

# SCIENTIFIC REPORTS



OPEN

## Matrisome Profiling During Intervertebral Disc Development And Ageing

Joana Caldeira<sup>1,2,3</sup>, Cátia Santa<sup>4,5</sup>, Hugo Osório<sup>1,3,6</sup>, Maria Molinos<sup>1,2,7</sup>, Bruno Manadas<sup>5</sup>, Raquel Gonçalves<sup>1,2</sup> & Mário Barbosa<sup>1,2,7</sup>

Intervertebral disc (IVD) degeneration is often the cause of low back pain. Degeneration occurs with age and is accompanied by extracellular matrix (ECM) depletion, culminating in nucleus pulposus (NP) extrusion and IVD destruction. The changes that occur in the disc with age have been under investigation. However, a thorough study of ECM profiling is needed, to better understand IVD development and age-associated degeneration. As so, iTRAQ LC-MS/MS analysis of foetus, young and old bovine NPs, was performed to define the NP matrisome. The enrichment of Collagen XII and XIV in foetus, Fibronectin and Prolargin in elder NPs and Collagen XI in young ones was independently validated. This study provides the first matrisome database of healthy discs during development and ageing, which is key to determine the pathways and processes that maintain disc homeostasis. The factors identified may help to explain age-associated IVD degeneration or constitute putative effectors for disc regeneration.

The intervertebral disc (IVD) is a complex structure capable of resisting spinal compression while allowing motion of intervertebral segments<sup>1,2</sup>. Besides water, it is mainly composed by extracellular matrix (ECM) molecules. These include collagens, proteoglycans (PGs) and other matrix proteins that contribute to the structural and mechanical function of the disc<sup>3,4</sup>. Matrix degrading enzymes are also present to regulate matrix breakdown, maintaining disc homeostasis<sup>5</sup>.

A young healthy disc consists of a highly plastic and hydrated region – the nucleus pulposus (NP) – and a network of collagen fibres oriented in sheets around the nucleus – the annulus fibrosus (AF), which provides tensile strength and confines the NP, limiting bulging<sup>6</sup>.

During disc degeneration and ageing, significant changes are observed in the IVD at both cell and tissue level. From birth, notochordal cells gradually disappear from the NP<sup>7</sup>. Loss of cell density is accompanied by a shift towards a chondrocyte-like cell population<sup>3</sup>, less effective in NP-specific matrix synthesis<sup>8</sup>. Ultimately this results in NP fibrous transformation, from a translucent gel to a more solid and cartilaginous tissue<sup>1</sup> making it difficult to distinguish between NP and AF<sup>6</sup>. Alterations in the composition and mechanical properties of the surrounding environment will in turn influence NP cell function and behaviour, in terms of differentiation, metabolism, proliferation and survival<sup>8</sup>.

Along with cellular changes, NP matrix remodeling is also an early step in the ageing process. Apart from overall matrix breakdown caused by MMPs (matrix metalloproteinases) and ADAMTS (a disintegrin and metalloprotease with thrombospondin motifs) overexpression<sup>7</sup>, PG and collagen synthesis patterns<sup>9,10</sup>, as well as fibre crosslinking<sup>1</sup> are also altered. This inhibits matrix turnover and, together with the already limited repair response, leads to dehydration and progressive ECM disorganization. Furthermore, it promotes mechanical failure, annular tears and many of the characteristic features of disc degeneration<sup>6</sup>. Over time, type II collagen is replaced by

<sup>1</sup>i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal. <sup>2</sup>INEB - Instituto de Engenharia Biomédica, Universidade do Porto, Rua Alfredo Allen, 208, 4200-180, Porto, Portugal. <sup>3</sup>IPATIMUP - Institute of Molecular Pathology and Immunology, University of Porto, Rua Júlio Amaral de Carvalho, 45, 4200-135, Porto, Portugal. <sup>4</sup>III - Institute for Interdisciplinary Research, University of Coimbra, Casa Costa Alemão – Pólo II, Rua Dom Francisco de Lemos, 3030-789, Coimbra, Portugal. <sup>5</sup>CNC – Center for Neuroscience and Cell Biology, University of Coimbra, 3004-504, Coimbra, Portugal. <sup>6</sup>Department of Pathology and Oncology, Faculty of Medicine, University of Porto, 4200-319, Porto, Portugal. <sup>7</sup>ICBAS - Instituto de Ciências Biomédicas de Abel Salazar, Universidade do Porto, Rua de Jorge Viterbo Ferreira n. 228, 4050-313, Porto, Portugal. Correspondence and requests for materials should be addressed to J.C. (email: [joana.caldeira@ineb.up.pt](mailto:joana.caldeira@ineb.up.pt))

type I collagen in the NP<sup>7</sup> and aggrecan content decreases<sup>3</sup>. Along with structural changes, soluble factors, and cytokines may also be released<sup>11</sup>, further affecting cell activity and tissue homeostasis<sup>7</sup>. The availability of oxygen, nutrients and growth factors<sup>12</sup>, and the acidity of the environment, as well as the removal of metabolites, are also influenced by ECM calcification and impermeabilization<sup>7</sup>. With increasing age, this imbalance of the normal homeostatic mechanism impairs normal disc function, particularly in the NP<sup>13</sup>, ultimately resulting in reduced disc height, hernia formation and spinal pain, as nerve roots become compressed<sup>2</sup>.

Low back pain (LBP) causes disability and life quality deterioration, constituting a tremendous social and economic burden<sup>14</sup>. In more than 40% of the cases, it is triggered by IVD degeneration, which mimics disc ageing but occurs at an accelerated rate<sup>1,7</sup>. Conventional therapies for LBP predominantly involve treatments based on pain modulators and invasive surgeries, like spine fusion or arthroplasty. However, spine surgeries have a high risk of complications associated<sup>15,16</sup>, and recurrent interventions are many times needed (15–30% of cases)<sup>17</sup>, increasing the personal and financial costs even further<sup>18</sup>. Of note, the underlying pathophysiology is not being addressed, nor is the restoration of IVD's function or the slowing down of disease progression. To date, promising strategies for disc regeneration, based on the maintenance and/or increase of matrix synthesis, are being explored *in vivo*: protein injection<sup>19,20</sup>, gene transfer<sup>21,22</sup> and cell implantation<sup>23,24</sup>. Although rapid advances are being made in understanding and regulating the degenerative process, many challenges remain<sup>2</sup>.

Understanding IVD pathophysiology (particularly in terms of IVD matrix constituents and their alterations in development and disease) is key to unveil molecular cues that might be used to slow, halt or reverse the age-associated degenerative cascade<sup>2,4</sup>.

In this report, we have investigated matrisome changes observed with development and ageing in healthy bovine NPs, with the purpose of validating candidate molecules that might constitute novel therapeutic alternatives to treat IVD degeneration.

## Methods

**Sample preparation and iTRAQ analysis.** Bovine caudal IVDs from foetus (around 7 months of gestation), young (12 months) and old animals (16 to 18 years old) were obtained from the local abattoir and dissected within 3–4 hours after slaughter. The NPs from 7–8 discs from Cd1 to Cd7 or Cd8 were collected as described by Molinos *et al.*<sup>25</sup> and stored at  $-80^{\circ}\text{C}$  in a batch of 500 to 800 mg. For protein extraction, 1100  $\mu\text{L}$  of guanidine extraction buffer were used. Further details on the iTRAQ analysis, including technicalities of protein extraction, precipitation and quantification, reduction, alkylation and trypsin digestion, as well as iTRAQ labelling, sample distribution, fractionation and LC-MS/MS analysis, database searching and protein identification, bioinformatics analysis and candidate selection criteria, can be found as Supplementary Data 1 and Supplementary Table 1. For cell extracts, cells were first isolated by Collagenase Type XI (2 mg/mL) treatment and posteriorly filtered to remove ECM contaminants, as previously reported<sup>25</sup>. Proteins were then extracted using the same guanidine hydrochloride based protocol that was used for whole tissue extracts (Supplementary Data 1). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD005616 and PXD004922<sup>26</sup>.

**Macroscopic characterization and scanning electron microscopy (SEM).** For qualitative macroscopic evaluation of the different age groups, IVDs were excised and photographed under an Olympus SZX16 stereomicroscope coupled with a DP71 camera (Olympus, Tokyo, Japan) at 10X magnification.

For SEM analysis, samples were fixed using 2.5% (v/v) glutaraldehyde (Agar Scientific) in 0.1 M sodium cacodylate solution (Sigma) and then stored in sodium cacodylate buffer 0.1 M at  $4^{\circ}\text{C}$  until further use. IVDs were then dehydrated in serial diluted ethanol solutions of 50, 60, 70, 80, 90, and 99% v/v, being incubated for 10 min in each dilution. Following critical point drying, samples were sputtered-coated with a Au/Pd thin film, using the SPI Module Sputter Coater equipment. Samples were examined at CEMUP (Materials Centre of the University of Porto), using a High resolution Scanning Electron Microscope with X-Ray Microanalysis - JEOL JSM 6301 F/Oxford INCA Energy 350 - at 300X and 5000X magnification. The following parameters (mean fibril diameter, mean pore area, number of pores and number of intersections) were obtained using “DiameterJ”, a plugin for ImageJ/FIJI software version 1.46r (NIH) for topographic comparison of samples under study<sup>27</sup>.

**Western Blotting.** Following denaturation for 10 min at  $65^{\circ}\text{C}$ , protein samples were separated by sodium dodecyl sulphate (SDS) 9% polyacrylamide gel electrophoresis (PAGE), and electroblotted onto a Hybond enhanced chemiluminescence (ECL) membrane (Amersham Biosciences). Antibodies for Collagen Type XII, Collagen Type XIV, Collagen Type XI alpha 2, Fibronectin and Prolargin were used. Further details on blocking solutions, primary and secondary antibodies used, as well as their respective host species, working dilutions and commercial suppliers can be found in Supplementary Table 2.

After ECL detection (Amersham Biosciences), bands were quantified using Quantity One 4.6.8 Software (Bio-Rad) and values were normalized to the total protein loading (density value of each complete lane, obtained after staining of the membrane following immunodetection with Page Blue Protein Staining Solution (ThermoFisher Scientific), using a protocol adapted from Welinder and Ekblad<sup>28</sup>). All samples were run in the same SDS-PAGE gel, and background signal was measured in several different areas of the membrane. Average background signal was then subtracted to the band intensity signal to minimize background variation and interference.

**Statistical Analysis.** Statistical analysis was performed using non-parametric Mann-Whitney test in GraphPad Prism software 5.0, to compare two groups of non-related samples (e.g. foetus *versus* young or young *versus* old NPs). The parametric distribution of the data was evaluated by D'Agostino and Pearson normality test. Results (from at least three independent biological samples) are expressed as median  $\pm$  Interquartile Range (IQR)

in box and whiskers plots or as mean  $\pm$  Standard Error of the Mean (SEM). Values from  $p \leq 0.05$  were considered statistically significant.

