOX trio) provides signals sufficient for induction of permanent cartilage. *Arthritis Rheum.* **50**, 3561–3573 (2004).
48. Akiyama, H., Chaboissier, M. C., Martin, J. F., Schedl, A. & de Crombrughe, B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev.* **16**, 2813–2828 (2002).
49. Sivakamasundari, V., Kraus, P., Jie, S. & Lufkin, T. Pax1(EGFP): new wildtype and mutant EGFP mouse lines for molecular and fate mapping studies. *Genesis* **51**, 420–429 (2013).
50. Dimeglio, A. & Canavese, F. The growing spine: how spinal deformities influence normal spine and thoracic cage growth. *Eur. Spine J.* **21**, 64–70 (2012).
51. Yu, J. et al. Microfibrils, elastin fibres and collagen fibres in the human intervertebral disc and bovine tail disc. *J. Anat.* **210**, 460–471 (2007).
52. Roughley, P. J. Biology of intervertebral disc aging and degeneration: involvement of the extracellular matrix. *Spine.* **29**, 2691–2699 (2004).
53. Roughley, P. J., Alini, M. & Antoniou, J. The role of proteoglycans in aging, degeneration and repair of the intervertebral disc. *Biochem. Soc. Trans.* **30**(Pt 6), 869–874 (2002).
54. Cappello, R., Bird, J. L., Pfeiffer, D., Bayliss, M. T. & Dudhia, J. Notochordal cell produce and assemble extracellular matrix in a distinct manner, which may be responsible for the maintenance of healthy nucleus pulposus. *Spine.* **31**, 873–882 (2006).
55. Bruehlmann, S. B., Hulme, P. A. & Duncan, N. A. In situ intercellular mechanics of the bovine outer annulus fibrosus subjected to biaxial strains. *J. Biomech.* **37**, 223–231 (2004).
56. Mouw, J. K., Ou, G. & Weaver, V. M. Extracellular matrix assembly: a multiscale deconstruction. *Nat. Rev. Mol. Cell Biol.* **15**, 771–785 (2014).
57. Frantz, C., Stewart, K. M. & Weaver, V. M. The extracellular matrix at a glance. *J. Cell Sci.* **123**(Pt 24), 4195–4200 (2010).
58. Hynes, R. O. & Naba, A. Overview of the matrisome—an inventory of extracellular matrix constituents and functions. *Cold Spring Harb. Perspect. Biol.* **4**, a004903 (2012).
59. Wu, J. J., Weis, M. A., Kim, L. S., Carter, B. G. & Eyre, D. R. Differences in chain usage and cross-linking specificities of cartilage type V/XI collagen isoforms with age and tissue. *J. Biol. Chem.* **284**, 5539–5545 (2009).
60. Aspberg, A. The different roles of aggrecan interaction domains. *J. Histochem. Cytochem.* **60**, 987–996 (2012).
61. Urban, J. P. & Winlove, C. P. Pathophysiology of the intervertebral disc and the challenges for MRI. *J. Magn. Reson. Imaging* **25**, 419–432 (2007).
62. Wang, Z., Weitzmann, M. N., Sangadala, S., Hutton, W. C. & Yoon, S. T. Link protein N-terminal peptide binds to bone morphogenetic protein (BMP) type II receptor and drives matrix protein expression in rabbit intervertebral disc cells. *J. Biol. Chem.* **288**, 28243–28253 (2013).
63. McKenna, L. A., Liu, H., Sansom, P. A. & Dean, M. F. An N-terminal peptide from link protein stimulates proteoglycan biosynthesis in human articular cartilage in vitro. *Arthritis Rheum.* **41**, 157–162 (1998).
64. Mecham, R. P. & Gibson, M. A. The microfibril-associated glycoproteins (MAGPs) and the microfibrillar niche. *Matrix Biol.* **47**, 13–33 (2015).
65. Craft, C. S. MAGP1, the extracellular matrix, and metabolism. *Adipocyte* **4**, 60–64 (2015).
66. Altshuler, D., Daly, M. J. & Lander, E. S. Genetic mapping in human disease. *Science* **322**, 881–888 (2008).
67. Khanshour, A. M. et al. Genome-wide meta-analysis and replication studies in multiple ethnicities identify novel adolescent idiopathic scoliosis susceptibility loci. *Hum. Mol. Genet.* **27**, 3986–3998 (2018).
