

Supplementary Information

Variability in proteoglycan biosynthetic genes reveals new facets of heparan sulfates diversity. A systematic review and analysis

Mohand Ouidir Ouidja, Denis S.F. Biard^{a,b}, Sandrine Chantepie^a, Xavier Laffray^a, Gael Le Douaron^a, Minh-Bao Huynh^a, Nicolas Rebergue^a, Auriane Maïza^a, Karla Rubio^{a,c,1}, Oscar González-Velasco^d, Guillermo Barreto^{a,c,1}, Javier De Las Rivas^d, and Dulce Papy-Garcia^{a,*}

HSPG biosynthetic genes review

The HSPG biosynthetic machinery genes were here reviewed following their allocation to Groups I, II, III, or IV. PRISMA guidelines¹ were used to search information on clinical variants for each human gene and for the corresponding null mouse. A second campaign was performed to retrieve information on each gene product functionally, including enzymes substrate selectivity.

	Page
GROUP I. Essential widely expressed HSPG biosynthetic genes	
Group I – Essential widely expressed HSPG core proteins coding genes	
<i>HSPG2</i>	3
<i>AGRN</i>	3
<i>COL18A1</i>	4
<i>GPC3</i>	4
<i>GPC4</i>	5
<i>GPC6</i>	5
Group I – Essential widely expressed GAG-CP linker biosynthetic genes	
<i>XYLT1</i>	6
<i>XYLT2</i>	7
<i>B4GALT7</i>	8
<i>FAM20B</i>	8
<i>B3GALT6</i>	9
<i>B3GAT3</i>	10
<i>PXYLP1</i>	11
<i>CANT1 and SLC10A7</i>	11
<i>SLC26A2, SCL35A2, and SLC35A3</i>	12

Group I – Essential widely expressed HS chain biosynthetic genes	
<i>EXTL3</i>	13
<i>EXT1</i> and <i>EXT2</i>	13
<i>NDST1</i>	14
<i>GLCE</i>	14
<i>HS2ST1</i>	15
<i>HS6ST1</i>	16
Group I – Ubiquitous HS chain remodeling machinery genes:	
<i>SULF2</i>	16
GROUP II. Essential tissue restricted HSPG biosynthetic genes	
Group II – Essential tissue restricted HS remodeling control:	
<i>HPSE2</i>	17
GROUP III. Non-essential widely expressed HSPG biosynthetic genes	
Group III - Non-essential widely expressed HSPG core proteins	
<i>GPC1</i>	18
<i>SDC1</i>	19
<i>SDC2</i>	19
<i>SDC3</i>	20
<i>SDC4</i>	20
Group III – Non-essential widely expressed HS chain remodeling genes	
<i>EXTL2</i>	20
<i>SULF1</i>	21
<i>HPSE</i>	21
GROUP IV. Non-essential tissue restricted HSPG biosynthetic genes	
Group IV – Non-essential tissue restricted HSPG core proteins	
<i>GPC2</i>	22
<i>GPC5</i>	22
Group IV – Non-essential tissue restricted specialized HS chain maturation	
<i>NDST2</i>	23
<i>NDST3</i> and <i>NDST4</i>	23
<i>HS6ST2</i>	24
<i>HS6ST3</i>	24
<i>HS3ST1</i>	25
<i>HS3ST2</i>	26
<i>HS3ST3A1</i>	26
<i>HS3ST3B1</i>	27
<i>HS3ST4</i>	27
<i>HS3ST5</i>	28
<i>HS3ST6</i>	28
Group IV – Tissue restricted HS chain remodeling genes	
<i>EXTL1</i>	29

GROUP I. Essential widely expressed HSPG biosynthetic genes

Group I – Essential widely expressed HSPG core proteins coding genes

HSPG2 (Chr.1; ID:3339)

Heparan sulfate proteoglycan 2 (*HSPG2*) is a widely expressed (Fig. 4) and highly conserved gene that codes for the secreted perlecan core protein, which can carry over four long chains of HS². Perlecan is the major HSPG in most basement membranes, which are the highly specialized extracellular matrices that tie the epithelium to connective, endothelial, muscular, neuronal, and other tissues³. During the early stages of embryogenesis, perlecan deposits in vascularized tissues such as cartilage primordia, pericardium, blood vessels walls, and heart². At later stages of development, perlecan is widely expressed in other tissues including lung, kidney, liver, gastrointestinal tract, and brain⁴. The highest deposition of perlecan occurs in tissues that are rich in extracellular matrix, as in cartilage undergoing endochondral ossification⁴. In human, *HSPG2* pathogenic variability leads to bone dysplasia and to several other defects including cardiovascular and pulmonary phenotypes, myotonia, and hyperthermia. These defects range from relatively mild conditions (Stuve-Wiedemann syndrome and the Schwartz–Jampel syndrome) to severe and/or lethal diseases (Lethal Kniest-Like Syndrome and Dyssegmental Dysplasia of Silverman-Handmaker type)², with a degree of severity inversely correlating with the amount of perlecan being deposited into the extracellular matrix⁵. Accordingly with the most severe phenotypes, *Hspg2*^{-/-} mice perinatally die with skeletal dysplasia and/or defective cephalic development and several other traits⁴. The perlecan phenotype in the knockout mice agrees its biological role in the pericellular space in where the proteoglycan binds, cross-links, and stabilizes many molecules, including extracellular matrix structural proteins, growth factors, morphogens, and cell-surface receptors⁵. Because of its essential role in normal embryonic development and its wild tissue distribution, *HSPG2* was HERE allocated to the Group I of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

AGRN (Chr.1; ID:375790)

Agrin (*AGRN*) is a widely expressed gene (Fig. 4) that codes for the secreted core protein agrin, predominantly present in the basement membranes that join muscle fibers and nerves at the neuromuscular junction³, although the protein is also present in other basement membranes and extracellular matrices^{3,6}. In agreement with its essential role in normal embryonic development, particularly at the neuromuscular junction, *AGRN* clinical variants cause congenital myasthenic syndromes including distal muscle weakness and atrophy⁷, and SNP has been associated to myasthenia gravis⁸, a rare autoimmune disorder affecting the neuromuscular junction. Accordingly, *Agrn*^{-/-} mice died at birth due to respiratory failure and central nervous system defects^{6,9}. The

involvement of agrin in the organization of basement membranes other than those in the neuromuscular junction was revealed by the selective re-expression of the gene in motor neurons of the *Agrn*^{-/-} mice, in which *Agrn* re-expression could prevent perinatal death, although mice exhibited severe postnatal growth retardation with markedly altered growth plate morphology, consistent with the implication of this HSPG in basal membranes during musculoskeletal development^{3,6}. Because of its essential role in embryonic development and wide distribution in basement membranes, particularly at the neuromuscular junction, *AGRN* was allocated to the Group I of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

COL18A1 (Chr 21; ID:80781)

Collagen XVIII is a secreted HSPG core protein coded by the ubiquitously expressed *COL18A1* gene, which is predominantly expressed in basement membranes of epithelial and vascular tissues¹⁰ (Fig. 4). In human, 3 collagen XVIII variants differing in their *N*-terminal non-collagenous sequences are known, all of them carry at least three HS chains. Mutations in *COL18A1* result in Knobloch syndrome, a genetic disorder characterized by severe eye defects (e.g. myopia, vitreoretinal degeneration, macular abnormalities) and occipital encephalocele¹¹. Accordingly, lack of the collagen XVIII in the *Col18a1*^{-/-} mice results in a deleterious phenotype affecting eye and encephalocele¹², although it also affects base membranes of many other tissues including kidney proximal tubules, heart valves, epidermis, choroid plexus¹³, as supported by increasing evidences in situations such as atherosclerosis, glomerulonephritis, and in other type of tissue damage^{10,11} in which collagen XVIII seems to participate to the regulation of cell survival, stem or progenitor cell maintenance and differentiation during inflammation. Because of its extensive tissue distribution and of the severe overt phenotypes observed in human clinical variants and knockout mice, *COL18A1* was allocated to the Group I of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

GPC3 (Chr.X; ID:2719)

GPC3 encodes the widely expressed (Fig. 4) membrane-associated glypican 3 core protein¹⁴, which clinical variants are responsible of the X-linked Simpson-Golabi-Behmel syndrome, characterized by pre and post-natal overgrowth with skeletal, facial, visceral, and neurological anomalies¹⁵. These abnormalities are recapitulated in the *Gpc3*^{-/-} mice, which exhibited perinatal death with developmental dysmorphisms including abnormal lung and kidney¹⁶. Other *GPC3* variants are responsible of Wilms tumor 1 (nephroblastoma), the most common kidney cancer in children, and with hepatocellular carcinoma¹⁷. Because of its high levels, glypican 3 has been proposed as biomarker and target for nephroblastoma and hepatocellular carcinoma¹⁸. Indeed, glypican 3 plays important biological roles by interacting and regulating the function of HBP including BMPs, WTNs, SHHs, and FGFs, among others¹⁷. As other glypicans, glypican 3 carries a *N*-terminus signal domain, a globular intermediary

domain containing multiple cysteine residues, and a C-terminus domain carrying the GPI that anchors the core protein to the cell surface. HS chains are attached to a C-terminus short domain that keeps the sugar chain close to cell membrane areas rich in lipids, allowing the concentration of HBP near to their high affinity cell surface receptors that allow cell signaling. Because of the overt phenotype of its genetic variants in human and knockout mice, *GPC3* was allocated to the Group I of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

GPC4 (Ch.X; ID:2239)

GPC4 encodes the membrane associated HSPG glypican 4 core protein, widely expressed during development and adulthood (Fig. 4). Interestingly, *Gpc4* was found enriched in the hippocampus of the developing mice, suggesting a central involvement in brain development¹⁹. In human, *GPC4* pathogenic variants cause Keipert Syndrome, characterized by prominent forehead, flat midface, hypertelorism, broad nose, digital abnormalities, and deafness, indicating a role of this HSPG in the regulation of morphogen activities and in perceptive impairment²⁰. Accordingly, the *Gpc4*^{-/-} mice reproduce the main traits of Keipert syndrome, including craniofacial and digital abnormalities²⁰, as well as behavioral alterations associated to abnormal levels of GluA1 and AMPARs at the synapsis²¹. In agreement with its implication in cognition and social behavior, an altered expression of *GPC4* has been linked to autism spectrum disorders (ASD)^{21,22}. Because of its wide distribution in tissues and the deleterious morphogenic and cognitive phenotypes in human clinical variants and knockout mice, *GPC4* was allocated to the Group I of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

GPC6 (Ch13; ID:10082)

GPC6 is a widely expressed gene that encodes the HSPG glypican 6 core protein¹⁹ (Fig. 4). In human, *GPC6* loss-of-function mutations result in an autosomal-recessive condition known as omodysplasia 1 (OMOD1), characterized by facial dysmorphism, short stature, and shortened limbs²³. In agreement with its requirement for normal embryonic development, *Gpc6*^{-/-} mice die at birth with severe facial dysmorphism and bones significantly shorter than those of their normal littermates. This phenotype has been related to a signaling deficit of the morphogenic HBP sonic hedgehog²⁴. Interestingly, a role of *GPC6* during adult bone homeostasis is currently supported by genetic wide association studies (GWAS) showing that SNP adjacent to the *GPC6* coding sequences is a risk factor for osteoporosis²⁵. Moreover, altered *GPC6* expression has been implicated in melanoma²⁶ and prostate cancer²⁷, indicating a role of glypican 6 in tumor progression possibly through its capacity to assure growth factors cell signaling¹⁷. Because of the deleterious phenotypes of its clinical variants in human and lethality of the knockout mice, *GPC6* was allocated to the Group I of genes involved in HSPG biosynthesis (Fig. 4 and Fig. 5).

Group I – Essential widely expressed GAG-CP linker biosynthetic genes

Based on their wide expression profiles and overt phenotypes caused by clinical variants in human and in knockout mice (Fig. 4), the GAG-CP linker tetrasaccharide biosynthetic genes *XYLT1*, *XYLT2*, *B4GALT7*, *FAM20B*, *B3GALT6*, *B3GAT3*, *PXYLP1*, *CANT1* and *SLC10A7*, were clustered in Group I. Because the GAG-CP linker tetrasaccharide supports both HS and CS chains in proteoglycans, it would be expected that defects in the GAG-CP linker biosynthetic machinery will lead to phenotypes associated to both GAGs, unless a preferred pathway favor the synthesis of a proteoglycan type over the other. To evaluate this possibility, we paid particular attention to literature reports on substrates specificities and products of the enzymes involved in the GAG-CP linker biosynthetic machinery. This analysis indicates that different alternative paths coexist during the GAG-CP biosynthesis, and that although some of these paths overlap, the biosynthesis of HSPG is preferably assured over that of CSPG. This can explain why the primarily observed phenotypes affect tissues classically rich in the non-preferred CSPG products. Interestingly, synthesis of the linker tetrasaccharide that will carry CS chains is itself spliced in different paths (Fig. 6A), likely assuring the biosynthesis of different CSPG types (not deeply treated in this analysis). The predominance of some pathways over others was inferred from data on substrate specificities and from the clinical traits associated to genetic variability on each biosynthetic enzyme, as reviewed in the next paragraphs.

XYLT1 (Chr.16; ID:64131)

Xylosyl transferase 1, coded by *XYLT1*, is a widely expressed ion binding enzyme (Fig. 4), which at the cis Golgi catalyzes the transfer of Xyl from the UDP-Xyl donor to specific serine residues of the core protein (Ser/CP)^{28,29}. During development, *XYLT1* is highly expressed, particularly in chondrocytes, which are the skeletal precursor cells that produce high levels of CSPG within the cartilage template that forms bones²⁸⁻³⁰. The critical role of the highly expressed *XYLT1* during cartilage template ossification is suggested by the severe skeletal phenotype observed in humans carrying *XYLT1* mutations, which are responsible of the skeletal Desbuquois dysplasia type 2 (DBQD-2), also called Baratela–Scott syndrome. DBQD-2 is characterized by short stature, advanced carpal ossification, distinct facial features, joint laxity, and intellectual disabilities^{30,31}. Analysis of GAGs from primary fibroblasts of DBQD-2 patients have shown an important reduction of large CSPG, as aggrecan, and unchanged HSPG levels, indicating maintenance of HSPG path over the secreted large CSPG-aggrecan (Fig. 6A). Reproduction of one of the patient's mutation in a cell line confirmed the reduction of large CSPG without affecting neither HS levels, nor small CSPGs (as decorin) levels. Indeed, *XYLT1* interacts with decorin with higher affinity than with aggrecan, suggesting that *XYLT1*

can more efficiently transfer Xyl to Ser residues in the small-CSPG core protein than in the large one^{32,33}. Accordingly, independent kinetic studies have shown a XYLT1 higher affinity for small-CSPG core proteins, whereas XYLT2 has similar affinity for small and large cores in both CSPG and for HSPG^{29,30}. This suggests that clinical variants resulting in reduced XYLT1 activity during development, when XYLT2 is not expressed, will mainly affect large-CSPG (as aggrecan), leading to a strong chondrocytic phenotype. Although the *Xytl1*^{-/-} mouse has not been generated possibly because of failure to develop, a pug mouse in which a homozygous missense mutation in *Xytl1* was produced by a *N*-ethyl-*N*-nitrosourea mutagenesis screen, showed a strong skeletal phenotype with premature maturation of chondrocytes, early ossification, and disproportionated dwarfism²⁸. Summarizing, XYLT1 can assure synthesis of both HSPG and CSPG, although with preferred production of small-CSPG and HSPG. As *XYLT1* is critically required for normal embryonic development and its genetic variability results in deleterious phenotypes in human and mice³¹, this gene was placed in the Group I of GAG-CP linker biosynthetic machinery genes common to HSPG and CSPG (Fig. 4 and Fig. 5).

XYLT2 (Chr.17; ID:64132)

Xylosyl transferase 2 (XYLT2), coded by *XYLT2*, is an ion binding protein expressed at post developmental stages. XYLT2 participates to the production of the Xyl-(Ser/CP) pool after birth²⁹, when XYLT1 expression decreases. Whereas *XYLT1* mutations cause severe developmental defects, mutations in *XYLT2* only cause post-developmental disorders, as spondyloocular syndrome, an autosomal recessive condition characterized by osteoporosis, cataracts, sensorineural hearing loss, learning difficulties, and atrial septal defect that appears after birth^{31,34}, translating deficiency of small-CSPG (decorin and neurocan paths affected) (Fig. 6A). Accordingly, *Xytl2*^{-/-} mice display no developmental defects, but show post-natal overt phenotypes particularly in organs in which *Xytl1* expression is the lower, as in liver and adipose tissues (Fig. 4)^{32,33}, but not only³⁵. XYLT2 shows a similar preference for the small and large CSPG and HSPG core proteins²⁸⁻³⁰ (Fig. 6A). Kinetically, XYLT2 glycosylates HSPG and CSPG core proteins with similar efficiency³⁰, suggesting that after birth XYLT2 equality assures the production of Xyl-(Ser/CP) pool, when *XYLT1* expression is reduced to basal levels (Fig. 6A). After birth, XYLT2 can then participate to HSPG synthesis in organs in which the XYLT1 expression is the lowest^{32,35}. Interestingly, XYLT2 activity can be measured in serum of patients with pathologies as systemic sclerosis and for the diagnosis of fibrotic remodeling processes³¹, suggesting a weak XYLT2 anchoring to the Golgi membrane, although the physiopathological implication of this event remains to be investigated. Currently, several single *XYLT2* pathogenic variants are reported, although none is lethal, suggesting that XYLT1 can assure CP-Ser xylosylation in most organs. The overt phenotype observed in mice lacking *Xytl2* and its pathogenic variants in

human place *XYLT2* in the Group I of HSPG (and CSPG) biosynthetic machinery genes essential for normal embryonic development and physiology (Fig. 4 and Fig. 5).