## Results

**Structural characterization of different aged IVDs by SEM.** At the macroscopic level, several differences could be identified (Fig. 1, left panel). Foetal intervertebral discs (IVDs) consisted of a well-distinguished annulus fibrosus (AF) formed of concentric lamellae delimiting a highly hydrated gel-like nucleus pulposus (NP). During ageing and degeneration, the boundary between AF and NP became less obvious. In addition, discs started looking increasingly dry and fibrous.

Scanning electron microscopy (SEM) has also been optimized, for a more detailed comparison of NP topography from the three different age groups. NP SEM images clearly showed a fine network of randomly oriented fibres (Fig. 1, central and right panels).

Further quantitative characterization of the SEM images obtained for different aged NP samples revealed age-associated changes in matrix architecture. Using DiameterJ, a plugin of ImageJ/Fiji software, we evaluated the following parameters: mean fibre diameter, number of pores, number of intersections and mean pore area (Fig. 1b). Mean fibre diameter of elder (8.826 pixels) NPs was statistically higher than that of the adults (7.482 pxels), whereas no differences were observed between foetal (7.803 pixels) and young NPs. Elder bovine NPs also presented significantly fewer (275.2 pores) but bigger (566.5 pixels<sup>2</sup>) pores in comparison to young animals (459 pores and 432.6 pixels<sup>2</sup>, respectively) and also less fibre intersections (927 vs 1391 intersections). Foetal NP mean pore area (602.6 pixels<sup>2</sup>) was also significantly bigger than that of young IVDs.

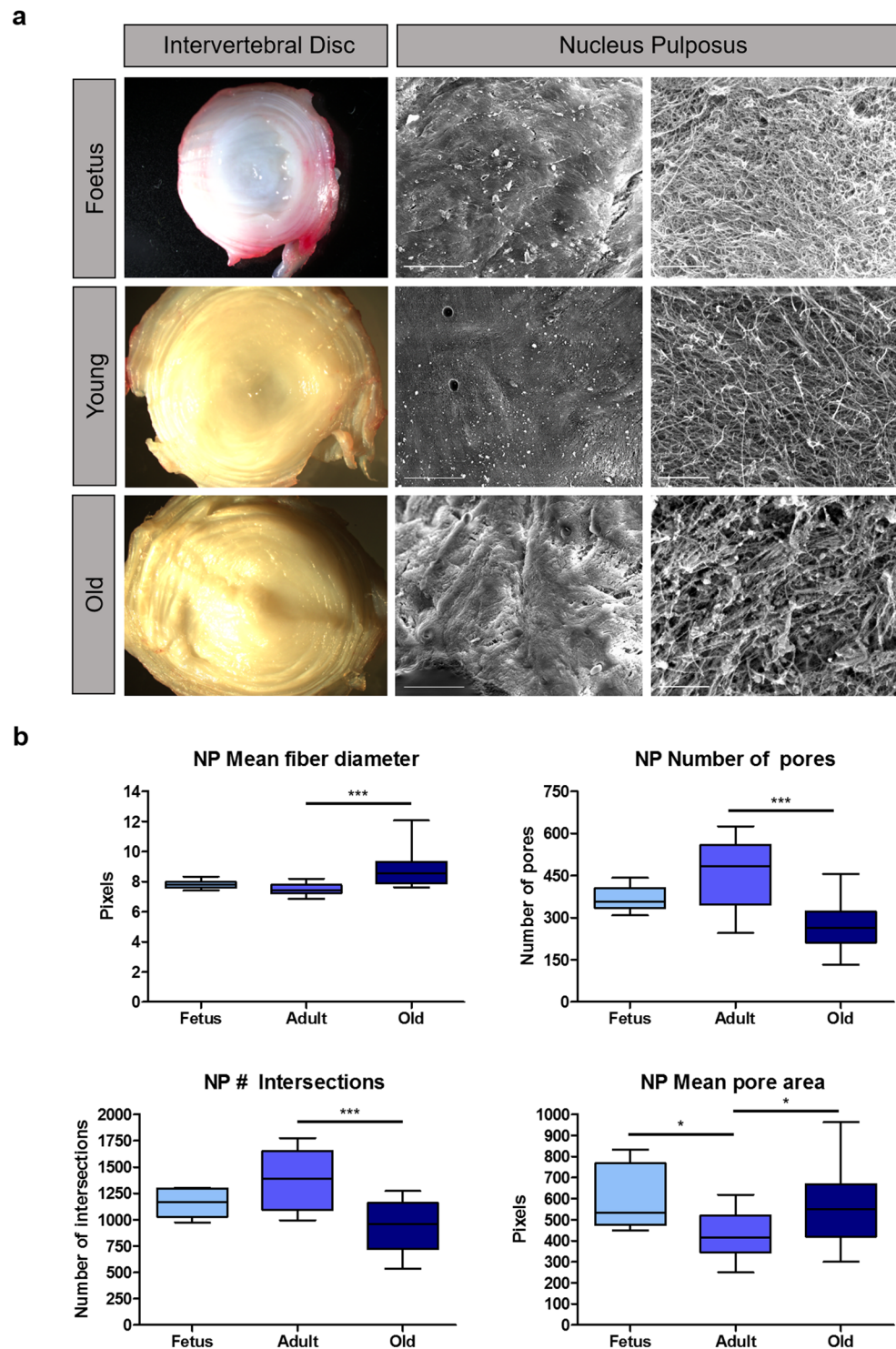
**Optimization of the proteomics workflow.** Age-associated protein expression profiles were identified in the first place by the distinct one dimension (1D) SDS-PAGE band signatures obtained. By comparing foetal, young and old NPs, we found that band intensities differed depending on disc age (Fig. 2a). Foetus and young samples presented an enrichment of high molecular weight proteins, between 150 and 250 kDa, where some of the molecules identified were tenascin, collagen type VI and biglycan (Fig. 2a–1F and Y and Supplementary Data 4). This trend was also visible among proteins around 50 kDa that corresponded to actin, link protein (HAPLN1), biglycan and decorin among others (Fig. 2a–6F,Y and O, Supplementary Table 3 and Supplementary Data 4). Noteworthy, we observed that young NPs had increased protein expression levels at around 30 kDa (Fig. 2a–7F,Y,O and 8F,Y,O). These bands corresponded to collagen type II and chondroadherin (Supplementary Table 3 and Supplementary Data 4). Moreover, above 100 kDa a band presented a trend to increase in Young and Old animals when compared to Foetus (Fig. 2a and 3F,Y,O). Fibromodulin, biglycan, aggrecan and cartilage oligomeric matrix protein (COMP) were identified within these bands (Supplementary Table 3 and Supplementary Data 4). In addition, prolargin (PRELP) was only identified in old samples (Fig. 2a–5O, Supplementary Table 3 and Supplementary Data 4) and not in the same molecular weight bands of Foetus or Young animals (Fig. 2a–5F and Y, Supplementary Table 3 and Supplementary Data 4). Nevertheless, the differences detected were only semi-quantitative and the fact that after MS identification of the molecules present in each of the gel bands, protein mixtures were found (Supplementary Table 3 and Supplementary Data 4), made it hard to dissect which were the ones accounting for the observed differences.

For a comprehensive age comparative analysis of the NP matrisome, we used a gel-free MS based high throughput proteomic approach. The first optimization step involved the decision on whether we could use the whole tissue extract or whether decellularization was needed. Comparing the protein band profiles of the whole tissue and of cell extracts obtained using the same buffer, more bands were observed within the cellular rather than the tissue extract. Nevertheless, MS/MS identification showed that the bands corresponding to ECM proteins (namely collagen type II, chondroadherin and mimecan) were mainly present in the tissue extract (Fig. 2b–1w, 2w and 3w, respectively, Supplementary Table 4 and Supplementary Data 5). In fact, most of the disc is made up of water, while ECM and cells represent a very small percentage in terms of tissue volume<sup>12</sup>. For this reason, we did not pursue further the laborious task of improving the decellularization step, given that it brought no added value.

We tested 3 different buffers taken from the literature<sup>29</sup>. For each of them, we evaluated the total number of bovine proteins identified, as well as the number of ECM-associated molecules obtained. For that we took advantage of the Functional Annotation Clustering Tool from DAVID (<https://david.ncicrf.gov/>). Qualitative results from Liquid Chromatography coupled to tandem Mass Spectrometry (LC-MS/MS) enabled us to select Guanidine Hydrochloride as the best buffer for our analysis, given that it enabled the identification of more bovine total proteins (146) and, particularly, more ECM molecules (26) than the chondroitinase-containing buffers (Fig. 2c and Supplementary Data 6).

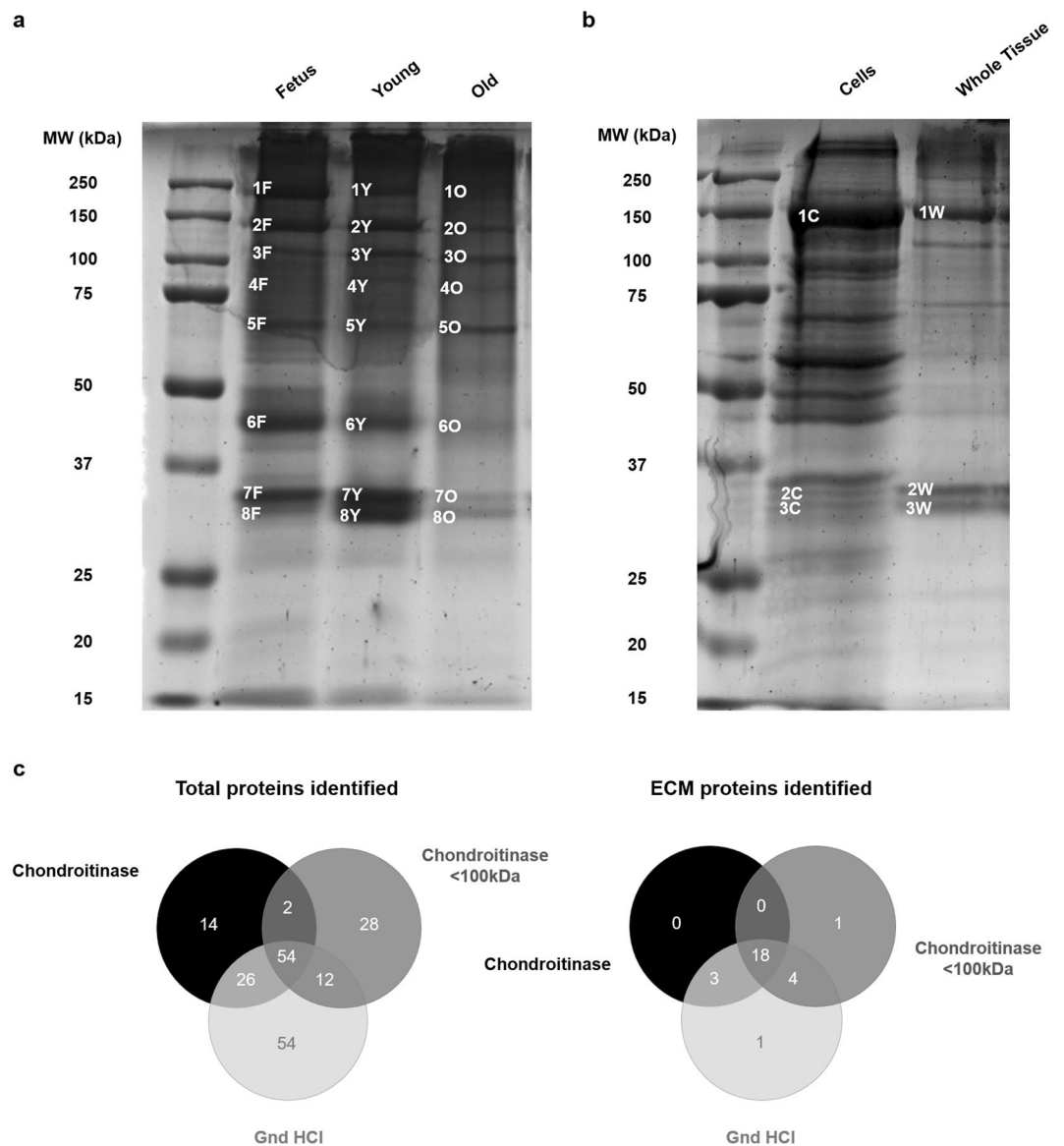
The optimized workflow that was used for sample processing is summarized in Fig. 3. Briefly, NPs were excised from foetal, young and old bovine tails and stored at  $-80^{\circ}\text{C}$  until further use. Samples were then snap frozen in liquid nitrogen and pulverized prior to the 24 hour protein extraction with the 4 M Guanidine Hydrochloride buffer. Protein was then precipitated, quantified, and digested. Following 8-plex isobaric tag for relative and absolute quantitation (iTRAQ) labelling, samples were pooled and fractionated by LC. Finally, the peptide mixture was analysed by LC-MS/MS and protein identification was performed using Protein Pilot. Of note, the labelled samples of each of the two batches were combined into one sample mixture (Batch 1: Y1, Y2, Y3, F1, F2, F3, O Pool, Total Pool; Batch 2: Y1, Y2, Y3, O1, O2, O3, F Pool, Total Pool – Supplementary Data 1 and Supplementary Table 1).

