Extracellular matrix composition of connective tissues: systematic review and meta-analysis

Turney McKee, Dentistry Division of Biomedical Sciences, McGill University, Montreal

Submitted April, 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Table of Contents

ABSTRACT

Objectives: The function of connective tissues depend on the physical and biochemical properties of their extracellular matrix (ECM), which are in turn dictated by ECM protein composition. We performed a systematic review and meta-analysis of publications containing quantitative information on the protein composition of human connective tissues. The primary objective was to obtain quantitative estimates for absolute and relative amounts of ECM proteins. The secondary objectives were to quantify the changes in ECM composition in distinct pathologies, and to build quantitative estimates of whole tissue composition.

Methods: Relevant articles were extracted from the Medline (OVID), EMBASE, and SCOPUS databases, screened by two co-authors, and included in meta-analysis if they contained absolute or relative quantification of proteins found in the ECM of human bone, adipose tissue, tendon, ligament, cartilage, as well as skeletal muscle.

Results: We generated absolute quantitative estimates for collagen in articular cartilage, intervertebral disc (IVD), skeletal muscle, tendon, and adipose tissue. In addition, quantifications for sulfated glycosaminoglycans were synthesized in articular cartilage, tendon and skeletal muscle; total proteoglycan in IVD and articular cartilage, fibronectin in tendon, ligament and articular cartilage, and elastin in tendon and IVD cartilage. We identified significant increases in collagen content in the annulus fibrosus of degenerating IVD and osteoarthritic articular cartilage, and in elastin content in degenerating disc. In contrast, IVD collagen content was decreased in the scoliotic IVD. Finally, we built quantitative whole-tissue breakdowns.

Conclusions: Quantitative estimates improve our understanding of composition of human connective tissues, providing insights into their function in physiology and pathology. This knowledge can further guide the development of tissue-mimetic scaffolds and implants.

ABSTRAIT

Objectifs: La fonction des tissus conjonctifs dépend des propriétés physiques et biochimiques de leur matrice extracellulaire (ECM), qui sont à leur tour dictées par la composition des protéines ECM. Nous avons effectué une revue systématique et une méta-analyse de publications contenant des informations quantitatives sur la composition protéique des tissus conjonctifs humains. L'objectif principal était d'obtenir des estimations quantitatives pour les quantités absolues et relatives de protéines ECM. Les objectifs secondaires étaient de quantifier les changements dans la composition de l'ECM dans des pathologies distinctes et de construire des estimations quantitatives de la composition tissulaire totale.

Méthodes: Les articles pertinents ont été extraits des bases de données Medline (OVID), EMBASE et SCOPUS, sélectionnés par deux co-auteurs et inclus dans la méta-analyse s'ils contenaient une quantification absolue ou relative des protéines trouvées dans l'ECM de l'os, tissu adipeux, tendon, ligament, cartilage, ainsi que le muscle squelettique humain.

Résultats: Nous avons généré des estimations quantitatives absolues pour le collagène dans le cartilage articulaire, le disque intervertébral (DIV), le muscle squelettique, le tendon et le tissu adipeux. De plus, les glycosaminoglycanes sulfatés ont étés quantifiés dans le cartilage articulaire, le tendon et le muscle squelettique — protéoglycane totale dans le DIV et le cartilage articulaire, fibronectine dans le tendon, le ligament et le cartilage articulaire, et l'élastine dans le tendon et le cartilage DIV. Nous avons identifié des augmentations significatives de la teneur en collagène dans l'anneau fibreux du DIV dégénératif et du cartilage articulaire arthrosique, et dans la teneur en élastine dans le disque dégénératif. En revanche, la teneur en collagène DIV a diminué dans la DIV scoliose. Enfin, nous avons construit des pannes quantitatives de tissus entiers.

Conclusions: Les estimations quantitatives améliorent notre compréhension de la composition des tissus conjonctifs humains, fournissant un aperçu de leur fonction en physiologie et en pathologie.

Acknowledgements

I would like to acknowledge the following individuals for their contributions and assistance in preparing this thesis.

Co-Supervisors and Committee Members

Dr. Svetlana Komarova

Dr. Dieter Reinhardt

Dr. Mari Kaartinen

Lab Members

Dr. Kerstin Tiedemann

Dr. Gulzhakhan Sadvakassova

Mr. Nicholas Mikolajewicz

Mr. George Perlman

McGill Librarian

Mr. Martin Morris

Author contributions: Introduction (Chapter 1): TJM, SVK. Review of the Literature (Chapter 2): TJM, SVK. Study design and methodology (Chapter 3): TJM, SVK, MM. Systematic search and analysis (Chapter 4): TJM, SVK, GP. Discussion (Chapter 5): TJM, SVK. All authors contributed to the critical revision and approval of the final manuscript.

1. Introduction and Objectives

Connective tissue is one of the four basic types of human tissue, and is primarily composed of fibrous extracellular matrix componentss [1]. Tendons, ligaments, adipose tissue, bone, cartilage, and intervertebral disc (IVD) are connective tissues involved in diverse physiological roles including nutrient storage, endocrine function, and providing structural integrity [2-4]. The physical properties of connective tissues are similarly varied, illustrated by the differences between adipose tissue, a loose tissue populated by large, lipid-containing adipocytes, and bone, a hard tissue consisting primarily of mineralized collagen fibrils [3, 5]. The protein composition of human connective tissues is described as differing substantially, however quantitative and comparative studies on the subject are few and far between.

The ECM is a composite of cell-secreted molecules that offers biochemical and structural support to cells, tissues, and organs [6]. In humans, the composition of the ECM can be broadly summarized as a combination of water, protein, and polysaccharide, with the precise balance of these three compartments reflecting functional requirements of the tissues. The structural requirements determine the mechanical properties of the ECM, which depend on the protein composition of the matrix, particularly the abundance of collagen and elastin [6]. The physiological relevance of these properties extends beyond simple structural integrity. The cells surrounded by the ECM are capable of sensing its rigidity by forming integrin-mediated interactions with the matrix and then applying and assessing a force [7]. The mechanical properties of the matrix are then interpreted and affect motility, proliferation, differentiation, and apoptosis [8-11]. Thus, the knowledge of the precise composition of different tissues is important for understanding their structure-function relationship.

Connective tissue pathologies are often closely associated with alterations in the composition and structure of the ECM. This is particularly well studied for disorders that affect cartilage, such as osteoarthritis, and IVD degeneration. It stands to reason that because of the matrix and tissue degradation observed in these disorders, there may be quantifiable differences between the protein composition of healthy and diseased connective tissues which may in turn provide insight into the pathology of the disorder itself and potentially guide the development of new diagnostic criteria.

The goal of this work was to perform a systematic review and meta-analytic synthesis of the published literature reporting quantitative information on the protein composition of human connective tissues. The primary objective was to obtain estimates for absolute and relative amount of ECM proteins in various tissues to serve as benchmarks for the development of preclinical models that better model the human environment. We set two additional secondary objectives: 1) to quantify the changes in ECM composition in distinct pathologies to better understand their molecular basis and provide avenues for developing potential diagnostic methods; 2) to combine the outcomes for ECM proteins with reference values for other components to build quantitative estimates of whole tissue composition.

2. Review of the Literature

2.1 Connective tissue, general introduction

Connective tissues are a primary component of the human body and account for a substantial portion of its total mass [12]. This diverse category includes tendons, ligaments, adipose tissue, bone, cartilage, and IVD. These durable connective tissues are well known for their ability to resist tension, compression, extension, and torsion, however advances in our understandings of their physiology have demonstrated an array of functions as diverse the tissues themselves. Together, connective tissues are known to provide mechanical support, facilitate motion, transport fluids and nutrients, anchor cell migration, promote healing processes, and send biochemical signals throughout the body [2-4, 13]. The precise functions and properties of individual connective tissues emerge from the proteomic composition of their underlying matrix, while the maintenance of this matrix is dependent on external influences including mechanical loading and biochemical signaling [14].

2.2 Extracellular matrix and its components

The connective tissue ECM is primarily composed of 3 types of macromolecules; collagenous and elastic fibers, proteoglycans, and glycoproteins [15]. Collagens, a family of proteins composed of 28 distinct types, are the most abundant proteins in the ECM, however also exist in transmembrane forms on the surface of cellular residents of the ECM [16]. Networks of collagen form the basis of the three-dimensional array that associates with other proteins and ECM macromolecular constituents [15]. Collagen type 1 is the most prominent of these, and while it is widely expressed across tissues, it is especially prominent in the dermis, bone, and tendon. Different forms of collagen are striated into seven categories depending on their homology and

function: fibrillary collagens, network-forming collagens, fibril-associated collagens with interrupted triple helices, anchoring fibrils, beaded-filament-forming collagens, and MULTIPLEXIN collagens. These collagens are composed of three α chains, which arrange in a triple-helix surrounded by ordered hydration networks [17]. Collagen type 1 and Collagen type 2 are fibrillary collagens that are known predominantly for their structural role in the ECM of connective tissues [18-20]. Collagen type 1 is the most abundant form of collagen, and assembles into fibers that form the mechanical scaffold for bone, skin, tendons, and blood vessel walls. Collagen type 2 is the primary protein component of articular cartilage, where it is thought to account for 60% of the dry weight of the tissue [20, 21]. Thus, these two collagens in particular are exceptionally relevant to the physical properties, and thus physiological role, of the ECM and overarching tissue.

Elastic fibers are additional ECM structures that provide elasticity to tissues and thus their abundance is tightly intertwined with the physical requirements of a tissue [22]. These fibers are formed in a regulated process by which tropoelastin monomers are secreted and deposited onto a scaffold of microfibrils, which are themselves structures composed predominantly of fibrillin-1 and constitute the periphery of the vertebrate elastic fiber [23, 24]. Cross-linking and polymerization of the tropoelastin monomers gives rise to elastin. and the resultant fiber further associates with an array of elastic fiber-associated proteins [25, 26]. Elastic fibers are especially important in tissues that undergo stretching, such as blood vessels, lungs, heart, bladder, and skin [25].

Proteoglycans constitute a major part of the ECM and are primarily organized into 4 families, two of which, lecticans and small leucine-rich family of proteoglycans (SLRPs), provide extracellular support [27, 28]. They are composed of core proteins covalently linked to at least one

glycosaminoglycan chain. These chains are long and unbranched polysaccharides characterized by a repeating disaccharide structure. There are 5 distinct glycosaminoglycan components of proteoglycans, and they are heparin sulfate, chondroitin sulfate, dermatan sulfate, hyaluronan, and keratin sulfate [27]. All connective tissue extracellular matrices are thought to contain proteoglycans of some form. The distribution and abundance of specific proteoglycans is variable, and dependent on the tissue in question, and can be further altered by external factors such as an active lifestyle and illness [29-31]. Aggrecan, a member of the lectican family, is known as the primary proteoglycan component of cartilage [29].

Glycoproteins are similarly ubiquitous in all connective tissues. They account for a small portion of the overall matrix, however they are critical in providing stability to the matrix as a whole and establishing a substrate for cellular adhesion [32]. There are a large number of glycoproteins, however fibronectin, laminin, thrombospondin, tenascin, osteopontin, and link protein and are among the most important and best understood. Fibronectin is particularly widespread and is critical in allowing cell adhesion through cell surface receptors such as integrin receptors, and adhesion is in turn one of the most vital processes in permitting cell-mediated remodeling and maintenance of healthy connective tissues [33]. Laminin is similarly involved in facilitating cell interaction with the ECM both directly by mediating adhesion and migration and indirectly by affecting differentiation of resident cell populations [33]. Thrombospondin is released by a variety of cell types during tissue growth, and interacts with cells to promote adhesion and cell growth [33]. Tenascin serves a complex array of functions including binding to proteoglycans as well as fibronectin, mediating cellular adhesion and anti-adhesion, and promoting wound healing [33]. Osteopontin sequesters calcium and promotes tissue calcification, serving a vital role in the ossification of bone [34]. Link protein is essential in stabilizing proteoglycan

aggregates, particularly in the cartilaginous matrix [35]. As a whole, glycoprotein constituents associate with the other macromolecular ECM components and are vital to the emergent tissue properties.

2.3 ECM remodeling and structural requirements

The ECM is a dynamic structure that is constantly being remodeled. The remodeling process is mediated by catabolic processes driven by proteolytic degradation, and cell-mediated anabolic matrix deposition. Proteolytic degradation of the ECM is orchestrated by enzymes including matrix metalloproteases (MMPs), cathepsins, a disintegrin and metalloproteases (ADAMs), ADAMs with thrombospondin motifs (ADAMTSs), and plasminogen activators [15]. Cellular inhabitants of the ECM are in turn responsible for producing and secreting the macromolecular constituents. Fibroblasts in particular are the primary contributors towards secreting and organizing the fibrous protein and proteoglycan building blocks of the ECM. Chondrocytes are the only cellular resident in articular cartilage, and thus are charged with the deposition of the cartilaginous matrix [36]. Osteoblasts are the third primary cellular contributor to the connective tissue ECM, and are highly specialized cells that deposit calcium hydroxyapatite onto an existing matrix, thus forming bone [37].

The catabolic and anabolic processes underlying ECM remodeling can both contribute to alterations in the structural and biochemical functions of the matrix, which can in turn feedback onto the remodeling process [38]. Resident cells receive signals through specialized cell-surface receptors, namely integrins, discoidin domain receptors (DDRs), and syndecans [39-41]. These receptors associate with macromolecular constituents of the matrix and the resultant signals affect cell growth, survival, motility, invasion, and differentiation [8-11].

The ECM serves a vital role as a reservoir for numerous cytokines, such as TGF-β, members of the fibroblast growth factor family, and Wnt factors [42-44]. These factors associate with components of the ECM leading to their sequestration and eventual release. Proteolytic cleavage of ECM components can release bioactive fragments of the matrix itself, and thus the precise identify of the proteins included can directly impact the surrounding biochemical environment [45]. Thus, cells are capable of perceiving and reacting to the tissue-specific matrix that surrounds them, and understanding its composition may in turn further our understanding of cellular behavior.

2.4 Adipose tissue

Adipose tissue is a connective tissue that is functionally distinct from other members in the overarching family. Its primary functions are energy storage, endocrine regulation of energy metabolism, thermal insulation, and to provide cushioning from mechanical damage [46]. Adipocytes are the main cellular component of the tissue and are responsible for the majority of the deposition of the matrix as well as subsequent turnover and remodeling. Mature adipocytes are dominated by a single lipid droplet that is prone to rupture. A strong external matrix is vital to protect individual adipocytes, and thus is turn vital to the health and stability of the tissue as a whole [46].

Individual mature adipocytes are encapsulated in a dense ECM known as the basal lamina that is composed primarily of collagen IV [47]. Collagens type I and III are similarly known to be abundant in the tissue [48]. Adipose tissue features a diverse proteoglycan composition, however decorin and biglycan are noted as having especially significant structural and signaling roles [49, 50]. Numerous glycoproteins have been identified in human adipose tissue including Laminin, fibronectin, tenascin, and thrombospondin [46].

The ECM of developing adipose tissue is markedly different from that of mature tissue. The thick protective basal lamina that is most prominent in mature adipose tissue is thought to restrict the growth and development of preadipocytes. The matrix of developing adipose tissue is fibrillary in nature and consists primarily of collagens type I and III. Synthesis of laminar collagens is increased throughout preadipocyte development and joins the existing network of fibrillary collagens that surround the cell [46]. Thus, the proteomic composition of the adipose tissue matrix at a given point in time can be used to infer information regarding the state of the tissue as a whole.

2.5 Tendon and ligament

Tendons and ligaments are key connective tissues that connect muscles to muscle and muscles to bone, respectively. Their function results in some of the highest forces experienced by any tissue, and thus their strength, conferred by their ECM, is critical to normal physiology [51]. The ECM of these connective tissues is primarily collagen type I, and features small amounts of collagen types III, V, XI, XII, and XIV [52]. The non-collagenous constituents of this ECM are extremely varied. Decorin is the prominent proteoglycan constituent of the tendon and ligament ECM, and can account for as much as 80% of the total proteoglycan present [53]. Versican and Aggrecan are prominent as well, particularly in the areas where the tissues wrap around joints and bone [54]. These proteoglycan components serve to increase water content and thus provide resistance to compression [55]. Lubricin, tenascin-c. and collagen oligomeric matrix protein are the primary glycoproteins in tendon and ligament,and are thought to localize to the sheath that surrounds the tissue in addition to interacting with the other macromolecular components. [56, 57].

2.6 Bone

Bone as a tissue is an extremely critical contributor to the structural integrity of the human body. Its extracellular matrix is incredibly diverse and features hundreds of distinct proteins. This

matrix provides a scaffold for the deposition of the mineral that provides bone with its characteristic hardness and strength [58]. The bone ECM is primarily composed of collagen type I [59]. Bone contains distinct compartments, each with subtle variations in ECM composition that extend beyond the universally high levels of collagen I. The marrow stroma is a loose connective tissue that provides a center for hematopoiesis and cellular differentiation. It's ECM is described as irregular and poorly characterized [60]. The endosteum is a highly cellularized compartment whose ECM is thought to be dominated by osteoid, or yet-unmineralized collagen I [61]. The periosteum features uniquely high levels of periostin [62]. The osteocyte perilacunar matrix is unmineralized and features high levels of dentin matrix protein 1 and matrix extracellular phosphoglycoprotein [63]. SLRPs constitute the majority of the proteoglycan abundance in bone. Bone includes significant amounts of biglycan, fibromodulin, lumican, and osteoaderin while also including some non-SLRPs such as Aggrecan [64].

2.7 Articular Cartilage

Articular cartilage is a specialized tissue that provides a smooth and strong surface for the transfer of mechanical loads. It's unique in that it is devoid of blood vessels and nerves, and due to its limited capacity for healing and repair the preservation of its extracellular matrix is of critical importance. Collagen type II is the predominant protein component of the cartilaginous ECM, while collagen type IX is also present to a substantial degree [65].