B4GALT7 (Chr. 5; GeneID:11285)

B4GALT7 encodes the ubiquitous cis-Golgi manganese ion binding β -1,4-galactosyltransferase 7 (*B4GALT7*) (Fig. 4), also known as galactosyltransferase I (GALT-I), which shows high UDP-galactose (UDP-Gal) substrate specificity³⁶. *B4GALT7* transfers the first galactose (Gal-I) to the growing Xyl-(Ser/CP) to form Gal-Xyl-(Ser/CP) (Fig. 6A). Because *B4GALT7* is the only enzyme that uses the Xyl-(Ser-CP) pool, synthesized by *XYLT1* from early developmental stages, the *B4galt7* knockout mice cannot be generated and genetic variability in human can be expected to result in a *XYLT1*-like developmental phenotype, although with stronger effect in the small-CSPG compared to *XYLT1*-variants, particularly after birth, when *XYLT1* decreases. Accordingly, in human, compound heterozygous mutations in *B4GALT7* cause *B4GALT7*-related Ehlers-Danlos syndromes (*B4GALT7*-EDS), previously known as spondylodysplastic EDS type 1 (spEDS-1), characterized by alterations of connective tissue with joint hypermobility, skin hyperextensibility, and tissue fragility. In severe phenotypes, *B4GALT7*-EDS additionally shows short stature (skeletal tissue affections), muscle hypotonia, and other conditions which severities depend on the clinical variant^{37,38}. One of the more severe forms is Reunion Island Larsen-like syndrome (RILLS), characterized by sdEDS-1 traits with additional traits including multiple dislocations, facial dysmorphism, ocular abnormalities, motor and cognitive delays³⁷⁻³⁹, suggesting developmental alterations in both small- (decorin and neurocan) and large-CSPG paths, affecting HSPG at lower extent. Indeed, the clinical significances of *B4GALT7* variants range from benign to pathogenic^{37,38} depending on the different catalytic efficiencies of the mutated enzyme^{36,40}, with the more severe traits likely correlating with decreased HS levels, although this remains to be confirmed. Thus, because of its requirement for normal embryonic development, *B4GALT7* was placed in the Group I of HSPG (and CSPG) biosynthetic machinery genes (Fig. 4 and Fig. 5).

FAM20B (Chr.1; ID: 9917)

The golgian xylosylkinase *FAM20B* (*FAM20B*) is a member of the ubiquitously expressed 'Family with sequence similarity 20' (*FAM20*) proteins (Fig. 4). *FAM20B* adds a phosphate group on the C2 of Xyl in Gal-Xyl-(Ser/CP) to form Gal-Xyl(P)-(Ser/CP), which is the preferred, although not exclusive, *B3GALT6* substrate that leads to Gal-Gal-Xyl(P)-Ser/CP, the preferred *B3GAT3* substrate^{41,42} (Fig. 4A). Phosphorylation of Xyl occurs in proteoglycans that follow the common HSPG/CSPG path, whereas the unphosphorylated Xyl is characteristic of the small-CSPG decorin path. Accordingly, the Xyl phosphorylated linker tetrasaccharide has only been observed in HSPGs and in some CSPGs including aggrecan and neurocan, but not in decorin^{40,42} (Fig. 6A). Therefore, *FAM20B* has been

proposed to act as a molecular switch that outline the identity of the GAG chain that will be elongated in the GAG-CP linker⁴². To date, there is only one reported *FAM20B* compound heterozygous variant, which affected two siblings causing the death of one of them shortly after birth, the second one died at early infancy⁴³. The two siblings showed a severe phenotype like *XYLT1*-Desbuquois dysplasia without joint laxity and very short stature, multiple dislocations of the large joints, mid-face, and thoracic hypoplasia with respiratory failure⁴³. This phenotype evokes a defective HSPG/large-CSPG path (as aggrecan and neurocan)⁴⁰, but not defective small-CSPG decorin path⁴⁰, as proposed in Fig. 6A. More severe phenotypes evoking disruption of HSPG in the HSPG/large-CSPG path are evoked by the embryonal lethality of *Fam20b*^{-/-} mouse, which showed severe stunted growth, multisystem organ hypoplasia and a delayed development most evident in the skeletal system, eyes, lung, gastrointestinal tract, and liver. Therefore, *FAM20B* was here allocated to the Group I of HSPG (and CSPG) biosynthetic machinery genes essential for normal embryonic development and tissue homeostasis (Fig. 4 and Fig. 5).

B3GALT6 (Chr.1; ID: 126792)

The β -1,3-galactosyltransferase 6 (*B3GALT6*), also known as galactosyltransferase II (*GALTII*), encoded by the *B3GALT6* intron-less gene, is an enzyme present in the medial Golgi, where it catalyzes the transfer of the second Gal (Gal II) to Gal-Xyl(\pm P)-(Ser/CP) to produce Gal-Gal-Xyl(\pm P)-(Ser/CP), with higher selectivity for the phosphorylated substrate^{41,42}, favoring the HSPG/large-CSPG path (Fig. 6A). As *B3GALT6* can produce both the phosphorylated and non-phosphorylated GAG-CP linker intermediates, it can act in both the HSPG/CSPG path and the small-CSPG decorin path (Fig. 6A). Accordingly, cultured fibroblasts from patients carrying biallelic *B3GALT6* mutations showed reduced production of both HSPG and CSPG^{44,45}. In human, *B3GALT6* mutations cause a wide phenotypic spectrum similar to that associated to *B4GALT7*, so called *B3GALT6*-related spondylodysplastic Ehlers-Danlos syndrome (*B3GALT6*-spEDS2 or spEDS-2) or spondyloepimetaphyseal dysplasia with joint laxity⁴⁶. As other EDS, these conditions affect connective tissues and skeletal tissues, including bone, skin, cartilage, tendons, and ligaments. Indeed, depending on the genetic variant, *B3GALT6*-spEDS2 shows overt phenotypes with variable severities involving joint hypermobility, short stature, muscle hypotonia, kyphoscoliosis, craniofacial defects, hyperextensible skin, delayed motor and/or cognitive development, joint contractures, ascending aortic aneurysm, and spondyloepimetaphyseal dysplasia^{44,45}. Accordingly, analysis of patients with compound heterozygous mutations showed reduced levels of the *B3GALT6* transcript or protein, while others show protein mislocalisation. Interestingly, deficient *B3GALT6* activity was recently shown to produce a non-canonic linker tetrasaccharide lacking the second Gal moiety but able to elongate GAG chains, although with altered structures⁴⁷. Although *B3galt6* knockout mice have not been produced, possibly because a defective HSPG/CSPG path incompatible with development and life, and because

of the deleterious phenotype of *B3GALT6* genetic variants in human, this gene was placed in the Group I of HSPG (and CSPG) biosynthesis machinery genes (Fig. 4 and Fig. 5).

B3GAT3 (Chr.11; ID: 26229)

B3GAT3 codes for β -1,3-glucuronyltransferase 3 (B3GAT3), also known as glucuronosyltransferase I (GLCATI), a ubiquitous metal ion binding enzyme which at the cis Golgi catalyzes GlcA transfer from UDP-GlcA into Gal(\pm S)-Gal-Xyl(\pm P)-(Ser/CP) to form GlcA-Gal(\pm S)-Gal-Xyl(\pm P)-(Ser/CP) in both the HSPG/CSPG and the small-CSPG paths (Fig. 6A). The critical role of this enzyme in development was confirmed by the early severe lethal phenotype observed during the *B3gat3*^{-/-} mouse generation, which resulted in embryonic lethality before the 8-cell stage due to failed cytokinesis with strong reduction of the synthesis of both HS and CS chains^{46,48}. In human, several *B3GAT3* clinical variants have been associated to *B3GAT3*-related syndromes affecting connective, skeletal, cardiac, and other tissues, with severities depending on the variant. As B3GAT3 is involved in the different GAG-CP linker biosynthetic paths, *B3GAT3* variants present overlapping overt phenotypes depending on the variant, including traits of Larsen, Antley-Bixler, Shprintzen-Goldberg, Geroderma osteodysplastica, and spEDS syndromes^{46,48}. Disease's traits combined to data on substrate selectivity and B3GAT3 affinities to different core protein, agree the overlaying GAG-CP linker synthetic paths (Fig. 6A) and corroborates that linkeropathies are a phenotypic continuum bridging EDS and skeletal disorders⁴⁸. For instance, it has been shown that phosphorylation of Xyl increases the affinity of B3GAT3 for its substrate in the common HSPG/CSPG pathway, while sulfation at C6 of Gal-II increases the affinity of the enzyme for its substrate in the small-CSPG pathway. As B3GAT3 is predominantly associated to PXYLP1 in the HSPG/CSPG path, dephosphorylation of the GlcA-Gal-Gal-Xyl(P)-(Ser/CP) linker immediately after its formation produces the substrate required for the HS chain initiation by EXTL3 (Fig. 6A and 6B). Alternatively, in the same path, a PXYLP1-free B3GAT3 fraction can also glycosylate Gal-Gal-Xyl(P)-(Ser/CP) to produce GlcA-Gal-Gal-Xyl(P)-(Ser/CP), which initiates CS elongation through the here called 'neurocan path', which was evidenced by the phosphorylated GAG-CP linker observed in neurocan⁴⁰. In a less privileged HSPG/CSPG-derived path, GlcA-Gal-Gal-Xyl(P)-(Ser/CP) is dephosphorylated and subsequently glycosylated by CS *N*-acetyl galactosamine transferase 1 (CSGALNACT1), promoting the here called 'aggrecan path', evidenced by the non-sulfated/non-phosphorylated GAG-CP linker that prevails in aggrecan⁴⁰. Interestingly, studies in substrate selectivity in the HSPG/CSPG indicate that the HSPG path is preferred over the 'aggrecan path'⁴⁰⁻⁴², in agreement with the predominant skeletal phenotype observed in most linkeropathies⁴⁸. Moreover, substrate selectivity at the FAM20B and BGALT6 involving steps indicates that as the HSPG/CSPG path is favored vs the 'CSPG-decorin' path. This can explain the predominant connective tissue phenotype characteristic of B3GAT3-related linkeropathies. Because of the central involvement of the

enzyme in normal embryonic development, *B3GAT3* was allocated to the Group I of HSPG (and CSPG) biosynthetic machinery genes (Fig. 4 and Fig. 5).

PXYLP1 (Chr.3; ID: 92370)

Phosphoxylose phosphatase 1 (*PXYLP1*) is a type II transmembrane phosphatase that dephosphorylate Xyl(P) residues in the GlcA-Gal-Gal-Xyl(P)-(Ser/CP) linker tetrasaccharide before HS or CS polymerization starts (Fig. 6A). *PXYLP1* transcripts are ubiquitously expressed (Fig. 4), suggesting the importance of this gene in homeostasis. In human, *PXYLP1* pathogenic variants have not been reported for the single gene, and *Pxylp1* knockout mice have not being generated. This might be due to lethality at early embryonic stages due to a blocked HS synthesis, or to non-pathologic phenotypes (less probable). Indeed, in the Golgi, *PXYLP1* interacts with either *B3GAT3* or with the *CSGALNACT1* to respectively promote initiation of HS or CS synthesis^{49,50} (Fig. 6A). Accordingly, HS initiation by *EXTL3* is blocked in the absence of *PXYLP1*⁵⁰, while CS elongation can still occur⁴⁹ (Fig. 6A). As abundancies of *B3GAT3* and *CSGALNACT1* determine their interactions with *PXYLP1*^{49,50}, the phosphatase might thus play important roles in determining whether a GAG chain will elongate in part-time proteoglycans⁴⁹. *PXYLP1* was here included in the Group I of ubiquitously expressed HSPG biosynthetic genes required for healthy life (Fig. 4 and Fig. 5).

Nucleotide regulators

CANT1 (Chr.17; ID:124583) and *SLC10A7* (Chr.4; ID: 84068)

CANT1 codes for the ubiquitous calcium activated nucleotidase-1 (*CANT1*) (Fig. 4), a calcium dependent enzyme that hydrolyses uridine diphosphate nucleotide (UDP) to produce uridine monophosphate²⁷ and inorganic phosphate (Pi) in the ER and in the Golgi (Fig. 6A). UDP nucleotides (UDP-Xyl, UDP-Gal, UDP-GlcA, UDP-GlcNAc, etc.) are the glycosylation reaction donors that are synthesized in the cytoplasm and transported into the Golgi by bidirectional transporters⁵¹. In the Golgi, after its production by glycosylation reactions, UDP must be hydrolyzed by Golgi nucleotidases, including *CANT1*, to produce Pi and UMP. UMP is a strong inhibitor of glycosyltransferases and must then be redirected to the cytosol. Thus, by avoiding the inhibition of glycosyltransferases including *XYLTs*, *CANT1* promotes the synthesis of the GAG-CP linker⁵¹. Thus, *CANT1* genetic variability can be expected to affect both HS and CS paths, leading to clinical traits like those produced by variants in the genes involved in the first steps of the GAG-CP linker synthesis⁵². Accordingly, in human, *CANT1* genetic variations result in Desbuquois dysplasia (DBQD), which regroups traits observed in the *XYLT1*-related DBQD-2 and in *B4GATT7*-Larsen syndrome, including severe prenatal and postnatal growth retardation, joint laxity, short extremities, and progressive scoliosis⁵². This disease phenotype is recapitulated in *Cant1*^{-/-} mice⁵³, suggesting the inhibition of both CSPG and HSPG/CSPG paths. Moreover, since *CANT1* is calcium dependent, it can be expected that an altered calcium

homeostasis, as that due to mutations in Ca^{2+} transporter *SLC10A7* (Fig. 6A), results in traits overlapping with those in *CANT1* (DBQD) and *XYLT1* (DBQD-2). Accordingly, the *Slc10a7*^{-/-} mice showed skeletal dysplasia with such traits⁵⁴. Thus, *CANT1* and *SLC10A7* were both included in the Group I of ubiquitously expressed HSPG biosynthetic machinery genes required for normal embryonic development (Fig. 4 and Fig. 5).

Other Golgi transporters

By compromising the levels of sugar and sulfate donors in the Golgi, genetic variability of nucleotide sugar transporters and sulfate transporters can be expected to dramatically alter GAG synthesis and consequently global tissue homeostasis^{51,55}. The sulfate ion (SO_4^{2-}) transporter *SLC26A2* (*SLC26A2*), functions as a $\text{SO}_4^{2-}/2\text{OH}^-$, $\text{SO}_4^{2-}/2\text{Cl}^-$, and $\text{SO}_4^{2-}/\text{OH}^-/\text{Cl}^-$ exchanger at the cell membrane, and supplies SO_4^{2-} for the biosynthesis of the glycosaminoglycan sulfate donor 3'-phosphoadenosine-5'-phosphosulfate (PAPS) in the cytosol before its transfer to the Golgi. *SLC26A2* clinical variants can then be expected to decrease PAPS levels in the Golgi, affecting GAG from the first steps of their biosynthesis. Accordingly, *SLC26A2* clinical variants are responsible of inborn genetic skeletal disorders, which severities depend on the *SLC26A2* mutations and vary from mild recessive multiple epiphyseal dysplasia (malformations of hands, feet, and knees with possible scoliosis)⁵⁶ to perinatally lethal atelosteogenesis 2 (chondrodysplasia with dysmorphic syndrome) passing through diastrophic dysplasia (dwarfism with short legs and arms) and achondrogenesis 1B (chondrodysplasia with severe micromelia)⁵⁷. Similarly, genetic variability in any of the transporters that assure UDP-sugars levels in the Golgi, as *SLC35A* (UDP-GlcNAc and UDP-Gal transporter) and *SCL35A3* (UDP-GlcNAc transporter) (Fig. 6), can affect the biosynthesis of GAGs and of several other glycosylated molecules depending on the tissue n where they are expressed. For instance, *SCL35A2* (UDP-GlcNAc transporter) variants can cause congenital disorders with CNS-related symptoms such as epilepsy and autism⁵⁸.