**Identification of NP age related proteomic signatures by iTRAQ analysis.** PCA-DA plots, obtained with MarkerView, show significant clustering and differentiation among foetus (green), young (blue) and old (red) animals (Supplementary Figure 2a). Despite biological variability within the distinct age groups, and



**Figure 1.** Gross view and topographical characterization of different aged IVDs. (a) - Macroscopic pictures from IVD transversal sections (left panels) and microscopy images from scanning electron microscopy (SEM) (central and right panels, with scale bars representing 200  $\mu\text{m}$  and 10  $\mu\text{m}$ , respectively) of foetal, young and old IVDs are presented. On the right panels are magnifications of the NP, representing the organization of ECM fibres. NP parameters such as mean fibril diameter, mean pore area, number of pores and number of intersections were also plotted regarding each age group (b). Box-plot lines represent median and interquartile ranges of the different parameters evaluated. (\*) stands for  $p \leq 0.05$  (\*\*) for  $p \leq 0.01$  and (\*\*\*) for  $p \leq 0.001$ , using the non-parametric Mann-Whitney test.

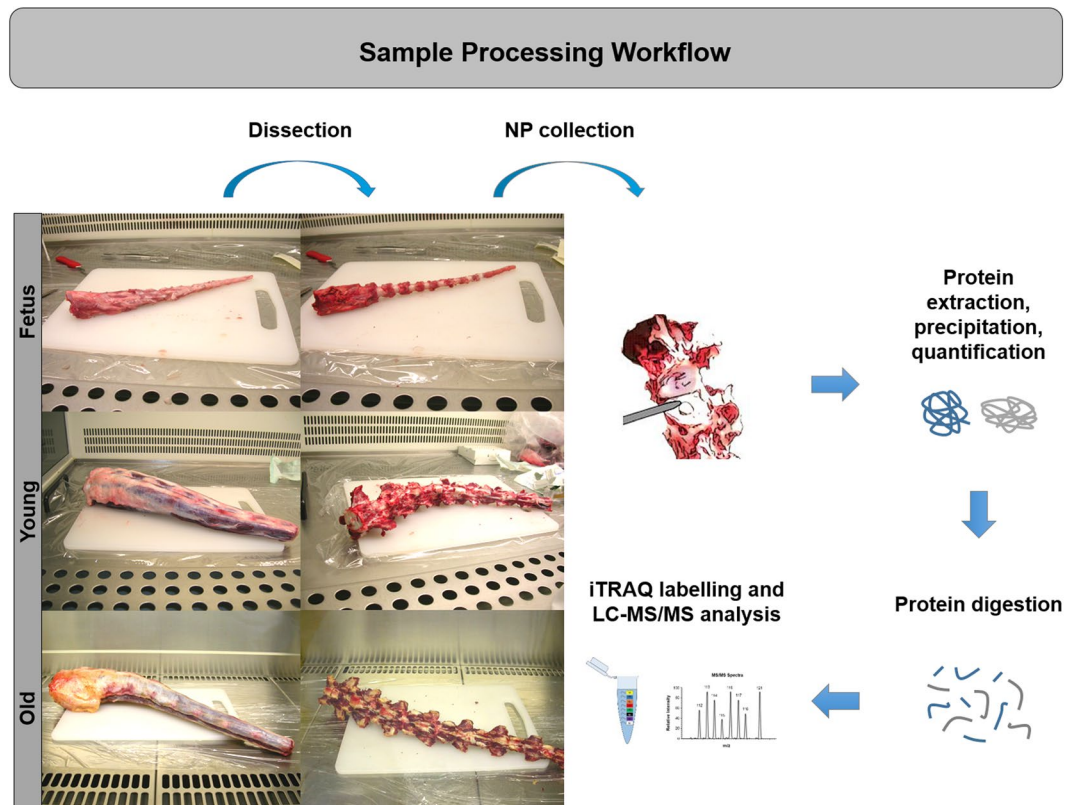




**Figure 2.** Sample complexity, age-associated SDS-PAGE band profile and extraction buffer evaluation. **(a)** - By comparing the band profiles observed in one dimension (1D) SDS-PAGE gels, for the three distinct age groups, there were already clear differences in the intensity of some of the bands identified by MS + MS/MS. Information on band identification from 1F to 8O can be found in Supplementary Table 3. **(b)** - With the same extraction buffer, we observed more bands within the cellular rather than the tissue extract. However, identification by MS + MS/MS (Supplementary Table 4) showed that the bands corresponding to ECM proteins were mainly present in the tissue extract (1C-COL6A1/COL6A2; 1W-COL6A1/COL2A1; 2C-ANXA1/ANXA2/GAPDH; 2W-COL2A1/CHAD/OGN; 3C-ANXA8/LDHA; 3W-CHAD). **(c)** - We compared the number of proteins identified by LC-MS/MS when using distinct buffers recommended from the literature: guanidine hydrochloride buffer (Gnd HCl), the same buffer with a prior chondroitinase 6 h treatment (chondroitinase) and this last condition followed by a centrifugal filtration, using a molecular weight cut-off of 100 kDa (chondroitinase <100 kDa). The extraction buffer with which more proteins in total were identified, and more ECM proteins, in particular, was Gnd HCl.

the fact that iTRAQ data was obtained from two independent runs, the three major groups of samples (Foetus, Young and Old) were well separated from each other, suggesting that NPs from different age sets presented distinct proteomes.

Averaged values of relative protein expression data from iTRAQ based LC-MS/MS (8-plex) assays were subjected to hierarchical clustering with Morpheus software. The heatmap and respective dendrogram generated are represented in Fig. 4a. This supervised analysis was performed by applying Spearman rank standardization. Foetuses were grouped together. Although they were part of the same cluster (showing that they both have similar protein expression profiles), Young and Old animals were also grouped according to their age status. Clustering data analysis demonstrated the ability to conduct global proteomics profiling on NP disc tissues, revealing that



**Figure 3.** Proteomic sample processing workflow. Bovine caudal IVDs from foetus, young and old animals were dissected within a few hours after slaughter. The NPs were collected and protein was further extracted, precipitated and quantified. After digestion, peptides were marked with isobaric tags for relative and absolute quantitation (iTRAQ) prior to liquid chromatography – tandem mass spectrometry (LC-MS/MS).

significantly different arrays of proteins were expressed depending on the age of the individuals, despite the fact that samples were collected from the same avascular organs in the bovine body.

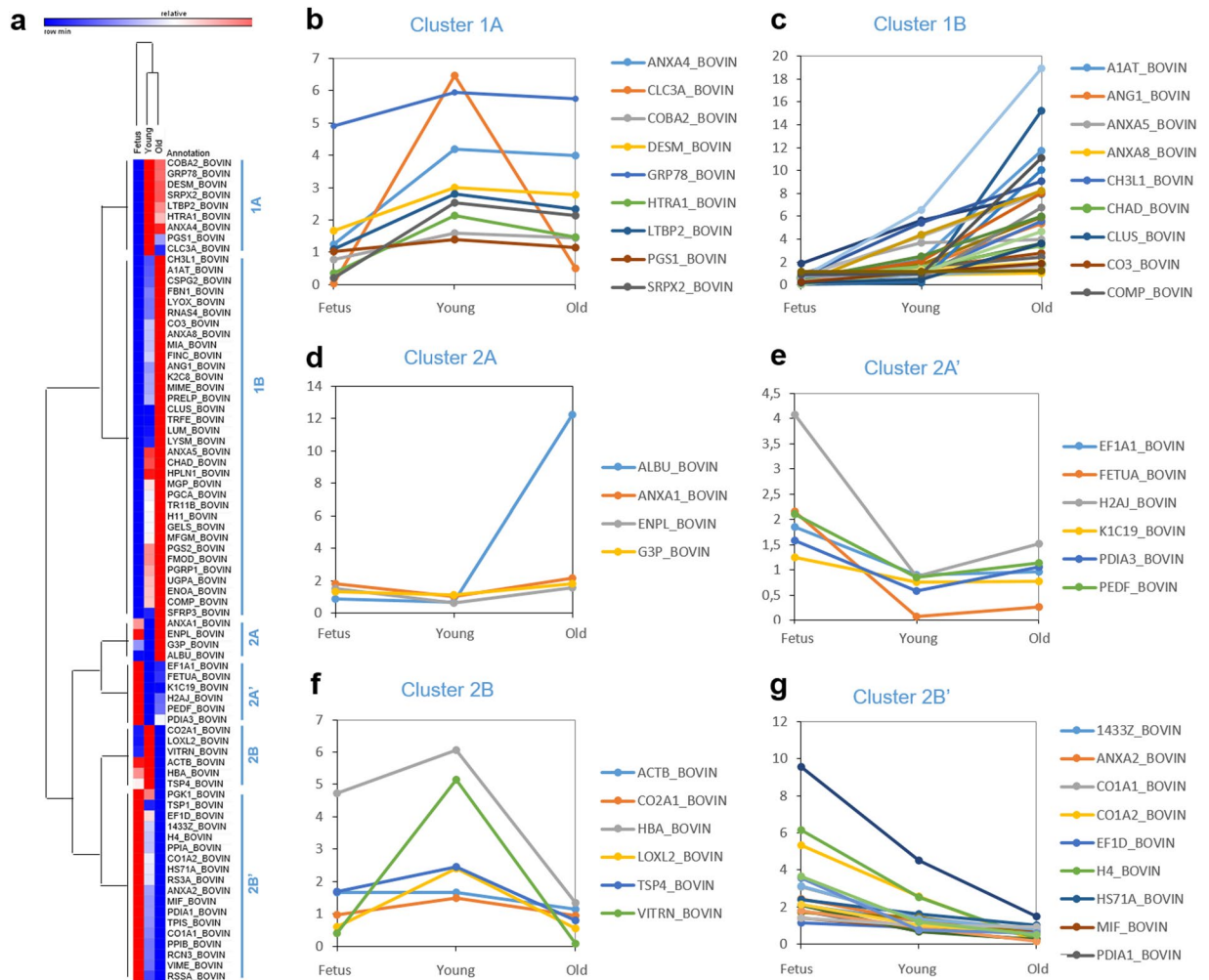
To explore the biological processes affected by the 77 (out of 161) common proteins identified in all the samples, and thus related to NP function, we performed Gene Ontology (GO) and Pathway term enrichment using the Functional Annotation Clustering Tool from DAVID Database. This allowed us to determine GO and Pathway terms that occurred more frequently than expected by chance. Proteins were then clustered according to functional similarity (Supplementary Tables 5 and 6).