Cartilage is composed of three distinct zones. The superficial zone is thin and accounts for approximately 10-20% of the overall thickness of the tissue. Densely packed collagen fibers are arranged in a regular pattern parallel to the articular surface of the tissue. This is the most cellularized layer of the tissue and accounts for the majority of the tensile strength of the tissue. The transitional zone accounts for 40-60% of the total tissue volume is primarily composed of

proteoglycans that supplement a collagenous network. The deep zone accounts for the remaining 30% of the volume of the tissue and has an organized network of collagen fibrils that are perpendicular to the articular surface, conferring a substantial amount of resistance to compression, and features the highest proteoglycan content [65].

The proximity of tissue to the resident chondrocytes further divides articular cartilage into distinct regions. Directly adjacent to the cells lies the pericellular matrix; it is mostly proteoglycan, glycoprotein, and other non-collagenous protein [66]. The territorial matrix is outside of the pericellular matrix and is primarily composed of collagen fibrils. Its primary purpose is thought to be to protect cellular residents of articular cartilage from mechanical stress [67]. The interterritorial region accounts for the remainder of the tissue and is thus the largest of the three regions [65]. Collagen fibrils dominate this region in arrangements that depend on the zone in which they're found [65].

Type II collagen accounts for 90-95% of the collagens of the articular cartilage ECM, which account for about 60% of its dry weight. Aggrecan is the predominant proteoglycan found in articular cartilage and thus accounts for a substantial portion of the remaining dry weight. Decorin, Biglycan, and fibromodulin are present in significant amounts. Fibronectin is thought to be the dominant glycoprotein in articular cartilage, however the specific functions of cartilaginous glycoproteins are not well characterized [65].

The composition of the articular cartilage extracellular matrix is known to be altered under pathological conditions. Osteoarthritis (OA) is a common disorder occurring in 60% of adults over the age of 65 in Europe and North America. OA leads to pathological changes in the structure and composition of the cartilage matrix, which are known to coincide with excessive ECM degradation [68]. The collagen II network is normally extremely low turnover, with the half-life of collagen

type II molecules exceeding 117 years [69]. Phenotypic shifts in the resident chondrocyte population promotes the production of matrix proteins and degrading enzymes that causes irreversible changes in the composition of articular cartilage [70].

2.8 IVD

The IVD is a structure that forms a cartilaginous cushion between the vertebrae. It's composed of three distinct regions, the nucleus pulposus, the annulus fibrosus, and the end-plates, which are arranged in a sandwich structure. The nucleus pulposus is surrounded by the annulus fibrosus, which are sandwiched between end-plates [71]. Collagens account for approximately 70% of the dry weight of the disc. The annulus fibrosus is substantially more collagen-rich than the nucleus pulposus [72]. Collagens are radially distributed, with collagen type I accounting for the majority of the collagens in the outer annulus while collagen type II is the primary constituent of the inner annulus and nucleus [73]. Aggrecan and Versican dominate the IVD during skeletal growth while SLRPs are more present in the mature disc. Aggregating proteoglycans such as Aggrecan and Versican are always more common in the annulus fibrosus, while non aggregating SLRPs tend to accumulate in the more fluid nucleus pulposus. Consistent with this preferential distribution, the nucleus pulposus tends to be more proteoglycan rich in the mature disc [71].

Cartilaginous tissue in the IVD is similar to articular cartilage in that it is a fairly low turnover tissue, and thus changes in tissue homeostasis can result in pathology. Degeneration of the IVD is a prominent cause of morbidity in elderly populations and is linked to excessive matrix catabolism [74]. A great deal of variance exists around reported incidence rates of disc degeneration, however some reports suggest that the rate is as high as 70% in some populations, and the disorder is canonically thought to be linked to back pain [75, 76]. Increased abundance of non-collagenous protein is thought to occur in both the nucleus pulposus and annulus fibrosus of

the degenerating disc [77]. Increased levels of matrix metalloproteinases, collagenases, aggrecanases, and ADAMTs proteases have all been demonstrated, and thus the disease is widely thought to be linked with aberrant IVD matrix remodeling [71]. Similarly, adolescent idiopathic scoliosis is suggested to thought to involve abnormalities in IVD composition and maintenance [78]. Quantitative methods have been suggested as diagnostic tools in the identification of both of the aforementioned conditions, and thus improved understanding of the proteomics in question provides an avenue through which clinicians can develop diagnostic criteria and improve understanding of the disease [79].

2.9 Relevance and importance of proteomic composition

The relevance of the proteomic composition of the ECM may extend beyond understanding physiology and pathology. Modern therapeutic development has demonstrated that in-vitro-based screening techniques and pre-clinical disease models are unable to accurately predict the success of a developing therapy [80]. Failure to replicate the complex architecture of the human cellular environment and surrounding proteomic architecture is thought to be a critical driver of this failure [81]. Thus, improving understanding of the proteomic composition of human connective tissues is essential in improving the success of the therapeutic development process.

3. Methods

3.1 Information Retrieval

With the assistance of a medical librarian (MM), we constructed a robust search strategy and performed a computerized bibliographic search of the Medline (OVID) database, provided in **Appendix 1**. We then adapted this search strategy to search EMBASE and SCOPUS. The searches were performed March 10, 2017 and returned 8,341 non-duplicate publications. The search of references and citations identified 7 additional publications.

3.2 Study Selection

All screening and selection was performed by 2 co-authors (TJM and GP). Abstract/title screening identified papers that contained information on ECM proteins derived from human tissue. During full text screening, the studies were selected if they contained 1) quantification of proteins and method of quantification; 2) the quantified proteins included those found in the ECM; 3) samples were taken from native human bone, adipose tissue, tendon, ligament, cartilage or skeletal muscle; 4) the articles reported complete statistical information including mean values, sample sizes, and either standard deviation or standard error. Exclusion criteria were 1) no other studies reported on the protein or tissue of interest, making synthesis impossible; 2) the reported units were incompatible with the transformation to μg protein/mg dry tissue. All conflicts that arose throughout the screening process were discussed by the two reviewers until consensus regarding the satisfaction of inclusion and exclusion criteria was reached.

3.3 Data Collection

The reported tissue type, protein name, mean, sample size, and standard deviation/standard error were extracted, together with information on subject age, sex, pathology, sub-location within the tissue as well as the methodological information. If the data within a paper was stratified by a variable such as age, sex, or region within a given tissue the data were entered as separate datasets. The complete data pool is available in **Appendix 3**.

3.4 Data Transformation

Data that were not presented in the form of μg protein/mg dry tissue was adjusted using the following assumptions and resources: 1) ECM proteins + proteoglycans account for the entirety of the decellularized dry weight of cartilage [65]; 2) Water content of tendon is 62.5% [52]; 3) Water content of skeletal muscle is 75% [82]; 4) If publications provided direct measurements of water content, that information was used in lieu of reference values for transformations within that study; 5) Molecular masses were taken from genecards.org for moles to mass transformations.

3.5 Data Synthesis

We assumed random effects model, and calculated individual dataset weights (w_i) using inverse variance weighting:

$$w_i = \frac{1}{se(\theta_i)^2 + \tau^2} \quad (1)$$

Standard errors ($se(\theta_i)$) were calculated for datasets that supplied standard deviations:

$$se(\theta_i) = \frac{sd(\theta_i)}{\sqrt{n_i}} \quad (2)$$

Inter-study variance (τ^2) was calculated using the DerSimonian and Laird method

$$\tau^2 = \frac{Q-(N-1)}{c} \quad (3)$$

Q statistics (Q) and concordance statistics (c) were calculated as follows:

$$Q = \sum_i \left(se(\theta_i)^{-2} \left(\theta_i - \frac{\sum_i se(\theta_i)^{-2}\theta_i}{\sum_i se(\theta_i)^{-2}} \right)^2 \right) \quad (4)$$

$$c = \sum_i se(\theta_i)^{-2} - \frac{\sum_i \left(se(\theta_i)^{-2}\right)^2}{\sum_i se(\theta_i)^{-2}} \quad (5)$$

Outcomes reported in individual datasets (θ_i) were synthesized into a weighted outcome ($\hat{\theta}$), according to individual dataset weights (w_i):

$$\hat{\theta} = \frac{\sum_i (\theta_i \cdot w_i)}{\sum_i (w_i)} \quad (6)$$

95% confidence intervals (CI) were calculated using a Z distribution:

$$\pm CI = \pm 1.96 \cdot se(\hat{\theta}) \quad (7)$$

3.6 Heterogeneity assessment

We report heterogeneity measures H^2 and I^2, which are in turn dependent on the total variation within our overall data pool (Q_{total}):

$$Q_{total} = \sum_{i=1}^{N} \left(w_i \cdot \left(\theta_i - \hat{\theta}_{FE} \right)^2 \right) \quad (8)$$

where $\hat{\theta}_{FE} = \frac{\sum_i se(\theta_i)^{-2}\theta_i}{\sum_i se(\theta_i)^{-2}}$ and $w_i = se(\theta_i)^{-2}$

$$H^2 = \frac{Q_{total}}{df} \quad (9)$$

$$I^2 = \frac{H^2-1}{H^2} \cdot 100\% \quad (10)$$

3.7 Cumulative and Single-Study exclusion plots

Cumulative exclusion plots were generated by iteratively removing the largest contributors to overall heterogeneity until a predefined homogeneity threshold was reached, assessed using a X^2 distribution (p = 0.05). Single study exclusion plots assessed the effect of removal of each individual dataset. It is important to note that both cumulative and single study exclusion plots were constructed for the analysis of heterogeneity only, and the studies remained incorporated into our overall estimates.

3.8 Funnel plots

Study level effects (θ_i) are plotted in relation to their inverse standard error. Theoretical 95% CIs are included to assist in visualizing an unbiased dataset.

3.9 Native tissue estimates

The estimates were generated for all tissues where 1) relative proteomic information was available; 2) an estimate for total, absolute collagen was calculated; 3) total collagen was included in the relative proteomic information. The estimated mass of a protein (M_p) was calculated as a function of its relative amount (R_p) to the relative amount of collagen ($R_{Collagen}$) and the absolute estimate for the mass of collagen in the tissue ($M_{collagen}$)

$$M_p = \frac{R_p}{R_{Collagen}} \times M_{collagen} \quad (11)$$

When the estimates for non-collagenous tissue constituents calculated during the study were based on at least 3 data sets, they were used in lieu of (11).

4. Results

4.1 Study Selection

The electronic search for reports containing quantification of ECM proteins in connective tissues returned 8,341 papers of which 203 were identified as potentially relevant after title/abstract screening. Full text screening yielded 37 articles containing quantitative information on at least one ECM protein. (**Fig. 1A**). From the selected articles, 12 reported the absolute quantitative information for proteins derived from articular cartilage [83-94]; 9 for intervertebral discs (IVD) [78, 95-102]; 3 each for tendon [103-105] and skeletal muscle [106-108]; 2 for ligament [103, 104]; and 1 for adipose tissue (**Fig 1B**) [109]. A single protein was quantified in 8 studies; multiple proteins in 19 studies; 5 studies quantified the entire tissue proteome; and 5 report on tissue-level composition (**Fig. 1C**). Absolute quantification was provided in 30 papers, while 7 studies reported relative values. In the studies reporting absolute protein quantification, several stratified data by age, sex, tissue sub-location and pathological state, which we extracted as 580 individual datasets. Pathology types covered in the selected papers included 173 datasets for the ECM composition in states of IVD degeneration; 54 for osteoarthritis; 26 for scoliosis; 12 for osteochondral lesions and 6 for other pathologies, including diabetes and obesity (**Fig. 1D**). In total, the absolute amount of 89 unique proteins was reported (**Appendix 4**). After data appraisal, we selected to perform meta-analysis for 5 most commonly reported proteins, which included collagen, sGAG, elastin, fibronectin and proteoglycan. Four studies, all reporting on articular cartilage, were excluded from meta-analysis since they did not provide information on these 5 proteins.

4.2 Data Distribution and Heterogeneity

Collagen quantification in healthy IVD (34 datasets) and articular cartilage (12 datasets) were the largest data pools and thus most suited for synthesis and meta-analysis. We constructed funnel plots to investigate publication bias in the reported datasets, which would result in an asymmetric distribution of points on the plot about our total mean. Data points are distributed evenly about the estimated effect size for IVD (**Fig 2A**), and for articular cartilage (**Fig 2B**). The weighted distribution and normal probability plots indicated that values reported in the IVD data pool are approximately normally distributed (**Fig. 2C**), however the articular cartilage data deviated from normality (**Fig. 2D**). The single study and cumulative exclusion analysis (**Fig. 2E**) for the IVD data pool demonstrated that the individual study contributions to overall heterogeneity were fairly consistent, and removal of 53% of the datasets generated a homogenous data pool. In the articular cartilage data pool, no individual study markedly contributed to the overall heterogeneity, and removal of 29% of the datasets generated a homogenous data pool (**Fig. 2F**).

4.3 Meta-analysis of collagen abundance in connective tissues

Using a random effects model, we estimated collagen abundance in connective tissues (**Fig. 3**). IVD (n = 34 datasets reporting N = 207 samples) was found to contain 385 μg collagen/mg dry tissue (95% confidence interval (CI): 350, 420); articular cartilage (n = 12 datasets reporting N = 182 samples) – 708 μg collagen/mg dry tissue (95% CI: 668, 748), skeletal muscle (n = 10 datasets reporting N = 65 samples) – 80 μg collagen/mg dry tissue (95% CI: 72, 88), and tendon (n = 2 datasets reporting N = 13 samples) 149 μg collagen/mg dry tissue (95% CI: 72, 226). We identified a single dataset for adipose tissue (6 samples) which reported collagen abundance to be 294 μg

collagen/mg dry tissue (95% CI: of 279, 309). Two cartilaginous tissues, IVD and articular cartilage, were found to contain significantly different amounts of collagen.

4.4 Investigation of methodological and biological contributors to heterogeneity

Since the degree of heterogeneity in the two most populated cartilage data pools was high (IVD: $H^2 = 708$, $I^2 = 99.9\%$; Articular cartilage: $H^2 = 9.6$, $I^2 = 89.6\%$), we next investigated whether factors related to methodological choices (**Fig. 4**) or biological factors (**Fig. 5**) contribute to the heterogeneity of the datasets. Different methodologies were employed to quantify collagen levels, including original hydroxyproline quantification using Stegemann technique (3 datasets) and its Kivirikko (24 datasets) and Woessner (2 datasets) modifications, as well as ELISA (5 datasets). Data stratification demonstrates that significantly lower collagen content was estimated using ELISA compared to other techniques (**Fig. 4A**). Studies varied in reporting the protein abundance in moles, micrograms, or a percent of total protein, related to the wet or dry weight of tissue. Therefore, the dataset values were transformed to micrograms per milligram of dry tissue weight prior to meta-analysis. We examined if data transformation systematically affected the outcome in IVD (**Fig. 4B,C**) and Articular cartilage (**Fig. 4D,E**) data pool, however data transformation did not significantly contribute to differences in outcome.

Several biological variables that could contribute to cartilage ECM composition were reported, including subject sex, age, and tissue sub-location. Subject sex or age did not significantly affect collagen abundance in IVD, even though an interesting age-dependent change was visually evident (not statistically significant by ANOVA, $p = 0.25$) (**Fig. 5A,B**). In IVD collagen abundance was significantly higher in the annulus fibrosus relative to both the endplate and the nucleus pulposus (**Fig. 5C**). Heterogeneity was reduced and data pools exhibited low bias and normal distribution (**Appendix 2**) when IVD data pools for annulus fibrosus ($n = 4$ reporting $N = 113$ samples, $H^2 =$

124.4, I^2 = 99.2%) and nucleus pulposus (n = 4 reporting N = 60 samples, H^2 = 43.5, I^2 = 97.7%) were separated. Collagen abundance in the lateral and medial condyle of articular knee cartilage was similar, however large variability in medial condyle reports was evident (**Fig. 5D**).

4.5 Estimated abundance for sGAG, fibronectin, elastin and proteoglycan

Robust analysis for non-collagenous ECM proteins was constrained by limited numbers of reports. Nevertheless, we estimated the content of sGAG, fibronectin, elastin, and proteoglycan in tendon, skeletal muscle, articular cartilage, ligament, and IVD (**Table 1**).

4.6 Changes in ECM composition in pathological samples

ECM composition in a pathological state was reported in 18 studies (**Table 2**). Disc degeneration was associated with a significant increase in elastin abundance, while scoliosis resulted in a significant decrease in collagen (**Fig. 6A**). Since we had identified IVD location as a significant determinant of collagen content, we further investigated if IVD pathologies differentially affect collagen content in different parts of IVD. We have found that IVD degeneration lead to significant increase in collagen in annulus fibrosus, but not in nucleus pulposus or intermediate zone (**Fig. 6B**). Scoliosis was associated with a reduction in collagen content in annulus fibrosus and in the endplate (**Fig. 6B**). In osteoarthritic articular cartilage, the collagen content was significantly increased, but articular cartilage degeneration or osteochondral lesions were not associated with changes in collagen or sGAG (**Fig. 6C**).

4.7 Protein and overall composition of connective tissues

Some of the selected studies reported the proteomic analysis of tissue samples. While these studies do not provide absolute quantification of the identified proteins, we compiled the relative quantification data to obtain an overall portrait of protein composition of different connective

tissues. The relative makeup of the proteomes of articular cartilage [87], skeletal muscle [110], tendon [104], ligament [104], and bone [111] were reported (**Fig. 7A**). To determine the relative composition of the whole tissues, we combined the proteomics data with the estimated amounts of water or unique constituents such as lipids, or mineral for articular cartilage [65, 92], IVD [78, 112], skeletal muscle [82], tendon [52], ligament [113], bone [114], and adipose tissue [115] (**Fig. 7B**). Finally, combining the relative quantifications with calculated estimates of absolute abundance of collagens, we calculated the estimated levels for all tissue constituents in articular cartilage **(Table 3)** skeletal muscle **(Table 4)**, and tendon (**Table 5**).

5. Discussion

In this work we synthesized the existing literature and generated robust estimates for collagen content in human connective tissues. In addition, within the IVD, we quantified the regional collagen content in the annulus fibrosus, nucleus pulposus, and endplate. Analysis of methodological techniques identified systematic differences between the estimates provided by ELISA and by the hydroxyproline-based assays. Collagen abundance was not significantly affected by sex or age. We demonstrate that osteoarthritis, disc degeneration and scoliosis lead to distinct changes in ECM composition, particularly in the abundance of collagen and elastin. Finally, we synthesized known information on absolute and relative protein content, as well as water, polysaccharide and unique tissue components such as lipid, to provide quantitative estimates for the native composition of human connective tissues.

