# PROTEIN AND TRANSCRIPTOME SIGNATURES OF CARTILAGE AGEING AND DISEASE

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

**Mandy Jayne Peffers**

**January 2013**

**Table of Contents**


**Acknowledgements** **i**

**List of abbreviations used in text** **iii**

**Abstract** **vii**

**General remarks and Introduction** **1**

Structure and function of normal cartilage **2**

Osteoarthritis **14**

Cartilage degradation **17**

Cartilage ageing **28**

Proteomics **34**

Cartilage secretomics **49**

Whole cartilage proteomics **52**

Next-generation sequencing **55**

**Manuscript 1** **58**

'Proteomic Characterisation and Quantification of an *In-vitro* Early Equine Inflammatory Model'

**Appendix to manuscript 1** **90**

**Manuscript 2** **97**

'Absolute Quantification of Cartilage Extracellular Matrix'

**Appendix to manuscript 2** **146**

**Manuscript 3** **153**

'Protein Signatures of the Human Osteoarthritic Secretome'

**Manuscript 4** **192**

'Characterisation of Neopeptides in Equine Articular Cartilage'

**Appendix to manuscript 4** **227**

**Manuscript 5** **232**

'MALDI Imaging Mass Spectrometry Identifies Markers of Ageing and Osteoarthritic Cartilage'

**Manuscript 6** **261**

'Transcriptomic Signatures in Cartilage Ageing'

**General Discussion** **290**

**References** **305**

**Acknowledgements**

The present work was carried out at the Department of Comparative Musculoskeletal Science Research Group and the Protein Function Group both at the University of Manchester. In addition some of the work was undertaken at the Wellcome Trust Centre for Cell Matrix Research, University of Manchester and the Biomolecular Imaging Mass Spectrometry Group, FOM Institute AMOLF, Amsterdam.

First of all, it is an honour for me to thank my main supervisor Professor Peter Clegg. Without his profound knowledge of cartilage biology, his encouragement, guidance, support and friendship from the beginning to the end, this work would never have been accomplished. I am also grateful to my co-supervisors Professor Rob Beynon for introducing me to the field of mass spectrometry and proteomics and Professor Dave Thornton for his help with cartilage extraction techniques. Finally I would like to thank the research staff at AMOLF and in particularly Dr. Berta Cillero-Pastor for their training, guidance and support in the MALDI-IMS work.

It is a pleasure to thank those who made this thesis possible, through their help and support in the laboratory in the Comparative Musculoskeletal Science Research Group especially Simon Tew, Liz Laird, Ben McDermott, Danae Zamboulis, Colette Redmond, John Collins and Alan Mueller. In addition I would like to thank Duncan Robertson, Deb Simpson, Lynn McClean and Phil Brownridge for their training in all things mass spectrometry.

I would like to express my sincere gratitude to Richard Parkinson, M.D., and the staff at Clatterbridge Hospital, Wirral, Merseyside, for provision of human joint tissue and Sarah Powell for providing the young equine cartilage.

None of the work in this thesis would have been possible without generous funding and support from the Wellcome Trust in the form of an Integrated Veterinary Research fellowship.

Last but not least, my greatest thanks go to Andy, my husband who has had to devote himself to extra childcare duties in order that I could travel to some excellent conferences, work in the laboratory at weekends and write undisturbed at home!

**Abbreviations**

| | |
|---|---|
| **2-DE** | Two-dimensional electrophoresis |
| **ACN** | Acetonitrile |
| **ACSL** | Acly-synthetase long chain family member |
| **ACTB** | Beta-actin |
| **ADAM-TS** | A disintegrin and metalloproteinase with thrombospondin |
| **AGE** | Advanced glycation end products |
| **AMBIC** | Ammonium bicarbonate |
| **AMRT** | Accurate mass retention time |
| **ANOVA** | Analysis of variance |
| **AQUA** | Label for absolute quantification |
| **ATOMS** | Amino-terminal oriented mass spectrometry of substrates |
| **ATP** | Adenosine Tri-Phosphate |
| **bp** | Base pair |
| **CCV** | Clathrin-coated vesicle |
| **CID** | Collision-induced spectra |
| **COFRADIC** | Combined fractional diagonal chromatography |
| **COl2A1** | Collagen 2A1 |
| **COMP** | Cartilage oligomeric matrix protein |
| **CONSeQuence** | Consensus predictor for quantotypic peptide sequence |
| **CPII** | C-terminal collagen type II propeptide |
| **CS** | Chondroitin sulphate |
| **CsCl** | Caesium chloride |
| **CTX-II** | Type II collagen C-telopeptide fragments |
| **Da** | Dalton |
| **DA** | Discriminant analysis |
| **DAF** | decay accelerating factor splice variant |
| **DDA** | Data dependant analysis |
| **DF** | Discriminant function |
| **DGE** | Differntial gene expression |
| **DIGE** | Difference gel electrophoresis |
| **DKK1** | Dickkopf homolog |
| **DMEM** | Dulbecco's eagles medium |
| **DMMB** | 1,9-dimethyl-methylene blue |
| **DNA** | Deoxyribonucleic acid |
| **DTT** | Dithiothreitol |
| **ECM** | Extra cellular matrix |
| **EDTA** | Ethylenediaminetetraacetic acid |
| **EIC** | Extracted ion chromatogram |
| **ELISA** | Enzyme linked immunosorbent assay |
| **emPAI** | Exponentially modified protein abundance index' |
| **ESI** | Electrospray ionization |
| **FA** | Formic acid |

| | |
|---|---|
| **FACIT** | Fibril Associated collagens with Interrupted Triple Helices |
| **FCD** | Fixed charge density |
| **FCS** | Foetal calf serum |
| **FDR** | False discovery rate |
| **FT** | Fourier transform ion cyclotron |
| **G** | Globular region |
| **GAG** | Glycosaminoglycan |
| **GAPDH** | Glyceraldehyde-3-phosphate dehydrogenase |
| **GdnHCl** | Guanidine hydrochloride |
| **GO** | Gene ontology |
| **H and E** | Hematoxylin and Eosin |
| **HA** | Hyaluronic acid |
| **HAC** | Human articular chondrocytes |
| **HCCA** | α-Cyano-4-hydroxycinnamic acid |
| **His-tag** | Hexa histidine-tag |
| **HMBG** | High mobility box group |
| **HPLC** | High-performance liquid chromatography |
| **HRP** | Horse-radish peroxidase |
| **IAA** | Iodoacetamide |
| **ICAT** | Isotope-coded affinity tagging |
| **IGD** | Interglobular domain |
| **IGF** | Insulin growth factor |
| **IL** | Interleukin |
| **IMS** | Imaging mass spectrometry |
| **IPA** | Ingenuity Pathway Analysis |
| **ITO** | Indium tin oxide |
| **ITRAQ** | Isobaric tagging for relative and absolute quantification |
| **JNK** | Jun N-terminal Kinase |
| **KBD** | Kashin-Beck disease |
| **KS** | Keratan sulphate |
| **LASP** | LIM and SH3 domain protein |
| **LC-MS/MS** | Liquid chromatography-tandem mass spectrometry |
| ***m/z*** | Mass-to-charge ratio |
| **MALDI** | Matrix-assisted laser desorption/ionization |
| **MCPRED** | Miscleave predictor |
| **MGF** | Mascot generated file |
| **miRNA** | MicroRNA |
| **MMP** | Matrix metalloproteinase |
| **$M_r$** | Molecular mass |
| **MRI** | Magnetic resonanace imageing |
| **MRM** | Multiple reaction monitoring |
| **mRNA** | Messenger ribonucleic acid |
| **MS** | Mass spectrometry |
| **MSC** | Mesenchymal stem cell |

| | |
|---|---|
| **MT-MMP** | Membrane type MMP |
| **MW** | Molecular weight |
| **MWCO** | Molecular weight cut-off |
| **NaCl** | Sodium chloride |
| **NCC** | Non-cleaved control |
| **NCS** | Non-cleaved spanning |
| **NFκB** | Nuclear factor kappa-light-chain-enhancer of activated B cells |
| **NGS** | Next-generation sequencing |
| **NPII** | N-terminal collagen type II propeptide |
| **OA** | Osteoarthritis |
| **OD** | Optical density |
| **PA** | Plasminogen activator |
| **PCA** | Principle component analysis |
| **PCR** | Polymerase chain reaction |
| **PDGF** | Platelet-derived growth factor |
| **PG** | Proteogylcan |
| **PIICP** | Procollagen type II C-propeptide |
| **PKLR** | Pyruvate kinase |
| **PPI** | Protein-protein interaction |
| **ppm** | Parts per million |
| **PSAQ** | Protein standard absolute quantification |
| **QconCAT** | Quantification concatamers |
| **q-peptide** | peptide |
| **Q-TOF** | Quadrupole mass spectrometry |
| **r** | Pearson's correlation coefficient |
| **RA** | Rheumatoid arthritis |
| **RetA** | Retenoic acid A |
| **RhoA** | Ras homolog gene family member A |
| **RNA** | Ribonucleic acid |
| **RNA-Seq** | RNA-sequencing |
| **ROCK** | Rho-associated protein kinase |
| **ROI** | Region of interest |
| **ROS** | Reactive oxygen species |
| **RPKM** | Reads per kilo base of exon model per million mappable reads |
| **SDS-PAGE** | Sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| **SFRP** | Secreted frizzled-related protein |
| **SID** | Stable isotope dilution |
| **SILAC** | Stable isotope labelling with amino acids in cell culture |
| **SIMS** | Secondary ion mass spectrometry |
| **siRNA** | Silencing RNA |
| **SLRP** | Small leucine-rich repeat protein family of proteoglycans |
| **snoRNA** | Small nucleolar RNA |
| **SRM** | Selected reaction monitoring |
| **SVM** | Support vector machine |

| | |
|---|---|
| **TBST** | Tris-buffered saline and tween 20 |
| **TBT** | TATA box binding protein |
| **$TC^A$** | Quarter fragment from collagenase cleavage of type II collagen |
| **$TC^B$** | Three quarter fragment from collagenase cleavage of type II collagen |
| **TFA** | Trifluoroacetic acid |
| **TGF** | transforming growth factor |
| **TIMP** | Tissue inhibitors of metalloproteinase |
| **TNF** | Tumour necrosis factor |
| **TOF** | Time-of-flight |
| **TQ** | Triple quadrupole |
| **TSP** | Thrombospondin motifs |
| **TTO** | Terminal-Tagging Oligo |
| **WIF** | Wnt inhibitory factor |
| **Wnt** | Wingless and Int signalling pathway |

**Abstract**

It is hypothesised that there are distinct mechanisms involved in cartilage ageing and disease which can be determined using next-generation technologies including mass spectrometry and RNA-sequencing. The aims of this thesis were to firstly characterise molecular mechanisms associated with age-related and arthritis-related changes in cartilage gene and protein signatures. Secondly the thesis developed new techniques to identify novel cleavage sites in matrix proteins and to quantify some known proteolytic events in articular cartilage using mass spectrometry-based proteomics platforms. Finally the levels of key proteinases and their inhibitors involved in the pathogenesis of OA were measured using mass spectrometry.

Osteoarthritis (OA) is an extremely common cause of morbidity in both man and animals. OA involves the biomechanical failure of articular cartilage, together with changes in the subchondral bone and inflammation of the joints and leads to a variety of symptoms including pain, stiffness and reduced mobility. Age is an important factor in the development of OA and represents a huge challenge for society as whilst life span increases, the quality of life faced by an ageing population is often poor. Articular cartilage is susceptible to age-related diseases such as OA, but it is not an inevitable result of ageing and is a consequence of a complex inter-relationship between age and further predisposing factors. There have been major advances in technologies used to interrogate proteins and genes due to genome sequencing enabling gene and protein sequences to be determined. These 'next-generation technologies' include mass spectrometry (MS) and next-generation sequencing. This thesis has used these technologies in an attempt to address important questions relating to cartilage ageing and disease.

The use of an inflammatory model of early OA in equine and human cartilage enabled the discovery and quantification of important proteins and pathways involved, using relative and absolute mass spectrometry techniques. In the equine secretome pathway enrichment analyses confirmed the up-regulation of glycolytic proteins. The novel proteins clathrin and LIM and SH3 domain protein-1 were

identified for the first time in cartilage proteomics. QconCAT technology and gene expression analysis enabled normal and OA cartilage extract to be interrogated. Absolute quantification values were demonstrated for the first time for aggrecan; first and third globular domains, biglycan, cartilage oligomeric matrix protein, decorin and fibromodulin. Whilst a novel MS based technique enabled previously identified and novel extracellular matrix cleavage sites derived from matrix metalloproteinase 3 and a disintegrin and metalloproteinase with thrombospondin motifs 4 digestion of cartilage to be determined. Some of these sites of degradation were also evident in OA but not normal cartilage using matrix assisted laser desorption ionization imaging MS (MALDI-IMS). Tentative markers of OA and ageing cartilage were also demonstrated. Finally an RNA sequencing study on ageing equine cartilage found an age-related failure of matrix, anabolic and catabolic cartilage factors together with a reduction in Wnt signalling.

This thesis developed novel proteomic methodologies to identify and quantify distinct differences between cartilage ageing and disease. Several proteins not previously described in cartilage were identified. In addition many novel cartilage degradation products were identified and age-related peptides were visualised in cartilage for the first time.

## 1.0. General Remarks and Introduction

The Arthritis Research Campaign web site states that more than 6 million people in the UK suffer from osteoarthritis (OA) of the knee and 650,000 have osteoarthritis of the hip. More than 1 million adults consult their GP each year with OA (http://www.arc.org.uk). In the future OA is projected to rank second for women and fourth for men in the developed countries in terms of years lived with a disability (Lohmander, 2000). Osteoarthritis is an extremely common cause of morbidity in both man and animals. OA involves the biomechanical failure of articular cartilage, together with changes in the subchondral bone and inflammation of the joints and leads to a variety of symptoms including pain, stiffness and reduced mobility. Age is an important factor in the development of OA and represents a huge challenge for society as whilst life span increases, the quality of life faced by an ageing population is often poor (Baek et al., 2011). Articular cartilage is susceptible to age-related diseases such as OA, although it is not an inevitable result of ageing and is a consequence of a complex inter-relationship between age and further predisposing factors.

Recently there have been major advances in the technologies used to interrogate proteins and genes. These 'next-generation technologies' include mass spectrometry and next-generation sequencing and have become available due our knowledge of whole genome and proteome sequences. These technologies have enabled new pathways involved in disease to be discovered as they allow the discovery and quantification of hundreds to thousands of genes and proteins to be identified in a single sample. The careful use of these technologies will allow important questions to be addressed relating to cartilage ageing and disease.

One way to provide new insights into the development and treatment of OA is to obtain an understanding of how cartilage responds to pathological degradation and physiological remodelling evident in ageing. Throughout life there is constant turnover of cartilage matrix by both synthesis and degradation. A prominent feature of OA is loss from the tissue of cartilage matrix proteins, a process that occurs as a consequence of proteolytic fragmentation by proteolysis. Whilst a

number of cleavage sites have been identified by use of specific neoepitope monoclonal antibodies (Chang et al., 2011) quantified using immunoblotting techniques studies focus on a small number of cleavage products. In order to test our hypothesis that there are distinct mechanisms involved in cartilage ageing and disease the project aims were to;

1. Characterise molecular mechanisms associated with age-related and arthritis-related changes in cartilage gene and protein signatures.
2. Develop new techniques to identify novel cleavage sites in matrix proteins and to quantify some known proteolytic events in articular cartilage using mass spectrometry-based proteomics platforms.
3. Measure levels of key proteinases and their inhibitors involved in the pathogenesis of OA using mass spectrometry.

### 1.1. The structure and function of normal cartilage

Cartilage is a specialised connective tissue, which consist of cells; the chondrocyte and extracellular components; the extracellular matrix (ECM) (Poole et al., 2001). Unlike other adult connective tissues it does not contain blood vessels and nerves and therefore receives nutrients via synovial fluid and subchondral bone. Depending on the composition of the matrix, cartilage is classified into elastic, fibro-cartilage and hyaline cartilage. The gliding surfaces of synovial joints are covered with hyaline cartilage, also known as 'articular cartilage'. Hyaline cartilage provides a low-friction gliding surface, which compared to bone has increases compressive strength and resistance to wear under normal physiological conditions (Voorhees et al., 2011). The primary function of articular cartilage is load-bearing. In low friction articulation it acts as a shock absorber and minimizes peak pressure on subchondral bone.

During embryonic development cartilage arises from mesenchymal condensations. Mesenchymal cells aggregate to form a blastema, the cells of which begin to secrete cartilage matrix and are then called chondroblasts. Further development pushes the chondrocytes apart due to ECM production. The ECM consist of, ground

substance (hyaluronan, chondroitin sulphates keratin sulphate and other cartilage molecules) and tropocollagen, which polymerises extracellularly into collagen fibres. Then mesenchymal tissue surrounding the blastema gives rise to a membrane called the perichondrium. After growth has ceased there is no detectable cell division of chondrocytes in healthy adult articular cartilage (Muir, 1995).

In diarthrodial joints the hyaline cartilage faces the joint cavity on one side and connects to the subchondral bone on the other by a narrow layer of calcified cartilage tissue. A capsule encloses the entire joint and retains the synovial fluid (Figure 1.1).