The most significant cluster of proteins, in comparison to *Bos taurus* proteome, included GO terms implicated in extracellular matrix and GAG binding, which was in accordance to what we expected. Interestingly, we also detected a statistically significant group of interactors involved in collagen fibril organization, cartilage and blood vessel development, and also proteins implicated in lipid binding and vesicles. Other statistically significant biological terms that appeared were glycolysis, regulation of phagocytosis, response to wounding, inflammatory response and calcium ion binding, among others. Redoing the analysis concerning pathway term enrichment, we obtained “ECM-receptor interaction”, “Focal adhesion” and “integrin signalling”.

To integrate known and predicted protein-protein interactions and better understand the relationships between the 77 distinct proteins commonly identified in all 3 age groups, we used STRING (Supplementary Figure 2b).

The interactome obtained highlighted six different clusters marked in same colour circles, broadly representing: cytoskeleton (light green), fibril/collagen organization and skeletal/cartilage development (beige), GAG binding/crosslinking (dark green), glycolysis (red) and vesicle-associated proteins (magenta). Large protein interaction networks illustrated the high degree of connectivity and the presence of promiscuous hub proteins.

Finally, to further highlight correlations between the proteins within each of the 6 main heatmap clusters (cluster 1A, cluster 1B, cluster 2A, cluster 2A', cluster 2B and cluster 2B' – Fig. 4), we performed gene ontology enrichment analysis based on the proteins identified in the different clusters. Among the proteins within cluster 1B (which increase with development and ageing), we found a significant enrichment of Gene Ontology terms related to extracellular matrix, glycosaminoglycan, polysaccharide, carbohydrate, hyaluronic acid and ion binding, as well as cell adhesion and membrane bound organelles. Among the proteins from cluster 2A, there was an enrichment of the term ion binding, whereas among the cluster 2B' (proteins which decrease with development and ageing) there was an enrichment of terms such as melanosomes, cytoplasmic vesicles, collagen fibril organization, blood vessel development, endoplasmic reticulum and also ion binding. No other significantly enriched



**Figure 4.** iTRAQ data analysis. **(a)** - Heatmap with the respective dendrogram representing sample-based hierarchical clusters. Average expression levels were represented by colour scale from blue (low) to red (high). In terms of protein expression, there were six main clusters. **(b–g)** - Graphical representation of iTRAQ relative protein expression profiles for the molecules within each of the 6 main clusters: cluster 1 A **(b)**, cluster 1B **(c)**, cluster 2A **(d)**, cluster 2A' **(e)**, cluster 2B **(f)** and cluster 2B' **(g)**.

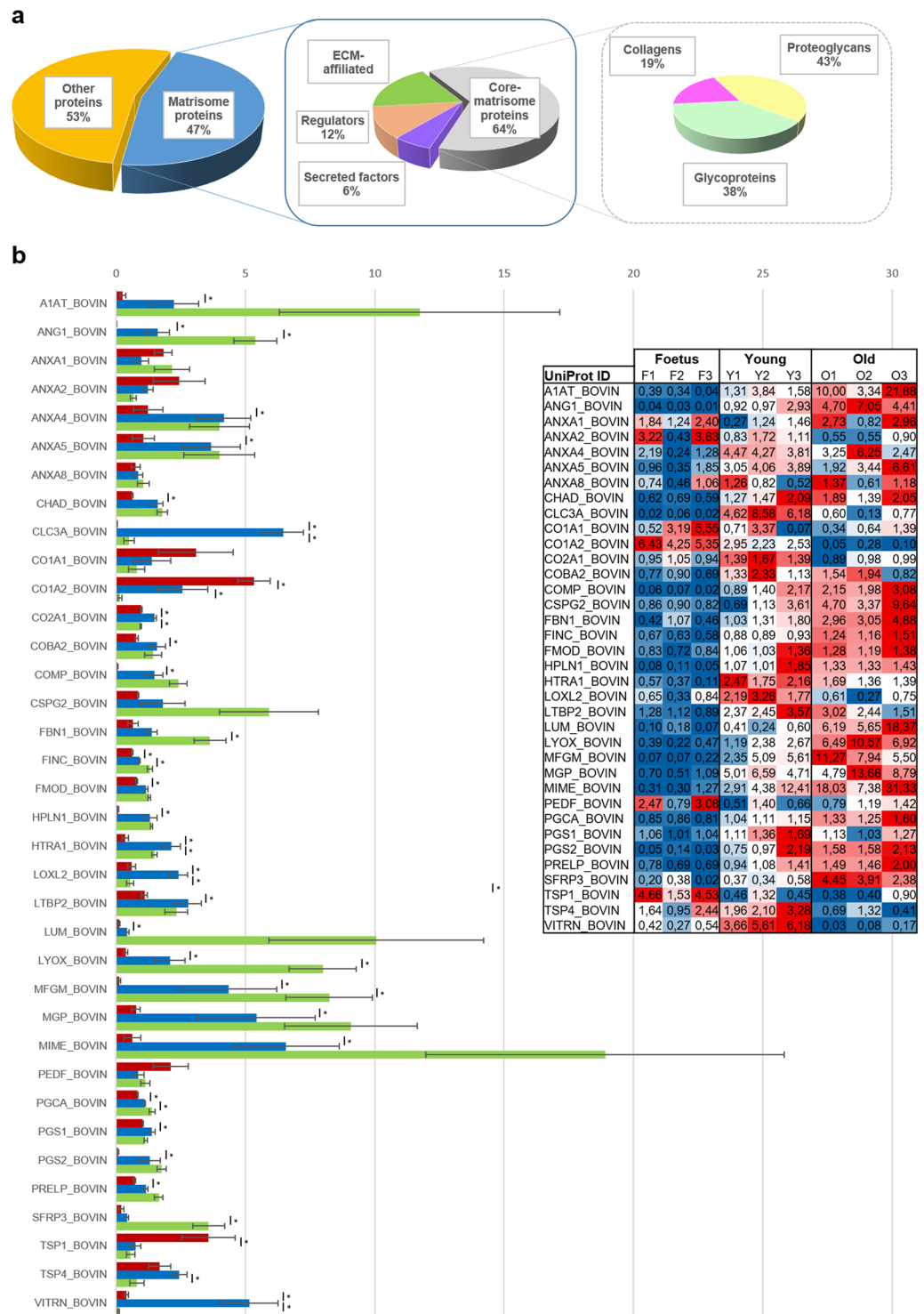
GO terms were found for other groups of samples. A detailed list of the enriched GO terms in each cluster can be found in the Supplementary Data 7.

**Definition of NP matrisome changes during development and ageing.** Given that GO analysis revealed enrichment of several ECM-related categories, we examined the overlap of the NP proteomic signature with the matrisome, a comprehensive list of genes coding for ECM molecules and regulators, which is significantly more comprehensive for data mining and for posing questions relevant to matrix biology than GO terms<sup>30</sup>.

Of note, 47% of the genes composing the NP signature encode for matrisomal proteins (Fig. 5a), and ECM-associated molecules. Core matrisomal proteins (64%) are over-represented, particularly in terms of proteoglycans (44%) and glycoproteins (39%). We further defined the NP matrisome as the 36 matrisomal proteins identified in the 3 age groups (relative protein quantification of such molecules is summarized in Fig. 5b).