68. Sharma, S. et al. A PAX1 enhancer locus is associated with susceptibility to idiopathic scoliosis in females. *Nat. Commun.* **6**, 6452 (2015).
69. Langenhan, T., Aust, G. & Hamann, J. Sticky signaling-adhesion class g protein-coupled receptors take the stage. *Sci. Signal.* **6**, re3 (2013).
70. Kou, I. et al. Genetic variants in GPR126 are associated with adolescent idiopathic scoliosis. *Nat. Genet.* **45**, 676–679 (2013).
71. Shukunami, C. et al. Chondrogenic differentiation of clonal mouse embryonic cell line ATDC5 in vitro: differentiation-dependent gene expression of parathyroid hormone (PTH)/PTH-related peptide receptor. *J. Cell Biol.* **133**, 457–468 (1996).
72. Ikegawa, S. Genomic study of adolescent idiopathic scoliosis in Japan. *Scoliosis Spinal Disord.* **11**, 5 (2016).
73. Monk, K. R. et al. G protein-coupled receptor is essential for Schwann cells to initiate myelination. *Science* **325**, 1402–1405 (2009).
74. Liebscher, I. et al. A tethered agonist within the ectodomain activates the adhesion G protein-coupled receptors GPR126 and GPR133. *Cell Rep.* **9**, 2018–2026 (2014).
75. Petersen, S. C. et al. The adhesion GPCR GPR126 has distinct, domain-dependent functions in Schwann cell development mediated by interaction with laminin-211. *Neuron* **85**, 755–769 (2015).
76. Philippova, M. et al. A guide and guard: the many faces of T-cadherin. *Cell Signal.* **21**, 1035–1044 (2009).
77. Fredette, B. J., Miller, J. & Ranscht, B. Inhibition of motor axon growth by T-cadherin substrata. *Development* **122**, 3163–3171 (1996).
78. Brohmann, H., Jagla, K. & Birchmeier, C. The role of Lbx1 in migration of muscle precursor cells. *Development* **127**, 437–445 (2000).
79. Gross, M. K., Dottori, M. & Goulding, M. Lbx1 specifies somatosensory association interneurons in the dorsal spinal cord. *Neuron* **34**, 535–549 (2002).
80. Miyake, A. et al. Identification of a susceptibility locus for severe adolescent idiopathic scoliosis on chromosome 17q24.3. *Plos One* **8**, e72802 (2013).
81. Gonzalez-Sandoval, A. & Gasser, S. M. On TADs and LADs: spatial control over gene expression. *Trends Genet.* **32**, 485–495 (2016).
82. Ogura, Y. et al. A functional variant in MIR4300HG, the host gene of microRNA MIR4300 is associated with progression of adolescent idiopathic scoliosis. *Hum. Mol. Genet.* **26**, 4086–4092 (2017).
83. Manolio, T. A. et al. Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
84. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
85. Chatterjee, S. & Ahituv, N. Gene regulatory elements, major drivers of human disease. *Annu Rev. Genomics Hum. Genet.* **18**, 45–63 (2017).
86. Verstraeten, A., Alaerts, M., Van Laer, L. & Loeys, B. Marfan syndrome and related disorders: 25 years of gene discovery. *Hum. Mutat.* **37**, 524–531 (2016).
87. Loeys, B. L. et al. The revised Ghent nosology for the Marfan syndrome. *J. Med. Genet.* **47**, 476–485 (2010).
88. Haller, G. et al. A polygenic burden of rare variants across extracellular matrix genes among individuals with adolescent idiopathic scoliosis. *Hum. Mol. Genet.* **25**, 202–209 (2016).
89. Spranger, J. The type XI collagenopathies. *Pediatr. Radiol.* **28**, 745–750 (1998).
90. Krakow, D. Skeletal dysplasias. *Clin. Perinatol.* **42**, 301–319 (2015).
91. Nebert, D. W. Z. L. SLC39A8 gene encoding a metal ion transporter: discovery and bench to bedside. *Hum. Genet.* **13**, 51 (2019).
92. Haller, G. et al. A missense variant in SLC39A8 is associated with severe idiopathic scoliosis. *Nat. Commun.* **9**, 4171 (2018).