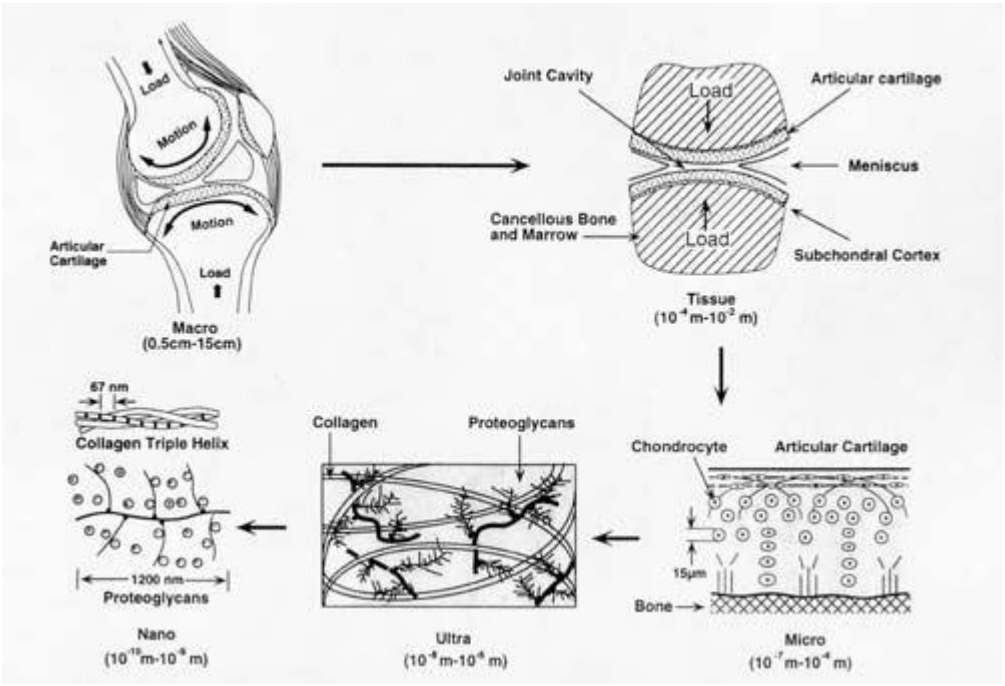



**Figure 1.1 Hierarchical structures of diarthrodial joints and articular cartilage** (Mow and Hayes, Basic Orthopaedic Biomechanics 1997). Clockwise from top left (following the arrows) illustrates the composite structure of diarthrodial joints. The next level indicates in more detail the actual bearing surface of the joint. The next level shows the existence of the structural features of articular cartilage including the chondrocytes and the organization of type II collagen fibrils.

## 1.2. Components of cartilage

### a. Chondrocyte

In mammals the chondrocytes fulfil two major roles. During development, most bones form through endochondral ossification in which bone is first laid down as cartilage precursors (Karsenty and Wagner, 2002). Whilst in the adult, chondrocytes are the sole cell type of articular cartilage and play crucial roles in joint function (Aigner et al., 2002).

Articular cartilage has the lowest cellular density of any tissue in the human body. In humans, chondrocytes contribute to only about 1% of the tissue volume and are situated in small cavities called lacunae within the cartilage and have an average size of 13µm. The spherical cells are found in a 'chondron', a structural unit comprising one or two chondrocytes and its pericellular microenvironment (Poole et al., 1988). Even in chondrons there is no cell-cell contact (Elfervig et al., 2001). Single cilia extend into the surrounding ECM (Buckwalter and Mankin, 1998). Recently it has been elucidated that the cilium act like switches, that when toggled by cyclical pulses of lacunocanalicular fluid or cartilage compression send signals (for instance by $Ca^{2+}$ signalling). This influxes into the cell to trigger a cascade of events that include appropriate gene activations to maintain and strengthen bone and cartilage (Whitfield, 2008)**.** The cells sense the structure and composition of the ECM (which they synthesise) and carry out their primary function which is to maintain it (Buckwalter and Mankin, 1998).

Irrespective of the size of a given animal, there is an inverse relationship between cell density and cartilage thickness. As cartilage is avascular, its nutrition depends on diffusion from outside and this may limit the total number of cells that can be sustained in a given volume (Stockwell, 1971). Moreover chondrocytes can exist under very low oxygen tensions and metabolise glucose primarily by glycolysis to produce lactate. This anaerobic pathway is maintained even under aerobic conditions (Buckwalter et al., 1994).

Articular chondrocytes have great longevity and normally live as long as their 'owners'. The metabolic state of the arrested cell division breaks down, however, whenever the integrity of the collagen network is compromised, as happens in the vicinity of lesions in OA. Here cell division appears to be reactivated, although any division is slow (Muir, 1995).

**b. Collagen**

Collagen accounts for two thirds of the dry weight of adult articular cartilage with the large aggregating proteoglycan aggrecan accounting for a large part of the remainder. The material strength of cartilage depends upon the extensive cross-linking of the collagen as well as the zonal changes in fibrillar architecture with tissue depth. Collagen concentration is highest at the surface where collagen is orientated parallel to the surface to resist shear forces. Currently, there are at least 28 members of the collagen super family, which function as structural components of the peri- and ECM in vertebrate tissue (Eyre et al., 2004). Articular cartilage contains at least eight collagens types; II, VI, IX, X, XI, XVI, XX and XXVII Types II, IX and XI form the characteristic basic architecture whilst the remainder are found in smaller amounts (Eyre, 2002). Type II collagen (a fibrillar collagen) constitutes 90-95% of collagen in the ECM and in association with type XI it forms a meshwork wherein type IX member of the collagen subgroup FACIT (Fibril Associated collagens with Interrupted Triple Helices) is covalently linked to the surfaces of the type II fibrils and further enables a cross-linked framework to aggrecan (Eyre, 1995). Thus collagen II, IX and XI consist as heterotypic copolymers. The non-fibrillating type VI forms elastic fibres and can be found pericellularly in middle zone and throughout the ECM in small amounts of up to 1% of overall collagen (Wu and Eyre, 1989). Furthermore it has been shown that Collagen VI interacts directly with the cell surface (Poole et al., 1988).

Collagen has a high level of structural organization and is represented as extended extracellular proteins composed of three polypeptide chains (α-chains), each possessing a characteristic tripeptide sequence (gly-x-y) that forms a left-handed helix. The three α-chains in each molecule are twisted tightly into a right-handed

helix to form a rope-like structure that is stabilised by hydrogen bonds. Glycine placed at every third residue of the tripeptide sequence, is small enough to occupy interior of the helix, while frequent other amino acids are proline and hydroxyproline. The collagen precursors, or procollagen are synthesized with large C- and N- terminal extensions which aid in chain assembly. These extension propeptides are cleaved by procollagen peptidases after secretion but before fibril formation. Furthermore collagen fibrils are stabilized by cross-links that involve lysine residues with fibrillar collagen, the biologically functional form, resulting from a series of post translational modifications (Muir, 1995).

Mature collagen fibres provide the capacity to withstand tensile and shear forces. Type II, specific to cartilage is often used as a marker of chondrocyte differentiation. Moreover the triple helix is composed of three identical alpha chains synthesized from the COL2A1 gene. Type II exists in two splice variants (IIA and IIB), in IIB, the dominant form found in mature cartilage, exon 2 is spliced out (encodes a 69 amino acid cystine-rich domain in the N-terminal propeptide). In IIA, a transient embryonic form found in prechondrogenic mesenchyme, perichondrium and vertebrae, this domain is retained (Sandell, 1994). Human OA chondrocytes also produce IIA (Zhu et al., 2001), which suggests a hypertrophic change in OA resembling the cartilage of a developing joint (Salminen et al., 2001).

Collagen is synthesized by chondrocytes as a precursor called procollagen. This consists of N and C terminal globular domains known as the N and C propeptides (NPII and CPII) which are attached to short linear sequences at each end; N and C telopeptides (Groutars et al., 2000) ( Figure 1.2). These telopeptides form covalent links with the triple helix. Hydroxylations and glycosylations, primarily O-linked, modify the protein prior to the formation of the triple helix. Winding of the chains in a C-N direction occurs due to disulphide bonds formed at the C-terminal region of three α chains. Glycine and hydroxyproline residues form hydrogen bonds within the helical centre, these stabilise the triple helix. The three α-chains in each molecule are twisted tightly into a right-handed helix to form a rope-like structure. The folded triple helix is secreted into the ECM. Tropocollagen is then formed

following the removal of the N and C propeptides by extracellular proteinases C and N. Aggregation of procollagens and fibril formation can then take place.

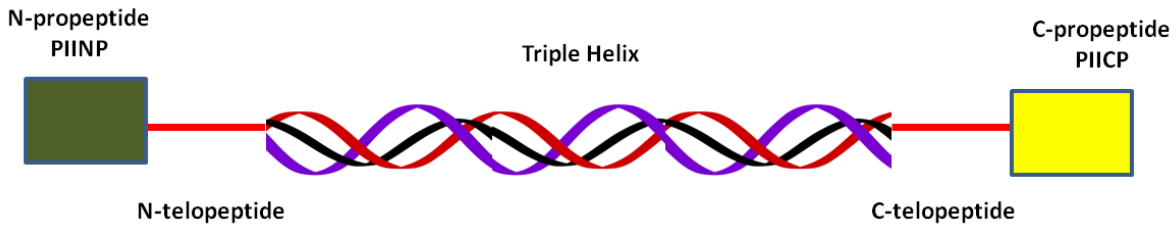



**Figure 1.2. Schematic representation of collagen secreted into cartilage ECM**

Propeptides may be released from the ECM into biological fluids, where their levels are believed to reflect type II collagen synthesis. Greatest concentrations of NPII and CPII are evident in foetal cartilage, three fold reductions are evident at birth which reduce further to very low levels in adults (3.9% of newborn). There is an increase in CPII levels in cartilage in OA, though this is not reflected in serum (Verzijl et al., 2000a). Interestingly by comparing sequence analyses of C-terminal regions, Van der Rest *et al*. 1986 (Van der Rest et al., 1986) demonstrated that the primary structure of chondrocalcin, a calcium-binding protein and CPII are identical. Thus CPII has functions in assembly of triple helical collagen II in cartrilage and cartilage calcification.

### c. Proteoglycans

Proteoglycans (PGs) are protein polysaccharide molecules that form 10-20% wet weight and provide a compressive strength to articular cartilage. There are two major classes of PG found in articular cartilage, large aggregating PG monomers or aggrecan and small proteoglycans including decorin, biglycan and fibromodulin (van den Bemt et al., 2000). The cartilage PG aggregate is a unique structure of macromolecules that, together with type II collagen and a number of minor accessory molecules, gives cartilage its specific biomechanical properties. Moreover, aggrecan is immobilized in the collagen network and the importance of

aggrecan in articular cartilage function is emphasized by altered metabolism and abnormal expression in animal models with arthritis (Hargrave et al., 2000).

### i. Aggrecan

The aggrecan family of proteoglycans contains aggrecan, versican, brevican and neurocan. All contain an amino-terminal domain capable of binding HA, a central domain containing covalently bound chondroitin sulphate (CS) chains and a carboxy-terminal containing C-type lectin domains. Recently it was demonstrated that aggrecan through the C-type lectin domain can interact with certain matrix proteins containing EGF-repeats. These include the fibrillins, fibulins and tenascin (Day et al., 2004). Aggrecan has a 220- to 250-kDa multiple domain protein core which is substituted with the glycosaminoglycans CS and keratan sulphate (KS) chains in addition to N- and O-linked oligosaccharides (Kiani et al., 2002). The core protein possesses two globular regions near the amino-terminus, known as G1 and G2, separated by an interglobular domain (IGD). A third globular region, G3, is found at the carboxy terminal end of the core protein, whilst an extended region containing KS and CS attachment sites is found between G2 and G3 domains, the CS1 and CS2 domains (Figure 1.3).

Extraction and purification techniques developed in the 1960's enabled a more detailed study of PGs. A novel extraction technique based on denaturing solvents such as 4M guanidine hydrochloride allowed most of the cartilage matrix molecules except cross-linked collagens to be extracted through the breakdown of all except covalent protein bonds. Further purification was achieved using caesium chloride gradient centrifugation (Sajdera and Hascall, 1969). This technique demonstrated that PGs form specific aggregates with a second protein, later identified as hyaluronan (HA) which together with link protein formed large aggregates (Hardingham and Muir, 1972). The binding of a ten sugar sequence of HA is mediated through G1 and stabilised by link protein (a 45- to 50-kDa glycoprotein) which joins HA and G1. HA is a polysaccharide having repeating disaccharide structure. Furthermore receptors on the surface of chondrocytes to HA function to provide a gel to which chondrocytes attached (Hardingham, 1981).

The GAG chains impart many of the physical properties to the PG. CS is composed of repeating disaccharide units of glucuronic acid and galactosamine, with a sulphate group per disaccharide. KS consist of repeating disaccharide units of glucosamine and galactose, also averaging a sulphate group per disaccharide. The sulphate and carboxy groups on the CS and KS chains become charged in solution and in-situ. The total fixed charge density (FCD) in cartilage ranges from 0.05 to 0.3 mEq/g wet weight of tissue (Mow et al., 1999) and it is this FCD that is responsible for the high Donnan equilibrium ion distribution in the interstitium. It is the Donnan osmotic pressure measured in cartilage (Lai et al., 1991) which contributes to the overall compressive stiffness of cartilage. The fixed negative charges of the PG serve to maintain a high degree of hydration in articular cartilage by generating a substantial osmotic pressure within the tissue (Hopewell and Urban, 2003). This explains why cartilage has a tendency to swell, but this is resisted by the collagen network, which is therefore under constant tension, even when unloaded. High transient loads are accommodated by changes in osmotic and hydrostatic pressure when fluid is forced from loaded to unloaded areas, while aggrecan remains immobilised within the collagen network provided it is intact and bound to the hyaluronan.

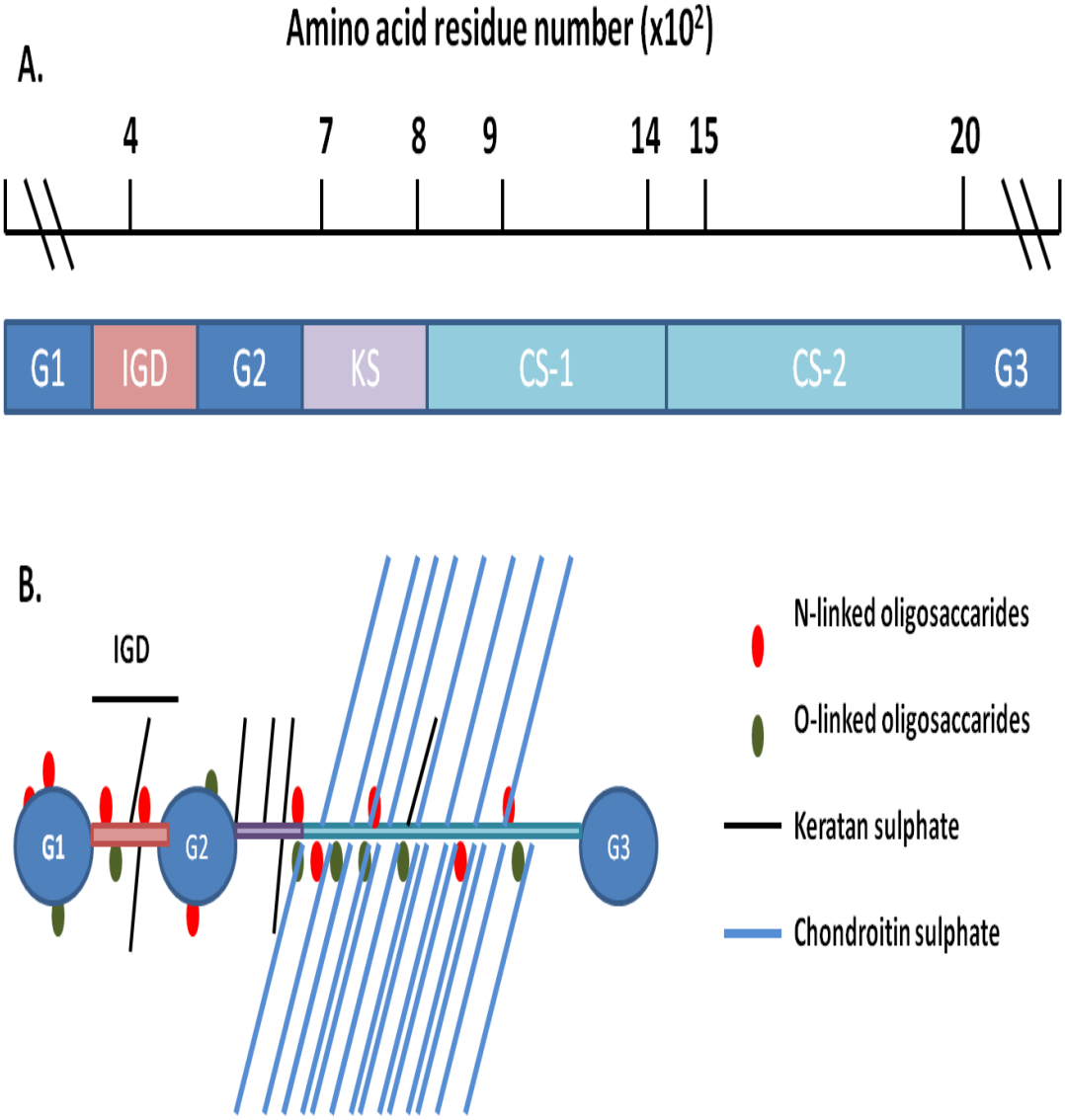



**Figure 1.3. Schematic representation of domain organisation and structure of cartilage aggrecan.** A. Diagram of the domains and GAGs within aggrecan; G1, G2, G3; globular regions 1, 2, 3, IGD; interglobular domain, KS; keratan sulphate-rich region, CS-1, -2; chondroitin sulphate-rich regions. Amino acid residue number line is not to scale. B. The post-translational modifications on aggrecan are shown.

### ii. **Small Leucine-Rich Proteoglycans**

Members of the small leucine-rich repeat protein family of proteoglycans (SLRP) include fibromodulin along with decorin, biglycan and lumican. SLRPs modulate tissue organization, cellular proliferation, matrix adhesion, growth factor and cytokine responses, and sterically protect the surface of collagen type I and II fibrils from proteolysis by matrix metalloproteinases (MMPs) (Geng et al., 2006). They are characterised by 40kda core proteins containing 6-10 leucine-rich repeats flanked by cysteine rich domains. The presence of these repeats accounts for their horse-shoe like conformation which allows their functional ability to interact with other proteins in the ECM. Furthermore their structure contains CS (decorin and biglycan), dermatan sulphate or KS (fibromodulin and lumican) side chains. Decorin has a role in the regulation of collagen fibrillogenesis and interacts with fibrils of type I and II collagen (Reed and Iozzo, 2002) as well as acting as a reservoir for transforming growth factor β (TGF-β) in the ECM (Finnson et al., 2012). Biglycan has similar functions but interacts with type VI collagen aiding in the formation of its network (Wiberg et al., 2003).