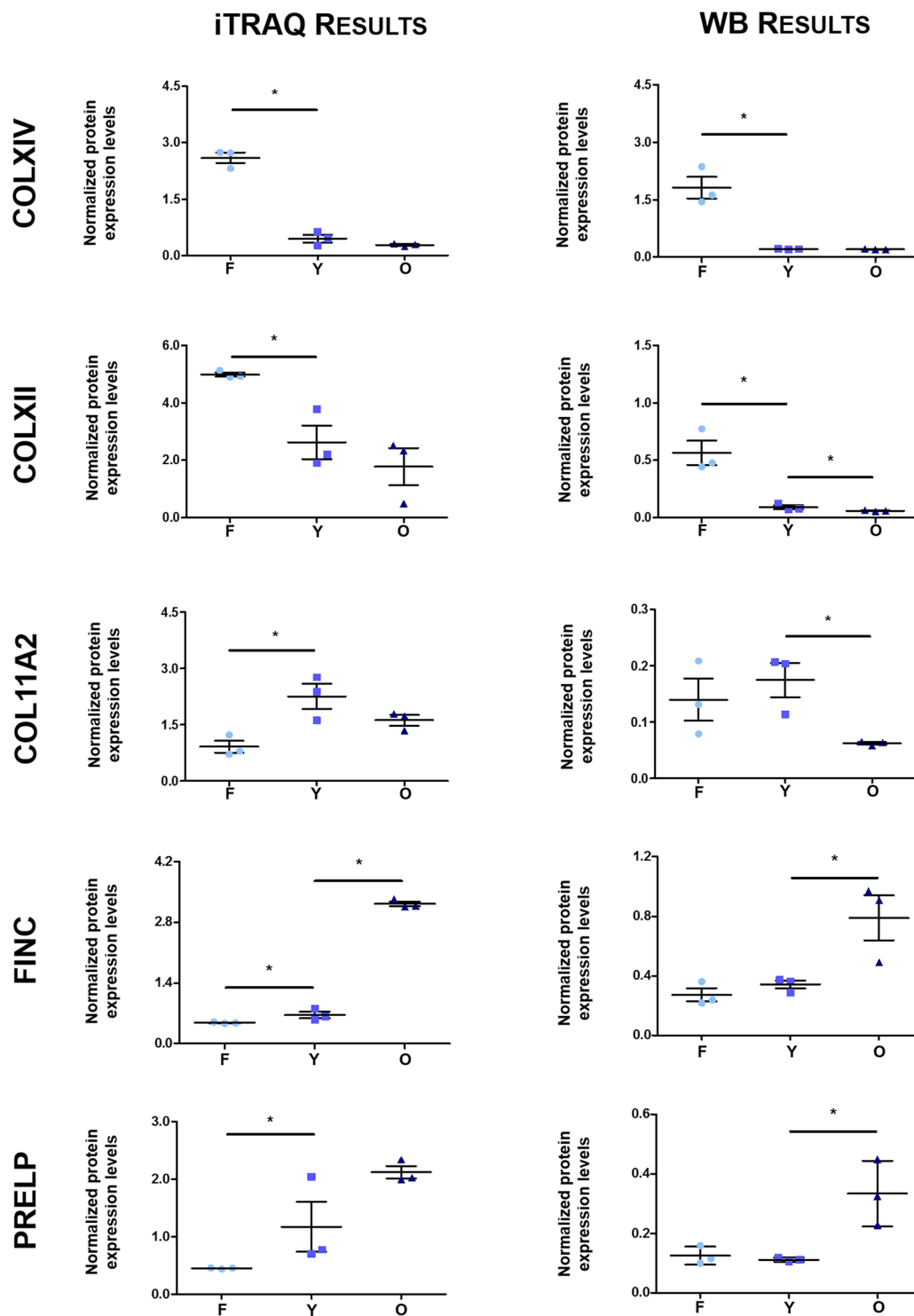
**Candidate validation by Western Blot.** The rationale behind the identification of potential age-related matrix components deregulated during NP development and choice of candidates for further investigation was: the analysis of iTRAQ data (e.g. differential protein expression ratio cut-off  $>1.3$  and  $<0.77$  and ProteinPilot protein identification confidence), literature mining concerning promising associations with IVD ageing and degeneration or lack of published data suggesting novel potential candidate biomolecules. Given that most proteins overexpressed in Foetus had not been identified with the highest confidence scores, we decided to further explore additional results obtained by searching the original MS/MS data against the proteomes from all organisms available on Swiss-Prot (results available upon request), instead of only using the *Bos taurus* complete proteome set.

Following these criteria, 5 proteins were selected for further validation: Collagen type XII, Collagen type XIV, Collagen type XI, Prolargin and Fibronectin. To confirm if these NP matrisome candidates were expressed in an age-dependent manner, Western blot analysis was performed (Fig. 6, right panel and Supplementary Figure 3).



**Figure 5.** Characterization of the NP matrisome. (a) - The pie charts exhibit percentages of identified proteins distributed by matrisome categories. (b) - Comparison of the different aged NP matrisome signatures. iTRAQ relative protein expression levels (x axis) are displayed for each of the molecules identified (y axis). Foetal samples are in red, young in blue and old in green. iTRAQ protein quantification scores for individual samples can be found on the embedded table (colour scale from blue – low expression – to red – high expression). In the case of Young animals, the values represent an average of the technical replicates. The non-parametric Mann-Whitney test was used to compare two groups of non-related samples. Standard error of the mean (SEM) is represented as the error bar. (\*) stands for  $p \leq 0.05$ , using the non-parametric Mann-Whitney test.





**Figure 6.** Validation of the candidates' protein expression levels. On the left are the graphics representing normalized protein expression levels from the iTRAQ LC-MS/MS analysis and on the right, of Western Blots (WBs), from Foetus, Young and Old animals. Collagen Type XII (COLXII) and XIV (COLXIV) expression levels are higher in NPs from foetus, Collagen Type 11 (COL11A2) is mostly expressed in Young animals, whereas Fibronectin (FINC) and Prolargin (PRELP) are typical from elder NPs. The non-parametric Mann-Whitney test was used to compare two groups of non-related samples (Foetus vs Young and Young vs Old). Graphs correspond to the average of protein expression levels obtained by band quantification and subsequent normalization to the total protein loading (Supplementary Figure 3). Standard error of the mean (SEM) is represented as the error bar. (\*) stands for  $p \leq 0.05$ , using the non-parametric Mann-Whitney test.

Collagen type XII and Collagen type XIV were significantly more expressed (6-fold and 9-fold, respectively) in NPs from Foetus rather than from Young animals (0.56 to 0.09 in Collagen type XII and 1.81 to 0.21 in Collagen type XIV). No other relevant protein expression differences were found for Collagen type XIV, when comparing Young and Old samples. Nevertheless, there was also a slight but significant 1.5-fold decrease of Collagen type XII from Young to Old age groups (0.09 to 0.06). Collagen type XI was shown to be expressed at significantly higher levels in Young NPs (0.17) in comparison to Old ones (0.06) and this enrichment at younger stages was also verified by immunofluorescence (Supplementary Figure 4). Fibronectin and Prolargin results reflected a significant (2-fold and 3-fold, respectively) overexpression (from 0.34 to 0.79 in Fibronectin and from 0.11 to 0.33 in Prolargin) only in older tissues. The results of this analysis were consistent with the trends obtained by exploring iTRAQ data (Fig. 6, left panel) and thus highlight the implication of Collagen type XII, Collagen type XIV, Collagen type XI, Prolargin and Fibronectin in age mediated IVD degeneration.

## Discussion

Disc degeneration and ageing are intimately associated. While disc ageing is a natural and gradually occurring process, disc degeneration involves more rapid and severe changes. In this study, we aimed to characterize NP proteomic changes that occur with development and ageing and that are closely related to disc degeneration, with a special focus on the ECM. For that, we have used iTRAQ labelling coupled to LC-based tandem mass spectrometry. This technique infers the relative abundance of individual proteins from peptide MS signal intensities and has emerged as an effective tool for quantitative proteomic profiling of complex tissue extracts, like cartilage<sup>31</sup>. SEM analysis and Western blotting also helped us to highlight and validate structural and molecular differences between the different age groups. The significance of our results is summarized as follows: several well-known effector proteins and a number of novel putative players were identified, 5 of which were independently validated.

Recently, proteomic-based studies have risen sharply and have started to be used to characterize normal and/or degenerated disc cells<sup>32,33</sup>, secreted factors<sup>34,35</sup> and tissue composition<sup>36</sup>, showing promising results. To date, however, limited studies have used proteomic strategies to study IVD matrix composition. One of the few existing studies focuses on the comparison of matrix proteomic signatures in different tissues<sup>37</sup>, while another is centred on the identification of cartilage matrix patterns of zonal distribution<sup>38</sup>. To the best of our knowledge, Yee *et al.* were the only ones that dissected ECM changes of human discs in age and degeneration<sup>39</sup>. Nevertheless, they used scoliotic samples as controls, which have been shown to present a gene expression profile that differs from healthy tissues, and signs of calcification that might reflect a premature degenerative process<sup>40,41</sup>. In addition, the age range of non-degenerated samples under study was very limited and not representative of elder individuals.

Unlike existing approaches, we used healthy samples from different age groups, and included, for the first time in this type of analysis, foetal tissue and NPs from very old animals. At the macroscopic level, we demonstrated that tissues, particularly the NP, tend to gradually become less translucent and AF lamellae appears increasingly disorganized, as has been previously reported for degeneration<sup>1</sup>. These observations were supported by SEM analysis through which we thoroughly characterized NP collagen fibres. Mean fibril diameter was reduced in younger samples, indicating that fibrillogenesis was somehow affected at later developmental stages. Fibril organization in elder individuals was also disrupted. Matrix was sparser, with less intersections and reduced numbers of pores, which were bigger in size. Foetal pore area was also larger, possibly due to imbibed water<sup>42</sup>. This may explain the gel-like appearance of foetal NPs and their greater capacity to absorb load and tension.

iTRAQ technology enabled the identification of 161 proteins in total, 77 of which were detected in all samples. From the bioinformatics analysis of the common hits, a few observations were predictable. The highest ranking functional cluster of genes was for those involved in extracellular matrix, which was consistent with the guanidine extraction method selected<sup>43</sup>. Additional cytosolic components equally important for disc homeostasis and function might not have been identified, but these were not the focus of this study. A significant enrichment of other proteins and protein classes like GAG, polysaccharide and carbohydrate binding, collagen fibril organization, blood vessel, skeletal system and cartilage development, as well as glycolysis, was also found and reflected a complex network of interactions, involving more than pure ECM biosynthesis. With regard to response to wounding, its link to disc de- and regeneration has been well established<sup>4</sup>. Interestingly, melanosome regulation and vesicle-mediated transport were also enriched and seem to have a role in embryonic elongation and spine morphogenesis<sup>44</sup>.

With respect to pathway term enrichment, ECM-receptor interaction, focal adhesion, integrin and TGF- $\beta$  signalling, were expected, given that most of these pathways are connected. In most of the tissues, integrin cell surface receptors mediate cell-matrix interactions, which are key to control adhesion, survival and differentiation, among others, in response to environmental cues, like mechanical stimuli<sup>45,46</sup>. Moreover, focal adhesion kinase (FAK) has been shown to be activated in response to strain in non-degenerate disc cells, in an integrin dependent manner<sup>47</sup>. TGF- $\beta$  can also be activated by integrin signalling, which, in turn, is affected by TGF- $\beta$ , whose bioavailability is controlled by ECM binding<sup>48</sup>. Most of these NP-associated functions might be deregulated not only in development but also during ageing and degeneration.

By assessing the overlap of the NP proteomic signature with a list of nearly 300 ECM and ECM-associated molecules generated by Naba and co-workers<sup>30</sup>, we defined the NP matrisome and how it is changed during development and ageing. Meta-analysis of such proteomic data was used to identify potential age-related components deregulated during NP development and select candidates for further investigation. Collagen Type XI, XII, XIV Fibronectin and Prolargin expression profiles were confirmed by Western Blot analysis. Collagen Type XI overexpression in young animals was also validated by immunofluorescence. Moreover, PRELP increase with age was additionally verified by SDS-PAGE band identification. Other proteins like Actin, Collagen Type II, Fibromodulin, Aggrecan and COMP, among others, presented the same trend in iTRAQ analysis as in gel band profiling.

Fibronectin (FN) mediates a wide variety of cellular interactions with the matrix and plays important roles in cell adhesion, migration, growth, differentiation and survival, particularly through integrin interactions<sup>4, 46</sup>. It interacts with a broad range of collagens (type I, II, III, IV, V and X)<sup>49</sup> and, in line with our results, FN has been shown to be upregulated in numerous models of disc ageing and degeneration<sup>35, 39, 50</sup>. Usually, 30 to 40%<sup>51</sup> of the protein content is in the form of fragments from enzymatic cleavage<sup>4, 52, 53</sup>, which in turn promote further degeneration<sup>54</sup>. While excessive FN deposition has been linked to fibrosis<sup>55</sup>, polymorphisms have also been associated with disc degeneration<sup>56</sup>, given that it mediates collagen deposition and thus preserves matrix structural integrity<sup>57</sup>. FN may also be involved in the clearance of tissue denatured collagen in age and degeneration, or of circulating fibrin after trauma or in inflammation<sup>58, 59</sup>.