Group I – Essential widely expressed HS chain biosynthetic genes

Elongation and maturation of a whole HS chain requires the action of glycosyl transferases, sulfotransferases, and epimerase. Without a priori and based on the phenotypes of their clinical variants and profiles of expression (Fig. 4), *EXTL3*, *EXT1*, *EXT2*, *GLCE*, *NDST1*, *HS2ST1*, and *HS6ST1* were included in the Group I of HSPG biosynthetic machinery genes. Interestingly, these ubiquitously expressed genes can by themselves assure the synthesis of a complete HS chain by assuring chain initiation, elongation, epimerization, and sulfation at relevant positions (*N*, 2-*O*- and 6-*O*-sulfation). It is then possible that, because of their ubiquitous expression, these genes assure

production of 'ubiquitous HS chains' (uHS) carrying the sulfation patterns dictated by NDST1, HS2ST1, and HS6ST21 substrate specificities, as detailed below.

Heparan chain initiation

EXTL3 (Chr.8; ID:2137)

The *Exostosin like glycosyltransferase 3* (*EXTL3*) gene encodes the type II integral membrane protein EXTL3 that catalyzes the transfer of a GlcNAc residue from UDP-GlcNAc to the non-reducing end of the GAG-CP linker tetrasaccharide GlcA-Gal-Gal-Xyl-(Ser/CP) to form GlcNAc-GlcA-Gal-Gal-Xyl-(Ser/CP), which initiates HS biosynthesis^{59,60} (Fig. 6A and 6B). EXTL3 is the only known initiator of HS chain in all organisms, and it has been suggested that its alternative splicing contribute to the regulation of HS vs CS levels^{60,61}. Indeed, in case of defective EXTL3 activity or when the enzyme is mis-localized due to genetic variability, GlcA-Gal-Gal-Xyl-(Ser/CP) accumulates [and possibly GlcA-Gal-Gal-Xyl(P)-(Ser/CP) in severe variants], leading to CS elongation^{49,60,61} (Fig. 6A). Accordingly, Human *EXTL3* genetic variability results in decreased HS synthesis with increased CS production⁶⁰. Clinical phenotypes associated to *EXTL3* mutations include skeletal dysplasia, developmental delay, severe epispondylo-metaphyseal dysplasia with immunodeficiency, and in most severe cases neuromotor and brain development delay⁶⁰. In accord with the fundamental role of EXTL3 in life, *Extl3*^{-/-} mice fail to produce HS chains and die during early embryonic development^{60,62}. Therefore, *EXTL3* was allocated to the Group I of ubiquitous HSPG biosynthetic genes essential for normal embryonic development and homeostasis (Fig. 4 and Fig. 5).

Heparan chain elongation

EXT1 (Chr.8; ID: 2131) and *EXT2* (Chr.11; ID: 2132)

Exostosin glycosyltransferase 1 (*EXT1*) and *exostosin glycosyltransferase 2* (*EXT2*) code for the ubiquitous type II integral membrane proteins EXT1 and EXT2, responsible of HS chain elongation in all tissues (Fig. 2). These enzymes act as a hetero-oligomeric complex that alternately transfer GlcA and GlcNAc to the growing HS chain. Both EXT1 and EXT2 are required for HS chain elongation and neither of the two proteins can substitute the other one. Indeed, it is likely that, while EXT1 assures the glycosyl transferase activity, EXT2 assists by assuring correct EXT1 folding and Golgian location^{59,63}. The essential role of EXT1 and EXT2 in normal development was shown in *Ext1/Ext2* deficient embryos, which failed to gastrulate and died in utero by embryonic day 7.5. However, compound heterozygous *Ext1*^{+/-};*Ext2*^{+/-} mice were viable but developed multiple osteochondromas by 2 months of age^{64,65}. Although chondrocytic defects were observed in that mice, several other defects also occur, as shown by the *Ext1* specific deletion in epithelial cells during lung development, which resulted in reduced branching number, expanded branching tips, and several mesenchymal

developmental defects⁶⁶. Indeed, defective EXT1/EXT2 leads to accumulation of the GlcNAc-GlcA-Gal-Gal-Xyl-(Ser/CP) substrate, which redirects the GAG synthesis through the CS pathway, enriching CS-aggreacan production (Fig. 6) and thus resulting in an increased CS/HS ratio, which results in an exacerbated differentiation of perichondrial progenitors into chondrocytes⁶³. Accordingly, in human, homozygous or compound heterozygous mutations in *EXT1* and/or *EXT2* are causative of hereditary multiple exostoses, a dominantly inherited genetic disorder characterized by the formation of cartilage-capped bony outgrowths at the epiphyseal growth plates, with a variable clinical phenotype depending on the variant⁶³. Interestingly, *EXT1*-related alterations of HS levels in tumor cells has been shown to participates to the progression of many cancers⁶⁷. Because of its severe clinical phenotypes and wide expression, *EXT1* and *EXTL2* were placed in the Group I of ubiquitous HSPG biosynthetic genes required for normal embryonic development and tissue homeostasis (Fig. 4 and Fig. 5).

HS chain maturation

NDST1 (Chr.5; ID: 3340)

N-deacetylase/N-sulfotransferase 1 (NDST1) codes for the widely expressed NDST1 enzyme that resides in the Golgi membrane and presents multiple isoforms resultant from alternative splicing⁶⁸. Among all *NDSTs*, only *NDST1* shows a ubiquitous expression (Fig. 4). NDST1 produces GlcNS residues by deacetylating GlcNAc and transferring a sulfate group from PAPS to GlcN. The enzyme can initiate *N*-sulfation at any internal GlcNAc residue, and sulfation proceeds from the nonreducing end toward the reducing end, to stop at least five saccharide units away from the GAG-CP linker or from previously existing GlcNS residues; this generates intermediary NA domains that separate GlcNS clusters⁶⁹. Interestingly, it has been established that the length of the *N*-sulfated clusters made by NDST1 is much shorter than that made by NDST2^{69,70}. In human, non-null homozygous *NDST1* mutations are known, in agreement with the phenotype of *Ndst1*^{-/-} mice, which perinatally die with immature lungs, delayed mineralization of the skeleton, vascular defects, craniofacial dysplasia, eye developmental defects, etc.⁷¹. Indeed, these mice produce HS with under sulfated structures, indicating that other NDSTs cannot compensate NDST1 function when the enzyme is completely missing. However, the possibility to live with partially active NDST1 is possible, as suggested by the varying pathologic phenotypes resultant from compound heterozygous *NDST1* mutations in human. For instance, combined heterozygous *NDST1* mutations have been associated to developmental delays, intellectual disability, muscular hypotonia, ataxia, history of a seizures, epilepsy, gastroesophageal reflux, and minor malformations^{72,73}. These predominant brain and muscle phenotypes can be explained by a partial compensation by NDST2 in most organs, but not in brain and skeletal muscle, in where NDST2 is absent or undetectable during adulthood (Fig. 4). For all these reasons, *NDST1* was allocated into the Group I of ubiquitous HSPG biosynthetic genes required for normal embryonic development (Fig. 4 and Fig. 5).

GLCE (Chr.15; ID: 26035)

D-glucuronyl C5-epimerase (GLCE) codes for C5-epimerase, the widely expressed type II integral membrane enzyme that epimerizes C5 in GlcA to form IdoA, which favors HS2ST recognition and confers to HS chains the flexibility required for their interaction with HBP ligands⁷⁴. C5-epimerase acts under a substrate control dictated by the distribution of GlcNS and GlcNAc residues flanking (+1 site) the GlcA that will be converted into IdoA^{75,76}. Enzyme kinetics have shown that C5-epimerisation can readily occur in GlcA residues flanked by GlcNS (-GlcNS-**GlcA**-GlcNS/GlcNAc-) to reversibly form IdoA (-GlcNS-**IdoA**-GlcNS/GlcNAc-), while presence of a GlcNAc residue at the non-reducing neighboring side (-GlcNAc-**IdoA**-GlcNS-**IdoA**-GlcNS/GlcNAc-) or of 2-O-sulfated IdoA (-GlcNS-**IdoA2S**-GlcNS-**IdoA**-GlcNS/GlcNAc-), avoids reaction reversibility and stabilizes the epimerized product^{75,76}. Indeed, this conversion occurs under a stringent substrate specificity in which 2-O-sulfation of IdoA, and/or 6-O- or 3-O-sulfation of GlcN units at subsite -2 to +2, leads to nonproductive substrate binding⁷⁵. Similarly, 6-O-sulfation in -GlcNS6S-**GlcA**-GlcNS6S- was shown to inhibit C-5 epimerase activity^{75,76}. In the other side, the association of C-5 epimerase and HS2ST, which assures 2-O-sulfation, allows the irreversible and thus stable generation of contiguous epimerized and 2-O-sulfated domains in the HS chain⁷⁷ (Fig. 6B). In agreement with an essential role in normal embryonic development, neither single *GLCE* homozygous null mutations nor heterozygote compound pathogenic variants are known in human, possible because lethal, as supported by the *Glce*^{-/-} mice, which neonatally die with severe developmental abnormalities affecting several organs including kidney, lung, skeleton, spleen, thymus, lymph node, etc.⁷⁴. However, SNP has been associated with hypertension and cerebrovascular events^{74,78}. The importance of *GLCE* in the regulation of cell behavior is also suggested by its potential tumor suppressor activity, for instance in breast and lung carcinogenesis⁷⁹. Because of its requirement for normal embryonic development and physiology, *GLCE* was placed in the Group I of ubiquitous HSPG biosynthetic genes essential for life (Fig. 4 and Fig. 5).

HS2ST1 (Chr.1; ID: 9653)

Heparan sulfate 2-O-sulfotransferase 1 (HS2ST1) codes for the ubiquitously expressed HS2ST type II integral membrane protein (Fig. 4), which transfers a sulfate group from PAPS to the C2 position of hexuronic acid residues in HS, with preference for IdoA over GlcA as a substrate^{80,81}. In the Golgi, HS2ST can interact with *GLCE* in a close functional relationship, which centrally contributes to the generation of IdoA2S residues in the HS chain (Fig. 6B). In human, *HS2ST1* clinical variants are unknown, possibly since not viable, as suggested by the *Hs2st1*^{-/-} mouse, which died at the neonatal period with developmental failure and an onset of abnormalities in several organs including kidney, which completely failed after mid-gestation⁸¹. Interestingly, although HS in organs of the *Hs2st1*^{-/-} mice lacked 2-O-sulfation, the overall HS charge density was maintained through increased *N*- and 6-O-

sulfations. However this did not rescue the mice developmental failure, in agreement with the importance of specific 2-O-sulfation in normal embryonic development. Interestingly, *HS2ST1* expression increases in adult tissues after injury, probably because the enzyme action favors the recovery of tissue homeostasis⁸². Thus, since HS2ST is a widely expressed and required for the synthesis of functional HS chains, *HS2ST1* was allocated into the Group I of ubiquitous HSPG biosynthetic genes required for normal embryonic development (Fig. 4 and Fig. 5).

HS6ST1 (Chr.2; ID: 9394)

Heparan sulfate 6-O-sulfotransferase 1 (HS6ST1) codes for the ubiquitously expressed HS6ST1, a type II integral membrane protein that transfers sulfate group from PAPS to C6 of GlcN in HexA-GlcNAc/NS residues to generate HexA-GlcNAc/NS6S along, and at the reducing extremities, the HS chain⁸³ (Fig. 6B). As 6-O-sulfation in -GlcNS6S-GlcA-GlcNS6S- inhibits C-5 epimerase activity on the GlcA motif^{75,76}, HS6ST1 favors formation of HS NA/NS domains, which characteristically display low content of IdoA2S residues. In human, *HS6ST1* clinical variants are unknown, possibly because carriers are not viable, as suggested by the *Hs6st1*^{-/-} mice, which failed to thrive or died at late embryonic and perinatal stages with growth retardation, lung defects, and abnormal angiogenesis, among other abnormalities except in kidney^{83,84}. Interestingly, the lack of noticeable abnormalities in the kidney of these mice suggests that HS6ST2, which is expressed in kidney at relatively high levels (Fig. 4), might compensate HS6ST1 lack^{83,84}, as recently suggested in mice with *Hs6st1* systemic inactivation⁸⁵. Moreover, *HS6ST1* SNPs have been associated to risk of delayed puberty and to idiopathic hypogonadotropic hypogonadism⁸⁶, and HS6ST1 and HS6ST2 were both proposed as candidate targets for ovarian failure during menopause⁸⁶. Irrespective to genetic variability, a critical role of *HS6ST1* is also suggested by a high *HS6ST1* mRNA observed in renal fibrosis associated with chronic rejection⁸⁷, idiopathic pulmonary fibrosis⁸⁸, and osteoarthritic cartilage⁸⁹. This suggest that while HS made by HS6ST1 are involved in the maintain of tissue homeostasis, they might not be suitable for an appropriate tissue response to injury, although this is speculative and remains to be investigated. Because of its ubiquitous tissue expression and critical role in normal embryonic development, *HS6ST1* was placed in the Group I of HSPG biosynthetic genes (Fig. 4 and Fig. 5).

Group I – Ubiquitous HS chain remodeling machinery genes

SULF2 (Chr.20; ID: 55959)

Sulfatase 2 (SULF2) encodes the broadly expressed extracellular HS endosulfatase 2 that is secreted to the extracellular matrix or localized in lipid rafts at the cell surface⁹⁰. SULF2 selectively removes 6-O-sulfate groups from HS in the extracellular space starting from the non-reducing end through the reducing end along highly sulfated NS domains in HS chains. Interestingly, by desulfating HS, SULF2 favors tumor growth, likely by enhancing cell signaling⁹⁰. The importance of *SULF2* in normal tissue

homeostasis was demonstrated in the *Sulf2*^{-/-} mouse, which showed partial embryonic lethality (~35% of mice) with developmental malformations particularly affecting the nervous system⁹¹. However, the double knockout *Sulf1*^{-/-}/*Sulf2*^{-/-} showed to be severely lethal, whereas the *Sulf1*-deficient mouse showed only mild phenotype. It remains to be clarified whether the partial lethality in *Sulf2*^{-/-} mice was associated to *Sulf1* compensatory mechanisms, functional co-operativity, limited redundancy, or all⁹². Interestingly, in human, genetic variations in *SULF2* have been reported as risk factors for altered regulation of lipoprotein metabolism⁹³ and therefore it has been proposed as a potential therapeutic target for atherogenic dyslipidemia in obesity and type 2 diabetes mellitus^{94,95}. Because of the lethality detected in *Sulf2*^{-/-} mice and for its ubiquitous expression, *SULF2* was placed in the Group I of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

Group II. Essential tissue restricted HSPG biosynthetic genes

Group II – Essential tissue restricted HS remodeling control

HPSE2 (Chr.10; ID: 60495)

Heparanase 2 (*HPSE2*) codes for the enzymatically inactive heparanase 2 that shows tissue restricted expression (Fig. 4) and close homology (40%) with the enzymatically active heparanase⁹⁶. Although heparanase 2 has no enzymatic activity, the protein displays higher affinity for HS than the enzymatically active heparanase and can function as its inhibitor⁹⁶. Under basal physiologic conditions, *HPSE2* is predominant expressed in the urinary track, followed by prostate and endometrium, with some expression in other tissues in which response to stimuli is central for proper function, including gastrointestinal and reproductive systems (Fig. 4). In human, biallelic mutations in the *HPSE2* gene have been associated with Urofacial or Ochoa Syndrome (UFS or UFOS), characterized by incomplete urinary bladder emptying and a facial grimace upon smiling, and which is also caused by mutation in the *LRG12* gene, which encodes leucine-rich-repeats and immunoglobulin-like-domains 2 and modulates growth factor signaling,^{97,98}. However, in up to 16% of UFS or UFOS patients, no *HPSE2* or *LRG12* associated mutations have been found⁹⁷, suggesting that *HPSE2* SNP contributes to a more complex polygenic variabilities leading to the altered condition. Accordingly, there is high clinical variability among patients with UFS and not all present *HPSE2* and *LRG12* significant variability. The involvement of *HPSE2* in larger phenomena is supported by genetic variability associations with presence of nocturnal lagophthalmos in some patients⁹⁶ and with the separation of migraine with and without depression⁹⁹. The essentialness of *Hpse2* in response to physiologic stimuli was confirmed in homozygous *Hpse2* gene trap mouse, which developed normally but showed deficient myogenic and neurogenic responses to chemical and electrical field stimulation.