The structure of decorin and biglycan can vary with age, indeed decorin synthesis increases with age (when foetal and adult cartilage are compared) (McAlinden et al., 2001). Structural changes are most evident in biglycan, where in aged ECM there appears to be a cleavage in the amino terminal domain resulting in a 'no-glycan' biglycan as the terminal peptide containing the GAG chain separates from the protein core. It has recently been demonstrated *in-vitro* that antibodies to the no-glycan biglycan have a role in rheumatoid arthritis (RA) (Antipova and Orgel, 2012). Further proteolytic modification of the core proteins in decorin and biglycan occurs in OA (Roughley et al., 1993). Work by Melching et.al. 2006 (Melching et al., 2006) found that biglycan was a substrate for ADAMTS-4 and -5 at a site within the fifth leucine rich region at an asparginine-cysteine bond. The group also demonstrated that human articular chondrocytes (HAC) from OA and RA joints contained the cleaved product within the ECM. In addition biglycan is also susceptible to cleavage by MMP-13 (at $^{177}$G -$^{178}$ V in human). This may represent an

early critical event in OA as it could result in collagen cleavage sites being exposed to MMP-13. Decorin can also undergo proteolysis by MMPs (Monfort et al., 2006).

Fibromodulin is a 59-kDa collagen-binding protein (Heathfield et al., 2004) which is tyrosine sulphated in its N-terminal extension (Antonsson et al., 1991). Binding of fibromodulin to both type I and type II collagen *in-vitro* has been demonstrated and provides an important role in regulating assembly of collagen fibrils (Hedbom and Heinegard, 1989). The binding appears to be independent of disulphide bridging and includes a region in the C terminus of fibromodulin (Font et al., 1998). Interestingly fibromodulin has been demonstrated to bind to the same site on the collagen fibril as lumican, whereas decorin appears to bind to a separate site (Hedbom and Heinegard, 1993; Svensson et al., 2000). Using neoepitope antibodies, cleavage products of fibromodulin during IL-1 stimulated cartilage explant studies *in-vitro* were found to be identical to those obtained following MMP-13 stimulated degradation (Monfort et al., 2006). ADAMTS-4 has also been identified as having activity against fibromodulin at the same site of cleavage as MMP-13 (Kashiwagi et al., 2004). As the observed cleavage occurs between the N-terminal negatively charged domain of fibromodulin and the region binding to collagen, the main fragment is retained for longer in the tissue, apparently by its binding to collagen. Degraded fragments of the core protein have been observed in OA cartilage (Cs-Szabo et al., 1995) and with age (Roughley et al., 1996) and removal of this portion of fibromodulin results in weaker interactions of a particular collagen fibre to surrounding structures.

### d. Cartilage Oligomeric Matrix Protein

Cartilage oligomeric matrix protein (COMP) is one of the five members of the thrombospondin matricellular (TSP) protein family. TSPs are extracellular calcium-binding proteins that function as adapter molecules to guide ECM and tissue organisation. COMP, also known as thrombospondin 5 is a 524kDa homopentameric, extracellular matrix glycoprotein with five identical subunits (Hedbom et al., 1992), enabling it to bind five collagen molecules simultaneously. The carboxy terminal globular domain binds to collagens I, II, and IX and fibronectin.

COMP also interacts with other ECM proteins including matrilin-3 and aggrecan. COMP can influence the fibril formation of collagens I and II by promoting early association of collagen molecules, thereby accelerating fibrillogenesis with a distinct organisation of the fibrils (Halasz et al., 2007). Thus it acts to bring the collagen fibres together in early fibril formation. A consequence of this is that if concentrations of COMP are elevated relative to collagen this will result in the occupancy of most binding sites on collagen by a single COMP, causing a reduction in collagen fibril formation. This is evident in OA when COMP production and presence is high, but collagen production is diminished (Chen et al., 2007).

COMP is considered a marker of cartilage breakdown, and has been studied as a biological marker (Tseng et al., 2009). Measurement of intact COMP and fragments thereof in synovial fluid or serum correlates to cartilage destruction in RA and OA patient studies (Mansson et al., 1995; Saxne and Heinegard, 1992). Interestingly COMP levels can be detected at ten times higher in synovial fluid than in serum indicating preferential release from the affected joints (Tseng et al., 2009). In synovial fluid COMP is present to some extent as an apparently intact protein, however the majority is found as several different fragments (Neidhart et al., 1997). Elevated synovial fluid COMP concentration in the joints of racehorses has been demonstrated as a useful marker for carpal joint osteochondral fragments (Abiola et al., 1990).

### 1.3. Organisation of the matrix

The thickness of normal articular cartilage is dependent on the joint, region within the joint and species. In the human medial femoral head it measures 2 to 3 mm thick, whilst on the patella it can be up to 5mm thick. In the horse the metacarpophalangeal joint it is 762 $\pm$ 131µm thick (Brommer et al., 2005) and is 1.5-2mm in the stifle joint (Frisbie et al., 2006). The organisation of the ECM and its distribution into zones differs between immature and mature cartilage; being thicker and less stratified in immature tissue. Chondrocytes are distributed in a random fashion and as the tissue matures the matrix becomes arranged in defined

zones, together with an increase in mechanical competence. The zones of articular cartilage are superficial (tangential), the middle (transitional), the deep (radial) and calcified cartilage. Where different zones are isolated and cultured, differences in terms of morphology, metabolism, phenotypic stability and responsiveness to signalling molecules, such as interleukin -1 (IL-1) are evident (Homandberg et al., 1992). In man the superficial zone is approximately 200µm thick and includes collagen fibr es tangential to the articular surface, with transition to more randomly orientated fibres in deeper regions. The parallel arrangement of fibrils is responsible for providing the great tensile and shear strength. Additionally chondrocytes in this zone (which are flattened ellipsoid) synthesize high concentrations of collagen and low concentrations of PG. The middle zone is approximately 1mm thick with randomly orientated collagen fibres. Collagen fibres are orientated perpendicular to the joint surface in the deep zone, which is approximately 600µm thick and contains the highest concentration of PG with the lowest cell density. It is at this point that a smoothly undulating tidemark separates the deep zone from the calcified cartilage. This is characterised by rounded chondrocytes arranged in columns, a high PG content and a radial collagen network (Figure 1.4).

## 1.4. Pathological perturbations

### 1.4.1. Osteoarthritis

Osteoarthritis (OA), the most common form of arthritis, is a chronic degenerative disease that affects diarthrodial joints. The disease, characterized by progressive destruction of articular cartilage, affects the entire joint, including the synovial membrane, joint capsule, ligaments, peri-articular muscles and tendons and subchondral bone (Altman et al., 1986). Additionally, there is a suggested role for subchondral bone adaptation in traumatic overload arthrosis in the racehorse (Barr et al., 2009) and this is in agreement with proposals that one of the mechanisms of initiation of joint failure may be steep stiffness gradient in the underlying subchondral bone (Radin and Rose, 1986).

Primary OA is a chronic degenerative disorder related to but not caused by ageing and is characterised by its late onset and no obvious cause. Secondary OA has an earlier onset and an identifiable cause such as injury or a developmental abnormality (Goldring and Goldring, 2007). Clinical manifestations include pain, stiffness and impairment of joint motion. There are, as yet no recognisable disease modifying treatments for OA, with most current treatments being entirely symptomatic.

The loss of the normal structure and function of articular cartilage is fundamental in OA. For instance there is shift in the collagen composition in OA cartilage with increased concentrations of collagen types I and III, which are considered as 'non cartilage' collagens (Aigner et al., 1993). In addition the expression of type X collagen, which is normally present at sites near cartilage mineralization containing hypertrophic chondrocytes (Grant et al., 1985), is increased at osteophytes and in fibrillated cartilage matrix (Hoyland et al., 1991). Major shifts are found in the turnover and structure of aggrecan with incorporation rate of aggrecan decreasing (Bulstra et al., 1989). Interestingly in experimental and human OA cartilage there is an increased release of both pre-existing and newly synthesized aggrecan (Inerot et al., 1991).

Following damage to cartilage an initial increase in water content is evident. Following disruption of the collagen the negatively charged glycosaminoglycans (GAGs) attract water resulting in swelling. The surface of the cartilage becomes fibrillated resulting in a reduction in resistance to shear forces. As OA progresses fibrillations become deep, vertical clefts, due to collagen orientation in the deeper layers, resulting in further ECM disruption (Libicher et al., 2005).

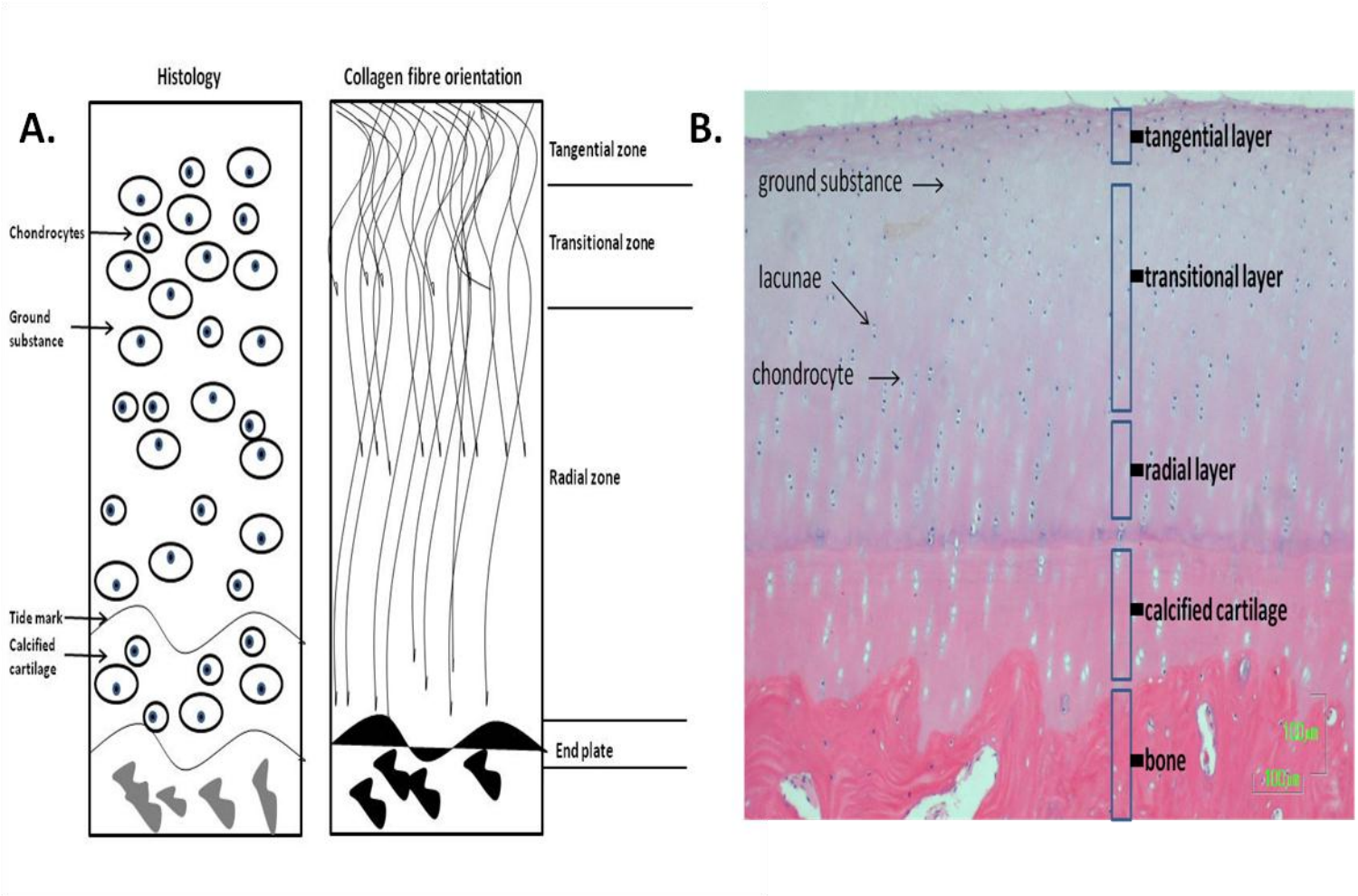



**Figure 1.4. Organisation of cartilage ECM.** A. Schematic representation of the zones within articular cartilage. B. Histological slide of equine cartilage from the metacarpus stained with haematoxylin and eosin. Cartilage components and zones are marked.

### 1.4.2. Osteoarthritis in equines

OA is a common cause of lameness with one study stating that it makes up 54% of lame horses (Todhunter and Lust, 1992) and it is estimated to account for 68% of lost training days in racehorses (Rossdale et al., 1985). Single event trauma or more insidious damage, has been suggested to be a common etiological factor in the occurrence of OA, with an increased incidence of OA in young horses used for athletic activity reported (Hoffman et al., 1984). This is due to prolonged 'wear and tear' from athletic training (Kidd et al., 2001) similar to that described in ageing (Aigner et al., 2004b). However a study of wild ponies suggested that OA in horses can also occur spontaneously (Cantley et al., 1999). Whilst joint loading is required in healthy joints to stimulate ECM protein production, excessive loading is harmful, suggesting that there is an injurious threshold of joint loading (Maly, 2008). Clinically OA in the horse is frequently seen as a slow progressive frequently bilateral lameness. Clinical signs are similar to those in man with commonly affected joints including the metacarpophalangeal, carpal joints, distal intertarsal and tarsometatarsal joints, and proximal interphalangeal joints (McIlwraith, 2012).

### 1.4.3. Cartilage degradation

Studies of the fragmentation of the ECM constituents have determined a series of cartilage protein degradations. Models were factors such as cytokines induce the breakdown of cartilage explants work by inducing a series of proteolytic enzymes. Initially aggrecan is fragmented and released followed by other molecules such as COMP, fibromodulin and collagens (Heathfield et al., 2004). The mechanisms by which the major components of cartilage ECM are degraded in OA have received a large amount of research. The mechanisms involved include both non-proteolytic and proteolytic. In the former, degradation by free radicals produced by neutrophils has been proposed (Roberts et al., 1987). However far more important are a number of specific proteinases which originate from the chondrocyte or the cells infiltrating inflamed synovium (as in RA and primary synovitis). There is evidence for the sequential degradation of matrix molecules by multiple proteinases. There are four distinct types of proteolytic enzymes which are classified according to their catalytic mechanisms. These groups are termed cysteine, aspartic, serine and

metalloproteinases, the former three will have a cysteine, aspartic acid or serine respectively within their active site (Burrage et al., 2006). Interestingly matrix fragments themselves can induce proteases. For example Johnson *et al.* 2004 (Johnson et al., 2004) demonstrated that upon stimulation with fibronectin fragments COMP was released from equine cartilage explants.

#### 1.4.3.1. Cysteine, aspartic and serine proteinases

The cathepsin papain-family of cysteine protease are distinguished by their structure (a 25kDa catalytic unit), catalytic mechanism, and which proteins they cleave. Cysteine proteinases include cathepsins B, H, K, L and S. Although cathepsin B and D work ideally at an acid pH, both are able to cleave aggrecan at pH as high as 6.5 (Handley et al., 2001). Cathepsin B has been implicated in various extracellular degradative processes (Mort and Buttle, 1997) and has been identified to be elevated in both RA and OA patients (Mort et al., 1984). Mature cathepsin B has the ability to degrade several ECM components including aggrecan (Fosang et al., 1992) and has been implicated in the progression of arthritis (Tortorella et al., 2000b). Studies in bovine cartilage have identified two cathepsin cleavage sites in link protein (LP) and five in aggrecan (Hou et al., 2003). Cathepsin K has activity against aggrecan, link protein and collagen. In addition the collagenolytic activity allows the cleavage of the triple helix of type I and II collagen, though at different sites to MMP's (Verzijl et al., 2002). In both OA and increasing age there is an increase in the neoepitope produced by the action of cathepsin K on collagen (Dejica et al., 2008).

Cathepsin D is a lysosomal aspartic protease of the pepsin family. Early studies in OA indicated that Cathepsin D enzyme activity was present in two to three times greater amounts in articular cartilage from patients with OA than in normal human cartilages (Sapolsky et al., 1973). Experiments using bovine cartilage have demonstrated that aggrecan digested with cathepsin D at pH 5.2 results in aggrecan core protein fragments due to cleavage at 5 sites within the core protein (Handley et al., 2001).

Members of the serine protease family include trypsin, elastase, chymotrypsin and plasminogen. The plasminogen activator (PA) system is a general proteolytic enzyme system suggested to play an important role in the degradation of the ECM during the development of RA (Belcher et al., 1996). Plasminogen is the zymogen of the serine protease plasmin. However, plasmin can directly degrade ECM components including the glycoproteins fibronectin, laminin, and elastin, as well as proteoglycans (Kwaan, 1992). Furthermore plasmin indirectly contributes to matrix degradation by proteolytically activating MMPs and proteoglycanases (Williams et al., 2010).

#### 1.4.3.2. Metalloproteinase Enzymes

##### a. Matrix metalloproteinases

Two families of metalloproteinases are capable of degrading cartilage specific collagen and aggrecan. These are the MMPs and the disintegrin and metalloproteinase with thrombospondin motifs (ADAMTs). Many metalloproteinases have been identified as being over-expressed in OA (Struglics et al., 2006). Using proteomics, metalloproteinase cleavage products in HAC have been identified including a wide variety of peptide cleavage sites from matrix components collagens I, II, III, biglycan, fibromodulin, prolargin, fibronectin, decorin, COMP and aggrecan (Zhen et al., 2008).

The MMPs form a multigene family and can be classified into subfamilies on the basis of domain structure and substrate selectivity. There are currently 23 identified MMPs (Kevorkian et al., 2004) which are categorized into the following groups: collagenases, gelatinases, stromelysins, matrilysins, metalloelastases, and membrane-type matrix metalloproteinases (MT-MMPs). The MMPs are considered the one of the main enzymes responsible for degradation of aggrecan and collagens in cartilage (Okada and Hashimoto, 2001). The expression of several MMPs is elevated in cartilage and synovial tissues of patients with OA (Okada and Hashimoto, 2001; Tetlow et al., 2001). Those over-expressed in cartilage are key enzymes in the development of OA (Dean et al., 1989). In OA inflammatory

cytokines such as IL-1β and tumor necrosis factor-alpha (TNFα) stimulate the production of MMPs which degrade many components of the ECM.