Our primary goal was to synthesize existing research to build robust estimates of ECM protein abundance. This question may appear outdated, as many publications and reviews present certain facts about tissue composition as common knowledge, such as cartilage containing somewhere between 20 and 35% protein, with the majority of that fraction being collagen and proteoglycans

[65, 116]. However, it is very difficult to follow the citation trail to find the origins of these estimates. We performed a systematic review in order to identify the primary papers reporting direct quantification of ECM proteins in human connective tissues. Based on these studies we were able to perform meta-analysis on a limited number of tissues and proteins. Importantly, for some tissues, such as bone, no study has passed the inclusion criteria, while for other tissues, such as adipose tissue and ligament, only 1-2 studies were identified, demonstrating significant gaps in information regarding quantification of absolute amounts of ECM proteins in human tissues.

The datasets describing collagen quantification in articular and IVD cartilage were sufficiently large and of high enough quality to obtain robust estimates and perform more in depth meta-analysis. While heterogeneity was high, the majority of the datasets for both IVD and articular cartilage data pools were evenly distributed about the random effects estimate within the expected 95% CI. No single study was found to account for a significant portion of overall heterogeneity. Methodologically, collagen abundance was assessed by measuring hydroxyproline content using Stegemann, Kivirikko, or Woessner methods in 29 of 35 studies [117-119] or by ELISA in 6 studies. The mean of the estimates generated by ELISA were significantly lower than the overall estimate, suggesting that systematic methodological differences need further investigation. We have found that the sex or age of the subject did not significantly affect IVD collagen abundance. It is known that articular cartilage contains proportionally more total protein than IVD [65, 112]. Our estimates suggest that, in addition, collagen accounts for different proportions of total protein in these two tissues. Within the IVD we demonstrate that collagen abundance differs between the annulus fibrosus and nucleus pulposus, with the annulus having significantly more collagen. This is consistent with the current literature, which asserts that the dry-weight of the nucleus is shifted towards proteoglycan [120]. Importantly, pathological conditions were associated with significant

and disease-specific changes in cartilage collagen content. Degeneration of IVD was associated with a significant increase in collagen abundance in the annulus fibrosus. Of interest, collagen content in nucleus pulposus of degenerating discs tended to decrease, providing biochemical basis for location-specific changes in degenerating IVD, where altered collagen distribution was proposed to contribute to the loss of structural integrity and horizontal bulging of the disc [120]. Similar to disc degeneration, the articular cartilage of osteoarthritis patients was found to contain more collagen than healthy tissue. Since increased deposition of ECM proteins was previously linked to fibrosis and injury repair of cartilage [121], high collagen content in osteoarthritic and degenerating cartilage may be related to its repair and regeneration. Tears in degrading discs tend to occur in the annulus of IVD, consistent with higher collagen content in the annulus rather than nucleus of degenerating discs [120]. In contrast to cartilage degeneration, scoliosis was associated with a reduction in collagen abundance, in particular in the annulus and the endplate, supporting the theory that abnormalities in the IVD ECM are involved in the pathophysiology of scoliosis [122]. Thus, although individual studies reported heterogeneous estimates, given sufficient number of primary studies, the data still can be successfully synthesized to provide significant insights into the physiology and pathology of connective tissues.

In contrast to collagen, reports on other ECM proteins were much less common, which limited data potential for synthesis. We built estimates for the abundance of sGAG in tendon, skeletal muscle, and articular cartilage; fibronectin in tendon, ligament, and articular cartilage; elastin in tendon and cartilage; and total proteoglycan in IVD and articular cartilage. Limited number of reports prevented further analysis of these data. Our results suggest that the abundance of these proteins can be affected in pathological conditions, such as an increase in elastin in the

degenerating disc. Thus, it is important to further investigate the abundance of non-collagenous ECM proteins, how they vary between tissues, and how they change in disease.

We generated relative protein breakdowns for 5 tissues (articular cartilage, skeletal muscle, tendon, ligament, and bone) and relative component proportions for 7 tissues (articular cartilage, IVD, skeletal muscle, tendon, ligament, bone, and adipose tissue). For three tissues, including articular cartilage, skeletal muscle and tendon, we had an absolute estimate for the total collagen content, a proteomic breakdown that includes collagen and the estimates for other components, allowing us to generate numerical estimates accounting for 100% of the native tissue (1000 μg constituent / 1 mg wet tissue). The summation of our individual estimates for articular cartilage (1085 ± 82) was slightly larger than expected. Our overall estimate for skeletal muscle (913 ± 41) was slightly lower than expected. Our 95% confidence interval for tendon (926 ± 99) includes 1000 μg constituent/1 mg tissue mass. The small amount of deviation from expected values in articular cartilage could result from systematic overestimation of collagen content by hydroxyproline based quantification methods. In regards to skeletal muscle, the reference proteomic study was not aimed at describing the complete proteome, and is missing some primary constituents of skeletal muscle such as its most abundant proteoglycan, decorin [123]. Protein and proteoglycan are thought to make up 20% of skeletal muscle, while we only account for 11.2%. Incorporation of the remaining ~9%, translating to 90 μg, would put our overall calculation right around 1000 μg constituent/1 mg tissue. An additional uncertainty is due to our use of static reference values for total water content, which may fail to take into consideration individual subject variations. Overall, the fact that for these 3 tissues the overall numerical estimates very closely (within 0.3-17%) account for an expected tissue mass strongly attests to the validity of underlying assumptions and calculations.

Outside of the scope of our findings themselves, academia is facing a systematic problem whereby flawed experimental design and selective reporting give rise to data that is erroneous, irreproducible, or cherry-picked [80]. The synthesis of published studies is often used in the context of clinical trial evaluation and is considered to be the gold standard of evidence [124]. In principle, synthesis can also be used to generate more robust and high powered estimates for virtually any quantitative scientific question, however the technique is seldom used in basic research. The generated estimates effectively employ a much larger sample size than any single study, and through analysis of the datasets themselves, researchers can identify systematic reporting biases and contributors to inter-study variations. It is in the interest of all researchers to use these tools to cost-effectively improve upon the existing bank of knowledge. It is important to note that the quality of these estimates, and thus the effectiveness of the technique, is dependent on the quality of the underlying publications. In our case, absolute quantification was not available for many tissues or proteins other than collagen. Relative quantification using proteomics techniques has been performed for the number of tissues, however the studies were rarely reproduced. In therapy development, limitations in modern pre-clinical disease models contribute a failure to translate in-vitro success to clinical trial success [80]. It is well understood that both the physical properties of the ECM conferred through proteomic composition, in addition to the proteins themselves, are powerful regulators of cellular behavior. Thus, failure to adequately replicate the extracellular environment of cells cultured in vitro may lead to altered behavior, and as a result, what is observed may not translate into the human patient [125]. Culturing on ECM-coated dishes, 3D culture on biomimetic scaffolds, and culture of organ-like organoids has become more and more commonplace [125-127]. By building a quantitative ingredient list for the

extracellular environment we will be able to refine and evaluate these models and further improve upon our repertoire of in vitro tools.

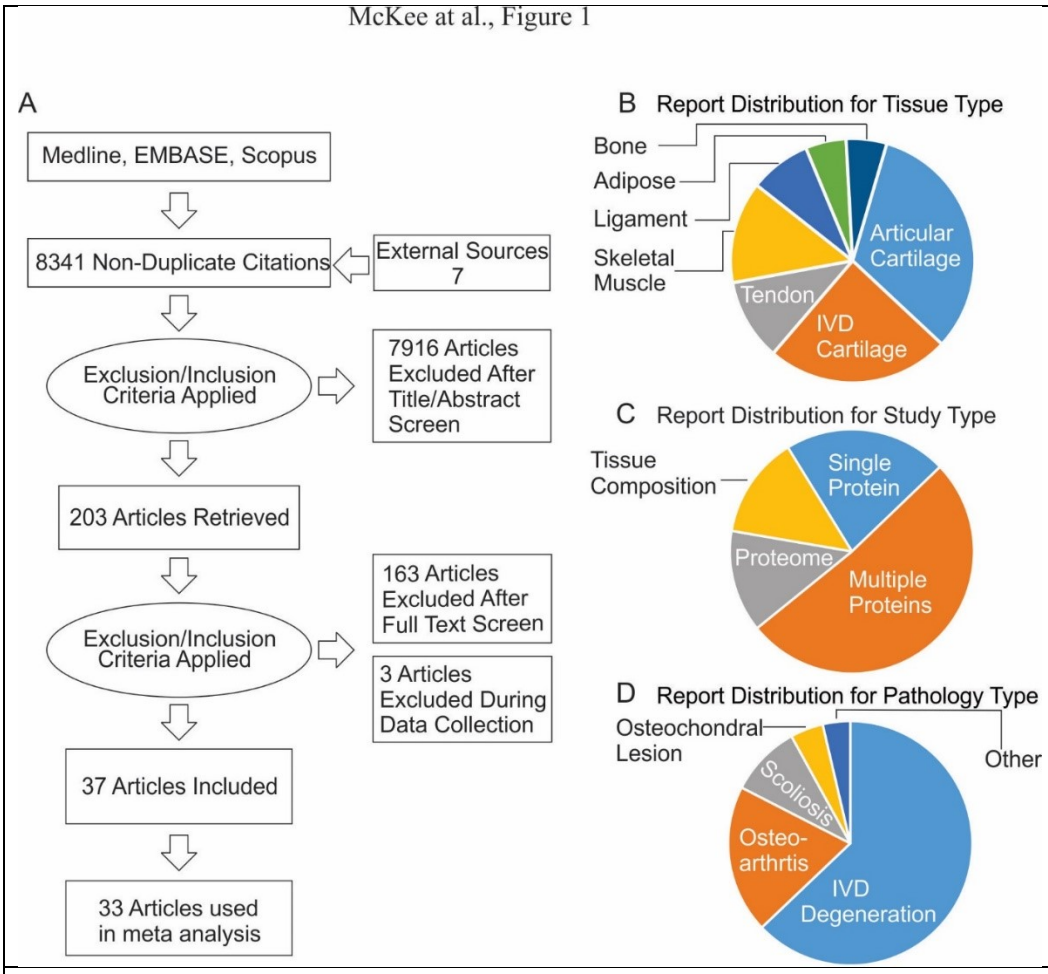


Figure 1. Search strategy and outcomes. (A) Prisma diagram for the information flow. B-D) The distribution of information within 32 accepted papers concerning (B) tissue type; (C) the number of proteins reported within each study; and (D) pathology type.

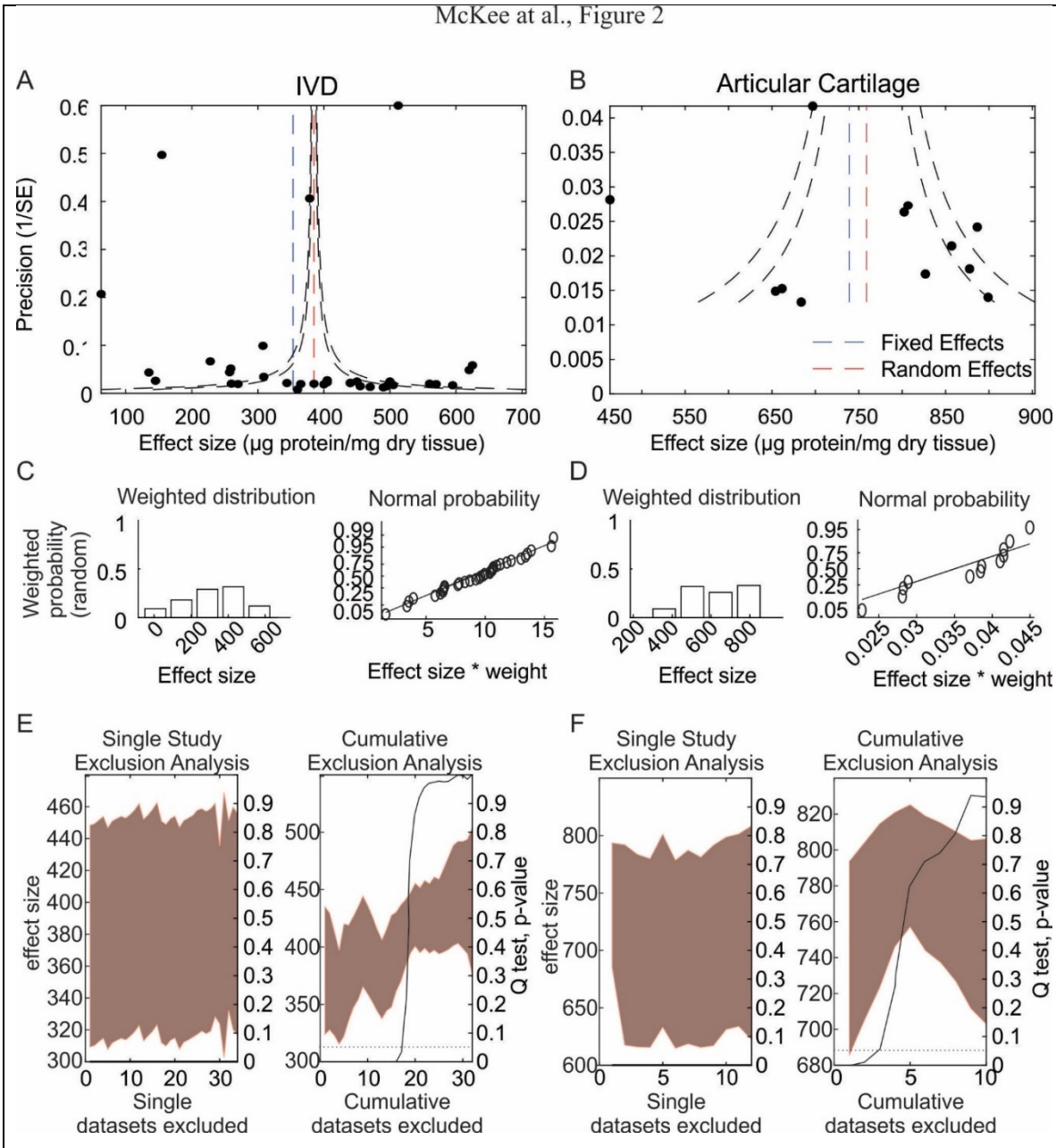


Figure 2. Data distribution and inter-dataset heterogeneity. A, B) Funnel plots indicating bias and heterogeneity for collagen estimates in IVD (A) and articular cartilage (B). *Blue lines*: fixed effect model estimates, *red lines*: random effects model estimates; *black lines:* expected 95% confidence interval in the absence of bias/heterogeneity. C, D) Histograms and normal probability plots for distribution of collagen estimates in IVD (C) and articular cartilage (D). E, F) Single study (left) and cumulative (right) exclusion analysis for collagen estimates in IVD (E) and articular cartilage (F).

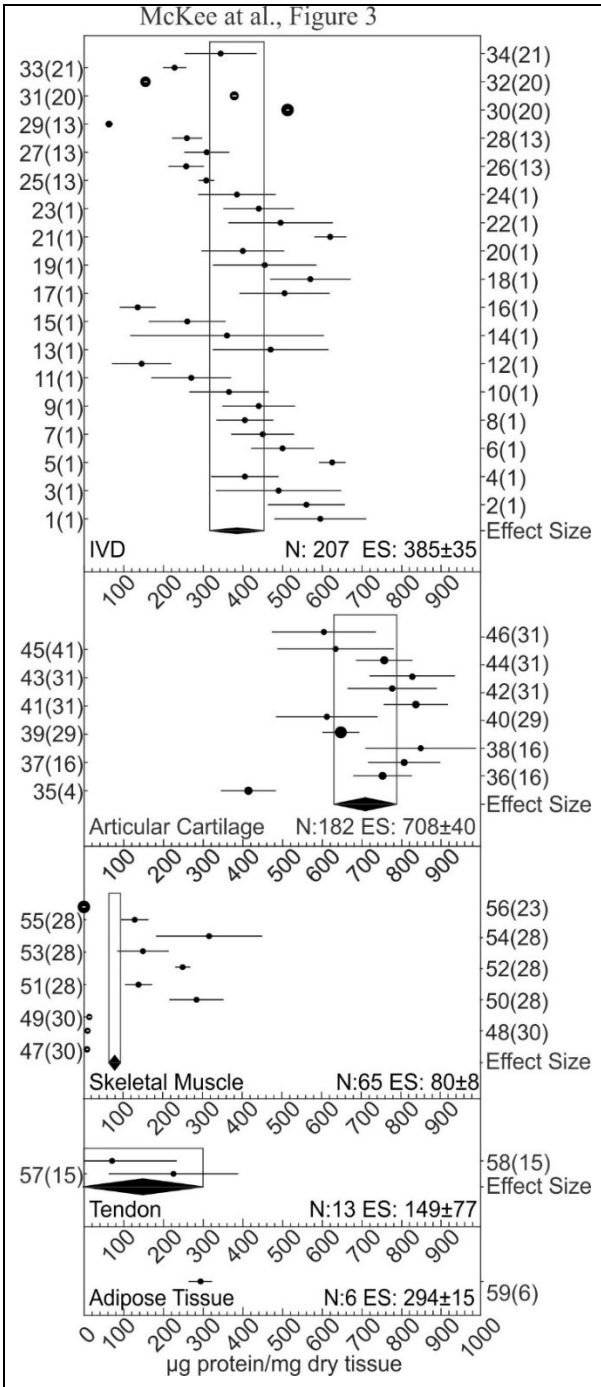


Figure 3. Collagen abundance in connective tissue. Forest plot for the estimated collagen abundance in IVD, articular cartilage, skeletal muscle, tendon, and adipose tissue. Shown are the effect sizes (black dots) with 95% CI (black lines) for each dataset included in the analysis and the overall effect sizes (ES) with 95% CI (black diamonds) for each tissue type. Y-axis labels indicate the included datasets and study numbers in the format dataset(study) according to Appendix 3 Table 1, where extended information can be found. The size of the circle is proportional to the number of datasets.