Prolargin (PRELP) is an ECM structural component that anchors the basement membrane to the underlying connective tissue<sup>4</sup>. Apart from binding GAG chains, it also binds type I and type II collagen<sup>60</sup>. PRELP overexpression in mice disrupts collagen fibres (which decrease in content and size), with no influence in fibril diameter<sup>61</sup>. In accordance, we have shown that the matrix of aged samples presented a lower density of fibres, resulting in a reduced number of fibre intersections, as well as in fewer but larger pores. Interestingly, PRELP is absent in neonatal articular cartilage ECM, in contrast to its abundance later in life<sup>62</sup>. This age-related accumulation agrees with our bovine data. In turn, human scoliotic NPs, also present PRELP increase with age<sup>39</sup>, whereas in dogs it has been associated with degeneration<sup>35</sup>.

Collagen Type XI is a fibril-forming collagen required for embryonic development and its abundance is inversely correlated with fibre diameter<sup>4, 63</sup>. This explains our observations (no differences between Foetuses and Young samples, where it is expressed, but an increase in fibre diameter in elder individuals, where it is absent). Apart from having a role in fibrillogenesis (controlling lateral growth of collagen type II fibrils) and in mineralization, Collagen XI also binds PGs, particularly at the cell surface, being important to maintain tissue integrity and cohesion, particularly during matrix remodeling<sup>4, 64</sup>. In line with our results, Collagen XI polymorphisms have been associated with disc degeneration<sup>56</sup>. COL11A1 expression levels tend to decrease with the severity of degeneration, at least in part due to MMP-mediated degradation<sup>6, 65</sup>. Recent data from Yee *et al.* also seem to indicate that collagen type XI declines in the human scoliotic NPs with ageing, in agreement with our bovine model<sup>39</sup>.

Collagen Type XII is a typical collagen-organizer molecule that binds to collagen I containing fibrils, as well as to other matrix proteins, like COMP, modulating fibril organization and mechanical properties<sup>4, 66</sup>. It has been suggested to take part in the development of stromal architecture and tissue cohesion<sup>67</sup>, particularly by promoting matrix bridges formation essential for network communication<sup>68</sup>. In addition, Collagen XII seems to have a pro-regenerative role, at least in other tissues<sup>69, 70</sup>.

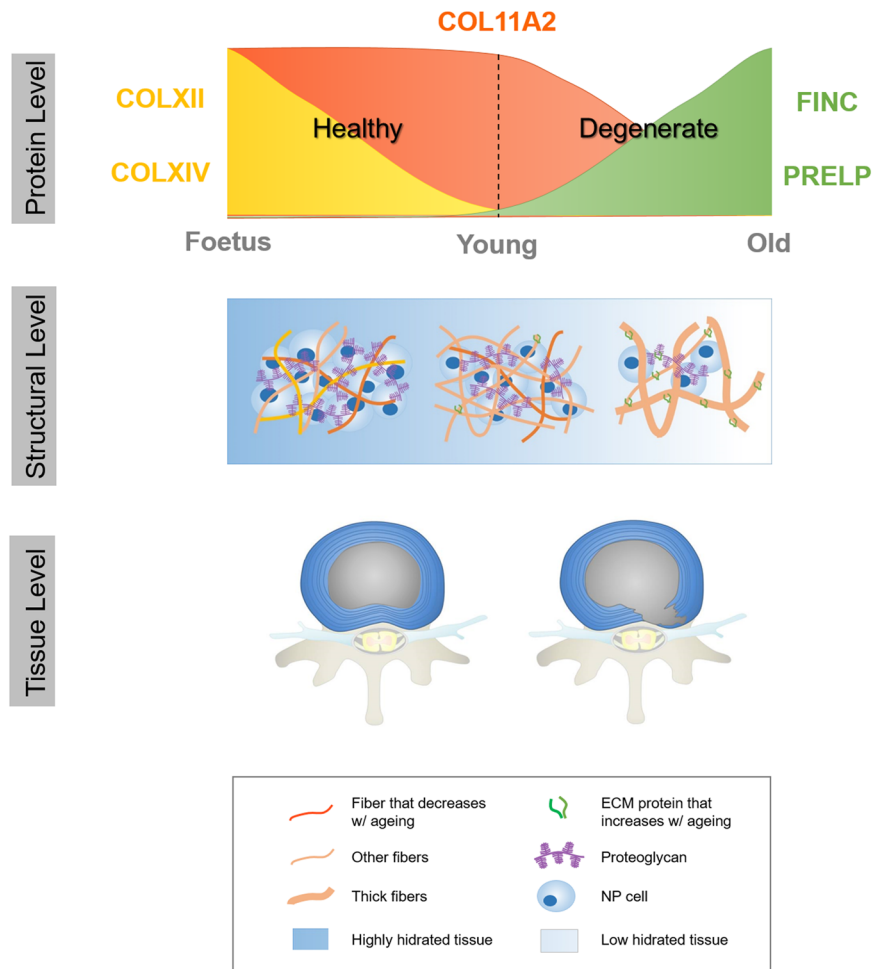
Collagen Type XIV is a fibril-associated collagen, transiently expressed in several epithelia, including those undergoing rapid remodeling. At later developmental stages, it only persists in the BM, where it co-localizes with Collagen XII<sup>71</sup>. Collagen XIV is thought to control collagen I fibrillogenesis during embryonic development (as has been supported by our results, it inversely correlates with fibril diameter and premature growth), as well as differentiation<sup>4, 66, 67</sup>. Interestingly, in the chick embryo it is expressed in a gradient around the spinal cord<sup>72</sup>. Like Collagen XII, it is key for the hydration and thickness (and therefore transparency) of tissues<sup>73</sup>, supporting our macroscopic observations of a gel-like appearance in foetal NPs. In other settings, type XIV collagen also appears to play a role in regeneration<sup>74, 75</sup>.

Overall, this study provides the first matrisome database of healthy discs during development and ageing, which is key to determine the pathways and processes required to maintain disc homeostasis. The integrated analysis of the proteomic datasets enabled us to discover novel components and characterize the developmental system in greater detail. The data herein presented may establish a solid foundation for better understanding the complex microenvironment of the IVD. Furthermore, they provide a starting point from which potential biomarkers and pathways that are altered during the dynamic disc degeneration process may be recapitulated or resumed, opening new possibilities for the development of novel therapeutic solutions for the disease. For instance, using these cues to modulate the ECM, recreating a microenvironment similar to early developmental stages, may enable the expansion and differentiation of autologous NP cells *in vivo*. In fact, others have shown that cell-free tissue engineering strategies are sufficient to promote disc regeneration<sup>76</sup>. Alternatively, the same approach could be combined with cell-based therapies currently under study<sup>23, 24</sup>.

Evidence about the use of bovine coccygeal discs as a model of ageing is still scarce. To date, existing studies have failed to use very old individuals. Nevertheless, a higher incidence of degenerative disc changes has been registered in elder specimens, indicating their clinical relevance<sup>77–79</sup>. In fact, ageing similarities between the two species were found, particularly in terms of extracellular matrix alterations<sup>77</sup>.

Moreover, although interspecies differences may naturally occur, bovine coccygeal discs are becoming increasingly accepted tissues for large animal organ culture, majorly because of their large size, low cost and availability. Given the similar aspect ratios, transport distances, and cell content, especially concerning the absence of notochordal cells during adulthood, between bovine and human lumbar discs, and contrarily to what happens in the vast majority of small animal models, they have been proposed as a suitable model to study several aspects of lumbar discs pathobiology<sup>12, 79–81</sup>. In particular, comparable types and distribution, synthesis and deposition of extracellular matrix molecules have been found in both human lumbar and bovine coccygeal discs<sup>82–86</sup>.

The work herein presented paves the way for future studies using human samples to validate the observed differences and investigating these constituents, thus further elucidating about their functional roles. Importantly, and since it is known that extractability of NP proteins reduces with age due to increased crosslinking and resistance to degradation, particularly concerning the collagen fibrillar network, chemical digestion proposed by Chan and co-workers should also be addressed to analyze the insoluble fractions obtained<sup>39</sup>.



**Figure 7.** Working model for the changes that occur in the IVD microenvironment with development and ageing. From the foetal stages to adulthood, collagen type XII and XIV expression is lost and only collagen type XI is maintained. With increasing age, there is an enrichment in fibronectin and prolargin. Proteomic alterations are accompanied by matrix remodeling (fibrillogenesis and fibril organization are both affected), concomitant with water loss and a cell population decline. This ultimately causes age-associated degeneration, hernia formation and low back pain.

In addition, reanalysis of this dataset focusing on the numerous cellular proteins identified will be of major importance, given that there are many other proteins equally important for disc homeostasis and function, which are still to be unveiled.

As a model (Fig. 7), we propose that remodeling of NP tissue architecture, which affects IVD mechanical properties and biological function of the tissue, reflects changes that occur in terms of matrix biochemical composition during development and ageing. Integrating information from the protein to the tissue level, taking into account the cell-matrix crosstalk, will provide helpful cues for disc regeneration<sup>26</sup>.

## References

- Adams, M. A. & Roughley, P. J. What is intervertebral disc degeneration, and what causes it? *Spine (Phila Pa 1976)* **31**, 2151–2161, <https://doi.org/10.1097/01.brs.0000231761.73859.2c> (2006).
- Anderson, D. G. & Tannoury, C. Molecular pathogenic factors in symptomatic disc degeneration. *Spine J* **5**, 260S–266S, <https://doi.org/10.1016/j.spinee.2005.02.010> (2005).
- Roughley, P. J. Biology of intervertebral disc aging and degeneration: involvement of the extracellular matrix. *Spine (Phila Pa 1976)* **29**, 2691–2699 (2004).
- Molinos, M. *et al.* Inflammation in intervertebral disc degeneration and regeneration. *J R Soc Interface* **12**, 20141191, <https://doi.org/10.1098/rsif.2014.1191> (2015).
- Le Maitre, C. L., Freemont, A. J. & Hoyland, J. A. Localization of degradative enzymes and their inhibitors in the degenerate human intervertebral disc. *J Pathol* **204**, 47–54, <https://doi.org/10.1002/path.1608> (2004).
- Galbusera, F. *et al.* Ageing and degenerative changes of the intervertebral disc and their impact on spinal flexibility. *Eur Spine J* **23**(Suppl 3), S324–332, <https://doi.org/10.1007/s00586-014-3203-4> (2014).
- Colombier, P., Clouet, J., Hamel, O., Lescaudron, L. & Guicheux, J. The lumbar intervertebral disc: from embryonic development to degeneration. *Joint Bone Spine* **81**, 125–129, <https://doi.org/10.1016/j.jbspin.2013.07.012> (2014).