All MMPs are expressed as pro-proteins, whilst most share a conserved domain structure of pro-domain, catalytic domain, hinge region and hemopexin domain. Furthermore all are synthesized with a signal peptide, which is cleaved during transport through the secretory pathway. The common classification of MMPs is based on substrate specificity and cellular location (Figure 1.5).

| Matrix metalloproteinase Group | Members |
|---|---|
| Collagenases | MMP-1, MMP-8, MMP-13 |
| Matrilysin | MMP-7, MMP-26 |
| Metalloelastase | MMP-12 |
| Gelatinases | MMP-2, MMP-9 |
| Enamelysin | MMP-20 |
| Stromolysins | MMP-3, MMP-10, MMP-11 |
| Membrane-type MMPs | MMP-14, MMP-15, MMP-16, MMP-17, MMP-24, MMP-25 |
| Other | MMP-19, MMP-21, MMP-23A, MMP-23B, MMP-27, MMP-28 |
| Inhibitors | TIMP1, TIMP2, TIMP3, TIMP4, α-2 macroglobulin |

**Figure 1.5. Table of the classification of MMPs.**

### i. Collagenases

The collagenases, which include MMP-1, MMP-8, and MMP-13, have the ability to initiate cleavage of triple helical collagens I, II, and III at neutral pH (Billinghurst et al., 1997). The site of cleavage of these fibrillar collagens has been shown to be at a single locus, a $^{775}$Gly–$^{776}$Ile/Leu bond in the collagen α chains, three-quarters of the distance from the amino terminal end of each chain, resulting in the distinctive ¾ ($TC^A$) and ¼ ($TC^B$) fragments (Miller et al., 1976; Mitchell et al., 1996). In cartilage MMP-13 preferentially cleaves type II collagen. Damage to the fibrillar network in cartilage, 90-95% of which is made up of type II collagen, is considered the critical event in OA due to the slow rate of collagen turnover in cartilage (McAnulty, 1990). The fragments produced spontaneously unwind (denature) at physiologic temperature, making them susceptible to further degradation by the gelatinases; MMP-2 and MMP-9 as well as further degradation by the collagenases and proteinases. However, cleavage of the collagen fibril itself requires removal of small proteoglycans and cleavage of interfibrillar cross-links in order for the collagenases to access triple helical regions. Additional elements in the gelatinases (the fibronectin-type-II repeats and the hemopexin domain) assist proteolysis by binding to the substrate and also enable the enzyme to attach to other components of the connective tissue matrix (Steffensen et al., 1998).

Collagenases have been detected in synovial fluid of patients with OA (Lohmander et al., 1993a) and rheumatoid arthritis (RA) (Clark et al., 1993). Interestingly IL-1 has been demonstrated to increase the expression of MMP-8 and MMP-13 in normal human articular chondrocytes (Chubinskaya et al., 1996; Reboul et al., 1996).

### ii. Gelatinases

Gelatinases are composed of the 72-kDa MMP-2 and the 92-kDa MMP-9. Both have fibronectin type II repeats, which mediate binding to collagens, inserted into the catalytic domain (Page-McCaw et al., 2007). These enzymes are known to cleave native type IV, V, VII, and X collagens and elastin, as well as the products of collagens types I, II and III after proteolysis by collagenases. MMP-2, reported to be

expressed in chondrocytes (Goessler et al., 2005) is capable of cleaving soluble triple helical type I collagen (Aimes and Quigley, 1995). MMP-9 mRNA expression is elevated in OA chondrocytes (Tsuchiya et al., 1996) and increased expression of MMP-9 precedes fibrillation of cartilage in OA (Tsuchiya et al., 1997). Thus MMP-2 and MMP-9 may have a role in the degradation of various matrix components.

### iii. Stromelysins

The stromelysin group is composed of MMP-3, MMP-10, and MMP-11. MMP-3 and MMP-10 have an identical broad spectrum of activity, a broad optimum pH range however MMP-3 is more potent. Their substrates include proteoglycan core protein, laminin, fibronectin, elastin, as well as non-helical regions of collagens (Nagase and Woessner, 1999). MMP-3 (stromelysin 1) is one of the most highly expressed proteases in cartilage and is significantly decreased in expression in OA in a number of studies (Davidson et al., 2006; Swingler et al., 2009) as well as in RA (So et al., 1999). The function of MMP-3 in cartilage homeostasis is not certain, although it is capable of degrading aggrecan and also of activating procollagenases. It is possible that MMP-3 has a maintenance function in cartilage that is lost in end-stage OA (Swingler et al., 2009).

### iv. Matrilysins and metalloelastases

The sole member of the matrilysin group is MMP-7, has greater activity than the other MMPs against versican, a chondroitin sulphate proteoglycan. Ohta *et al*. 1998 (Ohta et al., 1998) have reported over-expression of MMP-7 in human OA cartilage chondrocytes and suggested that cytokine-induced MMP-7 may play an important role in the degradation of ECM in OA cartilage.

MMP-12 is a metalloelastase with activity against elastin. It is also capable of degrading proteoglycan (Janusz et al., 1999), fibronectin, laminin, vitronectin, type IV collagen and heparin sulphate (Jormsjo et al., 2000).

### v. Membrane-type MMPs

The membrane type MMPs (MT-MMPs) are similar in structure to the soluble MMPs. Four different membrane-type MMPs have been identified MT-MMP 1 (MMP-14), MT-MMP 2 (MMP-15), MT-MMP 3 (MMP-16), and MT-MMP 4 (MMP-17)). MT-MMPs play a dual role in cell surface proteolysis. Firstly, they cleave a variety of ECM components *in-vitro*. These include gelatin, fibronectin, and laminin, vitronectin and dermatan sulphate proteoglycan (Kajita et al., 2001). Secondly, they are initiators of activation of MMP-2 (Murphy et al., 1999).

The expression of MMP-16 has demonstrated variable results. In one study of human OA cartilage, quantitative gene analysis revealed that there was no significant difference between OA and normal samples (Yamanaka et al., 2000). Work undertaken in equine OA cartilage suggested that MMP-16 may not have a role in matrix destruction in equine cartilage diseases (Garvican et al., 2008). However, in studies of naturally occurring human OA, MMP-16 was the only MT-MMP found to be significantly increased in OA cartilage and synovium compared with normal cartilage (Davidson et al., 2006; Kevorkian et al., 2004). Furthermore in a recent study also in human OA cartilage MMP-16 mRNA was found to be significantly increased in OA compared to normal cartilage (Swingler et al., 2009).

#### 1.4.3.3. Tissue inhibitors of metalloproteinases

MMP activity is also regulated by a family of tissue-specific inhibitors, of which there are four known tissue inhibitors of metalloproteinases (TIMPs); TIMP-1, TIMP-2, TIMP-3, and TIMP-4 (English et al., 2006). In addition α2-macroglobulin synthesised in the liver and by macrophages can inhibit the MMPs and the ADAMs (Baker et al., 2002). Besides their inhibitory role, TIMPs have other functions such as growth factor-like and anti-angiogenic activity (Gomez et al., 1997). The TIMPs are secreted by a variety of cells including chondrocytes and macrophages and their activity is increased by platelet-derived growth factor (PDGF) and TGF-β and either increased or decreased by different interleukins (Fabunmi et al., 1998). Disruption of this MMP–TIMP balance can result in disorders such as OA.

TIMP concentrations generally far exceed the concentration of MMPs in tissue and extracellular fluids, thereby limiting their proteolytic activity to focal pericellular sites. The four TIMPs act as a further level of extracellular regulation of MMP activity whilst also possessing specific patterns of gene regulation and tissue-specific expression (Folgueras et al., 2004). TIMP-3 appears to be the most potent inhibitor of ADAMTS-4 and ADAMTS-5 (Mengshol et al., 2002). It also appears to inhibit aggrecan degradation in cultured articular chondrocytes stimulated with IL-1 (Nagase and Brew, 2003). Interestingly TIMP-1 may have some inhibitory capacity against ADAMTS-5 (Tortorella et al., 1999). TIMP-4 is a weak inhibitor of ADAMTS-4 but not ADAMTS-5 (Mengshol et al., 2002). Both TIMP-2 and TIMP-3 are reported to inhibit ADAMTS-1 (Rodriguez-Manzaneque et al., 2002). Presently it is believed that the local balance of metalloproteinase and TIMP activities is crucial for cartilage homeostasis, with disturbances producing higher levels of MMPs over TIMPs resulting in pathological changes.

#### 1.4.3.4. Matrix metalloproteinase driven cartilage degradation

**a. Collagen degradation**

Fibrillar collagen in its native state is resistant to breakdown by proteolytic enzymes. However, it may be damaged by helical cleavage resulting in denaturation or by telopeptide cleavage, which can lead to the removal of cross-links and depolymerisation of the fibrillar network (Barrett, 1978). At neutral pH the breakdown of the helices may occur as a consequence of MMP-13 and to a lesser extent MMP-1 and MMP-8 (Knauper et al., 1996) at a specific site Gly-Leu/Ileu bond generating the characteristic ¾ and ¼ fragments. These fragments may then be degraded further by MMP-1 or by enzymes with gelanotolytic activity such as MMP-2, MMP-9, neutrophil elastase and plasmin at neutral pH and at acidic pH by cathepsin B, S, L (Kafienah et al., 1998). Additionally it has also been demonstrated that cathepsin K cleaves collagen II in the N-terminal region of the helical domain (Kafienah et al., 1998).

### b. Aggrecan degradation

Under pathological conditions aggrecan degradation occurs as an early event and is mediated by proinflammatory cytokines within the joint (Hubbard et al., 2000). Aggrecan degradation products are derived predominantly from two protease families the aggrecanases and the MMPs. Although aggrecan fragments found *in-vivo* are predominantly the products of aggrecanase activity, the MMPs are also involved (Fosang et al., 1996b; Struglics et al., 2006). Some studies have indicated a lack of MMP involvement in aggrecan degradation in short term studies of explants with cytokines (Durigova et al., 2008) others have demonstrated that they may be involved in later stages of OA (van Meurs et al., 1999). Indeed, Durigova *et al.* 2011 (Durigova et al., 2011) demonstrated MMP mediated aggrecan degradation within the IGD was only evident after day 12 of bovine cartilage explant culture.

Many *in-vitro* cartilage studies using agents such as IL-1, retinoic acid and TNF-α have been identified as promoting cartilage degradation (Ismaiel et al., 1992; MacDonald et al., 1992). These agents have been identified as causing the up-regulation of some MMPs in chondrocytes which suggested a link between MMP expression and cartilage degradation. Additionally specific MMP inhibitors prevent loss of aggrecan from cartilage explant cultures *in-vivo* (Cawston et al., 1999; Mort et al., 1993).

### b. Aggrecanases

The ADAMTSs family is made up of 19 proteases which are expressed in a wide range of tissues, but are more restricted in foetal tissues. The ADAMTS family are multi-domain metalloproteases that are secreted into the extracellular space as furin-active proteases (Bondeson et al., 2008). They are composed of a signal sequence, prodomain, catalytic domain, disintegrin-like domain, spacer region, thrombospondin motifs (TSP) and submotifs (Figure 1.6).

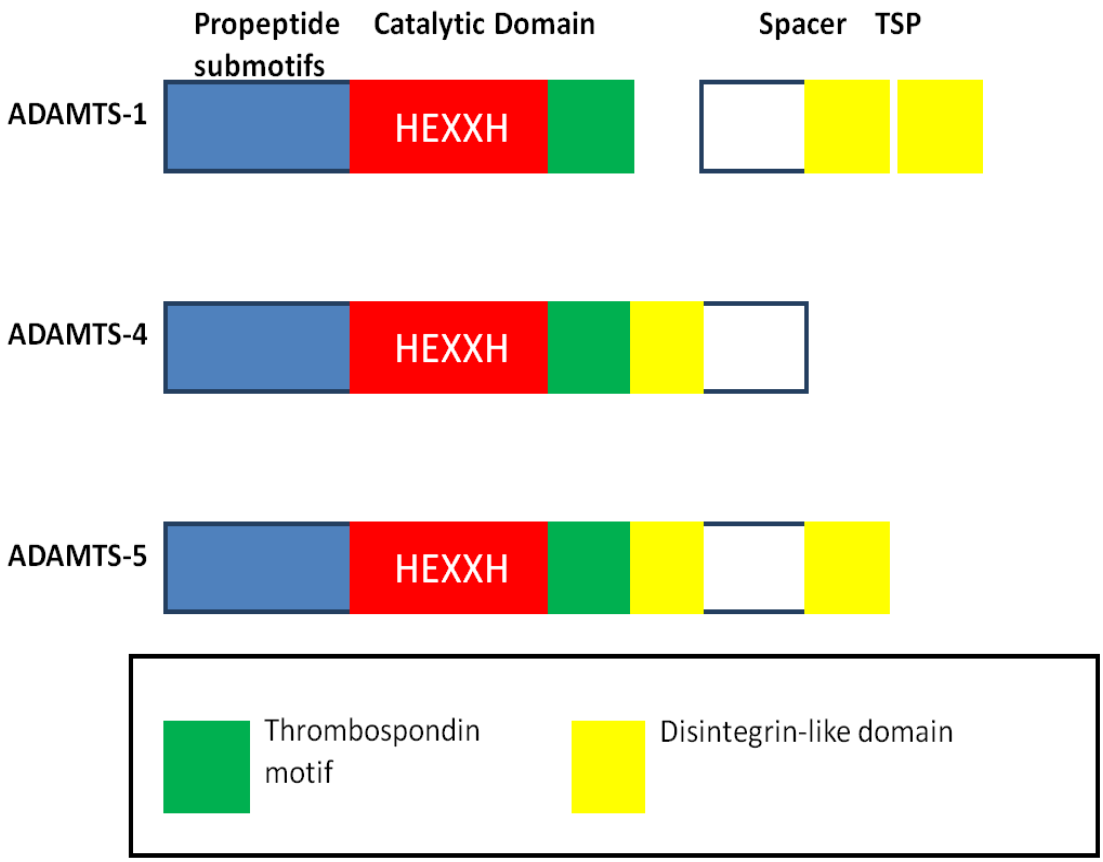



**Figure 1.6. Schematic representation of the domain structure of ADAMTS-1, -4, and -5 involved in aggrecan degradation.**

The latter regulate both their activity and substrate specificity (Bondeson et al., 2008). A subgroup of the enzymes, ADAMTS-1, 4, 5, 7, 8, 9, 12 and 15 have roles in the pathogenesis of arthritis with the ability to degrade aggrecan within the cartilage ECM (Hoch et al., 2011). Expression of ADAMTS-1, -4, -5, -9, and -15 has been found in normal HAC (Bau et al., 2002; Bondeson et al., 2008; Wachsmuth et al., 2004). However studies have demonstrated conflicting information on expression changes of these potential aggrecanases in OA compared with normal human cartilage. Both increased (Bau et al., 2002; Wachsmuth et al., 2004) and decreased (Kevorkian et al., 2004) levels being reported.

ADAMTS-4 and ADAMTS-5 are the pertinent enzymes in the pathogenesis of OA as demonstrated by their high *in-vitro* activity for aggrecan cleavage, expression in OA cartilage and localized expression in areas of aggrecan degradation (Fosang et al., 2008). Work in human chondrocytes and explants revealed that both ADAMTS-4

and -5 were important mediators of cytokine stimulated aggrecan loss in normal and OA cartilage. It was hypothesised that this may be due to a greater induction of ADAMTS-4 in man in the presence of catabolic stimuli (Song et al., 2007). Indeed it is still not clear if ADAMTS-4 or -5 is the major aggrecanases in man as ADAMTS-5 is constitutively expressed in human chondrocytes and synovial fibroblasts, were as previously mentioned in cytokine induced degradation ADAMTS-4 expression is predominant (Fosang et al., 2008). A recent study using equine cartilage revealed an significant increase in ADAMTS-5 in synovial tissue form OA joints whilst OA cartilage revealed a significant increase in ADAMTS-4, indicating it may be the principle aggrecanases in equine OA (Kamm et al., 2010). Furthermore the use of transgenic mice has enabled the analysis of aggrecanolysis from *in- vitro* cultures of mouse femoral head cartilage which has helped to determine further the role of these enzymes in aggrecan degradation (Stanton et al., 2011).

Aggrecan is the first matrix component to experience measurable loss in OA (Mankin and Lippiello, 1970). Aggrecanase-mediated aggrecan degradation is one of the significant early events in this disease (Huang and Wu, 2008). Degradation is due to increased proteolytic cleavage of the aggrecan interglobular domain (IGD) (Sandy, 2006). Its loss is attributed to its accelerated degradation within the two major proteolytic cleavage sites of the IGD; the $Asn^{341}$-$Phe^{342}$ and $Glu^{373}$-$Ala^{374}$ bonds. As discussed earlier it is known that MMPs can cleave the $Asn^{341}$-$Phe^{342}$ bond. The signature activity of the aggrecanases; primarily ADAMTS-4 and ADAMTS-5 is the cleavage of the $Glu^{373}$-$Ala^{374}$ bond (Tortorella et al., 2000b). The degradation products cleaved at the $Glu^{373}$-$Ala^{374}$ bond have been detected in cartilage explants and chondrocyte culture (Lark et al., 1995; Loulakis et al., 1992; MacDonald et al., 1992; Sandy et al., 1991) and in the synovial fluids of patients with joint disease (Sandy et al., 1992) An assay detecting cleavage at this site has recently been developed using an immunoaffinity based LC/MS/MS (Dufield et al., 2010a). In addition to the two major cleavage sites within the IGD, proteolysis of aggrecan *in-vivo* also occurs within the CS domain (Ilic et al., 1995; Loulakis et al., 1992) (Figure 1.7) and studies using recombinant ADAMTS-4 and ADAMTS-5 indicate preferential cleavage of aggrecan in the CS-2 domain (Tortorella et al.,

2002; Tortorella et al., 2000b). In bovine aggrecan the two most favoured cleavage sites correspond in human to FKEEE$^{1714}$-$^{1715}$GLGSV and ASELE$^{1545}$-$^{1546}$GRGTI. Cleavage then occurs at the signature site in the IGD and in the corresponding human sequences at PTAQE$^{1819}$-$^{1820}$AGEGP and TISQE$^{1919}$-$^{1920}$LGQRP. Similar cleavage preferences have been identified by native aggrecanases in chondrocytes culture (Sandy and Verscharen, 2001).