Accordingly, mutant mice failed to gain weight after weaning and die when maintained for more than 1 month of age¹⁰⁰. *HPSE2* is the only gene allocated to the Group II of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

Group III. Non-essential widely expressed HSPG biosynthetic genes

Group III – Non-essential widely expressed HSPG core proteins

GPC1 (Chr.2; ID: 2817)

GPC1 encodes the widely expressed glypican 1 core protein (Fig. 4). Absence of pathogenic mutations in this gene agrees with the phenotype of the *Gpc1*^{-/-} mouse, which develops normally, is viable and fertile¹⁰¹, although showed subtle alterations on brain patterning as those observed in *GPC3*^{-/-} mice¹⁰². The involvement of *GPC1* in neurologic adaptive behavior is suggested by role *GPC1* polymorphism correlation with risk to schizophrenia when using brain activation as a quantitative phenotype¹⁰³. Variability in *GPC1* has also been observed in people susceptible to biliary atresia susceptibility¹⁰⁴. Besides a possible role in the regulation of tissue homeostasis, which remains to be established, glypican 1 was also involved in growth factor signaling in several cancer cells¹⁰⁵, including in pancreatic, prostate, breast and cervical cancers¹⁷. Because its wide expression and non-requirement for normal embryonic development, *GPC1* was allocated to the Group III of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

Syndecans

The *syndecan* (*SDC*) family of genes, which includes *SDC1*, *SDC2*, *SDC3*, and *SDC4*, codes for type I transmembrane proteins cores formed by an ectodomain, a transmembrane region and an intracellular domain. All syndecans have been involved in the regulation of cell-matrix interactions and in cell signaling¹⁰⁶. Moreover, syndecans can act as endocytosis receptors and are also involved in the uptake of exosomes. Interestingly, the intracellular C-terminus in the four syndecans harbors a unique EFYA sequence that binds to the PDZ-containing proteins, which contribute to the anchoring of transmembrane proteins to the cytoskeleton, holding together large protein complexes along the cell membrane¹⁰⁷. In the extracellular space, all syndecans can be shed into the pericellular environment through the action of matrix metalloproteinases (MMPs). Although ubiquitously expressed, any pathogenic *SDC* clinical variant has been reported in human and the available knockout mice develop normally, are viable, fertile, and lives normally, signifying a non-essential role for development and homeostasis. However, syndecans have shown to play important roles during tissue response to physiological or pathological stimuli, as in wound healing, inflammation, host defense to instigating agents, and during tumor cell growth and invasiveness¹⁰⁸. Because their wide

expression and non-requirement for normal embryonic development and homeostasis, the four *SDCs* were included in the Group III of HSPG biosynthetic machinery genes.

SDC1 (Chr.2; ID: 6382)

SDC1 codes for syndecan 1, a widely expressed HSPG core protein that shows predominant expression in tissues exposed to external stimuli, as the gastrointestinal track, skin, liver, etc. (Fig. 4). In human, non-clinical variants are known for *SDC1*, in agreement with the non-deleterious phenotype observed in the *Sdc1^{-/-}* mouse, which develops, lives and reproduces normally¹⁰⁹. However, *SDC1* genetic variables have been associated to increase susceptibility to multiple sclerosis¹¹⁰ and with severity and vulnerability of coronary plaque in patients with coronary artery disease¹¹¹. In agreement with a possible role in the regulation of biological events related to stimuli, the *Sdc1^{-/-}* mouse displays altered capacities to recover when challenged with instigating agents, as microbial infection, or ischemia, suggesting its involvement in adaptative functions. *In vitro*, syndecan 1 regulates cell adhesion, cell migration, endocytosis, exosome formation, etc.¹¹². These processes are central to tissue defense to aggression and to the control of fibrosis, inflammation, and re-vascularization^{113,114}. Accordingly, *SDC1* altered expression often impacts the clinical course of cancers, as evidenced by its prognostic significance¹¹³. An emerging new role for shed syndecan-1 is related to its ability to reach the nuclei in a variety of cells, including tumor cells¹¹⁵. Because its wide expression and non-requirement for normal embryonic development and physiology, and for its involvement in tissue adaptative responses to injury or disease instigating conditions, *SDC1* was allocated to the Group III of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

SDC2 (Chr.8; ID: 6383)

Syndecan 2 core protein is coded by *SDC2*, which is widely expressed (Fig. 4). Although *Sdc2^{-/-}* mice have not been generated, the absence of reports on pathogenic genetic variants in human suggests a non-essential role in normal development and physiology. However, a GWAS has shown association to risk of posttraumatic stress disorder in a cohort of Iraq-Afghanistan era veterans¹¹⁶. Moreover, syndecan 2 has been involved in the processes of fibrosis¹¹⁷, inflammation¹¹⁸, and cancer progression¹¹⁹. This might be related to the syndecan 2 role that assures connections between the cell cytoskeleton and the extracellular environment, particularly at the epithelial-mesenchymal interface¹¹⁹, and during formation of specialized membrane domains requiring redefinition during adaptative responses or during tissue recovery¹²⁰. Similarly, in brain, syndecan 2 is involved in the continuous redefinition of specialized membrane domains in dendritic filopodia and dendritic spines in responsive neurons^{121,122}. Because its wide expression, its non-requirement for normal embryonic development and physiology, and its central involvement in tissue or cell responsiveness, *SDC2* was allocated to the Group III of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

SDC3 (Chr.1; ID: 9672)

SDC3 is a ubiquitously expressed gene that shows some predominant expression in the immune and adrenal systems and in brain (Fig. 4). Although no *SDC3* pathogenic variants are known, *SDC3* polymorphism has been associated to a reduced risk to obesity¹²³, positive association have been found with metabolic syndrome¹²⁴, with body composition in patients with schizophrenia under treatment with antipsychotic drugs¹²⁵, and with female hyperandrogenism¹²⁶. Accordingly, the *Sdc3*^{-/-} mouse develops normally, is viable and fertile, but exhibits altered feeding behavior with resistance to obesity¹²⁷. This suggests that syndecan 3 might be involved in fine events related to different behavioral response, as satiety, and possibly to other responses in tissues in which the gene is expressed. Besides, *SDC3* is part of the clustering of genes involved in cancer progression¹²⁸. Because its wide expression with non-requirement for normal embryonic development, *SDC3* was allocated to the Group III of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

SDC4 (Chr.20; ID: 6385)

SDC4 is a ubiquitously expressed gene and is the major syndecan expressed in most tissues (Fig. 4). To date, *SDC4* clinical variants have not yet been reported, in agreement with phenotype of the *Sdc4*^{-/-} mouse, which is viable, fertile, and macroscopically indistinguishable from wild-type littermates¹²⁹. However, polymorphism has been associated with regulation of whole-body energy metabolism¹³⁰ and with longevity and lipid profile in healthy elderly¹³¹. As other syndecans, *SDC4* expression contributes to tissue response to stimuli or trauma, as observed during wound healing in *Sdc4*^{-/-} mice, which showed statistically significant delayed healing of skin wounds and impaired angiogenesis¹²⁹. Indeed, as other syndecans, syndecan 4 can mediate numerous cellular processes through signaling pathways that affect cellular proliferation, migration, mechano-transduction, and endocytosis¹⁰⁸. Because its wide expression and its non-requirement for normal embryonic development, but required for proper response to stimuli or trauma, *SDC4* was allocated to the Group III of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

Group III – Non-essential widely expressed HS chain remodeling genes

EXTL2 (Chr.1; ID: 2135)

The widely expressed *EXTL2* gene codes for the exostosin like glycosyltransferase 2 (*EXTL2*) enzyme that acts as a 'quality control' when the proteoglycan biosynthetic machinery is over-activated during loss of tissue homeostasis¹³². *EXTL2* specifically transfers a GlcNAc residue to the GlcA-Gal-Gal-Xyl(P)-Ser/CP tetrasaccharide to generate GlcNAc-GlcA-Gal-Gal-Xyl(P)-Ser/CP, which blocks both HS and CS biosynthesis. *Extl2* knockout mice develops normally, are viable and live healthy and non *EXTL2* clinical variants have been reported in human. However, after tissue insult, *EXTL2* favors tissue recovery by avoiding excessive CS and HS biosynthesis, as shown by impaired tissue

regeneration in mice models of liver and kidney failure in which *Extl2* was knocked out^{133,134}. Because EXTL2 is not essential for normal embryonic development, and since its activity is relevant when a tissue needs to recover, *EXTL2* was allocated to the Group III of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

SULF1 (Chr.8; ID: 23213)

Sulfatase 1 (SULF1) encodes the broadly expressed extracellular HS endosulfatase-1, which selectively removes 6-O-sulfate groups from HS. SULF1 is secreted through the Golgi to be subsequently localized at the cell surface⁹⁰. Interestingly, *Sulf1*^{-/-} mice show non-overt phenotypes, whereas *Sulf2*^{-/-} mice show partial lethality and *Sulf1*^{-/-}/*Sulf2*^{-/-} show severe lethality, suggesting that SULF1 can be compensated by SULF2⁹¹. In human, SULF1 gene polymorphisms have been associated to increased risk to multiple sclerosis¹¹⁰, fetus failure in IVF technique¹³⁵, and Preeclampsia¹³⁶. SULF1 has been involved in events in which tissue homeostasis is lost^{90,94,137}, as inflammation, wound healing, and cancer. Interestingly, in these processes SULF1 protects possibly by avoiding excessive accumulation of the ubiquitous NA/NS domains in HS that promote cell signaling. On the opposite, SULF2 is associated with decreased survival⁹⁰, possibly by the abnormal accumulation of NA/NS-like HS domains resulting from 6-O-desulfation of NS domains. Since *SULF1* is widely expressed and *Sulf1*^{-/-} mice are viable and evolve healthy, unless homeostasis is lost, this gene was placed in the Group III of genes involved in the maintain of tissue homeostasis, particularly favoring tissue recovery when homeostasis is loss. *SULF1* was allocated to the Group III of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HPSE (Chr.4; ID: 10855) and *HPSE2* (Chr.10; ID: 60495)

Heparanase (HPSE) is the only enzyme in mammals known to exert glycolytic cleavage of HS chains in the extracellular space¹³⁸. *HPSE* is ubiquitously expressed (Fig. 4) and several transcript variants encoding different isoforms have been found for this gene. Interestingly, its activity can be regulated by its inactive close homolog *heparanase-2 (HPSE2)*, which shows tissue restricted expression (Fig. 4). HPSE cleaves $\beta(1,4)$ -glycosidic bonds between GlcN residues and GlcA at a limited number of sites in HS¹³⁹. This participates to the remodeling of extracellular matrices and basement membranes, allowing cell movement and release of HBP (cytokines, growth factors and several other biological signals) from HS scaffolds in the extracellular space¹³⁸. Although widely distributed, HPSE exhibits only very low levels of expression, possibly because of its principal role during tissue remodeling processes, as wound healing, angiogenesis, inflammation. Accordingly, *Hpse*^{-/-} mice develop normally, are viable, anatomically normal, and fertile¹⁴⁰ and non-pathologic variants for this gene are known in human. However, gene variability has been found to increase risk of graft-versus-host disease, which is a major complication following allogeneic hematopoietic cell transplantation (HCT)

and occurs when donor T cells respond to histo-incompatible antigens on the host tissues^{141,142}. In agreement with a role in matrix remodeling, HPSE is overexpressed in pathological conditions where active matrix remodeling is highly active, such as fibrosis, inflammation, and angiogenesis, and in a variety of pathologies including kidney diseases, bone osteolysis, preeclamptic placenta, multiple sclerosis, autoimmune disease, viral infection, thrombosis, atherosclerosis, sepsis, diabetes, cancer progression and metastasis¹⁴³. Since HPSE is widely expressed and the *Hpse*^{-/-} mouse is viable and fertile, *HPSE* was allocated to the Group III of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

Group IV. Non-essential tissue restricted HSPG biosynthetic genes

Group IV – Non-essential tissue restricted HSPG core proteins

GPC2 (Chr.20; ID: 221914)

GPC2 encodes the glypican 2 core protein, which is predominantly expressed in brain and in some reproductive organs, particularly in endometrium and testis, although the gene is also expressed in other tissues but at low extent (Fig. 4). In human, *GPC2* clinical variants have not been reported to this date, and the *Gpc2* knockout mouse has not yet been generated. However, *GPC2* polymorphism is part of the polygenic scores for risk of developing Alzheimer's disease¹⁴⁴. Moreover, some data suggest that *GPC2* expression is required for the adaptive responses of tissues in which it is expressed¹⁴⁵. For instance, in brain, both *GPC2* expression and glypican 2 protein levels increase when neurogenesis is stimulated in the hippocampus of the adult brain and decrease when neurogenesis is selectively ablated in transgenic models¹⁴⁵. Accordingly, glypican 2 was proposed as a potential marker to monitor adult neurogenesis in cerebrospinal fluid¹⁴⁵. Interestingly, *GPC2* has been correlated with poor survival in neuroblastoma⁶⁶. Because its tissue restricted expression, and although its role in reproductive organs is still unknown, *GPC2* was included in the Group IV of HSPG biosynthetic genes (Fig. 4 and Fig. 5).

GPC5 (Chr.13; ID: 2262)

Glypican 5 is a HSPG core protein coded by the *GPC5* gene, which expression is restricted to only some tissues in the adult human body, particularly kidney and testis (Fig. 4). Although there is not *GPC5* clinical variants in human and the *Gpc5* knockout mice has not yet been generated, genetic variants in association with other genes variants, have been associated to vulnerability to acquired nephrotic syndrome¹⁴⁶, multiple sclerosis¹⁴⁷, autoimmune thyroid disease¹⁴⁸. Moreover, *GPC5* seems to be required for the adaptive responses to injury in the tissues in which it is expressed. For instance, *GPC5* was associated to podocyte response to injury and enhanced glypican 5 urinary levels were associated to nephropathy progression in type 2 diabetes patients³³. Besides, *SDC5* has been

proposed as a tumor suppressor in lung and prostate cancer^{149,150}. Because its tissue restricted expression, with possible requirement for response to injury, *GPC5* was included in the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

Group IV – Non-essential tissue restricted specialized HS chain maturation

NDST2 (Chr.10; ID: 8509)

NDST2, a gene expressed in several but not all tissues (Fig. 4), codes for the Golgian NDST2 enzyme, which exhibits dual *N*-deacetylase and *N*-sulfotransferase activities and is involved in the production of the large NS domains in HS chains^{59,68}. Under physiological conditions, NDST2 is not active in the presence of the ubiquitous NDST1, likely because in the presence of NDST1, NDST2 remains distant from the GAGosome and thus from its substrate^{68,151}. Accordingly, *Ndst2*^{-/-} mice are viable, fertile and show normal HS structures in all tissues, with exception of mast cells, in which NDST2 is predominantly expressed and heparin is synthesized¹⁵¹. In agreement with a non-essential role during normal development, no pathogenic *NDST2* variants are known in human. However, NDST2 SNP was found potentially associated to risk of coronary artery and chronic kidney diseases¹⁵². Interestingly, during tissue injury or during inflammation NDST2 becomes active¹⁵³, possibly because under stimuli the enzyme is integrated into the GAGosome, resulting in the production of longer sulfated domains carrying characteristic NS structures¹⁵¹. Although more research is required to determine how injury, inflammation, or other stimuli can favor NDST2 expression and function, *NDST2* was allocated to the Groupe IV of HS biosynthetic machinery genes (Fig. 4 and Fig. 5).