8. Hwang, P. Y., Chen, J., Jing, L., Hoffman, B. D. & Setton, L. A. The role of extracellular matrix elasticity and composition in regulating the nucleus pulposus cell phenotype in the intervertebral disc: a narrative review. *J Biomech Eng* **136**, 021010, <https://doi.org/10.1115/1.4026360> (2014).
9. Takaishi, H. *et al.* Type-II collagen gene expression is transiently upregulated in experimentally induced degeneration of rabbit intervertebral disc. *J Orthop Res* **15**, 528–538, <https://doi.org/10.1002/jor.1100150408> (1997).
10. Cs-Szabo, G. *et al.* Changes in mRNA and protein levels of proteoglycans of the annulus fibrosus and nucleus pulposus during intervertebral disc degeneration. *Spine (Phila Pa 1976)* **27**, 2212–2219, <https://doi.org/10.1097/01.BRS.0000029249.02864.5B> (2002).
11. Gruber, H. E. & Hanley, E. N. Jr. Analysis of aging and degeneration of the human intervertebral disc. *Comparison of surgical specimens with normal controls. Spine (Phila Pa 1976)* **23**, 751–757 (1998).
12. Maroudas, A., Stockwell, R. A., Nachemson, A. & Urban, J. Factors involved in the nutrition of the human lumbar intervertebral disc: cellularity and diffusion of glucose *in vitro*. *J Anat* **120**, 113–130 (1975).
13. Deyo, R. A. & Weinstein, J. N. Low back pain. *N Engl J Med* **344**, 363–370, <https://doi.org/10.1056/NEJM200102013440508> (2001).
14. Waddell, G. Low back pain: a twentieth century health care enigma. *Spine (Phila Pa 1976)* **21**, 2820–2825 (1996).
15. Nasser, R. *et al.* Complications in spine surgery. *J Neurosurg Spine* **13**, 144–157, <https://doi.org/10.3171/2010.3.SPINE09369> (2010).
16. Swann, M. C., Hoes, K. S., Aoun, S. G. & McDonagh, D. L. Postoperative complications of spine surgery. *Best Pract Res Clin Anaesthesiol* **30**, 103–120, <https://doi.org/10.1016/j.bpa.2016.01.002> (2016).
17. Swartz, K. R. & Trost, G. R. Recurrent lumbar disc herniation. *Neurosurg Focus* **15**, E10 (2003).
18. Onesti, S. T. Failed back syndrome. *Neurologist* **10**, 259–264 (2004).
19. Masuda, K. *et al.* Osteogenic protein-1 injection into a degenerated disc induces the restoration of disc height and structural changes in the rabbit anular puncture model. *Spine (Phila Pa 1976)* **31**, 742–754, <https://doi.org/10.1097/01.brs.0000206358.66412.7b> (2006).
20. Walsh, A. J., Bradford, D. S. & Lotz, J. C. *In vivo* growth factor treatment of degenerated intervertebral discs. *Spine (Phila Pa 1976)* **29**, 156–163, <https://doi.org/10.1097/01.BRS.0000107231.67854.9F> (2004).
21. Leckie, S. K. *et al.* Injection of AAV2-BMP2 and AAV2-TIMP1 into the nucleus pulposus slows the course of intervertebral disc degeneration in an *in vivo* rabbit model. *Spine J* **12**, 7–20, <https://doi.org/10.1016/j.spinee.2011.09.011> (2012).
22. Yue, B. *et al.* Survivin-TGFB3-TIMP1 Gene Therapy Via Lentivirus Vector Slows the Course of Intervertebral Disc Degeneration in an *In Vivo* Rabbit Model. *Spine (Phila Pa 1976)* **41**, 926–934, <https://doi.org/10.1097/BRS.0000000000001474> (2016).
23. Okuma, M., Mochida, J., Nishimura, K., Sakabe, K. & Seiki, K. Reinsertion of stimulated nucleus pulposus cells retards intervertebral disc degeneration: an *in vitro* and *in vivo* experimental study. *J Orthop Res* **18**, 988–997, <https://doi.org/10.1002/jor.1100180620> (2000).
24. Sakai, D. *et al.* Differentiation of mesenchymal stem cells transplanted to a rabbit degenerative disc model: potential and limitations for stem cell therapy in disc regeneration. *Spine (Phila Pa 1976)* **30**, 2379–2387 (2005).
25. Molinos, M., Almeida, C. R., Goncalves, R. M. & Barbosa, M. A. Improvement of Bovine Nucleus Pulposus Cells Isolation Leads to Identification of Three Phenotypically Distinct Cell Subpopulations. *Tissue Eng Part A* **21**, 2216–2227, <https://doi.org/10.1089/ten.TEA.2014.0461> (2015).
26. Vizcaino, J. A. *et al.* The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res* **41**, D1063–1069, <https://doi.org/10.1093/nar/gks1262> (2013).
27. Hotaling, N. A., Bharti, K., Kriel, H. & Simon, C. G. Jr. Diameter]: A validated open source nanofiber diameter measurement tool. *Biomaterials* **61**, 327–338, <https://doi.org/10.1016/j.biomaterials.2015.05.015> (2015).
28. Welinder, C. & Ekblad, L. Coomassie staining as loading control in Western blot analysis. *J Proteome Res* **10**, 1416–1419, <https://doi.org/10.1021/pr1011476> (2011).
29. Wilson, R. *et al.* Changes in the chondrocyte and extracellular matrix proteome during post-natal mouse cartilage development. *Mol Cell Proteomics* **11**(M111), 014159, <https://doi.org/10.1074/mcp.M111.014159> (2012).
30. Naba, A. *et al.* The matrisome: in silico definition and *in vivo* characterization by proteomics of normal and tumor extracellular matrices. *Mol Cell Proteomics* **11**(M111), 014647, <https://doi.org/10.1074/mcp.M111.014647> (2012).
31. Wilson, R., Whitelock, J. M. & Bateman, J. F. Proteomics makes progress in cartilage and arthritis research. *Matrix Biol* **28**, 121–128, <https://doi.org/10.1016/j.matbio.2009.01.004> (2009).
32. Gilson, A., Dreger, M. & Urban, J. P. Differential expression level of cytokeratin 8 in cells of the bovine nucleus pulposus complicates the search for specific intervertebral disc cell markers. *Arthritis Res Ther* **12**, R24, <https://doi.org/10.1186/ar2931> (2010).
33. Ye, D. *et al.* Comparative and quantitative proteomic analysis of normal and degenerated human annulus fibrosus cells. *Clin Exp Pharmacol Physiol* **42**, 530–536, <https://doi.org/10.1111/1440-1681.12386> (2015).
34. Purmessur, D. *et al.* Notochordal conditioned media from tissue increases proteoglycan accumulation and promotes a healthy nucleus pulposus phenotype in human mesenchymal stem cells. *Arthritis Res Ther* **13**, R81, <https://doi.org/10.1186/ar3344> (2011).
35. Erwin, W. M. *et al.* The biological basis of degenerative disc disease: proteomic and biomechanical analysis of the canine intervertebral disc. *Arthritis Res Ther* **17**, 240, <https://doi.org/10.1186/s13075-015-0733-z> (2015).
36. Sarath Babu, N. *et al.* Quantitative proteomic analysis of normal and degenerated human intervertebral disc. *Spine J*, <https://doi.org/10.1016/j.spinee.2016.03.051> (2016).
37. Onnerfjord, P., Khabut, A., Reinholt, F. P., Svensson, O. & Heinegard, D. Quantitative proteomic analysis of eight cartilaginous tissues reveals characteristic differences as well as similarities between subgroups. *J Biol Chem* **287**, 18913–18924, <https://doi.org/10.1074/jbc.M111.298968> (2012).
38. Muller, C. *et al.* Quantitative proteomics at different depths in human articular cartilage reveals unique patterns of protein distribution. *Matrix Biol* **40**, 34–45, <https://doi.org/10.1016/j.matbio.2014.08.013> (2014).
39. Yee, A. *et al.* Fibrotic-like changes in degenerate human intervertebral discs revealed by quantitative proteomic analysis. *Osteoarthritis Cartilage* **24**, 503–513, <https://doi.org/10.1016/j.joca.2015.09.020> (2016).
40. Grad, S. *et al.* in *Global Spine J*. W0029.
41. Hristova, G. I. *et al.* Calcification in human intervertebral disc degeneration and scoliosis. *J Orthop Res* **29**, 1888–1895, <https://doi.org/10.1002/jor.21456> (2011).
42. Riches, P. E., Dhillon, N., Lotz, J., Woods, A. W. & McNally, D. S. The internal mechanics of the intervertebral disc under cyclic loading. *J Biomech* **35**, 1263–1271 (2002).
43. Belluoccio, D. *et al.* Proteomic analysis of mouse growth plate cartilage. *Proteomics* **6**, 6549–6553, <https://doi.org/10.1002/pmic.200600191> (2006).
44. Ellis, K., Bagwell, J. & Bagnat, M. Notochord vacuoles are lysosome-related organelles that function in axis and spine morphogenesis. *J Cell Biol* **200**, 667–679, <https://doi.org/10.1083/jcb.201212095> (2013).
45. Le Maitre, C. L. *et al.* Altered integrin mechanotransduction in human nucleus pulposus cells derived from degenerated discs. *Arthritis Rheum* **60**, 460–469, <https://doi.org/10.1002/art.24248> (2009).
46. Kurakawa, T. *et al.* Functional impact of integrin alpha5beta1 on the homeostasis of intervertebral discs: a study of mechanotransduction pathways using a novel dynamic loading organ culture system. *Spine J* **15**, 417–426, <https://doi.org/10.1016/j.spinee.2014.12.143> (2015).
47. Gilbert, H. T., Nagra, N. S., Freemont, A. J., Millward-Sadler, S. J. & Hoyland, J. A. Integrin - dependent mechanotransduction in mechanically stimulated human annulus fibrosus cells: evidence for an alternative mechanotransduction pathway operating with degeneration. *PLoS One* **8**, e72994, <https://doi.org/10.1371/journal.pone.0072994> (2013).