The proteolysis within the CS attachment domain has been demonstrated in cartilage explant cultures treated with IL-1 or retinoic acid (Ilic et al., 1995; Loulakis et al., 1992) and in the synovial fluids from arthritic joints (Lohmander et al., 1993b). Sequencing analyses revealed that these cleavage sites are located in gap regions, which are relatively devoid of glycosaminoglycan chains, within the CS domain (Sandy et al., 1995).

### 1.5. Cartilage ageing

Although the OA disease process is distinctive from normal ageing, the relationship between age and OA is important. Not only is age a major risk factor for OA but treatments aimed at delaying cartilage ageing provide potential therapeutics. The prevalence of the OA increases with age with between 30 and 50% of adults over 65 years experiencing the condition (Felson, 2004). However, OA is not an inevitable consequence of ageing (Loeser, 2010). Consistent with the heterogenous nature of OA, ageing is just one (though the greatest) of the many risk factors involved (Suri et al., 2012). The interactions between other OA risk factors and age in ascertaining the sites and severity of OA are illustrated in Figure 1.8.

There have been a number of theories as to why ageing and in particular cartilage ageing plays such a major role in OA pathogenesis. Joint health is dependent on the normal structure and function of all the constituent tissues and OA is a disorder of the entire joint. Cartilage is the most susceptible tissue within the joint to damage and demonstrates the most intense age related changes.

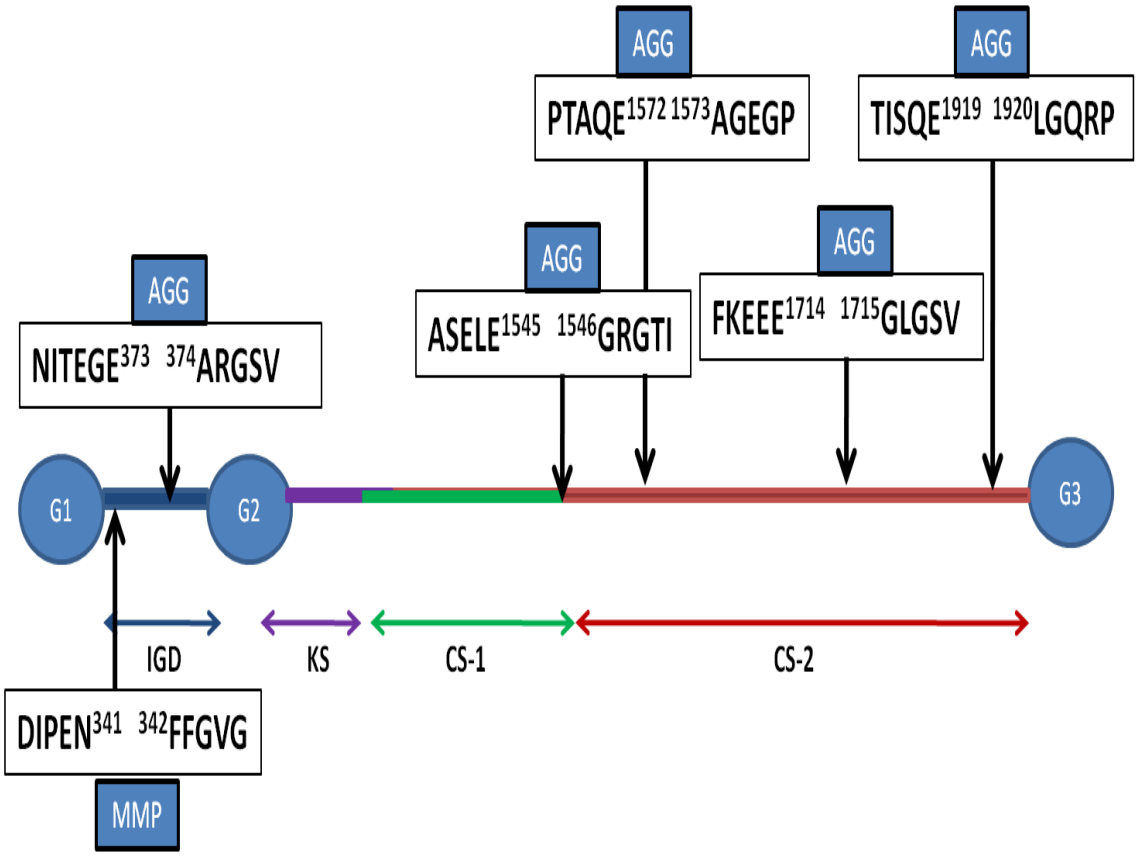



**Figure 1.7. Schematic diagram depicting major human aggrecan cleavage sites.** G1, G2, G3 represents the respective globular domains, IGD; interglobular domain, KS; keratan sulphate-rich region, CS-1, -2; chondroitin sulphate-rich regions. Aggrecanase (AGG) and MMP cleavage site sequences in the IGD and CS-2 domains are shown. Numbering of amino acids corresponds with the human sequence.

With advancing age changes occur in both in the ECM and chondrocytes. Mitotic activity and synthetic activity of chondrocytes alter with age whereas in OA chondrocytes are characterised by increased cell activation, proliferation and gene expression (Aigner et al., 2004b). A number of theories postulate as to why age contributes to cartilage degeneration. One established theory is that OA develops due to continuous accruing of repetitive load cycles which instigates constant microtrauma due to physiological loading, resulting in a loss of structure and function (Loeser, 2009). Another theory relates to ECM modifications including collagen and aggrecan with age. Increased cross-linking of collagen over time causes cartilage to stiffen, thus reducing flexibility during physiological deformation. Aggrecan structure also changes with age due to degradation and

impaired synthesis. The molecule becomes smaller and there is a reduction in sugar side chains (Glant et al., 1998). This affects the ability of aggrecan to bind water and it is this feature that gives aggrecan and therefore cartilage the compressive stiffness required for function. Finally glycation end products increase with age (known as advanced glycation end products (AGEs)) (Verzijl et al., 2002). These are non-enzymatic protein modifications that affect both matrix integrity and chondrocyte biology and therefore affect the mechanical properties of the cartilage.

### 1.5.1. Mechanisms of ageing in cells and tissues in the development of OA

Three main areas have been demonstrated as contributing to ageing in cartilage in relation to the development of OA. These are cell senescence, ECM ageing and age related oxidative stress.

#### a. Chondrocyte senescence

The senescence model for ageing theorises that chondrocytes become senescent due to proliferation and/or oxidative cell stress resulting in the inability of chondrocytes to maintain matrix turnover (Aigner et al., 2004b).

There is little evidence of chondrocyte turnover in adult articular cartilage (Martin and Buckwalter, 2001). Although they can divide occasionally, adult articular cartilage is classified as post-mitotic with trivial cell turnover. Consequently they are long lived and disposed to the accumulation of age-related changes over time. There is limited evidence for the existence of progenitor cells in cartilage which would allow senescent cells to be replaced. One study identified mesenchymal progenitor cells in human normal and OA cartilage (Alsalameh et al., 2004) whilst a study in young bovine tissue identified a progenitor cell population on the articular surface (Williams et al., 2010). Additionally equine derived cartilage progenitor cells are capable of functional cartilage repair (McCarthy et al., 2012). Should a local pool of progenitor cells exist they seem unable to replaced senescent chondrocytes.

Studies have demonstrated cellular degenerations including DNA damage in OA chondrocytes (Helmick et al., 2008) which normally results in apoptosis. The extent of apoptosis is debated. Although some believe it is rare in OA cartilage (Horton et al., 1998) there is evidence for both an age-related (Adams and Horton, 1998) and OA (Morgenroth et al., 2012) induced loss of chondrocytes from apoptosis. There does, in humans at least, appear to be a reduction in chondrocyte number with age progression (Dieppe, 1995).

A reduction in the chromatin protein high mobility box group B2 (HMGB2) in the cartilage superficial zone chondrocytes of animals and man with age occurs which may contribute to chondrocytes death (Taniguchi et al., 2009). The protein is a transcriptional regulator which maintains superficial zone chondrocytes survival through β-catenin signalling and in addition controls gene expression profile of the superficial zone cells. HMGB2 null mice display early-onset OA-like alterations associated with increased susceptibility to cell death (Taniguchi et al., 2009).

Telomere shortening, unrelated to replicative senescence, has been identified in chondrocytes with advancing age (Martin et al., 1997). This is possibly due to oxidative damage or inflammation (Grahame and Schlesinger, 2012). These finding may contribute to chondrocyte senescence *in-vivo* (Aigner et al., 2004b). There are a number of concepts evident in cell senescence. In classic replicative senescence there is an inability of the cells to experience further cell division. However there is also evidence for phenotypic alterations; the 'senescent secretory phenotype' (Campisi, 2005). Accumulation of these cells, which secrete increased amounts of MMPs and cytokines, contributes to cell ageing. Given the enhanced production of cytokines and MMPs in OA this provides a direct link between ageing and OA (Loeser, 2010).

Finally cell senescence has been linked to a reduction in the capacity of chondrocytes to respond to growth factors with age and OA (Campisi, 2005) including insulin growth factor-1 (IGF-I) and TGF-β, which may contribute to the imbalance between anabolic and catabolic activity in OA.

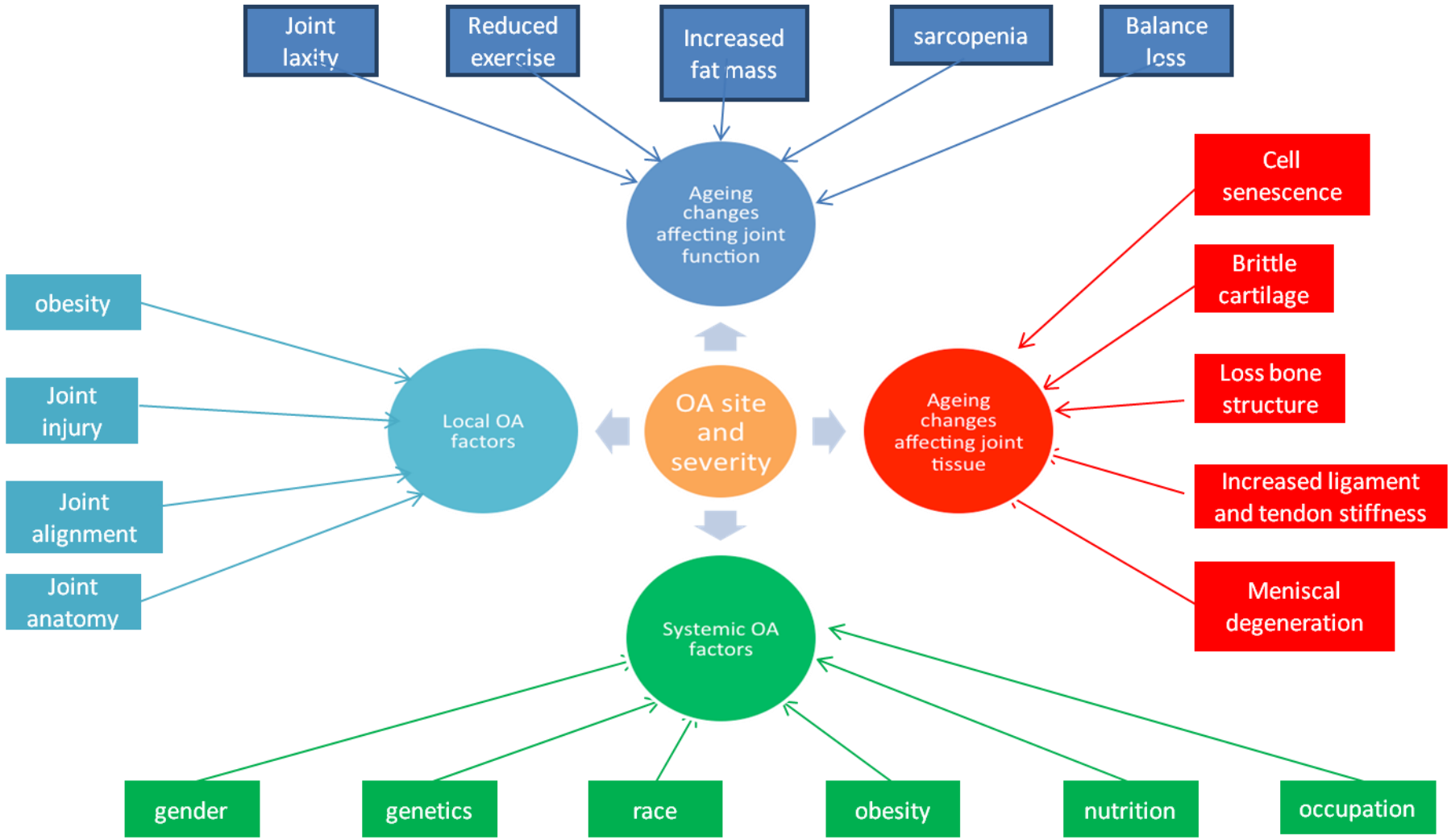



**Figure 1.8. The relationship between OA risk factors and ageing changes in OA.**

Since IGF-I has an important autocrine survival role in cartilage this age related reduction may be pivitol in age-related cell death (Loeser and Shanker, 2000).

### b. Cartilage matrix ageing

There is evidence of a gradual loss of cartilage matrix with ageing as demonstrated in knee magnetic resonanace imaging (MRI) studies (Connie et al., 2011). This is due to a loss of chondrocytes, reduced growth factor activity, and cartilage water loss. The water content of cartilage is largely controlled by the presence of aggrecan. This changes in its structure, glycosylation extent and size with age (Buckwalter et al., 1994). The age-related matrix protein modifications AGEs have a role in OA development (Loeser, 2009). As increased collagen cross-links are evident as a result of AGEs, cartilage biomechanical properties are effected resulting in susceptibilty to failure (Verzijl et al., 2000b). Thus overall these changes result in collagen that is less flexible, aggrecan that is smaller and less able to 'hold' water and altered phenotype leading to alterations in the anabolic/catabolic balance.

### c. Age-related oxidative stress

Oxidative damage from chronic formation of reactive oxygen species (ROS) results in age related tissue changes (Carlo and Loeser, 2003). Both ROS and reactive nitrogen species such as nitric oxide are produced by HAC. Furthermore, rat studies identified increased ROS with age (Jallali et al., 2005). Studies have also found increased oxidised glutathione (an intracellular anti-oxidant) with age in human chondrocytes (Jallali et al., 2005). Other anti-oxidants are detected at reduced levels with age (Jallali et al., 2005) and in OA. Whilst increased ROS results in deoxyribonucleic acid (DNA) damage leading to reduced chondrocyte viability and matrix production. Interestingly ROS interferes with IGF-I signalling leading to reduced matrix production (Yin et al., 2009). Further ROS may be stimulated by cytokines such as IL-1 and TNF-α which are elevated in OA. This increase in ROS can then increase MMP production (Forsyth et al., 2005).

## 1.6. What is proteomics?

Proteins are the molecules that are responsible for the majority of functions within the cell. Until recently it has been difficult to study global protein expression as proteins cannot be replicated easily in the laboratory. Furthermore, proteins do not have complimentary sequences like DNA and RNA that can be used as probes. Thus, historically scientists have relied on gene expression analysis to indicate how a protein is being regulated. As new technologies develop the examination of proteins has become more accessible and the study of proteins, known as proteomics, has become a discipline itself. Proteomics is the large scale study of proteins including their structure and functions. It is the systematic analysis of proteins expressed by a genome at a given time, including the set of protein isoforms and modifications (de Hoog and Mann, 2004). The concept was first proposed by Wilkins (Wilkins et al., 1996) and is analogous to genomics, the study of the gene. However, whereas a species genome is more or less constant, the proteome differs from cell to cell and from time to time depending upon the age, environment and diseases to which the animal is exposed. One of the challenges when studying the proteome is the number of proteins that need to be identified. The 20,322 genes in the equine genome (Wade et al., 2009) can code for at least ten times as many proteins. Measurement of transcript; messenger ribonucleic acid (mRNA) levels do not give complete information on cellular regulations. This is because there are a myriad of post translational modifications that occur following gene expression which results in many times more proteins than the coding potential of the organism (Cobon et al., 2002). As the complete genome sequence of many organisms has been identified, a shift in emphasis is taking place towards making genomics functional. Molecular biology has provided powerful tools for high-throughput DNA analysis (see 1.6 next generation sequencing) and emerging techniques that also allow the equivalent in protein analysis are becoming available. Previously these tools have resulted in an emphasis on message (mRNA or cDNA) rather than product of that message (protein). One of the advantages of proteomics is the potential to detect changes that occur after the mRNA ('message') step, thus giving quantitative analysis of protein expression profiles.

### 1.6.1. Methodologies for studying proteomics

Many elegant tools have been designed to study proteomics. These include light and electron microscopy for the imaging of cells, protein arrays, and mass spectrometry (MS). Affinity based assays cannot detect and quantify all proteins in a given species and consequently MS provides a powerful additional method to analyse complex protein samples. The technique is capable of identifying and quantifying thousands of proteins in complex samples.

Proteomics can be divided into expression proteomics and functional proteomics. Expression proteomics is the large scale study of variations in protein expression and is analogous to differential gene expression. It is based on the technique of two-dimensional gel electrophoresis followed by the characterisation of the protein spots by mass spectrometry. A good gel can separate several thousand proteins. One of the limitations is that certain classes of proteins such as membrane proteins do not readily enter gels and because abundance varies over a wide range enrichment strategies are required. Functional proteomics is the study of protein complexes and signalling pathways in order to understand how proteins interact to form cellular machines. It is not two-dimensional gel and image analysis based but uses inherent enrichment of proteins of interest by affinity purification. Affinity purification can be approached through the optimisation of various affinity reagents, such as antibodies, with specificity towards the proteins within a complex of interest. In a recent study an immunoaffinity-based liquid chromatography-tandem mass spectrometry (LC-MS/MS) method was used to detect cleavage at the two sites in the cartilage protein aggrecan (Dufield et al., 2010a).