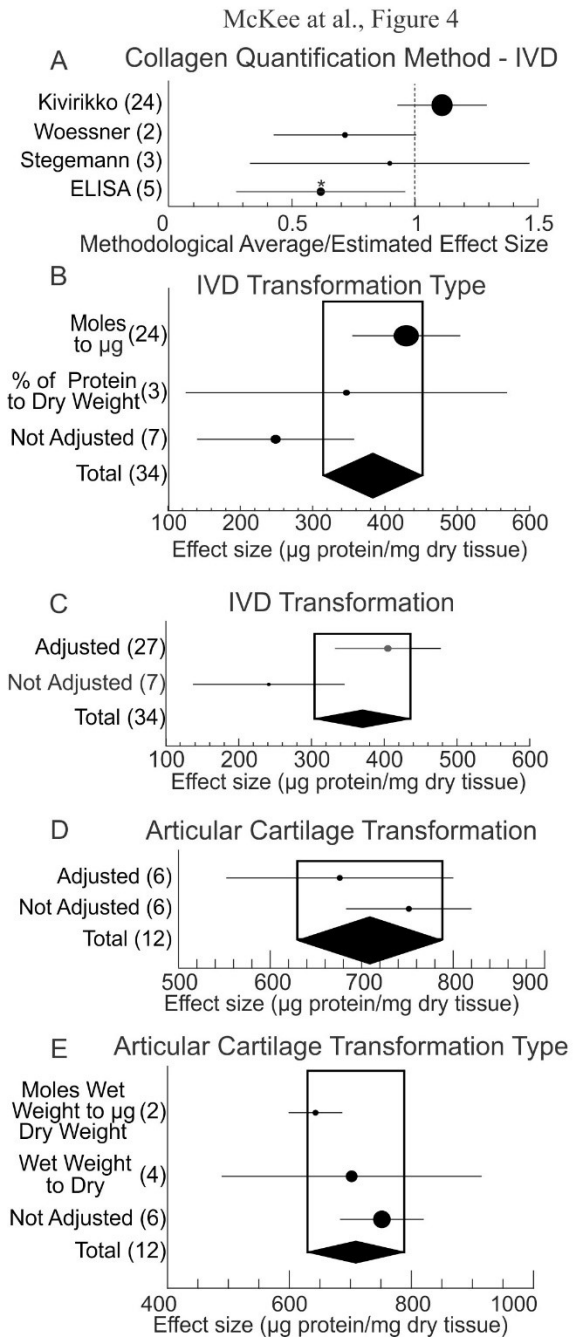


Figure 4. Technical contributors to inter-study differences. Forest plots for the estimated effect size of collagen abundance as a function of (A) quantification methods, presented as ratio to the overall estimate; (B, D) data transformation for IVD (B) or articular cartilage (D); (C, E) the type of data transformation for IVD (C) or articular cartilage (E). Shown are subgroup averages and 95% CI as black dots/black lines and the overall effect sizes and 95% CI as black diamonds. The size of the circle is proportional to the number of datasets, *indicates statistical significance.

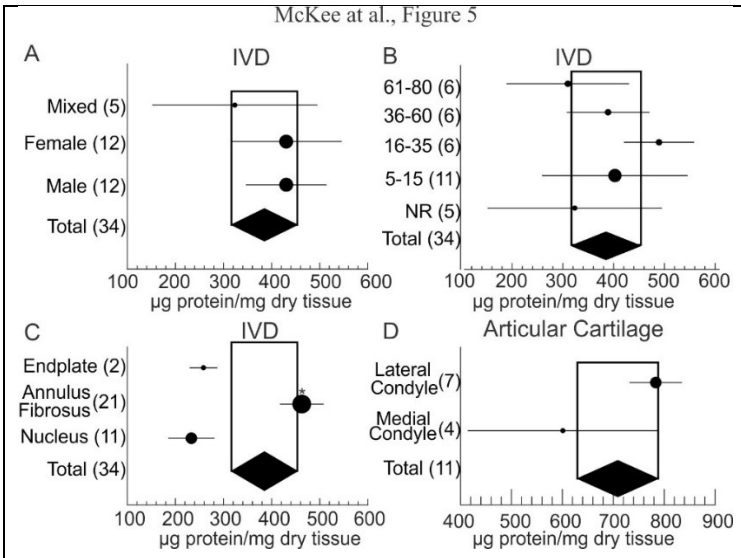


Figure 5. Estimating the contribution of biological factors to collagen abundance in IVD and articular cartilage. Forest plot for the estimated collagen abundance in IVD (A-C) and articular cartilage (D) as a function of sex (A), age (B), and tissue sub-location (C-D). Shown are subgroup averages and 95% CI as black dots/black lines, the overall effect sizes and 95% CI as black diamond, the number of datasets included within each subgroup is indicated in parenthesis. The size of the circle is proportional to the number of datasets; *indicates statistical significance.

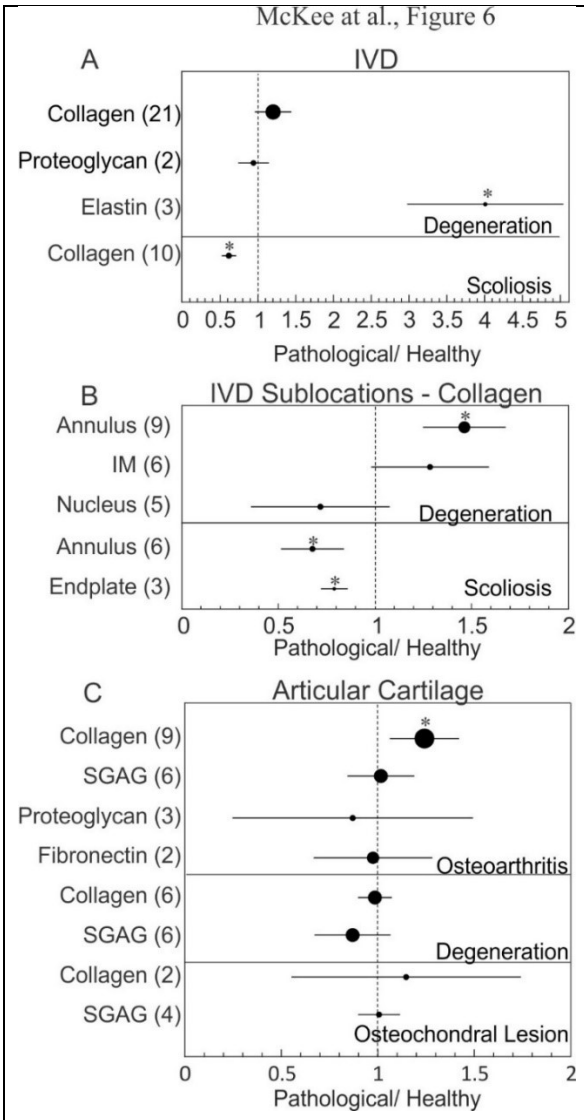


Figure 6. Forest plot for the estimated effect of pathologies on the abundance of ECM proteins. (A, B) Effect of IVD degeneration and scoliosis on abundance of collagen, proteoglycans and elastin (A) and collagen content in IVD sub-locations (B). (C) Effect of osteoarthritis, cartilage degeneration and osteochondral lesion on abundance of collagen, sGAG, proteoglycans and fibronectin in articular cartilage. Shown are the estimated effect sizes (black circles) and 95% CI (black lines) for the abundance of indicated proteins in pathological tissues relative to healthy estimates (1 indicates equal amounts, dashed vertical line); the number of datasets included within each group is indicated in parenthesis. The size of the circle is proportional to the number of datasets, *indicates statistical significance.

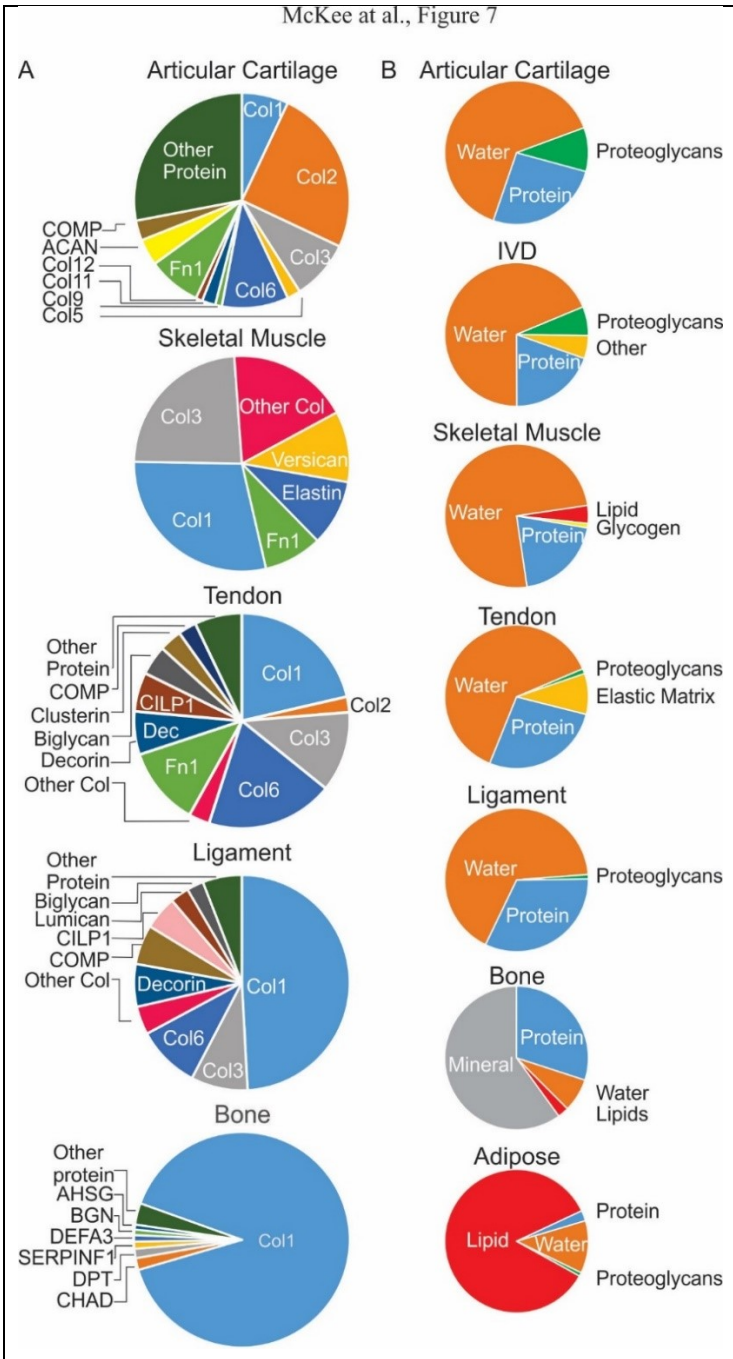


Figure 7. Comparative proteomic and overall composition of different connective tissues. A) Relative abundance of ECM proteins derived from single studies that investigated the large numbers of proteins or complete proteomes in articular cartilage, skeletal muscle, tendon, ligament, and bone. B) Relative quantity of whole tissue constituents in articular cartilage, skeletal muscle, ligament, tendon, bone, adipose tissue, and IVD.

Table 1. The estimated abundance of sulfated glycosaminoglycans (sGAG), fibronectin, elastin, and total proteoglycan in indicated tissues. For each protein/tissue, n is the number datasets that reported in total N samples.

Protein	Tissue	n	N	Amount (µg/mg dry tissue)	CI
sGAG	Tendon	2	13	34	± 57
	Skeletal Muscle	6	40	196	± 22
	Articular Cartilage	8	158	56	± 23
Fibronectin	Tendon	2	6	1.6	± 0.3
	Ligament	2	6	11	± 7.0
	Articular Cartilage	2	5	2.3	± 2.6
Elastin	Tendon	3	16	186	± 300
	IVD Cartilage	3	24	18	± 4.1
Proteoglycan	IVD Cartilage	2	20	144	± 16
	Articular Cartilage	4	25	120	± 46

Table 2. The estimated effects of various pathologies (osteoarthritis, IVD degeneration, scoliosis, osteochondral lesion, obesity, diabetes) on the abundance of indicated proteins in pathological tissues relative to healthy estimates (1 indicates equal amounts). For each protein/tissue/pathology, n is the number of datasets that reported in total N samples. Grey indicates significant difference from physiological values.

Protein	Tissue	Pathology	n	N	Pathological / Healthy	CI
Collagen	IVD	Degeneration	21	291	0.79	± 0.25
		Scoliosis	10	147	0.51	± 0.08
	IVD Annulus	Degeneration	9	116	1.5	± 0.20
		Scoliosis	6	132	0.67	± 0.16
	IVD Intermediate Zone	Degeneration	6	98	1.3	± 0.30
	IVD Nucleus	Degeneration	5	63	0.72	± 0.36
	IVD Endplate	Scoliosis	3	45	0.54	± 0.04
	Articular Cartilage	Osteoarthritis	9	91	1.24	± 0.18
		Degeneration	6	72	0.98	± 0.09
		Osteochondral Lesion	2	17	1.1	± 0.6
	Skeletal Muscle	Diabetes	1	10	2.2	± 1.4
		Obesity	1	10	2.1	± 1.3
sGAG	Articular Cartilage	Osteoarthritis	6	78	1.06	± 0.16
		Degeneration	6	72	0.95	± 0.27
		Osteochondral Lesion	4	64	1.13	± 0.24
Fibronectin	Articular Cartilage	Osteoarthritis	2	13	0.99	± 0.16
Elastin	IVD	Degeneration	3	15	4.0	± 1.0
Proteoglycan	IVD	Degeneration	2	17	0.94	± 0.20
	Articular Cartilage	Osteoarthritis	3	20	0.87	± 0.63

Table 3. Estimated whole-tissue composition of articular cartilage.

Protein	Amount (ug/mg articular cartilage)	±
Col1	23.0	2.6
Col2	82.1	9.2
Col3	29.6	3.3
Col5	6.57	0.73
Col6	32.8	3.7
Col9	3.28	0.37
Col11	6.57	0.73
Col12	3.28	0.37
Total Collagen	187	21
Fibronectin	26.3	2.9
COMP	9.9	1.1
Other Protein	92	10
Total	315	35
Water	650	
Aggrecan	49.8	5.6
Other Proteoglycan	70	
Total Proteoglycan	120	46
Total	1085	82

Table 4. Estimated whole-tissue composition of skeletal muscle.

Protein	Amount (µg/mg skeletal muscle)	±
Col1	32.6	6.0
Col3	26.6	4.9
Other Collagen	20.6	3.8
Total Collagen	80	15
Versican	12.0	2.2
Elastin	11.1	2.1
Fibronectin	9.9	1.8
Total Protein	113	21
Water	750	20
Lipid	40	
Glycogen	10	
Total	913	41

Table 5. Estimated whole-tissue composition of tendon.

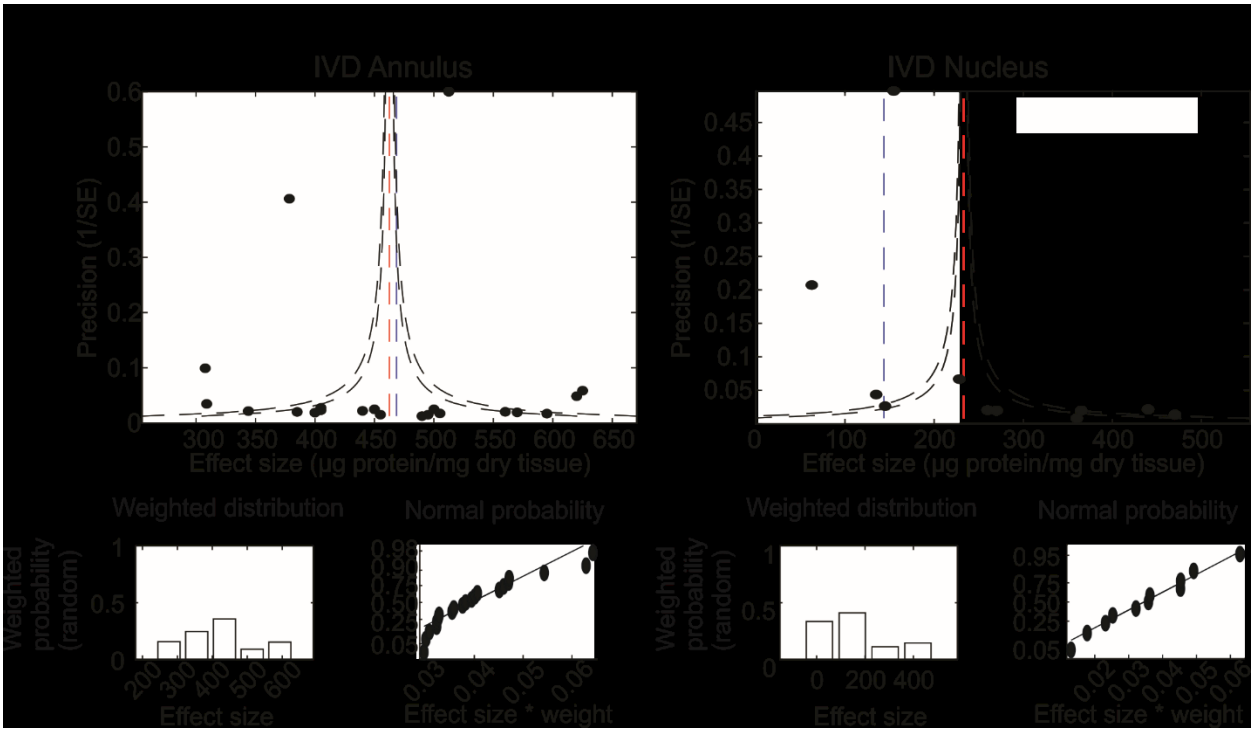
Component	**Amount (µg/mg tendon)**	±
Col1	103	38
Col3	17.8	6.6
Col6	19.8	7.3
Other Col	8.9	3.3
Total Collagen	149	77
COMP	12.4	4.6
CILP1	10.5	3.9
Other Protein	12.3	4.5
Total Protein	184	90
Decorin	13.4	4.9
Lumican	5.9	2.2
Biglycan	5.2	1.9
Total Proteoglycan	24.5	9.0
Water	625	
Other (elastic matrix)	92	
Total	926	99

Appendices

Appendix 1. Search strategy for identifying articles.