NDST3 (Chr.4; ID: 9348) and *NDST4* (Chr.4; ID: 64579)

N-deacetyl, *N*-sulfotransferases 3 and 4 (*NDST3* and *NDST4*) are both type II transmembrane enzymes that show highly restricted tissue expression, but that are particularly present in brain and spleen (Fig. 4), two organs highly receptive to immunological or behavioral stimuli, and/or, although at lower extent in other organs receptive to external stimuli, as stomach. *NDST3* and *NDST4* enzymes regulate HS fine structures with opposite actions, while *NDST3* exhibits high deacetylase and weak sulfotransferase activity (leading to GlcNH₂ carrying HS chains), *NDST4* shows weak deacetylase with high sulfotransferase activity (possibly to resulfate GlcNH₂). This might allow a molecular switch in high responsive tissues, as brain and spleen⁵⁹. In agreement with its restricted expression and non-direct involvement in central processes required for normal embryonic development and physiology, the *Ndst3*^{-/-} mouse develops normally, is fertile, and shows only subtle behavioral anomalies, with reduced anxiety and a slight hematological phenotype¹⁵⁴. In the other side, the *Ndst4*^{-/-} mouse exhibited no overt pathological phenotypes, although some anomalies were observed in histology images of the proximal colon in where the enzyme is not normally expressed; these images showed an increased number of goblet cells and a decreased number of colonocytes, in agreement with the

NDST4 proposed role in colorectal tumor growth¹⁵⁵. In human, non-pathogenic variants in *NDST3* or *NDST4* are known, but, in accord with their predominant expression in brain, *NDST3* genetic variants have been associated with schizophrenia and bipolar disorders^{156,157}, and variants in *NDST4* have been associated with reading disability and language impairment¹⁵⁸. However, confirmation of the involvement of *NDST3* and *NDST4* in these pathological conditions requires extensive research. Thus, *NDST3* and *NDST4* are not essential for normal embryonic development and homeostasis, but their lack in mice results in a slight cognitive and hematological phenotype and genetic variant might contribute to disease vulnerability. Because of their tissue restricted expression, *NDST3* and *NDST4* were included in the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HS6ST2 (Chr.X; ID: 90161)

HS6ST2, which codes for HS 6-O-sulfotransferase 2, depicts two alternative splicing variants in human, a full-length variant (*HS6ST2*) and a short length variant with 40 deleted amino acids (*HS6ST2S*). The major longer *HS6ST2* variant preferentially introduces sulfate groups at C6 of GlcNS motifs adjacent to IdoA2S (IdoA2S-GlcNS), generating the sulfate rich IdoA2S-GlcNS6S disaccharide present in HS NS domain^{84,159-161}. In the other side, the minor shorter variant shows no substrate specificity and can participate to the synthesis of both NS and NA/NS domains¹⁵⁹⁻¹⁶¹. The partially restricted expression of *HS6ST2* in tissues suggests specialized roles, particularly in kidney and in reproductive organs, as ovary, placenta, and endometrium, in which it is predominantly expressed (Fig. 4). Although the *Hs6st2* knockout mouse has not been generated, non-severe *HS6ST2* pathogenic variants are known in human. However, genetic variability has been reported as risk factor for intellectual disability linked to chromosome X¹⁶², in where the *HS6ST2* is located, suggesting that the NS domains produced by the enzyme can be involved in the regulation of discrete physiologic events enhanced fibroblast growth factor-mediated chondrocyte growth and differentiation¹⁶³, while *HS6ST2* downregulation by miR-23-3p enhanced matrix degradation in osteoarthritis¹⁶⁴, in which *HS6ST1* expression is enhanced⁸⁹. This supports different biological roles of the two *HS6STs* in homeostasis and during pathology. Interestingly, *HS6ST2* is found in body fluids, although the reason, mechanisms, and consequences of *HS6ST2* secretion remains to be elucidated. Moreover, *HS6ST2* overexpression has been associated with poor prognosis in patients with gastric cancer¹⁶⁵. Because of its partial tissue restricted expression, this gene was included in the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HS6ST3 (Chr.13; ID: 266722)

HS6ST3 codes for the tissue restricted HS 6-O-sulfotransferase 3 (*HS6ST3*), which expression is predominant in brain, although also found expressed at lower extent in some other tissues including the urinary bladder (Fig. 4). Interestingly, *HS6ST3* shows non substrate selectivity, as it acts on HexA-

GlcNAc/S residues to generate HexA-GlcNAc/S6S¹⁶⁰. Although the *Hs6st3* knockout mouse has not yet been generated, the absence of pathogenic variants in human suggests that the gene is not essential for embryonic development and physiology. However, the importance of discrete roles of HS6ST3 in the maintain of tissue homeostasis was suggested by GWAS associating polymorphism in loci near *HS6ST3* with risk to develop obesity¹⁶⁶, diabetic retinopathy¹⁶⁷. Moreover, HS6ST3 has also been involved in the growth and progression of breast cancer cells¹⁶⁸. Since *HS6ST3* shows high tissue restricted expression and non-clinical variants are known in human, this gene was allocated to the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HS3ST1 (Chr.4; ID: 9957)

HS3ST1 codes for the HS glucosamine 3-O-sulfotransferase 1 (HS3ST1), an enzyme that introduces a sulfate group from the sulfate donor PAPS into C3 of GlcN in GlcA-GlcNS±6S to generate GlcA-GlcNS**3S**±6S in HS chains. The enzyme tolerates IdoA, but 2-O-sulfation specifically prevents its action¹⁶⁹. HS3ST1 shows partial tissue distribution (Fig. 4), with transcripts predominantly expressed in urinary bladder, ovary, stomach, and in mastocytes, in where it participates to the production of the endogenous heparin¹⁷⁰. The addition of the 3-O-sulfate groups to GlcN units is a relatively rare modification, present in only a limited number of chains of HS in the body, except for heparin, which is a richly 3-O-sulfated HS (3S-HS) used in clinics as anticoagulant and mainly extracted from porcine, bovine, or ovine intestinal mucosa¹⁷¹. As the heparin anticoagulant activity is due to its high affinity for antithrombin, anticoagulant 3S-HS are called HS^{AT}¹⁷¹. However, although the *Hs3st1*^{-/-} mouse shows decreased HS^{AT} levels, this mouse is entirely viable and fertile and its coagulation is normal, although it showed very light intrauterine growth retardation that did not affect birth and normal life. Moreover, when the *Hs3st1*^{-/-} mice were submitted to pathologic stimuli, they exhibited a strong antithrombin-mediated proinflammatory phenotype¹⁷². Indeed, HS3ST1 controls the HS^{AT} production in endothelial cells, in where the enzyme plays a role in the regulation of the inflammo-modulatory activities of antithrombin¹⁷². In human, none *HS3ST1* pathogenic variant has been reported, in agreement with a non-essential role for normal development and physiology. However, HS3ST1 is centrally implicated in appropriate response to stimuli or injury in tissues or organs that express it¹⁷². Interestingly, in the human brain, genetic variability in *HS3ST1* carrying loci has been correlated with neurocognitive endophenotypes in a population at risk to develop mild cognitive impairment (MCI) and Alzheimer's disease¹⁷³⁻¹⁷⁵. Although the role of 3-O-sulfation in physiology is yet unknown, studies on 3S-HS binders have allowed identification of neuropilin-1, a heparin binding protein implicated in angiogenesis and axon guidance¹⁷⁶. Interestingly, HS3ST1 activity seems not to be compensated by other HS3STs¹⁷⁷ likely because a requirement of specific substrate precursor pools in cells expressing several HS3STs¹⁷⁷. Because the *HS3ST1* tissue restricted expression and as non-pathogenic variants are known in human, and as *Hs3st1*^{-/-} mice live healthy and fertile (their light intrauterine growth

retardation lacked pathological outcome) and requires stimuli to observe traits, *HS3ST1* was allocated to the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HS3ST2 (Chr.16; ID: 9956)

HS3ST2 codes for HS glucosamine 3-O-sulfotransferase 2 (*HS3ST2*), an enzyme predominantly expressed in brain and in less extent in lung and some other tissues¹⁶⁹ (Fig. 4). This protein belongs to the *HS3ST* family of enzymes that act during the last step of HS biosynthesis by transferring a sulfate group from PAPS to the C3 of GluA2S/I doA2S-GlcNS \pm 6S motifs to generate rare GluA2S/I doA2S-GlcNS3S \pm 6S motifs, which are different from those generated by *HS3ST1*¹⁷¹. 3S-HS carrying the rare I doA2S-GlcNS3S \pm 6S motif are called HS^{gD} since they serve as receptors for HSV1 entry into cells through interaction with the gD protein on the virus capsid^{171,178}. In human, non *HS3ST2* pathogenic variants have been reported, and *Hs3st2*^{-/-} mice are viable, healthy, and fertile (under submission), suggesting the non-requirement of this enzyme for normal development and physiology. However, the absence of deleterious phenotypes in human and in the knockout mouse could be due to transcriptional and functional redundancy with other *HS3STs* that also produce HS^{gD}, including *HS3ST3A*, *HS3ST3B*, *HS3ST4*, *HS3ST5*, and *HS3ST6*¹⁷¹, or because the enzyme activity is only relevant when an adequate stimuli is applied, as observed for other *HS3STs*¹⁷⁷. Redundancy is supported by the fact that HS^{gD} is made by several *HS3STs*, although the few data available in *HS3ST2* function suggest its involvement in discrete stimuli-related biological events. For instance, it has been reported that *Hs3st2* expression increases in the rat pineal gland during daylight and decreases at night, when it can be re-induced by light stimuli through adrenergic/cAMP signaling¹⁷⁹. Consistent with a role in brain, in where the enzyme is predominantly expressed, we have found that *HS3ST2* is implicated in tauopathy development¹⁸⁰. Interestingly, *HS3ST2* expression is regulated by methylation on the 5'UTR regions of its promoter, and controversy exists between a possible role as oncogene or as tumor suppressor¹⁸¹. Since *HS3ST2* shows tissue restricted expression, and as *Hs3st2* knockout mice develop normally, are viable and fertile, *HS3ST2* was allocated to the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HS3ST3A1 (Chr.17 ; ID: 9955)

HS3ST3A1 codes for HS glucosamine 3-O-sulfotransferase 3A1, a Golgian enzyme that transfers sulfate groups from PAPS to I doA \pm 2S-GlcN \pm S \pm 6S to generate I doA \pm 2S-GlcN \pm S3S \pm 6S in HS^{171,169}. *HS3ST3A1* shows partial tissue restricted expression with predominance in spleen (Fig. 4) and, although at lower extent, in reproductive tissues and in the glomerular basement membrane¹⁷⁷. *Hs3st3a1*^{-/-} mice have not yet been generated and non-pathogenic clinical variants have been reported in human, suggesting a non-essential role of this enzyme in normal development and homeostasis. However, *HS3ST3A1*, as other tissue restricted genes, might be implicated in the regulation of discrete

physiological events that might occur after stimuli in the tissues in which the gene is expressed, for instance in placenta, in where *HS3ST3A1* expression increases during normotensive and pre-eclamptic pregnancies¹⁸². As other HS3STs, *HS3ST3A1* has also been involved in HSV1 infection and genetic variability in the *HS3ST3A1* carrying locus has been associated to resistance to infection by *Plasmodium falciparum*^{183,184}, in agreement with the implication of 3S-HS during cells interactions with pathogens¹⁷¹. Similarly, as other HS3STs, *HS3ST3A1* expression is regulated by DNA methylation, and its association to malignant phenotypes has been proposed¹⁸¹. Since *HS3ST3A1* shows tissue restrictedly expression and no clinical variants are known in human, *HS3ST3A1* was allocated to the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HS3ST3B (Chr.17; ID: 9953)

HS glucosamine 3-O-sulfotransferase 3B1 (*HS3ST3B1*) is a Golgian enzyme that transfers 3-O-sulfates IdoA2S-GlcN±S±6S moieties in HS, generating IdoA2S-GlcN±S3S±6S sequences with non-anticoagulant activity, but that favor HSV1 infection¹⁷¹. *HS3ST3B* is expressed in several tissues with prevalence in liver, spleen, and lymph node, but lack in skeletal muscle, in where the highly restricted *HS3ST5* is expressed (Fig. 4)¹⁷¹. *Hs3st3b1* knockout mice have not yet been generated, but a non-deleterious phenotype might be supposed as non-pathogenic variants have been reported in human. However, similarly to *HS3ST3A1*, *HS3ST3B1* expression confers resistance to infection by *Plasmodium falciparum*¹⁸³⁻¹⁸⁵ and HSV1 infection¹⁷¹. Moreover, a role of *HS3ST3B1* during specific adaptative response to stimuli is suggested by its expression in peripheral T lymphocytes and Jurkat T cells. In these cells, cyclophilin B interacts with HS^{9D} on responsive peripheral blood T lymphocytes, triggering migration and integrin-mediated adhesion¹⁸⁶. Finally, as for *HS3ST3A1*, the enzyme affects tumor growth possible though an DNA methylation-related mechanism¹⁸¹. Because its restricted tissue expression, absence of clinical variants in human, and requirement for responsiveness to stimuli, *HS3ST3B1* was allocated to the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HS3ST4 (Chr.16; ID: 9951)

Heparan sulfate-glucosamine 3-O-sulfotransferase 4 (*HS3ST4*) is a Golgian enzyme that transfers sulfate groups to C3 of IdoA2S-GlcNS±6S moieties in HS, generating IdoA2S-GlcNS3S±6S sequences with non-anticoagulant activity but which, thought production of HS^{9D} sequences, favors HSV1 infection¹⁷¹. *HS3ST4* is one of the HS biosynthetic enzymes with the higher tissue restricted expression¹⁷⁸, predominantly in brain, in spleen, and in some reproductive organs, as testis and prostate (Fig. 4). Although *Hs3st4* knockout mice have not yet been reported, its non-deleterious phenotype is suggested by the fact that in human non-pathogenic variants are known. However, GWAS suggest that genetic variability in *HS3ST4* represent susceptibility risk for age-related macular degeneration¹⁸⁷. Similarly, in nondemented older people have shown *HS3ST4* SNP associations with

verbal declarative memory, suggesting a role in synaptic function^{188,189}. Accordingly with involvement in brain, variability on non-coding regions near the *HS3ST4* gene showed low but significant association with neuropathologic traits associated to tau protein burden in Alzheimer's disease¹⁸⁸. Interestingly, we recently found that *HS3ST4* is increased in Alzheimer's disease hippocampus¹⁹⁰. Although the physiologic role *HS3ST4* is still unknown, its tissue restricted expression suggest that it might play discrete functions, particularly in brain, in where it is predominantly expressed. As for other *HS3STs*, *HS3ST4* expression has also been observed in some cancer cells and its role as either tumor suppressor or oncogene is controverted¹⁸¹. Because of its strong tissue restricted expression and nonpathogenic variability in human, *HS3ST4* was allocated to the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HS3ST5 (Chr.6; ID: 222537)

HS3ST5 is a HS biosynthetic enzyme of the *HS3ST* family that, as *HS3ST4*, shows high tissue restricted expression (Fig. 4), predominantly in the central nervous system¹⁹¹, suggesting a role in neurologic functions. Interestingly, *HS3ST5* is the only tissue restricted sulfotransferase expressed in skeletal muscle, although in very low extent. *HS3ST5* produces HS^{AT} and HS^{9D} and has also been implicated in HSV1 infection¹⁷¹. *Hs3st5* knockout mice have not yet been reported and non-pathogenic *HS3ST5* mutations are known in human. However, GWAS have shown *HS3ST5* genetic polymorphism association with risk of microcephaly, cerebellar hypoplasia, and altered gray matter volume in patients with schizophrenia¹⁹². Because of its strong tissue restricted expression and non-pathogenic clinical variants, *HS3ST5* was allocated to the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

HS3ST6 (Chr.16; ID: 64711)

HS3ST6 is member of the type II membrane-bound protein *HS3ST* family of enzymes that show highly restricted expression. *HS3ST6* predominantly expressed in skin, and in considerably lower extent in some reproduction organs, lung, urinary bladder, and colon (Fig. 4)¹⁹³. Recently, whole exome sequencing studies have given evidence of *HS3ST6* SNP with traits characteristic of hereditary angioedema, which is a condition in which minor trauma or stress results in recurrent attacks of severe swelling of arms, legs, face, intestinal tract, and airways. Hereditary angioedema is characteristically due to mutations in *SERPING1*, a gene is involved in the regulation of the complement cascade, or in the *F12*, involved in the coagulation cascade. 3S-HS produced by *HS3ST6* was recently proposed to be involved in the response to stimuli that triggers angioedema formation¹⁹⁴. As *HS3ST6* is involved in response of tissue to stimuli, as trauma or stress, this gene was allocated to the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

Group IV – Tissue restricted HS chain remodeling

EXTL1 (Chr.1; ID: 2134)

EXTL1 encodes the α -1,4-*N*-acetylglucosaminyl transferase 1, also called exostosin like glycosyltransferase 1 (EXTL1), which catalyzes the addition of a GlcNAc to the non-reducing end of the growing HS chain, blocking HS chain elongation¹³². EXTL1 is the only EXTs/EXTL-type glycosyltransferase that shows tissue restricted expression, particularly in brain, spleen, muscle, and endometrium (Fig. 4). This might guaranty a stimuli-related discreet function in the tissue in which the gene is expressed. For instance, in spleen, the enzyme regulates the levels of HS, which likely function as a switch that controls the early maturation of B-cells, as shown in cultured B-cells of the transgenic mice overexpressing *Extl1*¹⁹⁵; in that cells, the stimuli responsive effect was attributed to reduced levels on membrane associated HS in the overexpressing *Extl1* B-cell lineage (heparin addition to cells had no effect). Moreover, genetic variability in this gene has been associated with muscle lipid composition suggesting a role in the regulation of lipid metabolism¹⁹⁶. Besides, EXTL1 was also involved in the control of the growth of some cancer cells¹⁹⁷. Although the knockout mouse has not been generated, no clinical variant has been reported for this gene, which is more likely involved in the adaptative response to stimuli, *EXTL1* was allocated to the Group IV of HSPG biosynthetic machinery genes (Fig. 4 and Fig. 5).