### 1.6.2. Mass spectrometry

#### a. Separation techniques

The foundation of MS proteomic research requires the separation of a large number of proteins prior to identification by MS. The separation stage is a crucial step. In order to detect anything but the most abundant proteins in a sample and to simplify MS spectra so that they can be interpreted, a proteome must be

fractionated prior to MS. This may be achieved by gel electrophoresis or gel-free techniques which normally involve LC, so-called “shotgun-proteomics”.

### i. Gel-based separation

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was the first separation method developed. Following pre-treatment involving solubilisation, denaturation and reduction, to completely breakdown interactions between the proteins (Rabilloud, 1996), cellular or tissue extracts are separated on a polyacrylamide gel. Proteins then migrate according to their molecular weight when a current is applied. In two-dimensional electrophoresis (2-DE) (O'Farrell, 1975) proteins are separated in two dimensions, first by the pH at which the net charge is zero; the isoelectric point and then by size. Both types of gels are subsequently stained. To identify protein spots of interest these are cut from the gel prior to in-gel digestion using a protease of choice, normally trypsin. Various MS techniques can then be employed in order to identify the protein.

### iii. Gel-free separation

Gel-free proteomics consists of multidimensional fractionation, commonly strong cation exchange and reverse phase C18 high-performance liquid chromatography (HPLC) (Wang et al., 2003). These techniques are high-throughput since they can be automated and run in-line with the mass spectrometer.

## b. Protein identification using mass spectrometry

MS determines the mass-to-charge ratio (*m/z*) of gas phase ions is increasingly becoming the method of choice for the analysis of complex protein samples. MS-based proteomics is made possible by the availability of gene and genome sequence databases. In the most simplistic model a mass spectrometer is a sensitive weighing scale. The basic principle involves the ionisation of peptides produced following protein digestion. In addition from a pool of detected peptides specific precursor ions may be selected for fragmentation. Collisional activities result in product ion mass spectra which are recorded and used to identify the

amino acid sequence of the peptides selected. Data dependant analysis (DDA) is the most common implementation of the method. Here the precursor ions are selected automatically from ions detected in a survey scan preceding the ion selection. If DDA is combined with stable-isotope labelling workflows, these protein identification methods allow relative (compared to a reference) or absolute quantification of the proteins identified.

### 1.6.3. Mass spectrometry instrumentation

A mass spectrometer consists of an ion source, a mass analyser that measures *m*/*z* of the ionized analyte, and a detector that registers the number of ions at each *m*/*z* value (Figure 1.9).

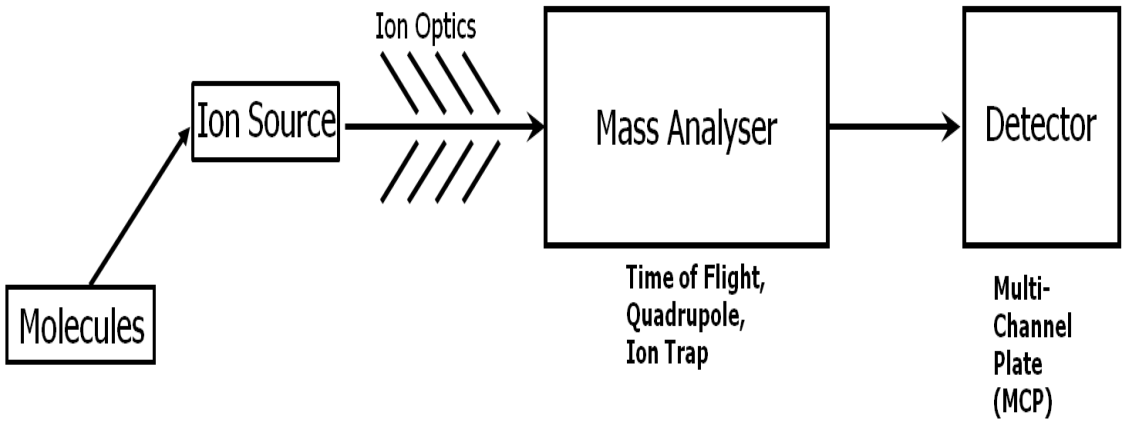



**Figure 1.9. Schematic of a mass spectrometer.**

The two most commonly used techniques to volatize and ionize the proteins or peptides for mass spectrometric analysis are electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) (Fenn et al., 1989; Karas and Hillenkamp, 1988). Both methods are known as "soft" ionisation methods because the sample is ionised by the addition or removal of a proton. Additionally very little extra energy remains to cause fragmentation of the sample ions. ESI is readily coupled to liquid based separation tools as it ionizes the analyte out of solution enabling the analysis of complex samples when integrated with LC; LC-MS. It is usually coupled to ion traps and triple quadrupole instruments. The fragment ion

spectra (collision-induced spectra or CID) of selected precursor ions are then produced (Aebersold and Mann, 2003). Interestingly samples with molecular weights greater than 1200 Da tend to produce multiply charged ions. Here the *m/z* values can be expressed as:

$m/z = (MW + nH^+)/n$

*m/z* = the mass-to-charge ratio of the spectrum;
MW = the molecular weight of the sample
n = the number of charges on the ion
H = the mass of a proton = 1.008Da

This is summarised in Figure 1.10.

| Notation | Charge | Observed m/z |
|---|---|---|
| $[M+H]^+$ | single | M + 1Da |
| $[M+2H]^{2+}$ | double | (M + 2Da)/2 |
| $[M+3H]^{3+}$ | triple | (M + 3Da)/3 |

**Figure 1.10. Table illustrates the *m/z* notation used in MS.**

Proteins have many suitable sites for protonation. Theoretically all of the backbone amide nitrogen atoms could be protonated. In addition, certain amino acids such as lysine and arginine contain primary amine functionalities which can also carry charge.

The mass analyser plays a central role in MS as its primary function is to resolve the ions formed in the ionisation source of the mass spectrometer according to their *m/z*. It provides varying degrees of the key parameters; sensitivity, mass accuracy, resolution and the ability to produce mass spectra from peptide fragmentation, so called tandem MS. In proteomic research three types of mass analysers are used primarily; the ion trap, time-of-flight (TOF), quadrupole and Fourier transform ion cyclotron (FT-MS). The compatibility of mass analysers with the numerous

ionisation methods varies. Whilst all types of analyser can be used in conjunction with electrospray ionisation, MALDI is rarely coupled to a quadrupole analyser.

The final part of the mass spectrometer is the detector which monitors ion current. The ion current is amplified and the signal transmitted to a data system which is recorded in the form of a mass spectra. The number of components in the sample, the molecular mass of each component, and the relative abundance of the various components in the sample are determined by plotting the *m/z* values of the ions against their intensities. More common detectors include the photomultiplier, the electron multiplier and the micro-channel plate detectors.

### 1.6.4. Protein identification

The two principle methods used for protein identification are peptide mapping (Henzel et al., 1993) and peptide sequencing. In the former the proteins are digested with a proteolytic enzyme (usually trypsin) to produce a set of tryptic fragments unique to that protein. The *m/z* of peptides are then determined by MS in the gas phase by observing their flight in electric and/or magnetic fields. Once ions are formed they can be separated according to their *m/z* and detected. Historically the most common system employed for this is matrix-assisted laser desorption/ionisation time-of-flight mass spectrometer (MALDI-TOF). MALDI uses a solid matrix and a laser to produce short burst of ions. Digested peptides are mixed with a matrix substance, normally α-cyanocynamic acid. Small ions move at a higher velocity and are detected prior to larger ions, thus producing a 'time-of-flight' (El-Aneeda et al., 2009). The molecular weight values of trypsinised peptides obtained by MALDI-TOF are then used to identify the predicted proteins using web-based search engines such as MASCOT (Perkins et al., 1999).

An alternative method employed for protein identification is peptide sequencing. Sequencing peptide spectra produced are used to search an *in-silico* digested proteome sequence database. Peptides are fragmented in the collision cell of the mass spectrometer to generate fragment ions of progressively lower mass to aid identification by fingerprinting. This workflow is called tandem mass spectrometry

(MS/MS) and uses instrumentation such as quadrupole mass spectrometry (Q-TOF) (Morris et al., 1997).

The fragment ions observed in an MS/MS spectrum will only be detected if they carry at least one charge. If this charge is retained on the N-terminal fragment, the ion is classed as either a, b or c. If the charge is retained on the C-terminal fragment the ion type is either x, y, or z (Figure 1.11). The difference in mass between adjacent y or b ions corresponds to that of an amino acid. Mass spectrometry does not fully sequence the sample but creates sufficient information that will identify the unknown protein by searching appropriate databases.

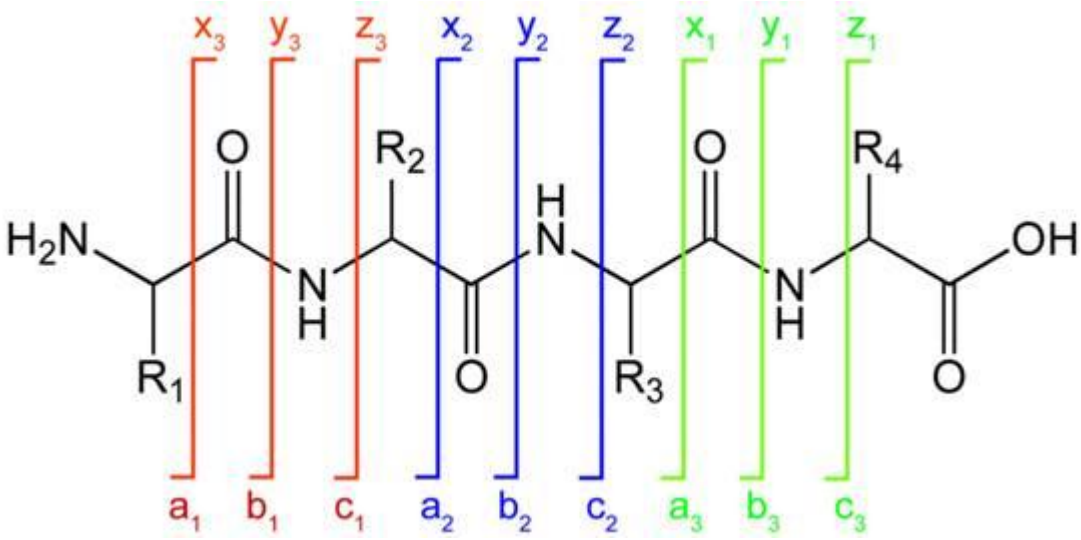



**Figure 1.11. The fragment ions generated upon peptide dissociation (Zhang et al., 2006a).**

### 1.6.5. Protein quantification using mass spectrometry

In discovery experiments large numbers of proteins are identified with no relevance to the biological question posed, whilst some proteins which may be relevant can be missed. Furthermore the absence of a protein from the identification list does not mean the protein is absent from the sample. It has been suggested that if reproducible and quantitatively accurate data can be generated for all proteins that are involved in a particular process then proteomics will make a greater impact in

biology (Ren et al., 2006). Thus a quantitative approach to proteomics identifying and quantifying sets of defined proteins will progress our understanding of biological systems and disease states.

There are three principle methods used to compare amounts of proteins using MS. These are relative quantification, absolute quantification and protein kinetics or turnover. In addition two approaches can be used for quantitative; the utilisation of stable isotope labelling and' label-free' techniques.

### a. Relative Quantification

Relative quantification measures changes in expression levels. Numerous quantitative methodologies have been dedicated to the extensive comparison of multiple proteomes commonly using isotopic-labelling approaches including stable isotope labelling with amino acids in cell culture; SILAC (Aebersold and Mann, 2003), isotope-coded affinity tagging; ICAT (Shiio and Aebersold, 2006). These methods involve labelling peptides, with reagents or labels that are chemically identical but vary in their isotope composition, prior to the sample being run on LC-MS. In addition isobaric tagging reagents have been used known as iTRAQ (isobaric tagging for relative and absolute quantification) (Wiese et al., 2007)and 'label-free protocols' have also been implemented (Wong and Cagney, 2010).

Label-free techniques have some advantages over labelling approaches. Although they require high mass precision instruments they do not require expensive isotope labels, or particular software and they can analyse many proteins in a single sample and many samples in a single experiment concurrently. Label-free quantitation uses either area under the curve, relying on signal intensity founded on precursor ion spectra or spectral counting which uses the number of peptides assigned to a protein in an MS/MS study (Zhu et al., 2010). The spectral counting approach uses data processing involving applications such as 'exponentially modified protein abundance index' or emPAI (Ishihama et al., 2005).

### b. Absolute Quantification

Systems biology approaches, which require the ability to make quantitative measurements have furthered the demand for absolute protein abundance values for input into models (Otto et al., 2012). Absolute quantification is a 'targeted approach' as it focuses on a specific set of peptides. Thus prior information is required in order to produce MS assays for the detection and quantification of predetermined analytes in a mixture. As the method is hypothesis driven, a subset of peptides uniquely associated with the proteins of interest are targeted. In absolute quantification copies per cell of a protein from a cell lysate or molar quantities of a protein from a secretome or media can be determined. One advantage of this method is that it allows the real comparison of data between laboratories. Such precise determination of concentrations requires internal standards which must perform in the linear range of the system. The three types of isotope-labelled quantification standards are AQUA (for absolute quantification) (Gerber et al., 2003), QconCAT (quantification concatamers) (Beynon et al., 2005) and PSAQ (protein standard absolute quantification) (Dupuis et al., 2008) standards. In the latter a full-length protein with homologous biochemical properties of the target protein are spiked in at the start of the analytical process.

#### i. AQUA

In 2003 a strategy known as Protein-Aqua was presented for absolute quantification by using isotopically labelled peptides for downstream analysis by LC-MS (Gerber et al., 2003). This method, also known as stable isotope dilution (SID) is regarded as the gold standard for quantitation by LC-MRM-MS (Pan et al., 2009). Commonly isotope-labelled internal standards are used to measure small molecules such as hormones in analytical chemistry. The method employs de novo chemically synthesised, isotopically labelled internal standards. Interestingly due to the chemical synthesis of the peptide standards, AQUA has enabled quantification of post-translational modifications and biomarkers (Cummings et al., 2008). Peptides are quantified by monitoring either accurate mass retention time (AMRT) (Silva et

al., 2005) or monitoring multiple reaction monitoring (MRM) transitions for each peptide (Figure 1.12).

### i. QconCAT

A method using artificial concatamers of a set of standard peptides (Q-peptides) was introduced by Beynon *et al.* in 2005 (Beynon et al., 2005); QconCAT. This extends the number of proteins that can be quantified in parallel. A synthetic gene is produced which is then inserted into plasmid DNA. At least two peptides per protein are incorporated into the QconCAT in order to reduce the risk associated with poorly ionizing or difficult peptides (Pratt et al., 2006).

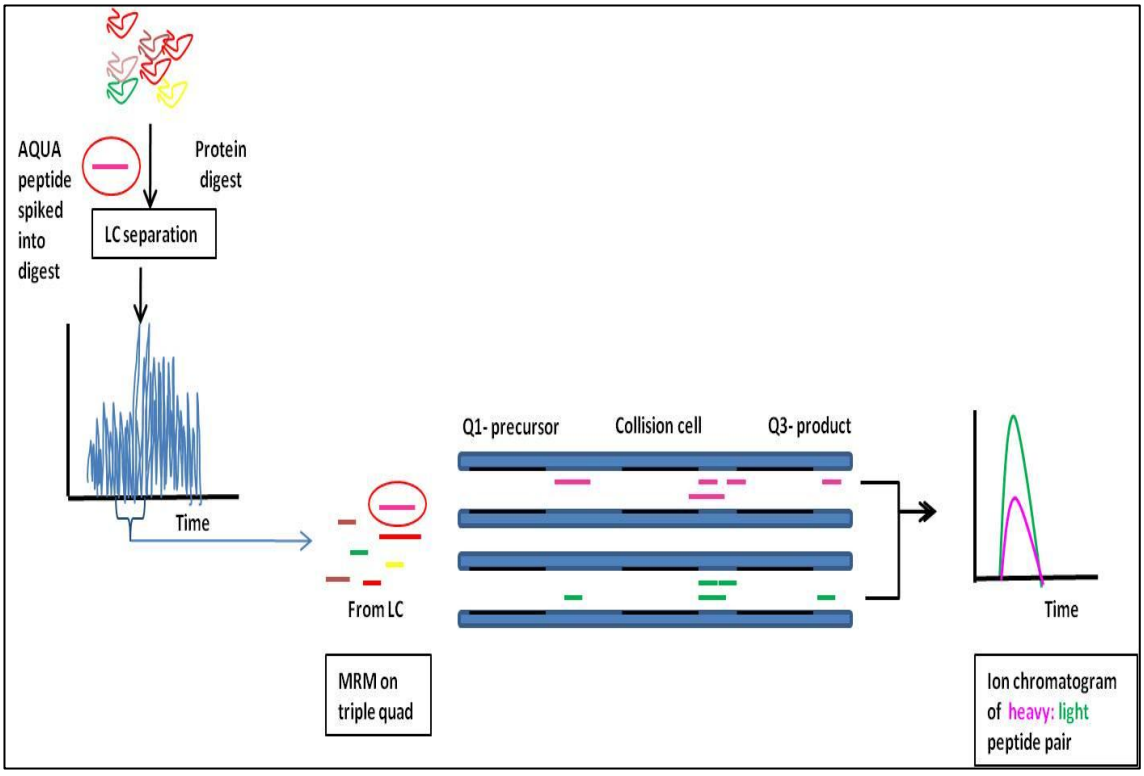



**Figure 1.12. Mass spectrometry methodology using AQUA and multiple reaction monitoring (MRM).** MS is undertaken using a triple quadrupole. At a set elution time the first quadrupole (Q1) is used to target and detect the parent ion of a given peptide (both heavy; AQUA peptide and light; native peptide co-elute at the same time from the LC column). This is fragmented by the collision cell, product ions are detected by the third quadrupole (Q3).