1. Extracellular Matrix/
2. Microfibrils/
3. (extracellular matrix or ECM).ti,ab,kf.
4. 1 or 2 or 3
5. exp Extracellular Matrix Proteins/
6. (alcam or activated-leucocyte or (cd166 adj1 (antigen* or ligand*)) or kg-cam or neurolin or ADAMTS* or adam metallopeptidase? with thrombospondin or aggrecan? or cartilage-specific proteoglycan core or connective tissue growth factor or hypertrophic chondrocyte specific protein 24 or insulin like growth factor binding protein 8 or cysteine-! rich protein 61 or collagen? or procollagen? or tropocollagen? or endostatin? or elastin? or tropoelastin? or fibrillin* or fibronectin or cold-insoluble globulin? or (sialoprotein adj3 (bone or integrin binding)) or laminin or merosin or matrilin or osteopontin or tenascin* or cytotactin or hexabrachion or j1-200-220 or versican? or vitronectin or ((sirp or small leucine rich) adj3 proteoclycan?) or ((alpha 2 surface binding or gp 2 or matrix or opsonic) adj3 glycoprotein) or ((cartilege oligomeric matrix or cyr ctgf nov matricellular or (ccn adj3 (intercellular signal?ing or family or matricellular)) or thrombospondin 5 or ccn1 or ccn3 or cyr61 or igfbp-8 or igfbp-rp2 or igfbp9 or novh or igfbp10 or matn? or nephroblastoma overexpressed or latent tgf beta binding or cartilage matrix or igf binding or insulin-like growth factor binding or (ECM adj5 regulat*)) adj3 protein?)).ti,ab,kf.
7. 5 or 6
8. exp Adipose Tissue/
9. Myocardium/ or Myocytes, Cardiac/
10. exp Muscle, Skeletal/
11. Endothelium, Vascular/
12. exp "Bone and Bones"/
13. exp Cartilage/
14. (((cardiac or skeletal) adj1 muscle?) or bone? or adipose tissue or myocardium or vascular endotheli* or cartilage).ti,ab,kf.
15. or/8-14
16. 4 and 7 and 15
17. (characteri?e or assess or composition* or analy?e or analysis or isolate? or isolation or differentiat* or quantif* or characteri* or quantitativ* or proteomic* or protein content* or qconcat or component*).ti,ab,kf.
18. exp Proteomics/
19. 17 or 18
20. (decellulari* or acellular*).ti,ab,kf.
21. Extracellular Matrix/ph
22. Microfibrils/ph
23. 21 or 22
24. 4 and 7 and 15 and 19
25. 23 and 15 and 19
26. 4 and 15 and 20
27. or/24-26
28. limit 27 to animals/
29. limit 27 to humans/
30. 28 not (28 and 29)
31. 27 not 30

Appendix 2. Data distribution and inter-study heterogeneity – IVD Sublocations.



A,B) Funnel plots indicating bias and heterogeneity for collagen estimates in IVD annulus fibrosus (A) and IVD nucleus pulposus (B). *Blue lines*: fixed effect model estimates, *red lines*: random effects model estimates; *black lines:* expected 95% confidence interval in the absence of bias/heterogeneity.

C,D) Histograms and normal probability plots for distribution of pooled collagen estimates in IVD annulus (C) and nucleus pulposus (D).

Appendix 3. List of protein reported in 32 included articles that contained absolute quantification. Shown are tissues in which the protein was reported, the number of times each protein was reported on in healthy and pathological conditions and corresponding references. References correspond to the paper legend available in supplementary table 3.

Protein	Count Healthy	Count Pathological	Tissue Type	Ref
a1 Antitrypsin	4	0	3;7	14
Actin - Cytoplasmic 1/2	1	0	4	10
Actin - Total	1	0	4	10
Aggrecan	1	2	2	7;8
Agrin	1	0	4	10
Agrin (iso 2,3,4,5,&6)	1	0	4	10
Albumin	4	0	3;7	14

Alpha-1 Microglobulin/Bikunin Precursor	4	0	3;7	14
Annexin A2	1	0	4	10
Apolipoprotein A1	2	0	3	14
Apolipoprotein A4	2	0	3	14
Apolipoprotein E	2	0	3	14
asporin	4	0	3;7	14
Biglycan	1	0	4	10
Cartilage Intermediate Layer Protein	1	2	2	4
Cartilage Oligomeric Matrix Protein	2	4	2	4;7;8
Chondroadherin	2	0	3	14
Chondroitin Sulphate (846 Epitope)	5	43	2;8	12;13;24;29
CSPG Core Protein 2	2	0	3	14
Clusterin	4	0	3;7	14
Collagen (Col12A1)	5	0	3;4;7	10;14
Collagen (Col18A1)	1	0	4	10
Collagen (Col1A1)	5	0	3;4;7	10;14
Collagen (Col1A2)	5	0	3;4;7	10;14
Collagen (Col3A1)	4	0	3;7	14
Collagen (Col4A1)	1	0	4	10
Collagen (Col4A1/5)	1	0	4	10
Collagen (Col4A2)	1	0	4	10
Collagen (Col4A5)	1	0	4	10
Collagen (Col5A1)	1	0	4	10
Collagen (Col5A2)	1	0	4	10
Collagen (Col6A1)	5	0	3;4;7	10;14
Collagen (Col6A2)	5	0	3;4;7	10;14
Collagen (Col6A3)	5	0	3;4;7	10;14
Collagen (Insoluble)	1	2	4	19
Collagen (Soluble)	1	2	4	19
Collagen (total)	59	51	1;2;5;7;8	1;4;6;11;12;13;15;16;20;21;22;23;24;28;29;30;31
Collagen (type 1)	3	0	2;3;7	7;18
Collagen (type 11)	1	0	2	7
Collagen (type 12)	1	0	2	7
Collagen (type 2)	3	18	2;3;7;8	7;12;17;18
Collagen (type 3)	1	0	2	7

Collagen (type 5)	1	0	2	7
Collagen (type 6)	1	0	2	7
Collagen (type 9)	1	0	2	7
Collagen II Degradation Protein (C2C)	0	2	2	24
Complement Component 3	4	0	3;7	14
Complement Component 4 Binding Protein α	2	0	3	14
Complement Component 9	4	0	3;7	14
C-propeptide of type 1 procollagen	4	17	8	12;13
C-propeptide of type 2 procollagen	0	37	2;8	12;24
Decorin	1	0	4	10
Dermatopontin	1	0	4	10
Desmin	1	0	4	10
Elastin	9	5	1;3;4;7;8	6;15;18;19;20
Emilin-1	1	0	4	10
Fibrillin-1	1	0	4	10
Fibrinogen	2	0	3	14
Fibrinogen Beta	2	0	3	14
Fibrinogen Gamma	2	0	3	14
Fibromodulin	5	4	2;3;7	4;8;14
Fibronectin	8	4	2;3;4;7	4;5;7;10;14
Fibulin-5	1	0	4	10
Galactin-1	1	0	4	10
Laminin Gamma-1	1	0	4	10
Laminin Subunit Alpha-2	1	0	4	10
Laminin Subunit Alpha-5	1	0	4	10
Laminin Subunit Beta-1	1	0	4	10
Laminin Subunit Beta-2	1	0	4	10
LTBP-1	1	0	4	10
Lubricin (PRG4)	4	0	3;7	14
Lumican	5	0	3;4;7	10;14
Microfibrillar-associated protein 2	1	0	4	10
Mimecan/Osteoglycin	5	0	3;4;7	10;14
MMP-1	0	2	2	8
MMP2 (Pro-MMP2)	0	1	8	11
MMP-3	0	2	2	8
Myosin (Myosin-3,4,6,7)	2	0	4	10
Nidogen-1	1	0	4	10
Nidogen-2	1	0	4	10

Pentraxin-2 (SAMP)	4	0	3;7	14
Periostin	2	1	2;4	2;10
Perlecan - HSPG2	5	0	3;4;7	10;14
Phosphatidylglycerophosphate Synthase 1	4	0	3;7	14
Phosphatidylglycerophosphate Synthase 2	4	0	3;7	14
Phosphoglucomutase	2	0	3	14
Plasminogen	0	2	2	8
Prolargin (PRELP)	6	2	2;3;4;7	4;10;14
Protein (Total)	1	1	8	13
Proteoglycan (total)	7	6	2;8	4;16;21;24;26
Sulphated Glycosaminoglycan	19	59	1;2;4;5;7;8	6;9;12;15;17;19;22;28;29;31
Tenascin C	1	1	2	25
TIMP-1	0	3	2;8	8;11
TnxB Protein	1	0	4	10
TGFβ-induced-protein ig-h3	5	0	3;4;7	10;14
Transglutaminase 2	1	0	4	10
Transthyretin	4	0	3;7	14
Tubulin beta	1	0	4	10
Vimentin	1	0	4	10

Appendix 4: Complete list of protein quantification entries. Numerical entries for studyNumber, Protein, Tissue, TissueRegion, and Pathology correspond to the legend in Appendix 5. ID is a unique, datapoint specific identifier. X, se, and n are the effect size in μg/mg dry tissue, standard error, and sample size, respectively.

studyNumber	ID	x	se	n	Protein	Tissue	TissueRegion	Pathology Type
14	329	0.15	0.02	3	1	3	3	0
14	339	0.15	0.08	3	1	3	3	0
14	344	0.04	0.08	3	1	7	10	0
14	349	0.08	0.02	3	1	7	10	0
10	58	0.02	0.00	6	2	4	0	0
10	57	0.10	0.03	6	3	4	0	0
7	39	70.00		1	4	2	6	0
8	44	0.00	0.00	3	4	2	7	16
8	51	0.00	0.00	3	4	2	7	16
10	59	0.00	0.00	6	5	4	0	0

10	60	0.01	0.00	6	6	4	0	0
14	312	1.30	0.24	3	7	3	3	0
14	314	3.78	1.06	3	7	3	3	0
14	316	1.54	0.12	3	7	7	10	0
14	318	4.02	0.71	3	7	7	10	0
14	320	0.88	0.96	3	8	3	3	0
14	330	0.35	0.15	3	8	3	3	0
14	340	2.46	0.56	3	8	7	10	0
14	345	2.58	0.73	3	8	7	10	0
10	61	0.00	0.00	6	9	4	0	0
14	321	0.10	0.02	3	10	3	3	0
14	331	0.17	0.06	3	10	3	3	0
14	322	0.08	0.02	3	11	3	3	0
14	332	0.06	0.02	3	11	3	3	0
14	350	0.97	0.16	3	12	3	3	0
14	353	1.25	0.78	3	12	3	3	0
14	356	0.28	0.10	3	12	7	10	0
14	359	0.38	0.12	3	12	7	10	0
10	63	0.01	0.00	6	14	4	0	0
4	15	0.25	0.11	2	15	2	7	0
4	16	0.76	0.25	6	15	2	7	17
4	17	1.43		1	15	2	7	18
4	12	1.86	1.03	2	16	2	7	0
4	13	12.94	3.37	6	16	2	7	17
4	14	6.89		1	16	2	7	18
7	40	50.00		1	16	2	6	0
8	41	0.00	0.00	3	16	2	7	16
8	48	0.01	0.00	3	16	2	7	16
24	481	0.00	0.00	10	17	2	7	16
24	486	0.00	0.00	10	17	2	4	16
29	506	0.95	0.30	10	17	2	4	0
29	512	1.05	0.23	10	17	2	7	0
29	507	123.05	15.88	10	17	2	4	12
29	513	0.10	0.01	7	17	2	7	12
12	158	29.54	8.43	10	17	8	1	8
12	159	103.80	11.82	10	17	8	1	8
12	160	491.97	21.95	10	17	8	1	8
12	161	167.92	21.95	10	17	8	1	8
12	162	46.41	6.76	10	17	8	1	8

12	163	5.90		15	17	8	1	8
12	164	26.15		15	17	8	1	8
12	165	86.91		15	17	8	1	8
12	166	26.15		15	17	8	1	8
12	167	16.03		15	17	8	1	8
12	168	5.90		20	17	8	1	8
12	169	12.66		20	17	8	1	8
12	170	26.15		20	17	8	1	8
12	171	12.66		20	17	8	1	8
12	172	12.66		20	17	8	1	8
12	173	5.90		19	17	8	1	8
12	174	5.90		19	17	8	1	8
12	175	5.90		19	17	8	1	8
12	176	5.90		19	17	8	1	8
12	177	9.29		19	17	8	1	8
12	178	2.58	0.38	14	17	8	1	8
12	179	5.48	0.78	14	17	8	1	8
12	180	15.91	2.41	14	17	8	1	8
12	181	5.59	0.61	14	17	8	1	8
12	182	4.09	0.43	14	17	8	1	8
12	183	2.58	0.78	22	17	8	1	8
12	184	2.46	0.55	22	17	8	1	8
12	185	6.87	1.71	22	17	8	1	8
12	186	2.81	0.26	22	17	8	1	8
12	187	2.81	0.15	22	17	8	1	8
12	188	2.81	0.03	10	17	8	1	8
12	189	4.67	0.72	10	17	8	1	8
12	190	5.25	0.67	10	17	8	1	8
12	191	4.67	0.67	10	17	8	1	8
12	192	3.28	0.61	10	17	8	1	8
13	281	8.74	1.94	17	17	8	1	0
13	282	63.11	13.59	17	17	8	1	0
13	283	95.15	16.02	17	17	8	1	0
13	277	16.50	5.34	15	17	8	1	20
13	278	78.16	33.98	15	17	8	1	20
13	279	66.99	24.76	15	17	8	1	20
13	280	68.45	9.71	15	17	8	1	20
14	323	1.69	0.21	3	18	3	3	0
14	333	1.38	0.71	3	18	3	3	0

14	341	0.50	0.10	3	18	7	10	0
14	346	0.31	0.19	3	18	7	10	0
14	299	0.56	0.28	3	20	3	3	0
14	303	1.13	0.00	3	20	3	3	0
10	64	0.04	0.01	6	20	4	0	0
14	307	3.66	0.56	3	20	7	10	0
14	311	4.79	0.56	3	20	7	10	0
10	65	0.01	0.00	6	21	4	0	0
14	284	300.00	16.13	3	22	3	3	0
14	287	287.10	12.90	3	22	3	3	0
10	66	78.37	8.08	6	22	4	0	0
14	290	474.19	29.03	3	22	7	10	0
14	293	470.97	25.81	3	22	7	10	0
14	285	241.94	6.45	3	23	3	3	0
14	288	238.71	16.13	3	23	3	3	0
10	67	41.73	3.69	6	23	4	0	0
14	291	370.97	22.58	3	23	7	10	0
14	294	341.94	51.61	3	23	7	10	0
14	286	287.10	16.13	3	25	3	3	0
14	289	261.29	25.81	3	25	3	3	0
14	292	54.84	6.45	3	25	7	10	0
14	295	80.65	9.68	3	25	7	10	0
10	68	15.82	1.61	6	26	4	0	0
10	69	12.98	1.05	6	27	4	0	0
10	70	9.52	1.86	6	28	4	0	0
10	71	0.19	0.07	6	29	4	0	0
10	72	3.83	0.38	6	30	4	0	0
10	73	1.75	0.13	6	31	4	0	0
14	296	7.32	1.41	3	32	3	3	0
14	300	8.45	2.25	3	32	3	3	0
10	74	1.27	0.32	6	32	4	0	0
14	304	6.20	0.00	3	32	7	10	0
14	308	10.14	3.94	3	32	7	10	0
14	297	4.23	1.13	3	33	3	3	0
14	301	4.79	1.41	3	33	3	3	0
10	75	1.46	0.37	6	33	4	0	0
14	305	2.54	0.28	3	33	7	10	0
14	309	3.66	0.56	3	33	7	10	0
14	298	24.79	8.45	3	34	3	3	0

14	302	29.86	14.08	3	34	3	3	0
10	76	1.42	0.38	6	34	4	0	0
14	306	11.83	0.85	3	34	7	10	0
14	310	20.85	6.76	3	34	7	10	0
12	243	0.00	0.00		36	8	1	8
12	244	0.00	0.00		36	8	1	8
12	245	0.00	0.00		36	8	1	8
12	246	0.00	0.00		36	8	1	8
12	247	0.00	0.00		36	8	1	8
12	248	0.00	0.00		36	8	1	8
12	249	0.00	0.00	7	36	8	1	8
12	250	0.00	0.00	7	36	8	1	8
12	251	0.00	0.00	7	36	8	1	8
12	252	0.00	0.00	46	36	8	1	8
12	253	0.00	0.00	46	36	8	1	8
12	254	0.00	0.00	46	36	8	1	8
13	268	8.91	2.00	17	36	8	1	0
13	270	46.15	3.90	17	36	8	1	0
13	271	41.15	4.40	17	36	8	1	0
13	272	33.65	3.90	17	36	8	1	0
13	269	5.37	0.35	15	36	8	1	20
13	273	8.36	2.41	15	36	8	1	20
13	274	5.95	0.64	15	36	8	1	20
13	275	4.79	0.40	15	36	8	1	20
13	276	8.78	1.13	15	36	8	1	20
24	479	0.07	0.02	10	37	2	7	16
24	484	0.18	0.04	10	37	2	4	16
12	208	0.12	0.01	10	37	8	1	8
12	209	0.80	0.17	10	37	8	1	8
12	210	1.56	0.24	10	37	8	1	8
12	211	1.03	0.30	10	37	8	1	8
12	212	0.18	0.01	10	37	8	1	8
12	213	0.04	0.05	15	37	8	1	8
12	214	0.19	0.07	15	37	8	1	8
12	215	0.28	0.11	15	37	8	1	8
12	216	0.17	0.07	15	37	8	1	8
12	217	0.05	0.02	15	37	8	1	8
12	218	0.02	0.05	20	37	8	1	8
12	219	0.04	0.00	20	37	8	1	8