References

- 1 Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Rev Esp Cardiol (Engl Ed)* **74**, 790-799, doi:10.1016/j.rec.2021.07.010 (2021).
- 2 Martinez, J. R., Dhawan, A. & Farach-Carson, M. C. Modular Proteoglycan Perlecan/HSPG2: Mutations, Phenotypes, and Functions. *Genes* **9**, doi:10.3390/genes9110556 (2018).
- 3 McCarthy, K. J. The Basement Membrane Proteoglycans Perlecan and Agrin: Something Old, Something New. *Current topics in membranes* **76**, 255-303, doi:10.1016/bs.ctm.2015.09.001 (2015).
- 4 Arikawa-Hirasawa, E., Watanabe, H., Takami, H., Hassell, J. R. & Yamada, Y. Perlecan is essential for cartilage and cephalic development. *Nature genetics* **23**, 354-358, doi:10.1038/15537 (1999).
- 5 Gubbiotti, M. A., Neill, T. & Iozzo, R. V. A current view of perlecan in physiology and pathology: A mosaic of functions. *Matrix Biol* **57-58**, 285-298, doi:10.1016/j.matbio.2016.09.003 (2017).
- 6 Hausser, H. J., Ruegg, M. A., Brenner, R. E. & Ksiazek, I. Agrin is highly expressed by chondrocytes and is required for normal growth. *Histochemistry and cell biology* **127**, 363-374, doi:10.1007/s00418-006-0258-2 (2007).
- 7 Nicole, S. *et al.* Agrin mutations lead to a congenital myasthenic syndrome with distal muscle weakness and atrophy. *Brain : a journal of neurology* **137**, 2429-2443, doi:10.1093/brain/awu160 (2014).
- 8 Topaloudi, A. *et al.* Myasthenia gravis genome-wide association study implicates AGRN as a risk locus. *Journal of medical genetics*, doi:10.1136/jmedgenet-2021-107953 (2021).
- 9 Daniels, M. P. The role of agrin in synaptic development, plasticity and signaling in the central nervous system. *Neurochemistry international* **61**, 848-853, doi:10.1016/j.neuint.2012.02.028 (2012).

- 10 Seppinen, L. & Pihlajaniemi, T. The multiple functions of collagen XVIII in development and disease. *Matrix Biol* **30**, 83-92, doi:10.1016/j.matbio.2010.11.001 (2011).
- 11 Heljasvaara, R., Aikio, M., Ruotsalainen, H. & Pihlajaniemi, T. Collagen XVIII in tissue homeostasis and dysregulation - Lessons learned from model organisms and human patients. *Matrix Biol* **57-58**, 55-75, doi:10.1016/j.matbio.2016.10.002 (2017).
- 12 Fukai, N. *et al.* Lack of collagen XVIII/endostatin results in eye abnormalities. *The EMBO journal* **21**, 1535-1544, doi:10.1093/emboj/21.7.1535 (2002).
- 13 Rygh, C. B. *et al.* Image-based assessment of microvascular function and structure in collagen XV- and XVIII-deficient mice. *The Journal of physiology* **592**, 325-336, doi:10.1113/jphysiol.2013.263574 (2014).
- 14 Iglesias, B. V. *et al.* Expression pattern of glypican-3 (GPC3) during human embryonic and fetal development. *Histology and histopathology* **23**, 1333-1340, doi:10.14670/hh-23.1333 (2008).
- 15 Cottreau, E. *et al.* Phenotypic spectrum of Simpson-Golabi-Behmel syndrome in a series of 42 cases with a mutation in GPC3 and review of the literature. *American journal of medical genetics. Part C, Seminars in medical genetics* **163c**, 92-105, doi:10.1002/ajmg.c.31360 (2013).
- 16 Cano-Gauci, D. F. *et al.* Glypican-3-deficient mice exhibit developmental overgrowth and some of the abnormalities typical of Simpson-Golabi-Behmel syndrome. *The Journal of cell biology* **146**, 255-264, doi:10.1083/jcb.146.1.255 (1999).
- 17 Fico, A., Maina, F. & Dono, R. Fine-tuning of cell signaling by glypicans. *Cell Mol Life Sci* **68**, 923-929, doi:10.1007/s00018-007-7471-6 (2011).
- 18 Zhou, F., Shang, W., Yu, X. & Tian, J. Glypican-3: A promising biomarker for hepatocellular carcinoma diagnosis and treatment. *Medicinal research reviews* **38**, 741-767, doi:10.1002/med.21455 (2018).
- 19 Allen, N. J. *et al.* Astrocyte glypicans 4 and 6 promote formation of excitatory synapses via GluA1 AMPA receptors. *Nature* **486**, 410-414, doi:10.1038/nature11059 (2012).
- 20 Amor, D. J. *et al.* Pathogenic Variants in GPC4 Cause Keipert Syndrome. *American journal of human genetics* **104**, 914-924, doi:10.1016/j.ajhg.2019.02.026 (2019).
- 21 Dowling, C. & Allen, N. J. Mice Lacking Glypican 4 Display Juvenile Hyperactivity and Adult Social Interaction Deficits. *Brain plasticity (Amsterdam, Netherlands)* **4**, 197-209, doi:10.3233/bpl-180079 (2018).
- 22 Doan, R. N. *et al.* Mutations in Human Accelerated Regions Disrupt Cognition and Social Behavior. *Cell* **167**, 341-354.e312, doi:10.1016/j.cell.2016.08.071 (2016).
- 23 Campos-Xavier, A. B. *et al.* Mutations in the heparan-sulfate proteoglycan glypican 6 (GPC6) impair endochondral ossification and cause recessive omodysplasia. *American journal of human genetics* **84**, 760-770, doi:10.1016/j.ajhg.2009.05.002 (2009).
- 24 Capurro, M. *et al.* Glypican-6 promotes the growth of developing long bones by stimulating Hedgehog signaling. *The Journal of cell biology* **216**, 2911-2926, doi:10.1083/jcb.201605119 (2017).
- 25 Kemp, J. P. *et al.* Identification of 153 new loci associated with heel bone mineral density and functional involvement of GPC6 in osteoporosis. *Nature genetics* **49**, 1468-1475, doi:10.1038/ng.3949 (2017).
- 26 Li, Y. *et al.* Glypican 6 is a putative biomarker for metastatic progression of cutaneous melanoma. *PLoS one* **14**, e0218067, doi:10.1371/journal.pone.0218067 (2019).
- 27 Kumar, A. *et al.* Exome sequencing identifies a spectrum of mutation frequencies in advanced and lethal prostate cancers. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 17087-17092, doi:10.1073/pnas.1108745108 (2011).
- 28 Mis, E. K. *et al.* Forward genetics defines Xylt1 as a key, conserved regulator of early chondrocyte maturation and skeletal length. *Developmental biology* **385**, 67-82, doi:10.1016/j.ydbio.2013.10.014 (2014).
- 29 Ponighaus, C. *et al.* Human xylosyltransferase II is involved in the biosynthesis of the uniform tetrasaccharide linkage region in chondroitin sulfate and heparan sulfate proteoglycans. *The Journal of biological chemistry* **282**, 5201-5206, doi:10.1074/jbc.M611665200 (2007).
- 30 Bui, C. *et al.* XYLT1 mutations in Desbuquois dysplasia type 2. *American journal of human genetics* **94**, 405-414, doi:10.1016/j.ajhg.2014.01.020 (2014).

- 31 Gotting, C., Kuhn, J. & Kleesiek, K. Human xylosyltransferases in health and disease. *Cell Mol Life Sci* **64**, 1498-1517, doi:10.1007/s00018-007-7069-z (2007).
- 32 Condac, E. *et al.* Polycystic disease caused by deficiency in xylosyltransferase 2, an initiating enzyme of glycosaminoglycan biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 9416-9421, doi:10.1073/pnas.0700908104 (2007).
- 33 Sivasami, P. *et al.* Adipose tissue loss and lipodystrophy in xylosyltransferase II deficient mice. *International journal of obesity (2005)* **43**, 1783-1794, doi:10.1038/s41366-019-0324-1 (2019).
- 34 Taylan, F. *et al.* Spondyloocular Syndrome: Novel Mutations in XYLT2 Gene and Expansion of the Phenotypic Spectrum. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research* **31**, 1577-1585, doi:10.1002/jbmr.2834 (2016).
- 35 Ferencz, B. *et al.* Xylosyltransferase 2 deficiency and organ homeostasis. *Glycoconjugate journal* **37**, 755-765, doi:10.1007/s10719-020-09945-9 (2020).
- 36 Tsutsui, Y., Ramakrishnan, B. & Qasba, P. K. Crystal structures of beta-1,4-galactosyltransferase 7 enzyme reveal conformational changes and substrate binding. *The Journal of biological chemistry* **288**, 31963-31970, doi:10.1074/jbc.M113.509984 (2013).
- 37 Caraffi, S. G. *et al.* Severe Peripheral Joint Laxity is a Distinctive Clinical Feature of Spondylodysplastic-Ehlers-Danlos Syndrome (EDS)-B4GALT7 and Spondylodysplastic-EDS-B3GALT6. *Genes* **10**, doi:10.3390/genes10100799 (2019).
- 38 Mihalic Mosher, T. *et al.* Expansion of B4GALT7 linkeropathy phenotype to include perinatal lethal skeletal dysplasia. *European journal of human genetics : EJHG*, doi:10.1038/s41431-019-0464-8 (2019).
- 39 Malfait, F. *et al.* The 2017 international classification of the Ehlers-Danlos syndromes. *American journal of medical genetics. Part C, Seminars in medical genetics* **175**, 8-26, doi:10.1002/ajmg.c.31552 (2017).
- 40 Klein, J. A., Meng, L. & Zaia, J. Deep Sequencing of Complex Proteoglycans: A Novel Strategy for High Coverage and Site-specific Identification of Glycosaminoglycan-linked Peptides. *Mol Cell Proteomics* **17**, 1578-1590, doi:10.1074/mcp.RA118.000766 (2018).
- 41 Gulberti, S. *et al.* Phosphorylation and sulfation of oligosaccharide substrates critically influence the activity of human beta1,4-galactosyltransferase 7 (GalT-I) and beta1,3-glucuronosyltransferase I (GlcAT-I) involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *The Journal of biological chemistry* **280**, 1417-1425, doi:10.1074/jbc.M411552200 (2005).
- 42 Wen, J. *et al.* Xylose phosphorylation functions as a molecular switch to regulate proteoglycan biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 15723-15728, doi:10.1073/pnas.1417993111 (2014).
- 43 Kuroda, Y. *et al.* A novel gene (FAM20B encoding glycosaminoglycan xylosylkinase) for neonatal short limb dysplasia resembling Desbuquois dysplasia. *Clinical genetics* **95**, 713-717, doi:10.1111/cge.13530 (2019).
- 44 Malfait, F. *et al.* Defective initiation of glycosaminoglycan synthesis due to B3GALT6 mutations causes a pleiotropic Ehlers-Danlos-syndrome-like connective tissue disorder. *American journal of human genetics* **92**, 935-945, doi:10.1016/j.ajhg.2013.04.016 (2013).
- 45 Van Damme, T. *et al.* Biallelic B3GALT6 mutations cause spondylodysplastic Ehlers-Danlos syndrome. *Human molecular genetics* **27**, 3475-3487, doi:10.1093/hmg/ddy234 (2018).
- 46 Taylan, F. & Mäkitie, O. Abnormal Proteoglycan Synthesis Due to Gene Defects Causes Skeletal Diseases with Overlapping Phenotypes. *Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et métabolisme* **48**, 745-754, doi:10.1055/s-0042-118706 (2016).
- 47 Delbaere, S. *et al.* b3galt6 Knock-Out Zebrafish Recapitulate β 3GalT6-Deficiency Disorders in Human and Reveal a Trisaccharide Proteoglycan Linkage Region. *Frontiers in cell and developmental biology* **8**, 597857, doi:10.3389/fcell.2020.597857 (2020).
- 48 Ritelli, M. *et al.* Further Defining the Phenotypic Spectrum of B3GAT3 Mutations and Literature Review on Linkeropathy Syndromes. *Genes* **10**, doi:10.3390/genes10090631 (2019).
- 49 Izumikawa, T. *et al.* GlcUA β 1-3Gal β 1-3Gal β 1-4Xyl(2-O-phosphate) is the preferred substrate for chondroitin N-acetylgalactosaminyltransferase-1. *The Journal of biological chemistry* **290**, 5438-5448, doi:10.1074/jbc.M114.603266 (2015).

- 50 Koike, T., Izumikawa, T., Sato, B. & Kitagawa, H. Identification of phosphatase that dephosphorylates xylose in the glycosaminoglycan-protein linkage region of proteoglycans. *The Journal of biological chemistry* **289**, 6695-6708, doi:10.1074/jbc.M113.520536 (2014).
- 51 Parker, J. L. & Newstead, S. Gateway to the Golgi: molecular mechanisms of nucleotide sugar transporters. *Current opinion in structural biology* **57**, 127-134, doi:10.1016/j.sbi.2019.03.019 (2019).
- 52 Huber, C. *et al.* Identification of CANT1 mutations in Desbuquois dysplasia. *American journal of human genetics* **85**, 706-710, doi:10.1016/j.ajhg.2009.10.001 (2009).
- 53 Paganini, C. *et al.* Calcium activated nucleotidase 1 (CANT1) is critical for glycosaminoglycan biosynthesis in cartilage and endochondral ossification. *Matrix Biol* **81**, 70-90, doi:10.1016/j.matbio.2018.11.002 (2019).
- 54 Dubail, J. *et al.* SLC10A7 mutations cause a skeletal dysplasia with amelogenesis imperfecta mediated by GAG biosynthesis defects. *Nature communications* **9**, 3087, doi:10.1038/s41467-018-05191-8 (2018).
- 55 Orellana, A., Moraga, C., Araya, M. & Moreno, A. Overview of Nucleotide Sugar Transporter Gene Family Functions Across Multiple Species. *Journal of molecular biology* **428**, 3150-3165, doi:10.1016/j.jmb.2016.05.021 (2016).
- 56 Anthony, S., Munk, R., Skakun, W. & Masini, M. Multiple epiphyseal dysplasia. *The Journal of the American Academy of Orthopaedic Surgeons* **23**, 164-172, doi:10.5435/jaaos-d-13-00173 (2015).
- 57 Alper, S. L. & Sharma, A. K. The SLC26 gene family of anion transporters and channels. *Molecular aspects of medicine* **34**, 494-515, doi:10.1016/j.mam.2012.07.009 (2013).
- 58 Winawer, M. R. *et al.* Somatic SLC35A2 variants in the brain are associated with intractable neocortical epilepsy. *Annals of neurology* **83**, 1133-1146, doi:10.1002/ana.25243 (2018).
- 59 Li, J. P. & Kusche-Gullberg, M. Heparan Sulfate: Biosynthesis, Structure, and Function. *International review of cell and molecular biology* **325**, 215-273, doi:10.1016/bs.ircmb.2016.02.009 (2016).
- 60 Yamada, S. Specific functions of Exostosin-like 3 (EXTL3) gene products. *Cellular & molecular biology letters* **25**, 39, doi:10.1186/s11658-020-00231-y (2020).
- 61 Osman, N. M., Kagohashi, Y., Udagawa, J. & Otani, H. Alpha1,4-N-acetylglucosaminyltransferase encoding gene EXTL3 expression pattern in mouse adult and developing tissues with special attention to the pancreas. *Anatomy and embryology* **207**, 333-341, doi:10.1007/s00429-003-0348-z (2003).
- 62 Takahashi, I. *et al.* Important role of heparan sulfate in postnatal islet growth and insulin secretion. *Biochem Biophys Res Commun* **383**, 113-118, doi:10.1016/j.bbrc.2009.03.140 (2009).
- 63 Pacifici, M. The pathogenic roles of heparan sulfate deficiency in hereditary multiple exostoses. *Matrix Biol* **71-72**, 28-39, doi:10.1016/j.matbio.2017.12.011 (2018).
- 64 Lin, X. *et al.* Disruption of gastrulation and heparan sulfate biosynthesis in EXT1-deficient mice. *Developmental biology* **224**, 299-311, doi:10.1006/dbio.2000.9798 (2000).
- 65 Stickens, D., Zak, B. M., Rougier, N., Esko, J. D. & Werb, Z. Mice deficient in Ext2 lack heparan sulfate and develop exostoses. *Development (Cambridge, England)* **132**, 5055-5068, doi:10.1242/dev.02088 (2005).
- 66 Li, N., Fu, H., Hewitt, S. M., Dimitrov, D. S. & Ho, M. Therapeutically targeting glypican-2 via single-domain antibody-based chimeric antigen receptors and immunotoxins in neuroblastoma. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E6623-e6631, doi:10.1073/pnas.1706055114 (2017).
- 67 Ushakov, V. S. *et al.* Heparan Sulfate Biosynthetic System Is Inhibited in Human Glioma Due to EXT1/2 and HS6ST1/2 Down-Regulation. *International journal of molecular sciences* **18**, doi:10.3390/ijms18112301 (2017).
- 68 Kjellen, L. Glucosaminyl N-deacetylase/N-sulphotransferases in heparan sulphate biosynthesis and biology. *Biochemical Society transactions* **31**, 340-342, doi:10.1042/bst0310340 (2003).
- 69 Sheng, J., Liu, R., Xu, Y. & Liu, J. The dominating role of N-deacetylase/N-sulphotransferase 1 in forming domain structures in heparan sulfate. *The Journal of biological chemistry* **286**, 19768-19776, doi:10.1074/jbc.M111.224311 (2011).