In addition to the Q-peptides, the artificial protein contains a C-terminal hexahistidine-tag (His-Tag) sequence for purification, fibrinopeptide following the initiator methionine and glufibrinopeptide 'EGVNDNEEGFFSAR' for quantification of the QconCAT. The fibrinopeptide and glufibrinopeptide also allow assessment of expression of the QconCAT as if both are present in a QconCAT digest this means the full length protein has been expressed. The genes are subsequently expressed in Escherichia coli *(E.Coli*) grown in media containing the selected label. Where trypsin is the proteolytic enzyme of choice for digestion $_{L}$-lysine ($^{13}C_{6,}$ $^{15}N_2$) and $_{L}$-arginine ($^{13}C_{6,}$ $^{15}N_4$) are the labels of choice providing a 6Da mass shift between standard and analyte. Following proteolysis each Q peptide is released in stoichimetry of 1:1. Subsequently the quantification of each selected peptide by MS can then be undertaken (Pratt et al., 2006). Normally trypsin is used for endoproteolysis, in order to cleave proteins into peptides suitable for MS; however other endoproteinases can be used were trypsin digestion provides inappropriate surrogate peptides. Trypsin cleaves C-terminal to arginyl and lysyl residues except arginyl-propyl and lysyl-propyl sequences which are not cleaved. It produces a range of peptides between 600-4000Da which are ideal for analysis with MS. Additionally tryptic peptides are normally doubly charged ions $[M+2H]^{2+}$ which aids in MS/MS analysis used in MRM experiments. Because the standard is biologically synthesised one of the advantages of QconCAT over AQUA is that once conceived and validated a QconCAT gene may be used for repeated production of unlimited amounts of isotopically-labelled peptide standards (Rivers et al., 2007). A further advantage is that were appropriate, QconCAT may be co-digested with the analyte protein in order to overcome problems associated with differential susceptibility to proteolysis (Rivers et al., 2007). Moreover it has also been suggested that by putting the peptides in sequence where they are surrounded by their native flanking sequence this may avoid this problem (Kito et al., 2007).

Experiments using AQUA and QconCAT can be performed on two MS platforms; ion trapping (e.g. Orbitrap) and quadrupole based instruments (e.g. triple quadrupole, quadrupole-time of flight (TOF)). Ion trapping instruments use AMRT. Figure 1.13 demonstrates the quadrupole methodology. For each peptide, pairs of precursor

and fragment ion *m/z* values (called transitions) are monitored during the predicted elution time, allowing hundreds of peptides to be analyzed in a single experiment.

QconCAT methodology also has its shortcomings. These include the poor solubility of some QconCATs, variable expression, incomplete digestion and post translational modifications during digestion (Aebersold and Mann, 2003). In a study conducted by Mizaei et al. 2008 (Mirzaei et al., 2008) the group compared AQUA peptide and QconCAT peptide absolute quantification using the *Caenorhabditis elegans* proteome. Results indicated that although most QconCAT peptides were in equimolar ratios, digestion efficiencies were an important factor for some. Additionally solubilisation of some AQUA peptides was reflected in the overestimation of some QconCAT peptides. The conclusion was there was no superior method.

### 1.1.1. Mass spectrometric imaging of cartilage

A further way in which MS is useful in interrogating cartilage is through imaging mass spectrometry (IMS). Mass spectrometry has the capabilities to determine the mass of a large mass range of molecules; from large biomolecular complexes down to small organic molecules, even single atoms and their isotopes. Continual development has enabled improvements in terms of its sensitivity, resolution and mass range. The introduction of ESI and MALDI has allowed the ionisation of smaller biomolecules such as metabolites as well as larger biomolecules such as lipids, peptides and proteins (Leinweber et al., 2009).

Imaging mass spectrometry is a powerful tool to study molecular distributions at biomedical tissue surfaces. It allows 'label free' biomolecular imaging technique in life sciences by the rapid detection, localisation and identification of molecules from even complex, biological samples. The technique enables detailed appreciation of biological processes on different length scales, from subcellular to entire organs. In terms of OA, knowledge of the dynamic distribution and the variation of molecules in different stages of disease, from a spatial point of view,

are essential to understand the principal factors implicated in OA development in order to establish efficient treatments.

#### 1.1.1.1. Basic Principles

There are three main platforms for MSI are; physics secondary ion mass spectrometry (SIMS), matrix assisted desorption ionisation mass spectrometric imaging MALDI-IMS (Caprioli et al., 1997) and desorption electrospray ionization (DESI) (Takats et al., 2004). There are four stages to an IMS experiment; sample preparation, desorption and ionisation, mass analysis and image registration. Following sample preparation methods specific to the tissue analysed, the biomolecules are desorbed and ionized from the surface. The entire surface of the sample is examined in order to collect mass spectral information about the molecular composition and distribution of the analysed molecules at each point. These ionised molecules may be intact proteins, peptides, lipids or small molecules. Ion images are then presented from the resulting molecular ion distributions.

#### 1.1.1.2 MALDI-IMS

MALDI-IMS is the most common method for measuring intact peptide and protein due to its high sensitivity, tolerance for salts, wide mass range, little fragmentation and simple data interpretation as the majority of ions are singly protonated (Chughtai and Heeren, 2010). It also enables the analysis and detection of other biomolecules (such as lipids) directly from tissue sections. It is capable of producing intact higher MW ions by the implementation of pulsed laser beams combined with energy absorbing matrix molecules. The MALDI desorption and ionisation process is shown in Figure 1.14. In a MALDI-IMS experiment the laser beam is rastered across the surface of the treated tissue, enabling desorption and ionization of biomolecules. The ablation crater has a depth of 1µm or more (Heeren et al., 2006).

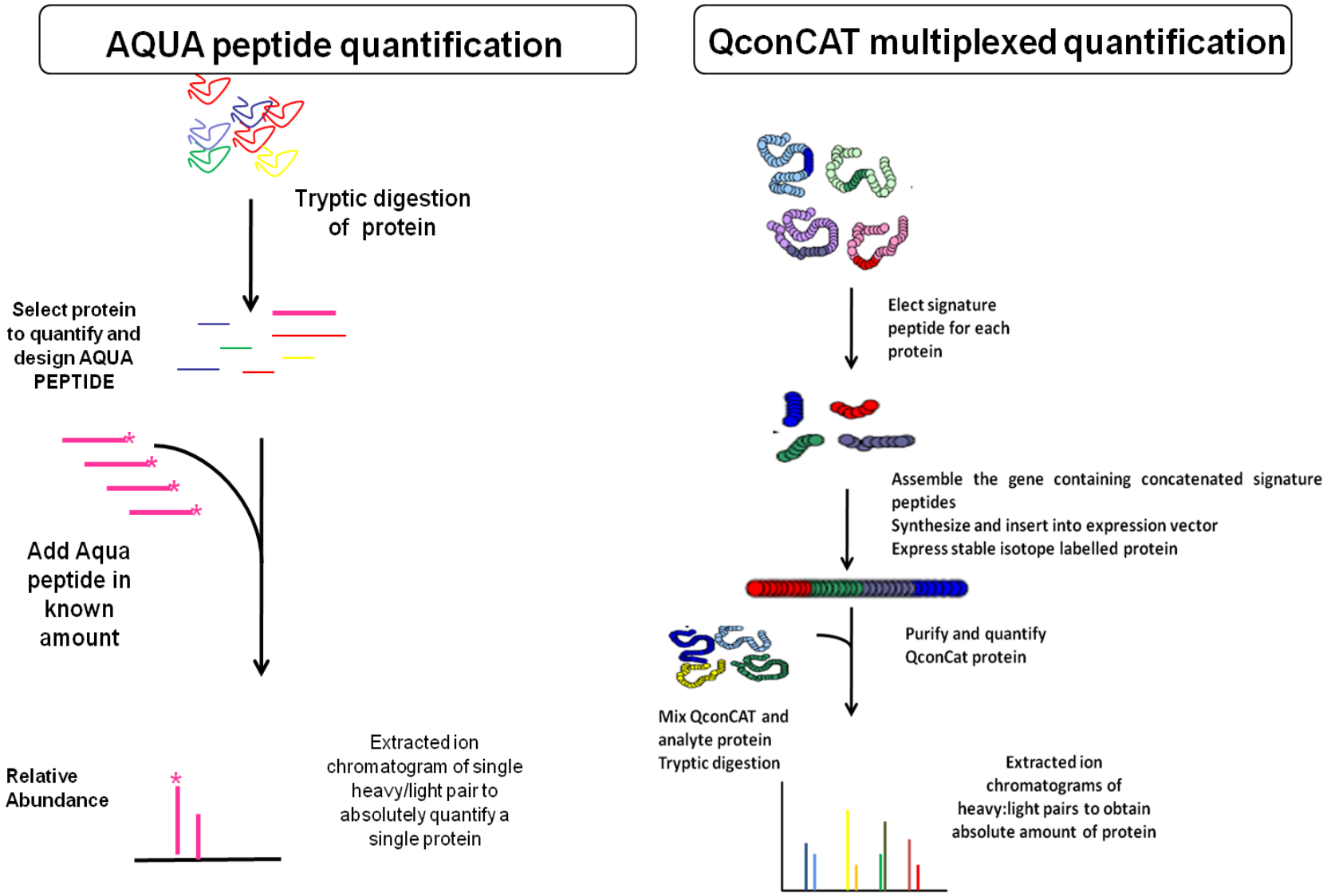



**Figure 1.13. Workflows used for absolute quantification with AQUA and QconCAT.**

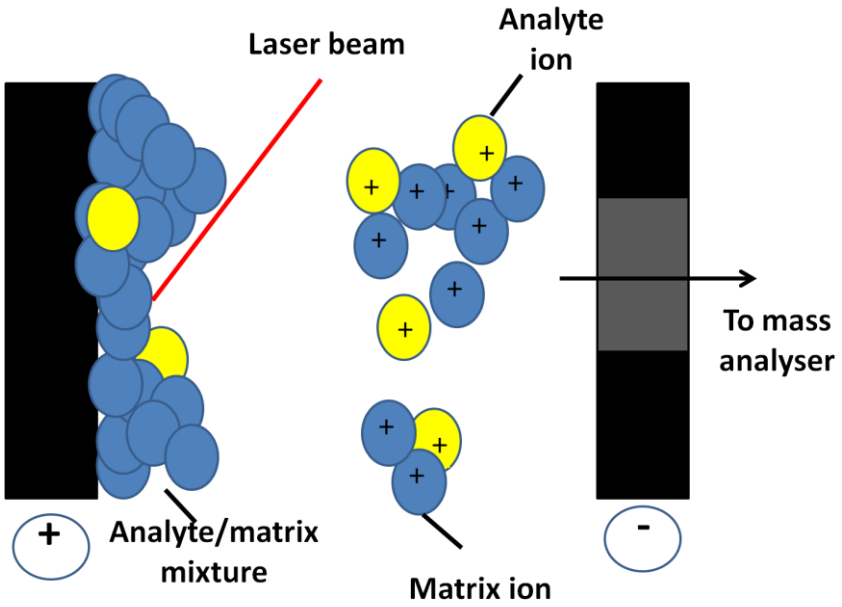



**Figure 1.14. Schematic representation of MALDI desorption and ionisation process.**

#### 1.1.1.2. Imaging mass spectrometry of cartilage

IMS has been used to identify the molecular distribution of peptides in many tissues including brain (Taban et al., 2007), liver (Lee et al., 2011) and kidney (Meistermann et al., 2006). Additional molecules including lipids (Hankin et al., 2011) and small molecules (Blatherwick et al., 2011) have been localised in numerous tissues. There have been very limited studies using IMS in musculoskeletal tissues. Indeed there are only two peer-reviewed citations. One study used a time of flight secondary ion mass spectrometry (TOF-SIMS) workflow in order to afford molecular-specific spatial distribution of lipids in normal and OA cartilage (Cillero-Pastor et al., 2012b). The study revealed lipid, calcium and phosphate distributions that distinguished human normal from OA cartilage. The other study investigated synovial membrane in rheumatoid arthritis with MALDI imaging in order to identify biomarkers (Kriegsmann et al., 2012). Recently a study identified potential biomarkers of OA using MALDI-IMS on OA cartilage (Cillero-Pastor et al., 2012c).

### 1.1.2. The role of proteomics in the study of cartilage and OA

Gene expression levels of many proteins involved in OA have been studied under experimental conditions (Aigner and Dudhia, 2003), but these do not always correlate to protein expression due to protein degradation, and alternative transcriptional and translational steps. Furthermore, post-translational modifications of proteins are not considered and it is proteins that regulate the cell phenotype important in OA. By studying proteomics a snap-shot of what is actually occurring in the tissue at a given time can be studied. Proteomic studies of cartilage are also important to provide insights into normal molecular pathways, discover disease specific biomarkers for disease, identify novel therapeutic targets in disease and identify factors that contribute to disease pathogenesis. In the case of biomarkers these can be monitor disease stage and progression. Although the study of cartilage and OA using proteomics has also used chondrocyte mitochondria(Ruiz-Romero et al., 2009), synovial fluid (Mateos et al., 2012), and serum (Yanagida et al., 2012), two types of samples have been used primarily; the cartilage secretome and cartilage tissue.

#### 1.1.2.1. Cartilage secretomics

The secretome is the global study of proteins that are secreted by a cell, tissue or organism at a given time, or under certain conditions. Secreted proteins are an important class of active molecules with roles in a number of physiological and pathological mechanisms. Protein secretion is a well regulated and balanced process required for normal physiologic function. Any aberrations in the secretory profile of a cell may lead to pathological conditions and so analysis of the complete secretome of a cell may also provide diagnostic biomarkers in disease.

Cartilage cells grown in monolayer rapidly lose their phenotype as demonstrated by a loss of the chondrocyte gene expression profile (Moskalewski et al., 1979) and a change in morphology to a more fibroblastic state (Kuettner et al., 1982). Although analysis of the secretome immediately after isolation minimizes the phenotypic changes, enzymatic isolation results in destruction of the ECM and associated cell-surface proteins. Proteomic analysis of the secretome using this system has limited

relevance to the cartilage proteome *in-vivo.* Cartilage explants allow chondrocytes to be retained within their ECM and are assumed to be less destabilized by their culture environment. The ECM also provides native substrates for proteolysis and protein release, equivalent to the shedding the proteins into the synovial fluid during cartilage degeneration. Proteins released into the media represent a combination of new protein synthesis and secretion as well as active and passive loss of pre-existing proteins within cartilage. Some of the proteins may be contained in cartilage explants via diffusion from the synovial fluid. Whilst this exchange of material between cartilage and synovium can produce cartilage derived proteins entering the serum, it has the potential to allow some serum derived proteins to enter the cartilage.

Pro-inflammatory cytokines such as IL-1β and TNF-α induce matrix degrading enzymes significant in the pathogenesis of OA (Goldring and Goldring, 2004). The use of catabolic stimulants on primary chondrocytes and explants, combined with proteomic analysis can offer an important approach to elucidate novel mechanisms in cartilage degradation. Studies have assessed the effect of IL-1β stimulation on the cartilage secretome (Catterall et al., 2006; De Ceuninck et al., 2004; Stevens et al., 2008) and more recently a SILAC study was carried out on IL-1β stimulated primary HAC (Polacek et al., 2010b). In a study undertaken in juvenile bovine cartilage explants secretome profiles were compared in groups treated with IL-1 β or TNF-α or traumatic mechanical compression and then cultured for 5 days (Stevens et al., 2008). The composition of proteins from the secretome from cytokine treated and untreated samples were similar and included proteins associated with innate immune, and stress responses including acute phase and complement proteins. In addition a number of novel cartilage proteins were discovered. In a recent study an IL-1 stimulated equine explants model was used to provide high throughput proteomic analysis of the secretome. Peptides identified included aggrecan, COMP, fibronectin, fibromodulin, thrombospondin-1, clusterin, cartilage intermediate layer protein-1, chondroadherin and MMP-1 and -3 (Clutterbuck et al., 2011). The explantation and cutting of articular cartilage for explant studies can activate intracellular inflammatory signalling pathways and

induces expression of mRNA for IL-1α and IL-1β (Gruber et al., 2004). This highlights the need to maintain explants in culture for at least 24h prior to experimentation.

The first application of 2-DE followed by LC-MS/MS to analyze the secretome of normal and OA articular cartilage was undertaken by Hermansson et al. 2004 (Hermansson et al., 2004). There was an increase in the synthesis of type II collagen in OA cartilage media along with the anabolic factors pro-inhibin βA and activin A. Soon after another group evaluated different methodologies to evaluate the OA cartilage secretome (De Ceuninck et al., 2005) including 2-DE, off-gel electrophoresis coupled to tandem MS, and antibody-based protein microarrays. 2-DE of explant culture media is difficult as it contains high concentrations of highly anionic compounds which interfere with isoelectric focalization. Off-gel electrophoresis provided a tool for discarding abundant proteins and concentrating less abundant ones. Antibody microarrays provided a method to identify low molecular weight proteins, minor and poorly stained proteins when electrophoretic techniques produced limited results. Using this combination of methods 43 proteins were identified.

Recently a quantitative proteomic analysis study was undertaken with primary HAC. In this SILAC study both the quantitative analysis of the IL-1β treated HACs proteome and the HAC secretome were undertaken. It revealed a global increase in cellular chaperones concurrent with a down-regulation of the actin cytoskeleton. Secretion of aggrecan, vitamin K-dependent proteins, and thrombospondin was reduced (Polacek et al., 2010a) .

Using 2-DE and ESI MS/MS a further study assessed the effects of IL-1β and oncostatin M on the secretome of human and bovine chondrocytes. It was determined that some of the secreted proteins for example MMP-1, MMP-3, YKL-40 and cyclophilin A were cleaved into smaller fragments by proteolysis (Catterall et al., 2006). Thus a proteomics approach is able to offer further insight into protein processing which cannot be determined by genomic studies.