12	220	0.04	0.01	20	37	8	1	8
12	221	0.04	0.00	20	37	8	1	8
12	222	0.03	0.01	20	37	8	1	8
12	223	0.01	0.01	19	37	8	1	8
12	224	0.01	0.01	19	37	8	1	8
12	225	0.01	0.01	19	37	8	1	8
12	226	0.01	0.01	19	37	8	1	8
12	227	0.01	0.01	19	37	8	1	8
12	228	0.01	0.00	14	37	8	1	8
12	229	0.02	0.00	14	37	8	1	8
12	230	0.03	0.01	14	37	8	1	8
12	231	0.02	0.00	14	37	8	1	8
12	232	0.01	0.00	14	37	8	1	8
12	233	0.00	0.00	22	37	8	1	8
12	234	0.01	0.00	22	37	8	1	8
12	235	0.01	0.00	22	37	8	1	8
12	236	0.01	0.00	22	37	8	1	8
12	237	0.01	0.00	22	37	8	1	8
12	238	0.00	0.00	10	37	8	1	8
12	239	0.01	0.00	10	37	8	1	8
12	240	0.01	0.00	10	37	8	1	8
12	241	0.01	0.00	10	37	8	1	8
12	242	0.00	0.00	10	37	8	1	8
19	434	18.06	4.97	5	38	4	0	0
19	435	106.94	25.00	4	38	4	0	1
19	436	91.67	11.81	4	38	4	0	2
19	437	25.40	4.83	5	39	4	0	0
19	438	6.98	2.06	4	39	4	0	1
19	439	4.13	1.11	4	39	4	0	2
6	27	294.00	14.74	6	40	1	11	0
4	3	413.39	35.54	2	40	2	7	0
4	4	473.87	46.63	6	40	2	7	17
4	5	291.81		1	40	2	7	18
16	419	752.35	37.93	5	40	2	7	0
16	420	806.99	46.63	7	40	2	7	0
16	421	848.90	71.42	10	40	2	7	0
24	477	801.20	76.56	10	40	2	7	16
24	482	720.33	58.93	10	40	2	4	16
29	508	646.98	23.98	10	40	2	4	0

29	514	611.60	65.53	10	40	2	7	0
29	509	950.25	134.26	10	40	2	4	12
29	515	525.67	61.13	7	40	2	7	12
31	521	836.48	41.36	23	40	2	7	0
31	524	776.74	57.54	23	40	2	7	0
31	527	827.60	55.20	23	40	2	7	0
31	530	756.71	36.61	23	40	2	7	0
31	533	633.69	75.15	23	40	2	7	0
31	536	604.13	67.17	23	40	2	7	0
31	522	818.34	89.47	12	40	2	7	8
31	525	842.11	48.88	12	40	2	7	8
31	528	703.47	45.36	12	40	2	7	8
31	531	779.14	66.06	12	40	2	7	8
31	534	660.09	89.53	12	40	2	7	8
31	537	652.63	61.18	12	40	2	7	8
31	523	914.85	128.11	13	40	2	7	16
31	526	846.55	95.54	13	40	2	7	16
31	529	805.59	71.96	13	40	2	7	16
31	532	805.47	117.55	13	40	2	7	16
31	535	820.75	106.47	13	40	2	7	16
31	538	763.09	56.60	13	40	2	7	16
23	474	2.08	0.31	10	40	5	12	0
23	476	4.62	1.31	10	40	5	12	9
23	475	4.36	1.23	10	40	5	12	15
28	494	285.71	34.92	10	40	5	14	0
28	495	139.68	17.46	10	40	5	15	0
28	496	250.79	9.52	5	40	5	14	0
28	497	150.79	33.33	5	40	5	15	0
28	498	317.46	68.25	5	40	5	14	0
28	499	130.16	17.46	5	40	5	15	0
30	518	10.00	1.20	5	40	5	16	0
30	519	11.30	1.40	5	40	5	16	0
30	520	15.60	1.80	5	40	5	16	0
15	411	226.19	83.33	9	40	7	5	0
15	414	71.43	83.33	4	40	7	5	0
1	557	595.00	59.18	2	40	8	1	0
1	558	560.00	49.51	4	40	8	1	0
1	559	490.00	80.55	3	40	8	1	0
1	560	405.00	43.40	4	40	8	1	0

1	561	625.00	17.10	2	40	8	1	0
1	562	500.00	40.57	2	40	8	1	0
1	563	450.00	40.57	4	40	8	1	0
1	564	405.00	36.62	5	40	8	1	0
1	565	440.00	46.83	2	40	8	1	0
1	566	365.00	51.11	4	40	8	1	0
1	567	270.00	51.39	3	40	8	1	0
1	568	145.00	38.16	4	40	8	1	0
1	569	470.00	74.54	2	40	8	1	0
1	570	360.00	124.50	2	40	8	1	0
1	571	260.00	49.51	4	40	8	1	0
1	572	135.00	22.98	5	40	8	1	0
1	573	505.00	58.02	2	40	8	1	0
1	574	570.00	51.89	4	40	8	1	0
1	575	455.00	66.68	3	40	8	1	0
1	576	400.00	53.28	4	40	8	1	0
1	577	620.00	20.63	2	40	8	1	0
1	578	495.00	67.29	2	40	8	1	0
1	579	440.00	45.57	4	40	8	1	0
1	580	385.00	49.96	5	40	8	1	0
11	107	33.72	2.41	14	40	8	1	7
12	143	881.16	37.04	10	40	8	1	8
12	144	603.38	40.14	10	40	8	1	8
12	145	183.66	15.42	10	40	8	1	8
12	146	529.31	33.98	10	40	8	1	8
12	147	813.29	55.56	10	40	8	1	8
12	148	788.56	18.52	20	40	8	1	8
12	149	554.03	40.14	20	40	8	1	8
12	150	214.49	37.04	20	40	8	1	8
12	151	572.55	37.04	20	40	8	1	8
12	152	819.44	21.62	20	40	8	1	8
12	153	566.34	33.98	19	40	8	1	8
12	154	288.56	30.88	19	40	8	1	8
12	155	177.45	49.40	19	40	8	1	8
12	156	356.48	30.88	19	40	8	1	8
12	157	615.74	30.88	19	40	8	1	8
13	255	307.74	10.09	17	40	8	1	0
13	257	257.29	22.70	17	40	8	1	0
13	263	308.95	28.96	17	40	8	1	0

13	264	259.06	19.31	17	40	8	1	0
13	265	62.75	4.83	17	40	8	1	0
13	256	156.21	5.62	15	40	8	1	20
13	258	139.79	9.30	15	40	8	1	20
13	259	154.32	10.28	15	40	8	1	20
13	260	136.51	10.96	15	40	8	1	20
13	261	141.99	13.02	15	40	8	1	20
13	262	47.44	6.85	15	40	8	1	20
20	452	512.50	1.67	8	40	8	1	0
20	453	378.46	2.46	8	40	8	1	0
20	454	154.66	2.01	8	40	8	1	0
20	455	513.46	3.34	5	40	8	1	8
20	456	255.05	7.46	5	40	8	1	8
20	457	183.04	5.95	5	40	8	1	8
21	462	228.00	15.00	9	40	8	1	0
21	463	344.00	46.43	11	40	8	1	0
21	464	187.00	23.33	9	40	8	1	7
21	465	285.00	16.26	8	40	8	1	7
22	470	394.74	17.76	12	40	8	1	20
22	471	471.71	13.82	12	40	8	1	20
22	472	193.42	19.74	12	40	8	1	20
22	473	169.74	19.74	21	40	8	1	20
7	30	30.00		1	41	2	6	0
18	431	125.52	101.75	6	41	3	13	0
18	428	601.75	125.52	3	41	7	2	0
7	31	80.00		1	42	2	6	0
17	423	783.55	360.68	18	42	2	7	11
17	424	547.24	323.37	18	42	2	7	12
17	422	833.30	547.24	18	42	2	7	13
18	429	125.52	114.74	6	42	3	13	0
18	432	0.00	0.00	3	42	7	2	0
12	193	200.29		10	42	8	1	8
12	194	581.22	40.63	10	42	8	1	8
12	195	250.71	46.22	10	42	8	1	8
12	196	771.72	161.05	10	42	8	1	8
12	197	457.98	46.22	10	42	8	1	8
12	198	183.49	15.38	20	42	8	1	8
12	199	424.37	37.82	20	42	8	1	8
12	200	261.89	21.01	20	42	8	1	8

12	201	525.21	43.40	20	42	8	1	8
12	202	457.98	23.82	20	42	8	1	8
12	203	133.07	23.78	19	42	8	1	8
12	204	228.28	29.41	19	42	8	1	8
12	205	211.47	12.61	19	42	8	1	8
12	206	312.31	40.63	19	42	8	1	8
12	207	245.08	29.41	19	42	8	1	8
7	32	20.00		1	43	2	6	0
7	33	1.00		1	44	2	6	0
7	34	40.00		1	45	2	6	0
7	35	1.00		1	46	2	6	0
7	36	1.00		1	47	2	6	0
7	37	10.00		1	48	2	6	0
24	480	0.00	0.00	10	49	2	7	16
24	485	0.01	0.00	10	49	2	4	16
14	324	0.04	0.06	3	50	3	3	0
14	334	0.04	0.00	3	50	3	3	0
14	325	0.21	0.02	3	51	3	3	0
14	335	0.27	0.21	3	51	3	3	0
14	342	0.08	0.10	3	51	7	10	0
14	347	0.17	0.10	3	51	7	10	0
14	326	0.17	0.06	3	52	3	3	0
14	336	0.04	0.00	3	52	3	3	0
14	343	0.04	0.00	3	52	7	10	0
14	348	0.02	0.02	3	52	7	10	0
10	77	0.08	0.01	6	53	4	0	0
10	79	0.10	0.01	6	54	4	0	0
10	78	0.03	0.01	6	55	4	0	0
6	29	195.00	12.11	6	57	1	11	0
18	433	922.08	188.31	6	57	3	13	0
19	440	138.10	18.46	5	57	4	0	0
19	441	50.79	11.11	4	57	4	0	1
19	442	55.56	7.94	4	57	4	0	2
15	412	857.14	425.60	9	57	7	5	0
15	415	345.24	62.50	4	57	7	5	0
18	430	38.96	12.99	3	57	7	2	0
20	446	20.00	0.77	8	57	8	1	0
20	447	18.46	2.31	8	57	8	1	0
20	448	13.08	2.31	8	57	8	1	0

20	449	68.46	6.15	5	57	8	1	8
20	450	95.38	11.54	5	57	8	1	8
20	451	50.00	5.38	5	57	8	1	8
10	80	0.03	0.01	6	58	4	0	0
10	81	14.54	2.59	6	59	4	0	0
14	327	0.08	0.06	3	60	3	3	0
14	337	0.08	0.00	3	60	3	3	0
14	328	0.23	0.15	3	61	3	3	0
14	338	0.15	0.10	3	61	3	3	0
4	18	17.65	3.57	2	62	2	7	0
4	19	15.88	0.88	6	62	2	7	17
4	20	32.65		1	62	2	7	18
8	42	0.00	0.00	3	62	2	7	16
8	49	0.00	0.00	3	62	2	7	16
14	362	0.25	0.06	3	62	3	3	0
14	365	0.30	0.00	3	62	3	3	0
14	368	0.74	0.28	3	62	7	10	0
14	371	0.61	0.11	3	62	7	10	0
4	9	0.92	0.71	2	63	2	7	0
4	10	5.28	1.63	6	63	2	7	17
4	11	10.88		1	63	2	7	18
5	26	3.55	0.33	2	63	2	6	0
5	25	3.50	0.46	3	63	2	6	16
5	24	11.68	1.55	3	63	2	6	19
7	38	60.00		1	63	2	6	0
14	313	10.16	4.61	3	63	3	3	0
14	315	11.34	5.55	3	63	3	3	0
10	82	0.99	0.26	6	63	4	0	0
14	317	1.54	0.00	3	63	7	10	0
14	319	2.48	1.18	3	63	7	10	0
10	62	0.04	0.01	6	64	4	0	0
10	89	0.00	0.00	6	65	4	0	0
27	491	8.40	0.26	5	66	2	9	0
27	493	10.20	0.17	12	66	2	9	20
10	88	0.46	0.04	6	68	4	0	0
10	84	0.12	0.03	6	69	4	0	0
10	85	0.19	0.04	6	70	4	0	0
10	86	0.12	0.02	6	71	4	0	0
10	87	0.18	0.04	6	72	4	0	0

10	90	0.06	0.01	6	74	4	0	0
14	363	0.44	0.06	3	75	3	3	0
14	366	0.61	0.33	3	75	3	3	0
10	91	0.04	0.01	6	75	4	0	0
14	369	0.17	0.00	3	75	7	10	0
14	372	0.41	0.03	3	75	7	10	0
10	92	0.07	0.01	6	76	4	0	0
14	378	0.48	0.21	3	77	3	3	0
14	380	0.75	0.65	3	77	3	3	0
10	96	0.01	0.00	6	77	4	0	0
14	382	0.12	0.00	3	77	7	10	0
14	383	0.18	0.00	3	77	7	10	0
8	45	0.00	0.00	3	78	2	7	16
8	52	0.00	0.00	3	78	2	7	16
11	106	0.00	0.00	15	79	8	1	7
8	46	0.00	0.00	3	80	2	7	16
8	53	0.00	0.00	3	80	2	7	16
10	93	0.01	0.00	6	81	4	0	0
10	101	0.03	0.01	6	81	4	0	0
10	94	0.06	0.01	6	82	4	0	0
10	95	0.07	0.01	6	83	4	0	0
2	1	0.30	0.22	4	84	2	7	0
2	2	1.44	0.61	6	84	2	7	16
10	98	0.26	0.11	6	84	4	0	0
14	351	1.25	0.37	3	85	3	3	0
14	354	1.62	0.35	3	85	3	3	0
14	357	0.40	0.10	3	85	7	10	0
14	360	0.50	0.14	3	85	7	10	0
14	352	2.37	0.09	3	86	3	3	0
14	355	2.38	0.30	3	86	3	3	0
14	358	2.19	0.02	3	86	7	10	0
14	361	2.45	0.37	3	86	7	10	0
8	43	0.00	0.00	3	87	2	7	16
8	50	0.00	0.00	3	87	2	7	16
4	21	7.32	4.63	2	88	2	7	0
4	22	3.93	0.43	6	88	2	7	17
4	23	7.92		1	88	2	7	18
14	364	2.76	0.58	3	88	3	3	0
14	367	3.20	0.97	3	88	3	3	0

10	97	0.02	0.00	6	88	4	0	0
14	370	1.10	0.08	3	88	7	10	0
14	373	1.30	0.33	3	88	7	10	0
13	266	215.79	200.00	17	89	8	1	0
13	267	586.84	73.68	15	89	8	1	20
14	384	0.56	0.12	3	90	3	3	0
14	386	0.73	0.21	3	90	3	3	0
10	83	2.08	0.31	6	90	4	0	0
14	388	0.26	0.02	3	90	7	10	0
14	390	0.35	0.07	3	90	7	10	0
4	6	50.76	17.04	2	91	2	7	0
4	7	76.27	3.72	6	91	2	7	17
4	8	58.09		1	91	2	7	18
16	416	99.24	11.56	5	91	2	7	0
16	417	106.87	12.25	7	91	2	7	0
16	418	106.40	15.63	11	91	2	7	0
24	478	126.40	17.42	10	91	2	7	16
24	483	46.55	21.43	10	91	2	4	16
21	458	160.00	25.67	9	91	8	1	0
21	459	134.00	19.90	11	91	8	1	0
21	460	138.00	34.00	9	91	8	1	7
21	461	150.00	24.75	8	91	8	1	7
26	489	496.03	39.31	37	91	8	1	0
6	28	3.00	0.49	6	93	1	11	0
17	426	131.05	110.89	18	93	2	7	11
17	427	98.79	108.87	18	93	2	7	12
17	425	173.39	82.66	18	93	2	7	13
29	510	147.75	16.37	10	93	2	4	0
29	516	162.44	7.41	10	93	2	7	0
29	511	171.35	25.57	10	93	2	4	12
29	517	146.58	19.67	7	93	2	7	12
31	539	46.99	12.26	23	93	2	7	0
31	542	89.16	9.47	23	93	2	7	0
31	545	122.24	15.27	23	93	2	7	0
31	548	70.02	12.06	23	93	2	7	0
31	551	135.96	21.26	23	93	2	7	0
31	554	144.16	18.98	23	93	2	7	0
31	540	40.46	10.13	12	93	2	7	8
31	543	113.47	16.89	12	93	2	7	8

31	546	154.86	22.44	12	93	2	7	8
31	549	41.79	2.63	12	93	2	7	8
31	552	123.93	35.40	12	93	2	7	8
31	555	137.16	22.18	12	93	2	7	8
31	541	43.46	10.86	13	93	2	7	16
31	544	91.82	10.82	13	93	2	7	16
31	547	165.03	5.70	13	93	2	7	16
31	550	56.44	16.58	13	93	2	7	16
31	553	123.27	15.02	13	93	2	7	16
31	556	159.33	19.56	13	93	2	7	16
19	443	4.86	0.25	5	93	4	0	0
19	444	4.00	1.14	4	93	4	0	1
19	445	5.96	1.41	4	93	4	0	2
28	500	210.55	22.63	10	93	5	14	0
28	501	178.08	23.61	10	93	5	15	0
28	502	201.69	16.73	5	93	5	14	0
28	503	220.39	30.50	5	93	5	15	0
28	504	219.40	26.56	5	93	5	14	0
28	505	137.74	29.52	5	93	5	15	0
15	410	71.43	29.76	9	93	7	5	0
15	413	11.90	2.98	4	93	7	5	0
9	55	15.52	3.91	3	93	8	1	0
9	56	9.48	1.50	3	93	8	1	8
12	108	84.75	3.39	10	93	8	1	8
12	109	315.25	16.95	10	93	8	1	8
12	110	658.75	22.61	10	93	8	1	8
12	111	405.66	24.85	10	93	8	1	8
12	112	129.93	39.56	10	93	8	1	8
12	113	116.37		20	93	8	1	8
12	114	450.85		20	93	8	1	8
12	115	690.41		20	93	8	1	8
12	116	450.85		20	93	8	1	8
12	117	206.78		20	93	8	1	8
12	118	120.92		17	93	8	1	8
12	119	351.42		17	93	8	1	8
12	120	346.88		17	93	8	1	8
12	121	342.37		17	93	8	1	8
12	122	161.59		17	93	8	1	8
12	123	84.75		19	93	8	1	8