93. Baschal, E. E. et al. Idiopathic scoliosis families highlight actin-based and microtubule-based cellular projections and extracellular matrix in disease etiology. *G3.* **8**, 2663–2672 (2018).
94. Baschal, E. E. et al. Exome sequencing identifies a rare HSPG2 variant associated with familial idiopathic scoliosis. *G3.* **5**, 167–174 (2014).

95. Viljoen, D. & Beighton, P. Schwartz-Jampel syndrome (chondrodystrophic myotonia). *J. Med. Genet.* **29**, 58–62 (1992).
96. Will, R. E., Stokes, I. A., Qiu, X., Walker, M. R. & Sanders, J. O. Cobb angle progression in adolescent scoliosis begins at the intervertebral disc. *Spine.* **34**, 2782–2786 (2009).
97. Long, F., Zhang, X. M., Karp, S., Yang, Y. & McMahon, A. P. Genetic manipulation of hedgehog signaling in the endochondral skeleton reveals a direct role in the regulation of chondrocyte proliferation. *Development.* **128**, 5099–5108 (2001).
98. Karner, C. M., Long, F., Solnica-Krezel, L., Monk, K. R. & Gray, R. S. Gpr126/Adgrg6 deletion in cartilage models idiopathic scoliosis and pectus excavatum in mice. *Hum. Mol. Genet.* **24**, 4365–4373 (2015).
99. Gorman, K. F. & Breden, F. Idiopathic-type scoliosis is not exclusive to bipedalism. *Med. Hypotheses* **72**, 348–352 (2009).
100. Boswell, C. W. & Ciruna, B. Understanding idiopathic scoliosis: a new zebrafish school of thought. *Trends Genet.* **33**, 183–196 (2017).
101. Gray, R. S. et al. Loss of col8a1a function during zebrafish embryogenesis results in congenital vertebral malformations. *Dev. Biol.* **386**, 72–85 (2014).
102. Hayes, M. et al. ptk7 mutant zebrafish models of congenital and idiopathic scoliosis implicate dysregulated Wnt signalling in disease. *Nat. Commun.* **5**, 4777 (2014).
103. Van Gennip, J. L. M., Boswell, C. W. & Ciruna, B. Neuroinflammatory signals drive spinal curve formation in zebrafish models of idiopathic scoliosis. *Sci. Adv.* **4**, eaav1781 (2018).
104. Grimbacher, B. et al. Hyper-IgE syndrome with recurrent infections—an autosomal dominant multisystem disorder. *N. Engl. J. Med.* **340**, 692–702 (1999).
105. Hawary, R. E., Zaaroor-Regev D., Floman Y., Lonner B. S., Alkhalife Y. I., Betz R. R. Brace treatment in adolescent idiopathic scoliosis: risk factors for failure—a literature review. *Spine J.* **19**, 1917–1925 (2019).
106. Sampara, P., Banala, R. R., Vemuri, S. K., Av, G. R. & Gpv, S. Understanding the molecular biology of intervertebral disc degeneration and potential gene therapy strategies for regeneration: a review. *Gene Ther.* **25**, 67–82 (2018).
107. Choi, H., Johnson, Z. I. & Risbud, M. V. Understanding nucleus pulposus cell phenotype: a prerequisite for stem cell based therapies to treat intervertebral disc degeneration. *Curr. Stem Cell Res. Ther.* **10**, 307–316 (2015).
108. Dunbar, C. E. et al. Gene therapy comes of age. *Science.* **359**, (2018). <https://doi.org/10.1126/science.aan4672>.
109. Loibl, M. et al. Controversies in regenerative medicine: should intervertebral disc degeneration be treated with mesenchymal stem cells? *JOR Spine* **2**, e1043 (2019).
110. Tibiletti, M., Gregar Velikonja, N., Urban, J. P. & Fairbank, J. C. Disc cell therapies: critical issues. *Eur. Spine J.* **23**(Suppl 3), S375–S384 (2014).
111. Kou, I. et al. Genome-wide association study identifies 14 previously unreported susceptibility loci for adolescent idiopathic scoliosis in Japanese. *Nat. Commun.* **10**, 3685 (2019).



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020