- 70 Dou, W., Xu, Y., Pagadala, V., Pedersen, L. C. & Liu, J. Role of Deacetylase Activity of N-Deacetylase/N-Sulfotransferase 1 in Forming N-Sulfated Domain in Heparan Sulfate. *The Journal of biological chemistry* **290**, 20427-20437, doi:10.1074/jbc.M115.664409 (2015).
- 71 Ringvall, M. & Kjellen, L. Mice deficient in heparan sulfate N-deacetylase/N-sulfotransferase 1. *Progress in molecular biology and translational science* **93**, 35-58, doi:10.1016/s1877-1173(10)93003-2 (2010).
- 72 Armstrong, L. *et al.* A girl with developmental delay, ataxia, cranial nerve palsies, severe respiratory problems in infancy-Expanding NDST1 syndrome. *American journal of medical genetics. Part A* **173**, 712-715, doi:10.1002/ajmg.a.37621 (2017).
- 73 Reuter, M. S. *et al.* NDST1 missense mutations in autosomal recessive intellectual disability. *American journal of medical genetics. Part A* **164a**, 2753-2763, doi:10.1002/ajmg.a.36723 (2014).
- 74 Li, J. P. Glucuronyl C5-epimerase an enzyme converting glucuronic acid to iduronic acid in heparan sulfate/heparin biosynthesis. *Progress in molecular biology and translational science* **93**, 59-78, doi:10.1016/s1877-1173(10)93004-4 (2010).
- 75 Debarnot, C. *et al.* Substrate binding mode and catalytic mechanism of human heparan sulfate d-glucuronyl C5 epimerase. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 6760-6765, doi:10.1073/pnas.1818333116 (2019).
- 76 Sheng, J., Xu, Y., Dulaney, S. B., Huang, X. & Liu, J. Uncovering biphasic catalytic mode of C5-epimerase in heparan sulfate biosynthesis. *The Journal of biological chemistry* **287**, 20996-21002, doi:10.1074/jbc.M112.359885 (2012).
- 77 Prechoux, A., Halimi, C., Simorre, J. P., Lortat-Jacob, H. & Laguri, C. C5-epimerase and 2-O-sulfotransferase associate in vitro to generate contiguous epimerized and 2-O-sulfated heparan sulfate domains. *ACS chemical biology* **10**, 1064-1071, doi:10.1021/cb501037a (2015).
- 78 Kunnas, T., Solakivi, T., Määttä, K. & Nikkari, S. T. Glucuronic Acid Epimerase (GLCE) Variant rs3865014 (A>G) Is Associated with BMI, Blood Hemoglobin, Hypertension, and Cerebrovascular Events, the TAMRISK Study. *Annals of human genetics* **80**, 332-335, doi:10.1111/ahg.12166 (2016).
- 79 Prudnikova, T. Y. *et al.* Antiproliferative effect of D-glucuronyl C5-epimerase in human breast cancer cells. *Cancer cell international* **10**, 27, doi:10.1186/1475-2867-10-27 (2010).
- 80 Dejima, K. *et al.* Analysis of Drosophila glucuronyl C5-epimerase: implications for developmental roles of heparan sulfate sulfation compensation and 2-O-sulfated glucuronic acid. *The Journal of biological chemistry* **288**, 34384-34393, doi:10.1074/jbc.M113.499269 (2013).
- 81 Merry, C. L. *et al.* The molecular phenotype of heparan sulfate in the Hs2st^{-/-} mutant mouse. *The Journal of biological chemistry* **276**, 35429-35434, doi:10.1074/jbc.M100379200 (2001).
- 82 Pratt, T., Conway, C. D., Tian, N. M., Price, D. J. & Mason, J. O. Heparan sulphation patterns generated by specific heparan sulfotransferase enzymes direct distinct aspects of retinal axon guidance at the optic chiasm. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **26**, 6911-6923, doi:10.1523/jneurosci.0505-06.2006 (2006).
- 83 Habuchi, H. *et al.* Mice deficient in heparan sulfate 6-O-sulfotransferase-1 exhibit defective heparan sulfate biosynthesis, abnormal placentation, and late embryonic lethality. *The Journal of biological chemistry* **282**, 15578-15588, doi:10.1074/jbc.M607434200 (2007).
- 84 Habuchi, H. & Kimata, K. Mice deficient in heparan sulfate 6-O-sulfotransferase-1. *Progress in molecular biology and translational science* **93**, 79-111, doi:10.1016/s1877-1173(10)93005-6 (2010).
- 85 Izvolsky, K. I., Lu, J., Martin, G., Albrecht, K. H. & Cardoso, W. V. Systemic inactivation of Hs6st1 in mice is associated with late postnatal mortality without major defects in organogenesis. *Genesis* **46**, 8-18, doi:10.1002/dvg.20355 (2008).
- 86 Howard, S. R. *et al.* HS6ST1 Insufficiency Causes Self-Limited Delayed Puberty in Contrast With Other GnRH Deficiency Genes. *The Journal of clinical endocrinology and metabolism* **103**, 3420-3429, doi:10.1210/jc.2018-00646 (2018).
- 87 Alhasan, A. A., Spielhofer, J., Kusche-Gullberg, M., Kirby, J. A. & Ali, S. Role of 6-O-sulfated heparan sulfate in chronic renal fibrosis. *The Journal of biological chemistry* **289**, 20295-20306, doi:10.1074/jbc.M114.554691 (2014).

- 88 Lu, J., Auduong, L., White, E. S. & Yue, X. Up-regulation of heparan sulfate 6-O-sulfation in idiopathic pulmonary fibrosis. *American journal of respiratory cell and molecular biology* **50**, 106-114, doi:10.1165/rcmb.2013-0204OC (2014).
- 89 Chanalaris, A. *et al.* Heparan Sulfate Proteoglycan Synthesis Is Dysregulated in Human Osteoarthritic Cartilage. *The American journal of pathology* **189**, 632-647, doi:10.1016/j.ajpath.2018.11.011 (2019).
- 90 El Masri, R., Seffouh, A., Lortat-Jacob, H. & Vives, R. R. The "in and out" of glucosamine 6-O-sulfation: the 6th sense of heparan sulfate. *Glycoconjugate journal* **34**, 285-298, doi:10.1007/s10719-016-9736-5 (2017).
- 91 Kalus, I. *et al.* Differential involvement of the extracellular 6-O-endosulfatases Sulf1 and Sulf2 in brain development and neuronal and behavioural plasticity. *Journal of cellular and molecular medicine* **13**, 4505-4521, doi:10.1111/j.1582-4934.2008.00558.x (2009).
- 92 Holst, C. R. *et al.* Secreted sulfatases Sulf1 and Sulf2 have overlapping yet essential roles in mouse neonatal survival. *PLoS one* **2**, e575, doi:10.1371/journal.pone.0000575 (2007).
- 93 Matikainen, N. *et al.* Genetic variation in SULF2 is associated with postprandial clearance of triglyceride-rich remnant particles and triglyceride levels in healthy subjects. *PLoS one* **8**, e79473, doi:10.1371/journal.pone.0079473 (2013).
- 94 Takashima, Y. *et al.* Heparan sulfate 6-O-endosulfatases, Sulf1 and Sulf2, regulate glomerular integrity by modulating growth factor signaling. *American journal of physiology. Renal physiology* **310**, F395-408, doi:10.1152/ajprenal.00445.2015 (2016).
- 95 Hassing, H. C. *et al.* SULF2 strongly predisposes to fasting and postprandial triglycerides in patients with obesity and type 2 diabetes mellitus. *Obesity (Silver Spring, Md.)* **22**, 1309-1316, doi:10.1002/oby.20682 (2014).
- 96 Pinhal, M. A. S., Melo, C. M. & Nader, H. B. The Good and Bad Sides of Heparanase-1 and Heparanase-2. *Advances in experimental medicine and biology* **1221**, 821-845, doi:10.1007/978-3-030-34521-1_36 (2020).
- 97 Osorio, S., Rivillas, N. D. & Martinez, J. A. Urofacial (ochoa) syndrome: A literature review. *Journal of pediatric urology* **17**, 246-254, doi:10.1016/j.jpuro.2021.01.017 (2021).
- 98 Woolf, A. S., Stuart, H. M. & Newman, W. G. Genetics of human congenital urinary bladder disease. *Pediatric nephrology (Berlin, Germany)* **29**, 353-360, doi:10.1007/s00467-013-2472-1 (2014).
- 99 Petschner, P. *et al.* A replication study separates polymorphisms behind migraine with and without depression. *PLoS one* **16**, e0261477, doi:10.1371/journal.pone.0261477 (2021).
- 100 Roberts, N. A. & Woolf, A. S. Heparanase 2 and Urofacial Syndrome, a Genetic Neuropathy. *Advances in experimental medicine and biology* **1221**, 807-819, doi:10.1007/978-3-030-34521-1_35 (2020).
- 101 Jen, Y. H., Musacchio, M. & Lander, A. D. Glypican-1 controls brain size through regulation of fibroblast growth factor signaling in early neurogenesis. *Neural development* **4**, 33, doi:10.1186/1749-8104-4-33 (2009).
- 102 Mian, M., Ranjitkar, S., Townsend, G. C. & Anderson, P. J. Alterations in mandibular morphology associated with glypican 1 and glypican 3 gene mutations. *Orthodontics & craniofacial research* **20**, 183-187, doi:10.1111/ocr.12170 (2017).
- 103 Potkin, S. G. *et al.* A genome-wide association study of schizophrenia using brain activation as a quantitative phenotype. *Schizophrenia bulletin* **35**, 96-108, doi:10.1093/schbul/sbn155 (2009).
- 104 Bai, M. R. *et al.* Association of common variation in ADD3 and GPC1 with biliary atresia susceptibility. *Aging* **12**, 7163-7182, doi:10.18632/aging.103067 (2020).
- 105 Kaur, S. P. & Cummings, B. S. Role of glypicans in regulation of the tumor microenvironment and cancer progression. *Biochem Pharmacol* **168**, 108-118, doi:10.1016/j.bcp.2019.06.020 (2019).
- 106 Gondelaud, F. & Ricard-Blum, S. Structures and interactions of syndecans. *The FEBS journal*, doi:10.1111/febs.14828 (2019).
- 107 Iozzo, R. V. & Schaefer, L. Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix Biol* **42**, 11-55, doi:10.1016/j.matbio.2015.02.003 (2015).

- 108 Choi, Y., Chung, H., Jung, H., Couchman, J. R. & Oh, E. S. Syndecans as cell surface receptors: Unique structure equates with functional diversity. *Matrix Biol* **30**, 93-99, doi:10.1016/j.matbio.2010.10.006 (2011).
- 109 Teng, Y. H., Aquino, R. S. & Park, P. W. Molecular functions of syndecan-1 in disease. *Matrix Biol* **31**, 3-16, doi:10.1016/j.matbio.2011.10.001 (2012).
- 110 Okolicsanyi, R. K., Bluhm, J., Miller, C., Griffiths, L. R. & Haupt, L. M. An investigation of genetic polymorphisms in heparan sulfate proteoglycan core proteins and key modification enzymes in an Australian Caucasian multiple sclerosis population. *Human genomics* **14**, 18, doi:10.1186/s40246-020-00264-6 (2020).
- 111 Nemoto, T. *et al.* Endothelial glycocalyx and severity and vulnerability of coronary plaque in patients with coronary artery disease. *Atherosclerosis* **302**, 1-7, doi:10.1016/j.atherosclerosis.2020.04.014 (2020).
- 112 Stepp, M. A., Pal-Ghosh, S., Tadvalkar, G. & Pajooohesh-Ganji, A. Syndecan-1 and Its Expanding List of Contacts. *Advances in wound care* **4**, 235-249, doi:10.1089/wound.2014.0555 (2015).
- 113 Gharbaran, R. Advances in the molecular functions of syndecan-1 (SDC1/CD138) in the pathogenesis of malignancies. *Critical reviews in oncology/hematology* **94**, 1-17, doi:10.1016/j.critrevonc.2014.12.003 (2015).
- 114 Tang, G. L. & Weitz, K. Impaired arteriogenesis in syndecan-1(-/-) mice. *The Journal of surgical research* **193**, 22-27, doi:10.1016/j.jss.2014.08.002 (2015).
- 115 Stewart, M. D. & Sanderson, R. D. Heparan sulfate in the nucleus and its control of cellular functions. *Matrix Biol* **35**, 56-59, doi:10.1016/j.matbio.2013.10.009 (2014).
- 116 Ashley-Koch, A. E. *et al.* Genome-wide association study of posttraumatic stress disorder in a cohort of Iraq-Afghanistan era veterans. *Journal of affective disorders* **184**, 225-234, doi:10.1016/j.jad.2015.03.049 (2015).
- 117 Tsoyi, K. *et al.* Syndecan-2 Attenuates Radiation-induced Pulmonary Fibrosis and Inhibits Fibroblast Activation by Regulating PI3K/Akt/ROCK Pathway via CD148. *American journal of respiratory cell and molecular biology* **58**, 208-215, doi:10.1165/rcmb.2017-0088OC (2018).
- 118 Hong, H. *et al.* Up-regulation of syndecan-2 in proximal colon correlates with acute inflammation. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, fj201900561R, doi:10.1096/fj.201900561R (2019).
- 119 Mytilinaiou, M. *et al.* Emerging roles of syndecan 2 in epithelial and mesenchymal cancer progression. *IUBMB life* **69**, 824-833, doi:10.1002/iub.1678 (2017).
- 120 Essner, J. J., Chen, E. & Ekker, S. C. Syndecan-2. *The international journal of biochemistry & cell biology* **38**, 152-156, doi:10.1016/j.biocel.2005.08.012 (2006).
- 121 Chen, C. Y., Lin, C. W., Chang, C. Y., Jiang, S. T. & Hsueh, Y. P. Sarm1, a negative regulator of innate immunity, interacts with syndecan-2 and regulates neuronal morphology. *The Journal of cell biology* **193**, 769-784, doi:10.1083/jcb.201008050 (2011).
- 122 Volta, M., Calza, S., Roberts, A. M. & Roberts, R. G. Characterisation of the interaction between syndecan-2, neurofibromin and CASK: dependence of interaction on syndecan dimerization. *Biochem Biophys Res Commun* **391**, 1216-1221, doi:10.1016/j.bbrc.2009.12.043 (2010).
- 123 Ha, E. *et al.* Positive association of obesity with single nucleotide polymorphisms of syndecan 3 in the Korean population. *The Journal of clinical endocrinology and metabolism* **91**, 5095-5099, doi:10.1210/jc.2005-2086 (2006).
- 124 Chang, B. C., Hwang, L. C. & Huang, W. H. Positive Association of Metabolic Syndrome with a Single Nucleotide Polymorphism of Syndecan-3 (rs2282440) in the Taiwanese Population. *International journal of endocrinology* **2018**, 9282598, doi:10.1155/2018/9282598 (2018).
- 125 Moons, T. *et al.* Clock genes and body composition in patients with schizophrenia under treatment with antipsychotic drugs. *Schizophrenia research* **125**, 187-193, doi:10.1016/j.schres.2010.10.008 (2011).
- 126 Schüring, A. N. *et al.* Role of syndecan-3 polymorphisms in obesity and female hyperandrogenism. *Journal of molecular medicine (Berlin, Germany)* **87**, 1241-1250, doi:10.1007/s00109-009-0529-1 (2009).