48. Margadant, C. & Sonnenberg, A. Integrin-TGF-beta crosstalk in fibrosis, cancer and wound healing. *EMBO Rep* **11**, 97–105, <https://doi.org/10.1038/embor.2009.276> (2010).
49. Steffensen, B. *et al.* Fragmentation of fibronectin by inherent autolytic and matrix metalloproteinase activities. *Matrix Biol* **30**, 34–42, <https://doi.org/10.1016/j.matbio.2010.09.004> (2011).
50. Yang, F., Leung, V. Y., Luk, K. D., Chan, D. & Cheung, K. M. Injury-induced sequential transformation of notochordal nucleus pulposus to chondrogenic and fibrocartilaginous phenotype in the mouse. *J Pathol* **218**, 113–121, <https://doi.org/10.1002/path.2519> (2009).
51. Oegema, T. R. Jr., Johnson, S. L., Aguiar, D. J. & Ogilvie, J. W. Fibronectin and its fragments increase with degeneration in the human intervertebral disc. *Spine (Phila Pa 1976)* **205**, 2742–2747 (2000).
52. Zhang, X., Chen, C. T., Bhargava, M. & Torzilli, P. A. A Comparative Study of Fibronectin Cleavage by MMP-1, -3, -13, and -14. *Cartilage* **3**, 267–277, <https://doi.org/10.1177/1947603511435273> (2012).
53. Ruel, N. *et al.* Fibronectin fragments and the cleaving enzyme ADAM-8 in the degenerative human intervertebral disc. *Spine (Phila Pa 1976)* **39**, 1274–1279, <https://doi.org/10.1097/BRS.0000000000000397> (2014).
54. Liu, H. F. *et al.* A novel rabbit disc degeneration model induced by fibronectin fragment. *Joint Bone Spine* **80**, 301–306, <https://doi.org/10.1016/j.jbspin.2012.07.009> (2013).
55. To, W. S. & Midwood, K. S. Plasma and cellular fibronectin: distinct and independent functions during tissue repair. *Fibrogenesis Tissue Repair* **4**, 21, <https://doi.org/10.1186/1755-1536-4-21> (2011).
56. Mayer, J. E. *et al.* Genetic polymorphisms associated with intervertebral disc degeneration. *Spine J* **13**, 299–317, <https://doi.org/10.1016/j.spinee.2013.01.041> (2013).
57. McDonald, J. A., Kelley, D. G. & Broekelmann, T. J. Role of fibronectin in collagen deposition: Fab' to the gelatin-binding domain of fibronectin inhibits both fibronectin and collagen organization in fibroblast extracellular matrix. *J Cell Biol* **92**, 485–492 (1982).
58. Antoniou, J. *et al.* The human lumbar intervertebral disc: evidence for changes in the biosynthesis and denaturation of the extracellular matrix with growth, maturation, ageing, and degeneration. *J Clin Invest* **98**, 996–1003, <https://doi.org/10.1172/JCI118884> (1996).
59. Pankov, R. & Yamada, K. M. Fibronectin at a glance. *J Cell Sci* **115**, 3861–3863 (2002).
60. Lewis, M. PRELP, collagen, and a theory of Hutchinson-Gilford progeria. *Ageing Res Rev* **2**, 95–105 (2003).
61. Grover, J., Lee, E. R., Mounkes, L. C., Stewart, C. L. & Roughley, P. J. The consequence of PRELP overexpression on skin. *Matrix Biol* **26**, 140–143, <https://doi.org/10.1016/j.matbio.2006.10.005> (2007).
62. Melching, L. I. & Roughley, P. J. A matrix protein of Mr 55,000 that accumulates in human articular cartilage with age. *Biochim Biophys Acta* **1036**, 213–220 (1990).
63. Hafez, A. *et al.* Coll1a1 Regulates Bone Microarchitecture during Embryonic Development. *J Dev Biol* **3**, 158–176, <https://doi.org/10.3390/jdb3040158> (2015).
64. Vaughan-Thomas, A., Young, R. D., Phillips, A. C. & Duance, V. C. Characterization of type XI collagen-glycosaminoglycan interactions. *J Biol Chem* **276**, 5303–5309, <https://doi.org/10.1074/jbc.M008764200> (2001).
65. Podrazky, V., Stovickova, J., Novotna, J. & Trnavsky, K. Evidence for the degradation of type XI collagen by bovine intervertebral disc- and articular cartilage extracts. *Biol Chem Hoppe Seyler* **374**, 497–500 (1993).
66. Agarwal, P. *et al.* Collagen XII and XIV, new partners of cartilage oligomeric matrix protein in the skin extracellular matrix suprastructure. *J Biol Chem* **287**, 22549–22559, <https://doi.org/10.1074/jbc.M111.335935> (2012).
67. Young, B. B., Zhang, G., Koch, M. & Birk, D. E. The roles of types XII and XIV collagen in fibrillogenesis and matrix assembly in the developing cornea. *J Cell Biochem* **87**, 208–220, <https://doi.org/10.1002/jcb.10290> (2002).
68. Izu, Y., Ezura, Y., Koch, M., Birk, D. E. & Noda, M. Collagens VI and XII form complexes mediating osteoblast interactions during osteogenesis. *Cell Tissue Res* **364**, 623–635, <https://doi.org/10.1007/s00441-015-2345-y> (2016).
69. El-Shabrawi, Y., Kublin, C. L. & Cintron, C. mRNA levels of alpha1(VI) collagen, alpha1(XII) collagen, and beta ig in rabbit cornea during normal development and healing. *Invest Ophthalmol Vis Sci* **39**, 36–44 (1998).
70. Tassava, R. A., Nace, J. D. & Wei, Y. Extracellular matrix protein turnover during salamander limb regeneration. *Wound Repair Regen* **4**, 75–81, <https://doi.org/10.1046/j.1524-475X.1996.40113.x> (1996).
71. Bader, H. L. *et al.* Zebrafish collagen XIV is transiently expressed in epithelia and is required for proper function of certain basement membranes. *J Biol Chem* **288**, 6777–6787, <https://doi.org/10.1074/jbc.M112.430637> (2013).
72. Fukushima, M., Hirano, S. & Tanaka, H. Dorsoroventral differential distribution of collagen type XIV around the spinal cord is regulated by the ectoderm. *Dev Growth Differ* **41**, 751–758 (1999).
73. Hemmavanh, C., Koch, M., Birk, D. E. & Espana, E. M. Abnormal corneal endothelial maturation in collagen XII and XIV null mice. *Invest Ophthalmol Vis Sci* **54**, 3297–3308, <https://doi.org/10.1167/iovs.12-11456> (2013).
74. Tono-Oka, S., Tanase, S., Miike, T. & Tanaka, H. Transient expression of collagen type XIV during muscle development and its reappearance after denervation and degeneration. *J Histochem Cytochem* **44**, 907–918 (1996).
75. Thorimbert, V., Konig, D., Marro, J., Ruggiero, F. & Jazwinska, A. Bone morphogenetic protein signaling promotes morphogenesis of blood vessels, wound epidermis, and actinotrichia during fin regeneration in zebrafish. *FASEB J* **29**, 4299–4312, <https://doi.org/10.1096/fj.15-272955> (2015).
76. Pfeiffer, M. *et al.* Intradiscal application of hyaluronic acid in the non-human primate lumbar spine: radiological results. *Eur Spine J* **12**, 76–83, <https://doi.org/10.1007/s00586-002-0478-7> (2003).
77. Demers, C. N., Antoniou, J. & Mwale, F. Value and limitations of using the bovine tail as a model for the human lumbar spine. *Spine (Phila Pa 1976)* **29**, 2793–2799 (2004).
78. Battie, M. C., Videman, T. & Parent, E. Lumbar disc degeneration: epidemiology and genetic influences. *Spine (Phila Pa 1976)* **29**, 2679–2690 (2004).
79. Boos, N. *et al.* Classification of age-related changes in lumbar intervertebral discs: 2002 Volvo Award in basic science. *Spine (Phila Pa 1976)* **27**, 2631–2644, <https://doi.org/10.1097/01.BRS.0000035304.27153.5B> (2002).
80. Illien-Junger, S., Walter, B. A., Mayer, J. E., Hecht, E. C. & Iatridis, J. C. In *The Intervertebral Disc* (eds I. M. Shapiro & M. V. Risbud) 353–371 (Springer Vienna, 2013).
81. Hansen, H. J. Studies on the pathology of the lumbosacral disc in female cattle. *Acta Orthop Scand* **25**, 161–182 (1956).
82. Oshima, H., Ishihara, H., Urban, J. P. & Tsuji, H. The use of coccygeal discs to study intervertebral disc metabolism. *J Orthop Res* **11**, 332–338, <https://doi.org/10.1002/jor.1100110304> (1993).
83. Roberts, S., Menage, J., Duance, V., Wotton, S. & Ayad, S. 1991 Volvo Award in basic sciences. Collagen types around the cells of the intervertebral disc and cartilage end plate: an immunolocalization study. *Spine (Phila Pa 1976)* **16**, 1030–1038 (1991).
84. Sakai, D., Nakai, T., Mochida, J., Alini, M. & Grad, S. Differential phenotype of intervertebral disc cells: microarray and immunohistochemical analysis of canine nucleus pulposus and annulus fibrosus. *Spine (Phila Pa 1976)* **34**, 1448–1456, <https://doi.org/10.1097/BRS.0b013e3181a55705> (2009).
85. Ishihara, H., McNally, D. S., Urban, J. P. & Hall, A. C. Effects of hydrostatic pressure on matrix synthesis in different regions of the intervertebral disk. *J Appl Physiol* (1985) **80**, 839–846 (1996).
86. Ohshima, H., Urban, J. P. & Bergel, D. H. Effect of static load on matrix synthesis rates in the intervertebral disc measured *in vitro* by a new perfusion technique. *J Orthop Res* **13**, 22–29, <https://doi.org/10.1002/jor.1100130106> (1995).

## Acknowledgements

We would like to acknowledge Daniela Silva for running the SEM, as well as Carnes Landeiro, SA and Matadouro e Carnes Linda Rosa, Lda for kindly providing the bovine tails. We are also grateful to Professor Katheleen Keller, Dr. Nick Morris and Dr. Gregory Lunstrum for the collagen XII and XIV antibodies. In addition, we thank Marta Pinto for the technical support on Western Blot troubleshooting, Ana Maria Magalhães and Stefan Mereiter for the tips on sample deglycosylation, as well as Catarina Leite Pereira, Morena Fiordalisi and Graciosa Teixeira for the help with tail dissection. Supported by: The Portuguese Foundation for Science and Technology (FCT) (Projects: PTDC/SAU-NEU/103728/2008, PTDC/NEU-NMC/0205/2012, UID/NEU/04539/2013; Post-Doc grants: SFRH/BPD/78187/2011-JC, SFRH/BD/88419/2012-CS and SFRH/BD/80577/2011-MM; and the Investigator Programme: IF/00638/2014. IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by FCT. This work was co-financed by FEDER - Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020 - Operational Programme for Competitiveness and Internationalisation (POCI), Portugal 2020, and by Portuguese funds through FCT—Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Inovação in the framework of the project “Institute for Research and Innovation in Health Sciences” (POCI-01-0145-FEDER-007274). This work also had the financial support of FCT / MEC through National Funds and, when applicable, co-financed by the FEDER via the PT2020 Partnership Agreement under the 4293 Unit I&D. Last but not least, it was also financed by QREN (Quadro de Referência Estratégica Nacional) and by The National Mass Spectrometry Network (RNEM) (REDE/1506/REM/2005).

## Author Contributions

Joana Caldeira—acquisition of data; analysis and interpretation of data; drafting of the manuscript. Cátia Santa—acquisition of data; analysis and interpretation of data. Hugo Osório—acquisition of data; analysis and interpretation of data. Maria Molinos—acquisition of data. Bruno Manadas—provided expertise in proteomics and critical revision of the manuscript for important intellectual content. Raquel Gonçalves—critical revision of the manuscript for important intellectual content. Mário Barbosa—provided expertise in the intervertebral disc field and supervision of the work.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-11960-0>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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

© The Author(s) 2017