12	124	193.22		19	93	8	1	8
12	125	179.66		19	93	8	1	8
12	126	188.71		19	93	8	1	8
12	127	93.80		19	93	8	1	8
12	128	108.03	19.39	16	93	8	1	8
12	129	394.94	37.97	16	93	8	1	8
12	130	523.22	39.65	16	93	8	1	8
12	131	432.08	32.91	16	93	8	1	8
12	132	182.28	2.53	16	93	8	1	8
12	133	87.77	9.29	25	93	8	1	8
12	134	256.53	24.48	25	93	8	1	8
12	135	276.78	31.24	25	93	8	1	8
12	136	256.53	16.05	25	93	8	1	8
12	137	114.76	10.99	25	93	8	1	8
12	138	108.03	17.72	10	93	8	1	8
12	139	202.53	36.28	10	93	8	1	8
12	140	162.03	34.61	10	93	8	1	8
12	141	172.15	22.78	10	93	8	1	8
12	142	111.39	19.42	10	93	8	1	8
22	466	132.24	13.82	12	93	8	1	20
22	467	140.13	17.76	12	93	8	1	20
22	468	483.55	27.63	12	93	8	1	20
22	469	505.26	25.66	12	93	8	1	20
25	488	0.71	0.18	7	94	2	7	0
25	487	5.79	1.11	7	94	2	7	16
8	47	0.00	0.00	3	96	2	7	16
8	54	0.00	0.00	3	96	2	7	16
11	105	0.00	0.00	15	96	8	1	7
10	102	0.07	0.01	6	98	4	0	0
14	396	1.29	0.45	3	99	3	3	0
14	402	1.55	0.65	3	99	3	3	0
10	99	0.13	0.02	6	99	4	0	0
14	406	0.39	0.19	3	99	7	10	0
14	409	0.52	0.19	3	99	7	10	0
10	100	0.07	0.02	6	100	4	0	0
10	103	0.01	0.00	6	101	4	0	0
27	490	7.60	0.21	5	102	2	9	0
27	492	9.50	0.16	12	102	2	9	20
10	104	0.01	0.00	6	104	4	0	0

14	374	0.39	0.31	3	106	3	3	0
14	376	0.13	0.09	3	106	3	3	0
14	375	0.55	0.15	3	107	3	3	0
14	377	0.58	0.09	3	107	3	3	0
14	379	0.03	0.02	3	108	3	3	0
14	381	0.11	0.10	3	108	3	3	0
14	385	0.72	0.16	3	109	3	3	0
14	387	1.06	0.19	3	109	3	3	0
14	389	0.55	0.13	3	109	7	10	0
14	391	0.96	0.62	3	109	7	10	0
14	392	0.77	0.97	3	110	3	3	0
14	398	0.19	0.13	3	110	3	3	0
14	393	0.19	0.06	3	111	3	3	0
14	399	1.42	2.26	3	111	3	3	0
14	404	0.13	0.06	3	111	7	10	0
14	407	0.06	0.06	3	111	7	10	0
14	394	4.58	2.65	3	112	3	3	0
14	400	2.97	0.77	3	112	3	3	0
14	395	0.39	0.06	3	113	3	3	0
14	401	0.58	0.19	3	113	3	3	0
14	405	0.32	0.26	3	113	7	10	0
14	408	0.39	0.26	3	113	7	10	0
14	397	0.19	0.13	3	114	3	3	0
14	403	0.19	0.06	3	114	3	3	0

Appendix 5: Legend for coded data entries corresponding to studyNumber, Tissue, TissueRegion, PathologyType, and Protein.

StudyNumber	
Tan	1
Attur	2
Yammane	3
Lorenzo	4
Chevalier	5
Choi	6
Hsueh	7
Peffers	8
Alkhatib	9
Johnson	10
Kozaci	11

Antoniou 1996	12
Antoniou 2001	13
Little	14
Steigman	15
Franz	16
Squires	17
Sato	18
Guyette	19
Cloyd	20
Cheng	21
Bibby	22
Berria	23
Aurich 2017	24
Patel	25
Benneker	26
Theocaris	27
Wilson	28
Aurich 2005	29
Cook	30
Temple-Wong	31
Tissue	
Adipose	1
Cartilage	2
Ligament	3
Myocardium	4
Skeletal Muscle	5
Small Intestine	6
Tendon	7
IVD Cartilage	8
TissueRegion	
IVD	1
Achilles Tendon	2
ACL	3
Ankle	4
Diaphragmatic Tendon	5
Hip	6
Knee	7

Medial Rectus	8
Nasal	9
Patellar Tendon	10
Stomach	11
Vastus Lateralus	12
Yellow Ligament	13
Rectus Femoris	14
Supraspinatus	15
Obturator Internus Muscle	16
Pathology Type	
Brain Death	1
Cardiac Death	2
Cardiac Infarction - 1 Week Post	3
Cardiac Infarction - 2 Weeks Post	4
Cardiac Infarction - 4 Weeks Post	5
Cardiac Infarction - 8 Weeks Post	6
Deformity (IVD)	7
Degradation	8
Diabetic	9
Intermittent Exotropia	10
Lesion - Adjacent	11
Lesion - Lesion	12
Lesion - Remote	13
Mechanical Injury - Acute	14
Obesity	15
Osteoarthritis	16
Osteoarthritis - Early	17
Osteoarthritis - Late	18
Osteonecrosis	19
Scoliosis	20
Protein	
a1 Antitrypsin	1
Actin - Cytoplasmic 1/2	2
Actin - Total	3
Aggrecan	4
Agrin	5
Agrin (iso 2,3,4,5,&6)	6

Albumin	7
Alpha-1 Microglobulin/Bikunin Precursor	8
Annexin A2	9
Apolipoprotein A4	10
Apolipoprotein E	11
asporin	12
Basic Fibroblast Growth Factor	13
Biglycan	14
Cartilage Intermediate Layer Protein	15
Cartilage Oligomeric Matrix Protein	16
Chondroitin Sulphate (846 Epitope)	17
Clusterin	18
Collagen (Col11A2)	19
Collagen (Col12A1)	20
Collagen (Col18A1)	21
Collagen (Col1A1)	22
Collagen (Col1A2)	23
Collagen (Col2A1)	24
Collagen (Col3A1)	25
Collagen (Col4A1)	26
Collagen (Col4A1/5)	27
Collagen (Col4A2)	28
Collagen (Col4A5)	29
Collagen (Col5A1)	30
Collagen (Col5A2)	31
Collagen (Col6A1)	32
Collagen (Col6A2)	33
Collagen (Col6A3)	34
Collagen (Col9A3)	35
Collagen (C-propeptide of type 1 procollagen)	36
Collagen (C-propeptide of type 2 procollagen)	37
Collagen (Insoluble)	38
Collagen (Soluble)	39
Collagen (total)	40
Collagen (type 1)	41
Collagen (type 2)	42
Collagen (type 3)	43
Collagen (type 5)	44
Collagen (type 6)	45

Collagen (type 9)	46
Collagen (type 11)	47
Collagen (type 12)	48
Collagen II Degradation Protein (C2C)	49
Complement Component 4 Binding Protein Alpha	50
Complement Component 3	51
Complement Component 9	52
Decorin	53
Dermatopontin	54
Desmin	55
Dioxy-Pyridinoline	56
Elastin	57
Emilin-1	58
Fibrillin-1	59
Fibrinogen Beta	60
Fibrinogen Gamma	61
Fibromodulin	62
Fibronectin	63
Fibulin-5	64
Galactin-1	65
Hexosamine	66
Hydroxyproline	67
Laminin Gamma-1	68
Laminin Subunit Alpha-2	69
Laminin Subunit Alpha-5	70
Laminin Subunit Beta-1	71
Laminin Subunit Beta-2	72
Laminin - Total	73
Latent Transforming Growth Factor Beta Binding Protein-1	74
Lumican	75
Microfibrillar-associated protein 2	76
Mimecan/Osteoglycin	77
MMP-1	78
MMP2 (Pro-MMP2)	79
MMP-3	80
Myosin (Myosin-3,4,6,7)	81
Nidogen-1	82
Nidogen-2	83
Periostin	84

Phosphatidylglycerophosphate Synthase 1	85
Phosphatidylglycerophosphate Synthase 2	86
Plasminogen	87
Prolargin (PRELP)	88
Protein (Total)	89
Perlecan - HSPG2	90
Proteoglycan (total)	91
Pyridinoline	92
Sulphated Glycosaminoglycan	93
Tenascin C	94
Thrombospondin-2	95
TIMP-1	96
Tissue Inhibitor of Metalloproteinases-2	97
TnxB Protein	98
Transforming Growth Factor-Beta-induced-protein ig-h3	99
Transglutaminase 2	100
Tubulin beta-4B chain (4b & 5 Chain)	101
Uronic Acid	102
Vascular Endothelial Growth Factor	103
Vimentin	104
Water	105
Phosphoglucomutase	106
Chondroitin Sulphate Proteoglycan Core Protein 2	107
Chondroadherin	108
Lubricin (PRG4)	109
Apolipoprotein A1	110
Transthyretin	111
Fibrinogen	112
Pentraxin-2 (SAMP)	113
Milk Fat Globule Membrane	114

References

1. Bhagavan, N. V. & Ha, C.-E. (2011) *Essentials of medical biochemistry : with clinical cases*, Elsevier/Academic Press, Amsterdam.
2. Halper, J. & Kjaer, M. (2014) Basic Components of Connective Tissues and Extracellular Matrix: Elastin, Fibrillin, Fibulins, Fibrinogen, Fibronectin, Laminin, Tenascins and Thrombospondins in *Progress in Heritable Soft Connective Tissue Diseases* (Halper, J., ed) pp. 31-47, Springer Netherlands, Dordrecht.
3. Choe, S. S., Huh, J. Y., Hwang, I. J., Kim, J. I. & Kim, J. B. (2016) Adipose Tissue Remodeling: Its Role in Energy Metabolism and Metabolic Disorders, *Front Endocrinol.* **7**, 30.
4. Benjamin, M., Kaiser, E. & Milz, S. (2008) Structure-function relationships in tendons: a review, *J Anat.* **212**, 211-228.
5. Gentili, C. & Cancedda, R. (2009) Cartilage and bone extracellular matrix, *Curr Pharm Des.* **15**, 1334-48.
6. Alberts, B. (2015) *Molecular biology of the cell*, Garland Science.
7. Plotnikov, S. V., Pasapera, A. M., Sabass, B. & Waterman, C. M. (2012) Force Fluctuations within Focal Adhesions Mediate ECM-Rigidity Sensing to Guide Directed Cell Migration, *Cell.* **151**, 1513-27.
8. Engler, A. J., Sen, S., Sweeney, H. L. & Discher, D. E. (2006) Matrix Elasticity Directs Stem Cell Lineage Specification, *Cell.* **126**, 677-689.
9. Lo, C. M., Wang, H. B., Dembo, M. & Wang, Y. L. (2000) Cell movement is guided by the rigidity of the substrate, *Biophys J.* **79**, 144-152.
10. Hadjipanayi, E., Mudera, V. & Brown, R. A. (2009) Close dependence of fibroblast proliferation on collagen scaffold matrix stiffness, *J Tissue Eng Regen Med.* **3**, 77-84.
11. Wang, H. B., Dembo, M. & Wang, Y. L. (2000) Substrate flexibility regulates growth and apoptosis of normal but not transformed cells, *Am J Physiol Cell Physiol.* **279**, C1345-50.
12. Coombe, D. (1998) EXTRACELLULAR MATRIX VOLUME 1: TISSUE FUNCTION VOLUME 2: MOLECULAR COMPONENTS AND INTERACTIONS, *Immunology And Cell Biology.* **76**, 114.
13. Culav, E. M., Clark, C. H. & Merrilees, M. J. (1999) Connective Tissues: Matrix Composition and Its Relevance to Physical Therapy, *Physical Therapy.* **79**, 308-319.
14. Hukins, D. W. L., Weston, S. A., Humphries, M. J. & Freemont, A. J. (1996) Chapter 8 Extracellular matrix in *Principles of Medical Biology* (Bittar, E. E. & Bittar, N., eds) pp. 181-232, Elsevier.
15. Theocharis, A. D., Skandalis, S. S., Gialeli, C. & Karamanos, N. K. (2016) Extracellular matrix structure, *Adv Drug Deliv Rev.* **97**, 4-27.
16. Kadler, K. E., Baldock, C., Bella, J. & Boot-Handford, R. P. (2007) Collagens at a glance, *J Cell Sci.* **120**, 1955-1958.
17. Brodsky, B. & Persikov, A. V. (2005) Molecular structure of the collagen triple helix, *Adv Protein Chem.* **70**, 301-39.
18. Heino, J. (2007) The collagen family members as cell adhesion proteins, *Bioessays.* **29**, 1001-10.
19. Henriksen, K. & Karsdal, M. A. (2016) Chapter 1 - Type I Collagen in *Biochemistry of Collagens, Laminins and Elastin* pp. 1-11, Academic Press.
20. Gudmann, N. S. & Karsdal, M. A. (2016) Chapter 2 - Type II Collagen in *Biochemistry of Collagens, Laminins and Elastin* pp. 13-20, Academic Press.
21. Makareeva, E. & Leikin, S. (2014) Chapter 7 - Collagen Structure, Folding and Function in *Osteogenesis Imperfecta* (Byers, P. H., Glorieux, F. H. & Sponseller, P. D., eds) pp. 71-84, Academic Press, San Diego.
22. ROSS, R. (1973) The Elastic Fiber: A Review, *J Histochem Cytochem.* **21**, 199-208.

23. Jensen, Sacha A., Robertson, Ian B. & Handford, Penny A. Dissecting the Fibrillin Microfibril: Structural Insights into Organization and Function, *Structure.* **20**, 215-225.
24. Sandberg, L. B., Weissman, N. & Gray, W. R. (1971) Structural features of tropoelastin related to the sites of cross-links in aortic elastin, *Biochemistry.* **10**, 52-6.
25. Wagenseil, J. E. & Mecham, R. P. (2007) New insights into elastic fiber assembly, *Birth Defects Res C Embryo Today.* **81**, 229-40.
26. Kielty, C. M., Sherratt, M. J. & Shuttleworth, C. A. (2002) Elastic fibres, *J Cell Sci.* **115**, 2817-28.
27. Couchman, J. R. & Pataki, C. A. (2012) An Introduction to Proteoglycans and Their Localization, *J Histochem Cytochem.* **60**, 885-897.
28. Merline, R., Schaefer, R. M. & Schaefer, L. (2009) The matricellular functions of small leucine-rich proteoglycans (SLRPs), *J Cell Commun Signal.* **3**, 323-335.
29. Aspberg, A. (2012) The Different Roles of Aggrecan Interaction Domains, *J Histochem Cytochem.* **60**, 987-996.
30. Reed, C. C. & Iozzo, R. V. (2002) The role of decorin in collagen fibrillogenesis and skin homeostasis, *Glycoconj J.* **19**, 249-55.
31. Rada, J. A., Cornuet, P. K. & Hassell, J. R. (1993) Regulation of corneal collagen fibrillogenesis in vitro by corneal proteoglycan (lumican and decorin) core proteins, *Exp Eye Res.* **56**, 635-48.
32. Sorokin, K. v. d. M. L. (2003) Adhesive Glycoproteins in *Connective Tissue and Its Heritable Disorders*.
33. Yamada, K. M. (1991) Fibronectin and Other Cell Interactive Glycoproteins in *Cell Biology of Extracellular Matrix: Second Edition* (Hay, E. D., ed) pp. 111-146, Springer US, Boston, MA.
34. Hao, C., Cui, Y., Owen, S., Li, W., Cheng, S. & Jiang, W. G. (2017) Human osteopontin: Potential clinical applications in cancer (Review), *International Journal of Molecular Medicine.* **39**, 1327-1337.
35. Neame, P. J. & Barry, F. P. (1994) The link proteins, *EXS.* **70**, 53-72.
36. Goldring, M. B. (2012) Chondrogenesis, chondrocyte differentiation, and articular cartilage metabolism in health and osteoarthritis, *Therapeutic Advances in Musculoskeletal Disease.* **4**, 269-285.
37. Caetano-Lopes, J., Canhao, H. & Fonseca, J. E. (2007) Osteoblasts and bone formation, *Acta Reumatol Port.* **32**, 103-10.
38. Frantz, C., Stewart, K. M. & Weaver, V. M. (2010) The extracellular matrix at a glance, *J Cell Sci.* **123**, 4195-4200.
39. Hynes, R. O. (2002) Integrins: bidirectional, allosteric signaling machines, *Cell.* **110**, 673-87.
40. Hynes, R. O. (2003) Structural biology. Changing partners, *Science.* **300**, 755-6.
41. Hynes, R. O. (2009) The extracellular matrix: not just pretty fibrils, *Science.* **326**, 1216-9.
42. Ornitz, D. M. & Itoh, N. (2015) The Fibroblast Growth Factor signaling pathway, *Wiley Interdiscip Rev Dev Biol.* **4**, 215-266.
43. Massagué, J. (2012) TGFβ signalling in context, *Nat Rev Mol Cell Biol.* **13**, 616.
44. Komiya, Y. & Habas, R. (2008) Wnt signal transduction pathways, *Organogenesis.* **4**, 68-75.
45. Adair-Kirk, T. L. & Senior, R. M. (2008) Fragments of Extracellular Matrix as Mediators of Inflammation, *Int J Biochem Cell Biol.* **40**, 1101-1110.
46. Mariman, E. C. M. & Wang, P. (2010) Adipocyte extracellular matrix composition, dynamics and role in obesity, *Cellular and Molecular Life Sciences.* **67**, 1277-1292.
47. Pierleoni, C., Verdenelli, F., Castellucci, M. & Cinti, S. (1998) Fibronectins and basal lamina molecules expression in human subcutaneous white adipose tissue, *Eur J Histochem.* **42**, 183-8.
48. Nakajima, I., Yamaguchi, T., Ozutsumi, K. & Aso, H. (1998) Adipose tissue extracellular matrix: newly organized by adipocytes during differentiation, *Differentiation.* **63**, 193-200.
49. Mariman, E. C. M. & Wang, P. (2010) Adipocyte extracellular matrix composition, dynamics and role in obesity, *Cell Mol Life Sci.* **67**, 1277-1292.