- 127 Reizes, O. *et al.* Transgenic expression of syndecan-1 uncovers a physiological control of feeding behavior by syndecan-3. *Cell* **106**, 105-116, doi:10.1016/s0092-8674(01)00415-9 (2001).
- 128 Su, J., Zhang, Y., Su, H., Zhang, C. & Li, W. A recurrence model for laryngeal cancer based on SVM and gene function clustering. *Acta oto-laryngologica* **137**, 557-562, doi:10.1080/00016489.2016.1247984 (2017).
- 129 Echtermeyer, F. *et al.* Delayed wound repair and impaired angiogenesis in mice lacking syndecan-4. *The Journal of clinical investigation* **107**, R9-r14, doi:10.1172/jci10559 (2001).
- 130 De Luca, M. *et al.* A conserved role for syndecan family members in the regulation of whole-body energy metabolism. *PLoS one* **5**, e11286, doi:10.1371/journal.pone.0011286 (2010).
- 131 Rose, G. *et al.* Metabolism and successful aging: Polymorphic variation of syndecan-4 (SDC4) gene associate with longevity and lipid profile in healthy elderly Italian subjects. *Mechanisms of ageing and development* **150**, 27-33, doi:10.1016/j.mad.2015.08.003 (2015).
- 132 Busse-Wicher, M., Wicher, K. B. & Kusche-Gullberg, M. The exostosin family: proteins with many functions. *Matrix Biol* **35**, 25-33, doi:10.1016/j.matbio.2013.10.001 (2014).
- 133 Purnomo, E. *et al.* Glycosaminoglycan overproduction in the aorta increases aortic calcification in murine chronic kidney disease. *Journal of the American Heart Association* **2**, e000405, doi:10.1161/jaha.113.000405 (2013).
- 134 Nadanaka, S., Kagiya, S. & Kitagawa, H. Roles of EXTL2, a member of the EXT family of tumour suppressors, in liver injury and regeneration processes. *The Biochemical journal* **454**, 133-145, doi:10.1042/bj20130323 (2013).
- 135 Taghizadeh, E. *et al.* SULF 1 gene polymorphism, rs6990375 is in significant association with fetus failure in IVF technique. *Iranian journal of reproductive medicine* **13**, 215-220 (2015).
- 136 Zhang, P. *et al.* Polymorphisms in Sex Hormone Metabolism Genes and Risk of Preeclampsia in Taiyuan, China. *Gynecologic and obstetric investigation* **83**, 179-186, doi:10.1159/000478931 (2018).
- 137 Oshima, K. *et al.* Loss of endothelial sulfatase-1 after experimental sepsis attenuates subsequent pulmonary inflammatory responses. *American journal of physiology. Lung cellular and molecular physiology* **317**, L667-l677, doi:10.1152/ajplung.00175.2019 (2019).
- 138 Masola, V., Bellin, G., Gambaro, G. & Onisto, M. Heparanase: A Multitasking Protein Involved in Extracellular Matrix (ECM) Remodeling and Intracellular Events. *Cells* **7**, doi:10.3390/cells7120236 (2018).
- 139 Peterson, S. B. & Liu, J. Multi-faceted substrate specificity of heparanase. *Matrix Biol* **32**, 223-227, doi:10.1016/j.matbio.2013.02.006 (2013).
- 140 Poon, I. K. *et al.* Mice deficient in heparanase exhibit impaired dendritic cell migration and reduced airway inflammation. *European journal of immunology* **44**, 1016-1030, doi:10.1002/eji.201343645 (2014).
- 141 Martin, P. J. *et al.* Replication of associations between genetic polymorphisms and chronic graft-versus-host disease. *Blood* **128**, 2450-2456, doi:10.1182/blood-2016-07-728063 (2016).
- 142 Ostrovsky, O., Shimoni, A., Rand, A., Vlodavsky, I. & Nagler, A. Genetic variations in the heparanase gene (HPSE) associate with increased risk of GVHD following allogeneic stem cell transplantation: effect of discrepancy between recipients and donors. *Blood* **115**, 2319-2328, doi:10.1182/blood-2009-08-236455 (2010).
- 143 Jin, H. & Zhou, S. The Functions of Heparanase in Human Diseases. *Mini reviews in medicinal chemistry* **17**, 541-548, doi:10.2174/1389557516666161101143643 (2017).
- 144 Sierksma, A. *et al.* Novel Alzheimer risk genes determine the microglia response to amyloid- β but not to TAU pathology. *EMBO molecular medicine* **12**, e10606, doi:10.15252/emmm.201910606 (2020).
- 145 Lugert, S. *et al.* Glypican-2 levels in cerebrospinal fluid predict the status of adult hippocampal neurogenesis. *Scientific reports* **7**, 46543, doi:10.1038/srep46543 (2017).
- 146 Okamoto, K. *et al.* Common variation in GPC5 is associated with acquired nephrotic syndrome. *Nature genetics* **43**, 459-463, doi:10.1038/ng.792 (2011).
- 147 Chorąży, M. *et al.* Analysis of chosen SNVs in GPC5, CD58 and IRF8 genes in multiple sclerosis patients. *Advances in medical sciences* **64**, 230-234, doi:10.1016/j.advms.2018.12.004 (2019).
- 148 Guan, Y. *et al.* The Role of Cell Growth-Related Gene Copy Number Variation in Autoimmune Thyroid Disease. *Biological trace element research* **195**, 409-416, doi:10.1007/s12011-019-01880-7 (2020).

- 149 Guo, L., Wang, J., Zhang, T. & Yang, Y. Glypican-5 is a tumor suppressor in non-small cell lung cancer cells. *Biochemistry and biophysics reports* **6**, 108-112, doi:10.1016/j.bbrep.2016.03.010 (2016).
- 150 Sun, Y., Xu, K., He, M., Fan, G. & Lu, H. Overexpression of Glypican 5 (GPC5) Inhibits Prostate Cancer Cell Proliferation and Invasion via Suppressing Sp1-Mediated EMT and Activation of Wnt/beta-Catenin Signaling. *Oncology research* **26**, 565-572, doi:10.3727/096504017x15044461944385 (2018).
- 151 Deligny, A. *et al.* NDST2 (N-Deacetylase/N-Sulfotransferase-2) Enzyme Regulates Heparan Sulfate Chain Length. *The Journal of biological chemistry* **291**, 18600-18607, doi:10.1074/jbc.M116.744433 (2016).
- 152 Chen, H., Wang, T., Yang, J., Huang, S. & Zeng, P. Improved Detection of Potentially Pleiotropic Genes in Coronary Artery Disease and Chronic Kidney Disease Using GWAS Summary Statistics. *Frontiers in genetics* **11**, 592461, doi:10.3389/fgene.2020.592461 (2020).
- 153 Samoszuk, M. *et al.* Inhibition of thrombosis in melanoma allografts in mice by endogenous mast cell heparin. *Thromb Haemost* **90**, 351-360, doi:10.1160/TH03-02-0088 (2003).
- 154 Pallerla, S. R. *et al.* Altered heparan sulfate structure in mice with deleted NDST3 gene function. *The Journal of biological chemistry* **283**, 16885-16894, doi:10.1074/jbc.M709774200 (2008).
- 155 Jao, T. M. *et al.* Alteration of colonic epithelial cell differentiation in mice deficient for glucosaminyl N-deacetylase/N-sulfotransferase 4. *Oncotarget* **7**, 84938-84950, doi:10.18632/oncotarget.12915 (2016).
- 156 Lencz, T. *et al.* Genome-wide association study implicates NDST3 in schizophrenia and bipolar disorder. *Nature communications* **4**, 2739, doi:10.1038/ncomms3739 (2013).
- 157 Wang, L. *et al.* Association study of NDST3 gene for schizophrenia, bipolar disorder, major depressive disorder in the Han Chinese population. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics* **177**, 3-9, doi:10.1002/ajmg.b.32573 (2018).
- 158 Eicher, J. D. *et al.* Genome-wide association study of shared components of reading disability and language impairment. *Genes, brain, and behavior* **12**, 792-801, doi:10.1111/gbb.12085 (2013).
- 159 Habuchi, H. *et al.* Biosynthesis of heparan sulphate with diverse structures and functions: two alternatively spliced forms of human heparan sulphate 6-O-sulphotransferase-2 having different expression patterns and properties. *The Biochemical journal* **371**, 131-142, doi:10.1042/bj20021259 (2003).
- 160 Habuchi, H. *et al.* The occurrence of three isoforms of heparan sulfate 6-O-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted N-sulfoglucosamine. *The Journal of biological chemistry* **275**, 2859-2868, doi:10.1074/jbc.275.4.2859 (2000).
- 161 Jemth, P. *et al.* Oligosaccharide library-based assessment of heparan sulfate 6-O-sulfotransferase substrate specificity. *The Journal of biological chemistry* **278**, 24371-24376, doi:10.1074/jbc.M212155200 (2003).
- 162 Paganini, L. *et al.* A HS6ST2 gene variant associated with X-linked intellectual disability and severe myopia in two male twins. *Clinical genetics* **95**, 368-374, doi:10.1111/cge.13485 (2019).
- 163 Wang, W. *et al.* Overexpression of heparan sulfate 6-O-sulfotransferase-2 enhances fibroblast growth factor-mediated chondrocyte growth and differentiation. *International journal of molecular medicine* **36**, 825-832, doi:10.3892/ijmm.2015.2272 (2015).
- 164 Guo, Y. *et al.* Downregulation of HS6ST2 by miR-23b-3p enhances matrix degradation through p38 MAPK pathway in osteoarthritis. *Cell death & disease* **9**, 699, doi:10.1038/s41419-018-0729-0 (2018).
- 165 Jin, Y. *et al.* Overexpression of HS6ST2 is associated with poor prognosis in patients with gastric cancer. *Oncology letters* **14**, 6191-6197, doi:10.3892/ol.2017.6944 (2017).
- 166 Wang, K. S., Wang, L., Liu, X. & Zeng, M. Association of HS6ST3 gene polymorphisms with obesity and triglycerides: gene x gender interaction. *Journal of genetics* **92**, 395-402 (2013).
- 167 Huang, Y. C. *et al.* Genome-wide association study of diabetic retinopathy in a Taiwanese population. *Ophthalmology* **118**, 642-648, doi:10.1016/j.ophtha.2010.07.020 (2011).
- 168 Iravani, O., Bay, B. H. & Yip, G. W. Silencing HS6ST3 inhibits growth and progression of breast cancer cells through suppressing IGF1R and inducing XAF1. *Experimental cell research* **350**, 380-389, doi:10.1016/j.yexcr.2016.12.019 (2017).

- 169 Shworak, N. W. *et al.* Multiple isoforms of heparan sulfate D-glucosaminyl 3-O-sulfotransferase. Isolation, characterization, and expression of human cdnas and identification of distinct genomic loci. *The Journal of biological chemistry* **274**, 5170-5184, doi:10.1074/jbc.274.8.5170 (1999).
- 170 Shworak, N. W., HajMohammadi, S., de Agostini, A. I. & Rosenberg, R. D. Mice deficient in heparan sulfate 3-O-sulfotransferase-1: normal hemostasis with unexpected perinatal phenotypes. *Glycoconjugate journal* **19**, 355-361, doi:10.1023/a:1025377206600 (2002).
- 171 Thacker, B. E., Xu, D., Lawrence, R. & Esko, J. D. Heparan sulfate 3-O-sulfation: a rare modification in search of a function. *Matrix Biol* **35**, 60-72, doi:10.1016/j.matbio.2013.12.001 (2014).
- 172 Shworak, N. W., Kobayashi, T., de Agostini, A. & Smits, N. C. Anticoagulant heparan sulfate to not clot--or not? *Progress in molecular biology and translational science* **93**, 153-178, doi:10.1016/s1877-1173(10)93008-1 (2010).
- 173 Desikan, R. S. *et al.* Polygenic Overlap Between C-Reactive Protein, Plasma Lipids, and Alzheimer Disease. *Circulation* **131**, 2061-2069, doi:10.1161/circulationaha.115.015489 (2015).
- 174 Espinosa, A. *et al.* Exploring Genetic Associations of Alzheimer's Disease Loci With Mild Cognitive Impairment Neurocognitive Endophenotypes. *Frontiers in aging neuroscience* **10**, 340, doi:10.3389/fnagi.2018.00340 (2018).
- 175 Witoelar, A. *et al.* Meta-analysis of Alzheimer's disease on 9,751 samples from Norway and IGAP study identifies four risk loci. *Scientific reports* **8**, 18088, doi:10.1038/s41598-018-36429-6 (2018).
- 176 Thacker, B. E. *et al.* Expanding the 3-O-Sulfate Proteome--Enhanced Binding of Neuropilin-1 to 3-O-Sulfated Heparan Sulfate Modulates Its Activity. *ACS chemical biology* **11**, 971-980, doi:10.1021/acscchembio.5b00897 (2016).
- 177 Girardin, E. P. *et al.* Synthesis of anticoagulant active heparan sulfate proteoglycans by glomerular epithelial cells involves multiple 3-O-sulfotransferase isoforms and a limiting precursor pool. *The Journal of biological chemistry* **280**, 38059-38070, doi:10.1074/jbc.M507997200 (2005).
- 178 Lawrence, R. *et al.* The principal neuronal gD-type 3-O-sulfotransferases and their products in central and peripheral nervous system tissues. *Matrix Biol* **26**, 442-455, doi:10.1016/j.matbio.2007.03.002 (2007).
- 179 Borjigin, J. *et al.* Diurnal pineal 3-O-sulphotransferase 2 expression controlled by beta-adrenergic repression. *The Journal of biological chemistry* **278**, 16315-16319, doi:10.1074/jbc.M300828200 (2003).
- 180 Sepulveda-Diaz, J. E. *et al.* HS3ST2 expression is critical for the abnormal phosphorylation of tau in Alzheimer's disease-related tau pathology. *Brain : a journal of neurology* **138**, 1339-1354, doi:10.1093/brain/awv056 (2015).
- 181 Denys, A. & Allain, F. The Emerging Roles of Heparan Sulfate 3-O-Sulfotransferases in Cancer. *Frontiers in oncology* **9**, 507, doi:10.3389/fonc.2019.00507 (2019).
- 182 Amraoui, F. *et al.* Placental expression of heparan sulfate 3-O-sulfotransferase-3A1 in normotensive and pre-eclamptic pregnancies. *Placenta* **36**, 1218-1224, doi:10.1016/j.placenta.2015.09.008 (2015).
- 183 Atkinson, A., Garnier, S., Afridi, S., Fumoux, F. & Rihet, P. Genetic variations in genes involved in heparan sulphate biosynthesis are associated with Plasmodium falciparum parasitaemia: a familial study in Burkina Faso. *Malaria journal* **11**, 108, doi:10.1186/1475-2875-11-108 (2012).
- 184 Nguyen, N. T. *et al.* Genetic and enzymatic characterization of 3-O-sulfotransferase SNPs associated with Plasmodium falciparum parasitaemia. *Glycobiology* **28**, 534-541, doi:10.1093/glycob/cwy038 (2018).
- 185 Mackinnon, M. J. *et al.* Environmental Correlation Analysis for Genes Associated with Protection against Malaria. *Molecular biology and evolution* **33**, 1188-1204, doi:10.1093/molbev/msw004 (2016).
- 186 Vanpouille, C. *et al.* The heparin/heparan sulfate sequence that interacts with cyclophilin B contains a 3-O-sulfated N-unsubstituted glucosamine residue. *The Journal of biological chemistry* **282**, 24416-24429, doi:M701835200 [pii] 10.1074/jbc.M701835200 (2007).
- 187 Spencer, K. L. *et al.* Dissection of chromosome 16p12 linkage peak suggests a possible role for CACNG3 variants in age-related macular degeneration susceptibility. *Investigative ophthalmology & visual science* **52**, 1748-1754, doi:10.1167/iops.09-5112 (2011).

- 188 Debette, S. *et al.* Genome-wide studies of verbal declarative memory in nondemented older people: the Cohorts for Heart and Aging Research in Genomic Epidemiology consortium. *Biol Psychiatry* **77**, 749-763, doi:10.1016/j.biopsych.2014.08.027 (2015).
- 189 Stacey, D., Ciobanu, L. G. & Baune, B. T. A systematic review on the association between inflammatory genes and cognitive decline in non-demented elderly individuals. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* **27**, 568-588, doi:10.1016/j.euroneuro.2015.12.017 (2017).
- 190 Huynh, M. B. *et al.* Glycosaminoglycans from Alzheimer's disease hippocampus have altered capacities to bind and regulate growth factors activities and to bind tau. *PloS one* **14**, e0209573, doi:10.1371/journal.pone.0209573 (2019).
- 191 Mochizuki, H. *et al.* Characterization of a heparan sulfate 3-O-sulfotransferase-5, an enzyme synthesizing a tetrasulfated disaccharide. *The Journal of biological chemistry* **278**, 26780-26787, doi:10.1074/jbc.M301861200 (2003).
- 192 Wang, Q. *et al.* Genome-wide association analysis with gray matter volume as a quantitative phenotype in first-episode treatment-naive patients with schizophrenia. *PloS one* **8**, e75083, doi:10.1371/journal.pone.0075083 (2013).
- 193 Xu, D. *et al.* Characterization of heparan sulphate 3-O-sulphotransferase isoform 6 and its role in assisting the entry of herpes simplex virus type 1. *The Biochemical journal* **385**, 451-459, doi:10.1042/BJ20040908 (2005).
- 194 Bork, K. *et al.* Novel hereditary angioedema linked with a heparan sulfate 3-O-sulfotransferase 6 gene mutation. *J Allergy Clin Immunol*, doi:10.1016/j.jaci.2021.01.011 (2021).
- 195 Duchez, S. *et al.* Glycotranscriptome study reveals an enzymatic switch modulating glycosaminoglycan synthesis during B-cell development and activation. *European journal of immunology* **41**, 3632-3644, doi:10.1002/eji.201140865 (2011).
- 196 Jiang, Z., Michal, J. J., Wu, X. L., Pan, Z. & MacNeil, M. D. The heparan and heparin metabolism pathway is involved in regulation of fatty acid composition. *International journal of biological sciences* **7**, 659-663, doi:10.7150/ijbs.7.659 (2011).
- 197 Mathysen, D. *et al.* Molecular analysis of the putative tumour-suppressor gene EXTL1 in neuroblastoma patients and cell lines. *European journal of cancer (Oxford, England : 1990)* **40**, 1255-1261, doi:10.1016/j.ejca.2004.01.013 (2004).