50. Bolton, K., Segal, D. & Walder, K. (2012) The small leucine-rich proteoglycan, biglycan, is highly expressed in adipose tissue of Psammomys obesus and is associated with obesity and type 2 diabetes, *Biologics.* **6**, 67-72.
51. Ker, R. F. (2007) Mechanics of tendon, from an engineering perspective, *International Journal of Fatigue.* **29**, 1001-1009.
52. Thorpe, C. T. & Screen, H. R. (2016) Tendon Structure and Composition, *Adv Exp Med Biol.* **920**, 3-10.
53. Samiric, T., Ilic, M. Z. & Handley, C. J. (2004) Characterisation of proteoglycans and their catabolic products in tendon and explant cultures of tendon, *Matrix Biol.* **23**, 127-40.
54. Vogel, K. G. (2004) What happens when tendons bend and twist? Proteoglycans, *J Musculoskelet Neuronal Interact.* **4**, 202-3.
55. Benjamin, M. & Ralphs, J. R. (1998) Fibrocartilage in tendons and ligaments--an adaptation to compressive load, *J Anat.* **193 (Pt 4)**, 481-94.
56. Funakoshi, T., Schmid, T., Hsu, H. P. & Spector, M. (2008) Lubricin distribution in the goat infraspinatus tendon: a basis for interfascicular lubrication, *J Bone Joint Surg Am.* **90**, 803-14.
57. Smith, R. K., Gerard, M., Dowling, B., Dart, A. J., Birch, H. L. & Goodship, A. E. (2002) Correlation of cartilage oligomeric matrix protein (COMP) levels in equine tendon with mechanical properties: a proposed role for COMP in determining function-specific mechanical characteristics of locomotor tendons, *Equine Vet J Suppl*, 241-4.
58. Alford, A. I., Kozloff, K. M. & Hankenson, K. D. (2015) Extracellular matrix networks in bone remodeling, *Int J Biochem Cell Biol.* **65**, 20-31.
59. Jiang, X., Ye, M., Jiang, X., Liu, G., Feng, S., Cui, L. & Zou, H. (2007) Method Development of Efficient Protein Extraction in Bone Tissue for Proteome Analysis, *Journal of Proteome Research.* **6**, 2287-2294.
60. Bennett, K. P., Bergeron, C., Acar, E., Klees, R. F., Vandenberg, S. L., Yener, B. & Plopper, G. E. (2007) Proteomics reveals multiple routes to the osteogenic phenotype in mesenchymal stem cells, *BMC Genomics.* **8**, 380.
61. Igor, M., G., M. B., Xi, W., A., D. N., L., W. D., W., R. D., Danka, G. & Ivo, K. (2016) Quiescent Bone Lining Cells Are a Major Source of Osteoblasts During Adulthood, *STEM CELLS.* **34**, 2930-2942.
62. Horiuchi, K., Amizuka, N., Takeshita, S., Takamatsu, H., Katsuura, M., Ozawa, H., Toyama, Y., Bonewald Lynda, F. & Kudo, A. (2009) Identification and Characterization of a Novel Protein, Periostin, with Restricted Expression to Periosteum and Periodontal Ligament and Increased Expression by Transforming Growth Factor β, *Journal of Bone and Mineral Research.* **14**, 1239-1249.
63. Dallas, S. L., Prideaux, M. & Bonewald, L. F. (2013) The Osteocyte: An Endocrine Cell … and More, *Endocrine Reviews.* **34**, 658-690.
64. Lamoureux, F., Baud'huin, M., Duplomb, L., Heymann, D. & Rédini, F. (2007) Proteoglycans: key partners in bone cell biology, *Bioessays.* **29**, 758-71.
65. Sophia Fox, A. J., Bedi, A. & Rodeo, S. A. (2009) The Basic Science of Articular Cartilage: Structure, Composition, and Function, *Sports Health.* **1**, 461-468.
66. Eggli, P. S., Herrmann, W., Hunziker, E. B. & Schenk, R. K. (1985) Matrix compartments in the growth plate of the proximal tibia of rats, *Anat Rec.* **211**, 246-57.
67. Patterson-Kane, J. C. & Firth, E. C. (2014) CHAPTER 13 - Tendon, ligament, bone, and cartilage: Anatomy, physiology, and adaptations to exercise and training in *The Athletic Horse (Second Edition)* (Hodgson, D. R., McKeever, K. H. & McGowan, C. M., eds) pp. 202-242, W.B. Saunders.
68. Martel-Pelletier, J., Barr, A. J., Cicuttini, F. M., Conaghan, P. G., Cooper, C., Goldring, M. B., Goldring, S. R., Jones, G., Teichtahl, A. J. & Pelletier, J.-P. (2016) Osteoarthritis, *Nat Rev Dis Primers.* **2**, 16072.
69. Verzijl, N., DeGroot, J., Thorpe, S. R., Bank, R. A., Shaw, J. N., Lyons, T. J., Bijlsma, J. W., Lafeber, F. P., Baynes, J. W. & TeKoppele, J. M. (2000) Effect of collagen turnover on the accumulation of advanced glycation end products, *J Biol Chem.* **275**, 39027-31.

70. Goldring, M. B. & Marcu, K. B. (2009) Cartilage homeostasis in health and rheumatic diseases, *Arthritis Res Ther.* **11**, 224.
71. Sivan, S. S., Hayes, A. J., Wachtel, E., Caterson, B., Merkher, Y., Maroudas, A., Brown, S. & Roberts, S. (2014) Biochemical composition and turnover of the extracellular matrix of the normal and degenerate intervertebral disc, *European Spine Journal.* **23**, 344-353.
72. Eyre, D. R., Matsui, Y. & Wu, J. J. (2002) Collagen polymorphisms of the intervertebral disc, *Biochem Soc Trans.* **30**, 844-8.
73. Eyre, D. R. & Muir, H. (1976) Types I and II collagens in intervertebral disc. Interchanging radial distributions in annulus fibrosus, *Biochemical Journal.* **157**, 267-270.
74. Roughley, P. J. (2004) Biology of intervertebral disc aging and degeneration: involvement of the extracellular matrix, *Spine.* **29**, 2691-9.
75. Battié, M. C., Videman, T. & Parent, E. (2004) Lumbar Disc Degeneration: Epidemiology and Genetic Influences, *Spine.* **29**.
76. Cheung, K. M. C. (2010) The relationship between disc degeneration, low back pain, and human pain genetics, *Spine.* **10**, 958-960.
77. Dickson, I. R., Happey, F., Pearson, C. H., Naylor, A. & Turner, R. L. (1967) Variations in the Protein Components of Human Intervertebral Disk with Age, *Nature.* **215**, 52.
78. Antoniou, J., Arlet, V., Goswami, T., Aebi, M. & Alini, M. (2001) Elevated synthetic activity in the convex side of scoliotic intervertebral discs and endplates compared with normal tissues, *Spine.* **26**, E198-206.
79. Mwale, F., Iatridis, J. C. & Antoniou, J. (2008) Quantitative MRI as a diagnostic tool of intervertebral disc matrix composition and integrity, *European Spine Journal.* **17**, 432-440.
80. Begley, C. G. & Ellis, L. M. (2012) Raise standards for preclinical cancer research, *Nature.* **483**, 531.
81. Ghallab, A. (2013) In vitro test systems and their limitations, *EXCLI J.* **12**, 1024-1026.
82. Reidy, P. T., Hinkley, J. M., Trappe, T. A., Trappe, S. W. & Harber, M. P. (2014) Protein composition of endurance trained human skeletal muscle, *Int J Sports Med.* **35**, 476-81.
83. Aurich, M., Hofmann, G. O. & Rolauffs, B. (2017) Differences in type II collagen turnover of osteoarthritic human knee and ankle joints, *Int Orthop.* **6**, 06.
84. Aurich, M., Squires, G. R., Reiner, A., Mollenhauer, J. A., Kuettner, K. E., Poole, A. R. & Cole, A. A. (2005) Differential matrix degradation and turnover in early cartilage lesions of human knee and ankle joints, *Arthritis Rheum.* **52**, 112-119.
85. Chevalier, X. (1993) Fibronectin, cartilage, and osteoarthritis, *Semin Arthritis Rheum.* **22**, 307-18.
86. Franz, T., Hasler, E. M., Hagg, R., Weiler, C., Jakob, R. P. & Mainil-Varlet, P. (2001) In situ compressive stiffness, biochemical composition, and structural integrity of articular cartilage of the human knee joint, *Osteoarth Cartil.* **9**, 582-92.
87. Hsueh, M. F., Khabut, A., Kjellstrom, S., Onnerfjord, P. & Kraus, V. B. (2016) Elucidating the Molecular Composition of Cartilage by Proteomics, *J Proteome Res.* **15**, 374-88.
88. Lorenzo, P., Bayliss, M. T. & Heinegard, D. (2004) Altered patterns and synthesis of extracellular matrix macromolecules in early osteoarthritis, *Matrix Biol.* **23**, 381-91.
89. Patel, L., Sun, W., Glasson, S. S., Morris, E. A., Flannery, C. R. & Chockalingam, P. S. (2011) Tenascin-C induces inflammatory mediators and matrix degradation in osteoarthritic cartilage, *BMC Musculoskel Disord.* **12**, 164.
90. Peffers, M. J., Beynon, R. J. & Clegg, P. D. (2013) Absolute quantification of selected proteins in the human osteoarthritic secretome, *Int J Mol Sci.* **14**, 20658-81.
91. Squires, G. R., Okouneff, S., Ionescu, M. & Poole, A. R. (2003) The pathobiology of focal lesion development in aging human articular cartilage and molecular matrix changes characteristic of osteoarthritis, *Arthritis Rheum.* **48**, 1261-70.

92. Temple-Wong, M. M., Bae, W. C., Chen, M. Q., Bugbee, W. D., Amiel, D., Coutts, R. D., Lotz, M. & Sah, R. L. (2009) Biomechanical, structural, and biochemical indices of degenerative and osteoarthritic deterioration of adult human articular cartilage of the femoral condyle, *Osteoarth Cartil.* **17**, 1469-1476.
93. Theocharis, A. D., Karamanos, N. K., Papageorgakopoulou, N., Tsiganos, C. P. & Theocharis, D. A. (2002) Isolation and characterization of matrix proteoglycans from human nasal cartilage. Compositional and structural comparison between normal and scoliotic tissues, *Biochim Biophys Acta.* **1569**, 117-26.
94. Yamane, T., Matsuo, T., Hasebe, S. & Ohtsuki, H. (2003) Clinical correlations of aggrecan in the resected medial rectus muscle of patients with intermittent exotropia, *Acta Medica Okayama.* **57**, 199-204.
95. Alkhatib, B., Rosenzweig, D. H., Krock, E., Roughley, P. J., Beckman, L., Steffen, T., Weber, M. H., Ouellet, J. A. & Haglund, L. (2014) Acute mechanical injury of the human intervertebral disc: link to degeneration and pain, *Eu Cell Mater.* **28**, 98-110.
96. Antoniou, J., Steffen, T., Nelson, F., Winterbottom, N., Hollander, A. P., Poole, R. A., Aebi, M. & Alini, M. (1996) 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 Inv.* **98**, 996-1003.
97. Benneker, L. M., Heini, P. F., Alini, M., Anderson, S. E. & Ito, K. (2005) 2004 Young Investigator Award Winner: vertebral endplate marrow contact channel occlusions and intervertebral disc degeneration, *Spine.* **30**, 167-73.
98. Bibby, S. R., Fairbank, J. C., Urban, M. R. & Urban, J. P. (2002) Cell viability in scoliotic discs in relation to disc deformity and nutrient levels, *Spine.* **27**, 2220-8; discussion 2227-8.
99. Cheng, K. K., Berven, S. H., Hu, S. S. & Lotz, J. C. (2014) Intervertebral discs from spinal nondeformity and deformity patients have different mechanical and matrix properties, *Spine.* **14**, 522-30.
100. Cloyd, J. M. & Elliott, D. M. (2007) Elastin content correlates with human disc degeneration in the anulus fibrosus and nucleus pulposus, *Spine.* **32**, 1826-31.
101. Kozaci, L. D., Guner, A., Oktay, G. & Guner, G. (2006) Alterations in biochemical components of extracellular matrix in intervertebral disc herniation: role of MMP-2 and TIMP-2 in type II collagen loss, *Cell Biochem Funct.* **24**, 431-6.
102. Tan, C. I., Kent, G. N., Randall, A. G., Edmondston, S. J. & Singer, K. P. (2003) Age-related changes in collagen, pyridinoline, and deoxypyridinoline in normal human thoracic intervertebral discs, *J Gerontol A Biol Sci Med Sci.* **58**, B387-93.
103. Little, D., Thompson, J. W., Dubois, L. G., Ruch, D. S., Moseley, M. A. & Guilak, F. (2014) Proteomic differences between male and female anterior cruciate ligament and patellar tendon, *PLoS ONE.* **9**, e96526.
104. Sato, N., Taniguchi, T., Goda, Y., Kosaka, H., Higashino, K., Sakai, T., Katoh, S., Yasui, N., Sairyo, K. & Taniguchi, H. (2016) Proteomic Analysis of Human Tendon and Ligament: Solubilization and Analysis of Insoluble Extracellular Matrix in Connective Tissues, *J Proteome Res.* **15**, 4709-4721.
105. Steigman, S. A., Oh, J. T., Almendinger, N., Javid, P., LaVan, D. & Fauza, D. (2010) Structural and biomechanical characteristics of the diaphragmatic tendon in infancy and childhood: an initial analysis, *J Pediatr Surg.* **45**, 1455-8.
106. Berria, R., Wang, L., Richardson, D. K., Finlayson, J., Belfort, R., Pratipanawatr, T., De Filippis, E. A., Kashyap, S. & Mandarino, L. J. (2006) Increased collagen content in insulin-resistant skeletal muscle, *Am J Physiol Endocrinol Metab* **290**, E560-5.
107. Cook, M. S., Bou-Malham, L., Esparza, M. C. & Alperin, M. (2016) Age-related alterations in female obturator internus muscle, *Int Urogynecol J.* **28**, 729-734.
108. Wilson, K., Terlouw, A., Roberts, K. & Wolchok, J. C. (2016) The characterization of decellularized human skeletal muscle as a blueprint for mimetic scaffolds, *J Mater Sci Mater Med.* **27**.

109. Choi, J. S., Kim, B. S., Kim, J. Y., Kim, J. D., Choi, Y. C., Yang, H. J., Park, K., Lee, H. Y. & Cho, Y. W. (2011) Decellularized extracellular matrix derived from human adipose tissue as a potential scaffold for allograft tissue engineering, *J Biomed Mater Res.* **97**, 292-9.
110. Tavares, R. A., Sennes, L. U., Mauad, T., Imamura, R., Da Silva, L. F. F. & Carrau, R. L. (2012) Extracellular matrix composition of the cricopharyngeus muscle, *Dysphagia.* **27**, 277-283.
111. Sawafuji, R., Cappellini, E., Nagaoka, T., Fotakis, A. K., Jersie-Christensen, R. R., Olsen, J. V., Hirata, K. & Ueda, S. (2017) Proteomic profiling of archaeological human bone, *Royal Soc Open Sci.* **4**.
112. Newell, N., Little, J. P., Christou, A., Adams, M. A., Adam, C. J. & Masouros, S. D. (2017) Biomechanics of the human intervertebral disc: A review of testing techniques and results, *J Mech Behav Biomed Mater.* **69**, 420-434.
113. Frank, C. B. (2004) Ligament structure, physiology and function, *J Musculoskel Neuronal Interact.* **4**, 199-201.
114. Clarke, B. (2008) Normal bone anatomy and physiology, *Clin J Am Soc Nephrol.* **3 Suppl 3**, S131-9.
115. Thomas, L. W. (1962) The chemical composition of adipose tissue of man and mice, *Q J Exp Physiol Cogn Med Sci.* **47**, 179-88.
116. Bhosale, A. M. & Richardson, J. B. (2008) Articular cartilage: structure, injuries and review of management, *Br Med Bull.* **87**, 77-95.
117. Stegemann, H. & Stalder, K. (1967) Determination of hydroxyproline, *Clin Chim Acta.* **18**, 267-273.
118. Kivirikko, K. I. & Myllylä, R. (1982) Posttranslational enzymes in the biosynthesis of collagen: Intracellular enzymes in *Methods Enzymol* pp. 245-304, Academic Press.
119. Woessner, J. F. (1961) The determination of hydroxyproline in tissue and protein samples containing small proportions of this imino acid, *Arch Biochem Biophys.* **93**, 440-447.
120. Choi, Y.-S. (2009) Pathophysiology of Degenerative Disc Disease, *Asian Spine J.* **3**, 39-44.
121. Wynn, T. A. (2008) Cellular and molecular mechanisms of fibrosis, *J Pathol.* **214**, 199-210.
122. Altaf, F., Gibson, A., Dannawi, Z. & Noordeen, H. (2013) Adolescent idiopathic scoliosis, *Br Med J.* **346**, 30-34.
123. Gillies, A. R. & Lieber, R. L. (2011) Structure and Function of the Skeletal Muscle Extracellular Matrix, *Muscle Nerve.* **44**, 318-331.
124. Haidich, A. B. (2010) Meta-analysis in medical research, *Hippokratia.* **14**, 29-37.
125. Edmondson, R., Broglie, J. J., Adcock, A. F. & Yang, L. (2014) Three-Dimensional Cell Culture Systems and Their Applications in Drug Discovery and Cell-Based Biosensors, *Assay Drug Dev Technol.* **12**, 207-218.
126. Cooke, M. J., Phillips, S. R., Shah, D. S., Athey, D., Lakey, J. H. & Przyborski, S. A. (2008) Enhanced cell attachment using a novel cell culture surface presenting functional domains from extracellular matrix proteins, *Cytotechnology.* **56**, 71-79.
127. Huch, M. & Koo, B.-K. (2015) Modeling mouse and human development using organoid cultures, *Development.* **142**, 3113-3125.