Human endemic OA was investigated using the cartilage cell culture media from normal and Kashin-Beck disease (KBD) (Du et al., 2010). KBD is an endemic deformed osteochondropathy prevalent in China, whose aetiology is uncertain. 2-D-DIGE (difference gel electrophoresis) and MALDI-TOF/TOF analysis revealed differences in 27 proteins mainly involving cellular redox homeostasis and stress response, glycolysis, cell motility and cytoskeletal organisation, indicating these processes are aberrant in KBD.

#### 1.1.2.2. Whole cartilage proteomics

In order to successfully analyse the cartailge proteome extraction techniques are required which are both effectual and reproducible. A number of reviews discuss the approaches to cartilage proteomics within different cartilage extracts (Ruiz-Romero and Blanco, 2009; Ruiz-Romero and Blanco, 2010; Wilson et al., 2009). Unfortunately none are suitable for the analysis of all proteins. The two major proteins in cartilage; proteoglycans, with their anionic nature, high density and size and highly abundant insoluble collagens are the greatest hurdles in proteomic analysis. In addition to the presence of small numbers of high abundance proteins masking the identification of less abundant proteins, their properties impede the biochemical behaviour of proteins such as isoelectric focussing in 2-DE potentially disguising other important proteins (Wilson et al., 2009). As a consequence of this most studies have been undertaken following proteoglycan removal using molecular weight cut-off filtration (Garcia et al., 2006) or cetylpyridinium chloride precipitation of anionic PG followed by 2-DE (Hermansson et al., 2007). A further approach to reduce sample complexity and problems associated with high concentrations of GAGs in the sample is to undertake secretome studies of chondrocytes or cartilage explants. It must be remembered that all techniques have disadvantages, as demonstrated by the large number of techniques employed in cartilage proteomics (Iliopoulos et al., 2010). A table illustrating the techniques employed historically are demonstarted in Figure 1.15. For instance the removal of PGs can result in the loss of other proteins both partially or totally (Wilson et al., 2008).

One of the first MS-based proteomic studies of OA whole cartilage, undertaken by Garcia et al. (Garcia et al., 2006) revealed a limited number of proteins due to the presence of a few highly abundant proteins and free sugars in the form of GAGs from proteoglycans. However a 1-D SDS-PAGE in-gel digest also performed was more revealing identifying 100 different proteins from OA cartilage.

A novel method of tissue dissection and sample prefractionation was used in the proteomic analysis of mouse growth plates using 4M guanidine extraction and a 100kDa molecular weight cut-off (MWCO) filters to prefractionate the sample. This was used as most proteins detected with 2-DE are between 10-100kDa. By the removal of higher molecular mass ($M_r$) this allowed the enrichment of low $M_r$ proteins resulting in more proteins identified with 2-DE (Belluoccio et al., 2006).

Solubility-based fractionation techniques have been used to increase reproducibilty and allow comparative experiments in neonatal mouse growth plates. Intracellular proteins were extracted using NaCl buffer (to enrich cellular proteins) followed by a guanidine extraction in order to interrupt matrix components and thus enrich them allowing additional proteins to be identified (Wilson et al., 2010b). Sequential extraction methods separated the proteome into distinct subgroups as analyzing all the proteins in one fraction (and expecting good coverage) was thought to be too ambitious and would produce a reduction in the performance of the chromatography. The analysis of sequential extracts indicated a transition in protein solubilty from readily soluble proteins in juvenile cartilage to a high proportion of poorly soluble in neocartilage. Proteins were evident as significantly enriched dependant upon whether the sample was neonatal or juvenile cartilage.

| Citation | Species | Tissue | Technique | Proteomic analysis |
|---|---|---|---|---|
| (Garcia et al., 2006) | human | whole cartilage | Cartilage minced, Liberase digestion, 100kDa centricon filter | Supernantant either 1D-SDS-PAGE or in solution tryptic digest; LC/MS/MS |
| (Hermansson et al., 2004) | human | whole cartilage | Cartilage microtome sliced, extraction buffer with NaCl, CPC precipitation, methanol/chloroform precipitation | 2DE, tryptic digestion,MALDI-TOF |
| (Guo et al., 2008) | human | whole cartilage | Cartilage powdered, extraction buffer with urea, high speedcentrifugation, acetone precipitation | 2DE, tryptic digestion, LC/MS/MS |
| (Wu et al., 2007) | human | whole cartilage | Cartilage minced, 4M guanidine extraction, centrifuge, caesium chloride gradient | 1D-SDS-PAGE, tryptic digestion, LC/MS/MS |
| (Belluoccio et al., 2006) | mouse | whole cartilage | Cryosection of cartilage, guanidine extraction, 100kDa filter, ethanol precipitation, resuspend IEF buffer | 2DE, tryptic digestion, LC/MS/MS and western blotting |
| (Wilson et al., 2008) | mouse | cartilage explants and media | Cartilage pulversied, chondrotinase ABC, guanidine extraction, 100Kda filter, ethanol precipitation, resuspend IEF buffer | 2DE, tryptic digestion, LC/MS/MS and westerern blotting |
| (Guo et al., 2008) | any | whole cartilage | Cartilage pulverised, chondroitinase ABC, guanidine extraction, 100Kda filter, concentrate on 3kDa filter to remove disaccarides released, retenate ethanol precipitated | 2DE, tryptic digestion, LC/MS/MS |
| (Wilson et al., 2010b) | mouse | whole cartilage | Cartilage pulverised, chondroitinase ABC, NaCl extraction, guanidine extraction, 100kDa filter, ethanol precipitation, resuspend in solubilsation buffer(7M urea, 2M thiourea, 4% CHAPS). Insoluble fraction incubated with pepsin to release collagen polypeptides | Novel in-solution tryptic digest with methanol, LC/MS/MS |

**Figure 1.15. Table of the techniques used in cartilage proteomic studies**

## 1.5. Next-generation sequencing

Whilst mass spectrometry has enabled massive steps forward in the discovery and quantification of proteins, the advent of next-generation sequencing has become the premier tool in genetic and genomic analysis. Next-generation sequencing (NGS) otherwise known as high-throughput sequencing allows the parallel sequencing of millions of sequences at once (Hall, 2007). In principle, the bases of a small fragment of DNA are sequentially identified from signals emitted as each fragment is re-synthesized from a DNA template strand. NGS enables this process across millions of reactions in a parallel fashion. This enables rapid sequencing of large stretches of DNA base pairs spanning entire genomes. One NGS platform is ribonucleic acid (RNA) sequencing. This enables deep coverage and base-level resolution, providing information on differential expression of genes, including gene alleles and differently spliced variants; non-coding RNAs; post-transcriptional mutations; and gene fusions.

### 1.6.1. Deep transcriptome sequencing (RNA-Seq)

In order to comprehend development and disease and define the molecular constituents of cells and tissues, it is essential to understand the transcriptome. Various technologies have been developed to deduce and quantify the transcriptome, including hybridization such as microarrays (Clark et al., 2002) and sequence-based approaches. Microarray systems employ glass slides containing up to millions of anchored oligonucleotides designed to hybridise to predefined transcripts. Microarrays though high-throughput are hypothesis driven requiring knowledge of the genes expressed in a given set of conditions. In addition they have limits in the dynamic range of detection and specificity. In contrast sequence-based approaches directly determine cDNA sequence using tagged libraries of short cDNAs (Figure 1.16). RNA-seq is a technique that enables deep sequencing of the transcriptome of any species with single base resolution and without the need for a known reference genome sequence. It provides massive and valuable information about functional elements in the genome including identification and quantification of all RNA species, of any size or abundance and it can provide

greater resolution and accuracy than microarrays (Marioni et al., 2008). However, RNA-seq is a complicated, multistep process involving sample preparation, amplification, fragmentation, purification and sequencing (Mortazavi et al., 2008).

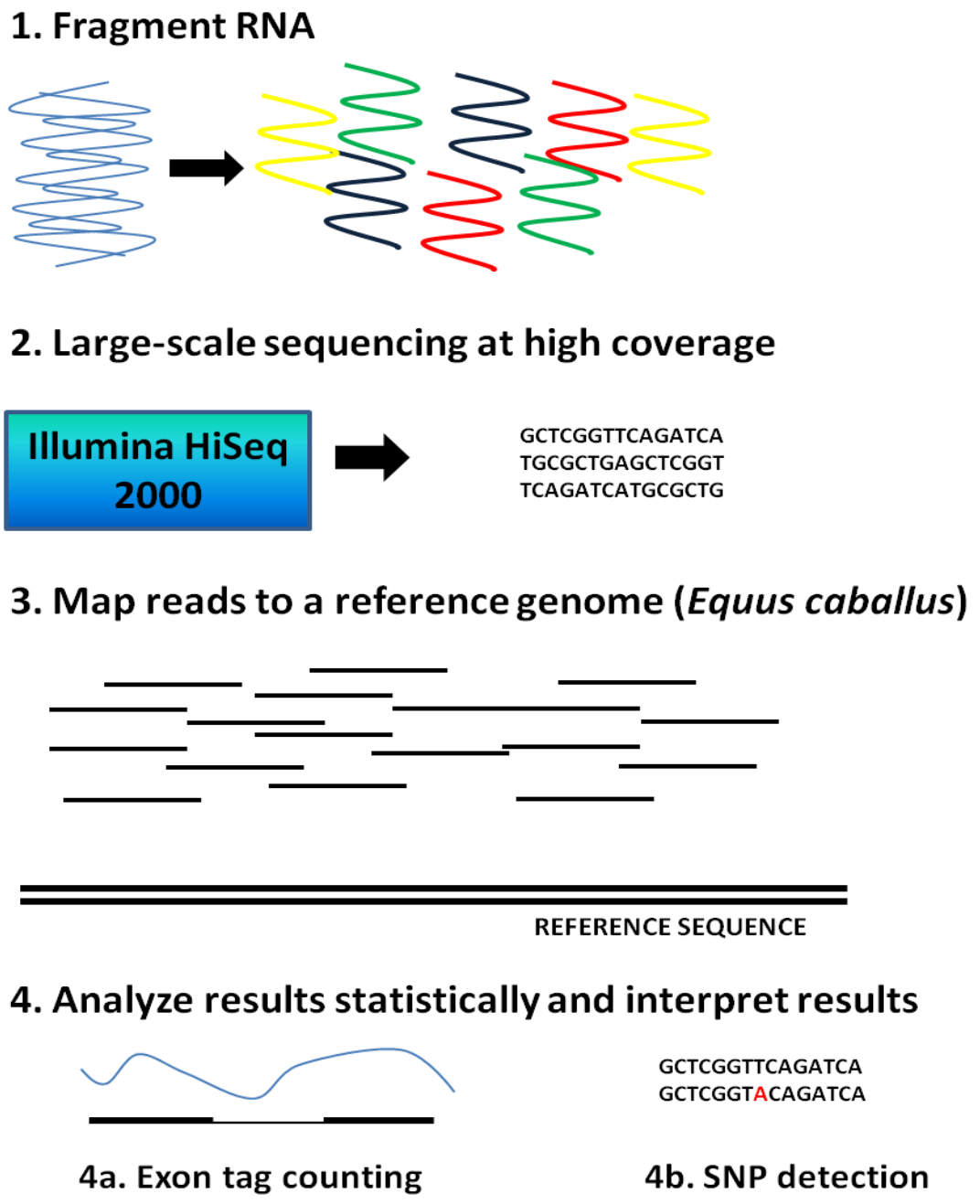



**Figure 1.16. Diagrammatic scheme of an RNA-Seq pipeline.** RNA is fragmented into smaller pieces (1). Libraries are constructed from the fragments and sequenced at a high coverage (2). The sequenced reads are aligned to a reference genome (3) and the results are analyzed statistically and interpreted (4). Depending on the specific application, reads may be counted across genes (4a), SNPs detected (4b) or other analyses carried out such as splice variant detection.

### 1.6.2. RNA-Seq in the study of cartilage

Whilst a number of studies have used microarrays to investigate cartilage physiology (Diaz-Prado et al., 2012) and arthritis (Geyer et al., 2009; Loeser, 2010) there are currently no published studies using RNA-Seq to interrogate cartilage.

# Manuscript 1

MJ Peffers[1*], RJ Beynon[2], DJ Thornton[3], JS Selley[3], PD Clegg[1]

**Corresponding Author**: M J Peffers[*]

[1]Dept.of Musculoskeletal Biology, Institute of Ageing and Chronic Disease, University of Liverpool, Leahurst, Chester High Road, Neston, Wirral, CH64 7TE; peffs@liv.ac.uk, pclegg@liv.ac.uk

[2] Proteomics Group, Institute of Integrative Biology, University of Liverpool, Biosciences Building, Crown Street, Liverpool, L69 7ZB; rbeynon@liv.ac.uk

[3] Wellcome Trust Centre for Cell Matrix Research, Michael Smith Building, Faculty of Life Sciences, Oxford Road, Manchester, M13 9PT , UK; dave.thornton@manchester.ac.uk, j.selley@manchester.ac.uk

**Title: Proteomic Characterisation and Quantification of an *In-vitro* Early Equine Inflammatory Model**

**Running Title**: Proteomic Quantification of Equine Cartilage Secretome

**Key Words:** osteoarthritis, cartilage, proteomics, equine, glycolysis, secretome

**Proteomic Characterisation and Quantification of an *In-vitro* Early Equine Inflammatory Model**


## Abstract

**Introduction:** Osteoarthritis (OA) is characterized by a loss of extracellular matrix (ECM) which is driven by catabolic cytokines. Proteomic analysis of the OA cartilage secretome, enables the global study of this important class of molecules with roles in numerous pathological mechanisms. Previous studies have identified profiles of secreted proteins, but there has been a lack of quantitative proteomics techniques implemented which would enable further biological questions to be addressed. The aim of the study was to identify and quantify secreted proteins in an equine cartilage explant model of inflammation.

**Methods:** An *in-vitro* model of cartilage degradation using stimulation of cartilage explants (n=8) with interleukin-1 (IL-1) enabled us to identify and quantify reproducible data sets of defined proteins using a label-free liquid chromatography mass spectrometry (LC-MS/MS) based strategy. Cartilage explant supernatants were analyzed by one dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis (1D-SDS-PAGE) followed by in-gel tryptic digestion to assess quantitative/qualitative differences in protein profiles. In-solution tryptic digestion of explant supernatants were analyzed using an LTQ-Orbitrap Velos and Progenesis™ LC-MS software enabled label-free quantification. The resulting gene list was used for gene ontology, pathway enrichment analysis and protein network analysis for up-regulated and down-regulated proteins.

**Results:** A total of 248 proteins were identified; 100 differentially expressed in the model. 13% of down-regulated proteins were attributed to 'ECM' and 15% of up-regulated proteins to 'glycolysis' using gene ontology. Novel proteins clathrin and LIM and SH3 domain protein-1 were identified for the first time in cartilage proteomics.

**Conclusion:** A highly sensitive proteomic comparison together with insightful data mining enabled us to identify proteins and pathways involved in early OA which could aid the development of early OA diagnostic markers and therapeutics.

## Introduction

The surfaces of long bones within diarthrodial joints are lined with articular cartilage, an avascular connective tissue that provides a nearly frictionless bearing surface for transmitting and distributing mechanical loads between the bones of the skeleton (Mow et al., 1992). Articular cartilage is composed of a single cell type, the chondrocyte (Archer and Francis-West, 2003), embedded within an extracellular matrix (ECM) and its unique load bearing properties are dependent upon its structural composition and organisation, particularly the interactions between collagens and proteoglycans (Poole et al., 2001). Progressive degeneration of articular cartilage leads to joint pain and dysfunction that is clinically identified as osteoarthritis (OA). Under normal circumstances, there is equilibrium between matrix deposition and degradation which is disrupted in OA leading to the excessive

digestion and progressive loss of important matrix components, especially aggrecan and collagens.

The degradation of the ECM in OA is driven by catabolic cytokines released by chondrocytes and synoviocytes which include IL-1β (Goldring and Goldring, 2004), nitric oxide (Chapman et al., 2001) and tumour necrosis factor (TNFα) (Fernandes et al., 2002). These induce chondrocyte apoptosis (Nelson et al., 1998), and the proteolytic enzymes matrix metalloproteinases and aggrecanases (A Disintegrin and metalloproteinases with Thrombospondin Motifs; ADAM-TSs) which favour degradation, whilst matrix synthesis is limited and so cannot compensate accordingly (Sandell and Aigner, 2001).

Proteomic studies have allowed the investigation of the functional molecules of cartilage to elucidate the pathogenesis of arthritis (Iliopoulos et al., 2010; Wilson et al., 2009). Even so, knowledge of the underlying pathological processes still requires greater research efforts as treatment is primarily symptomatic in the form of pain relief, with no effect on disease progression. Additionally, in this highly prevalent disease there are limited early diagnostic techniques, whilst those present lack sensitivity. There have been proteomic analyses of the osteoarthritic cartilage secretome (Catterall et al., 2006; Polacek et al., 2010a; Stevens et al., 2008). The cartilage secretome is composed of proteins in the media surrounding the chondrocyte or explants and may be proteins that are secreted or shed from the cell surface, in addition to intracellular proteins released into the media due to cell lysis, apoptosis or necrosis. In cartilage explant studies proteins released into media, by chondrocytes and ECM are similar to proteins released *in-vivo* in cartilage

degradation (Wilson et al., 2009) and data from these studies has improved our understanding of OA pathogenesis (Fosang et al., 2008). Our work has focused on the cartilage explant secretome as this approach preserves the three dimensional background allowing chondrocytes to be retained within the ECM; essential for the chondrocyte phenotype. Additionally, the ECM provides native substrates for proteolysis and protein release, equivalent to the shedding of the proteins into the synovial fluid during cartilage degeneration. Initial discovery experiments identified proteins released into the media, which represent a combination of new protein synthesis and secretion, as well as active and passive loss of pre-existing proteins within cartilage. Furthermore, catabolic stimulation of the explants, an accepted method of studying matrix metabolism in experimental investigations of OA *in-vitro* (Arner et al., 1998; Goldring and Goldring, 2004) was employed.

Relative quantification of the cartilage secretome have been undertaken in studies incorporating sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and analysis by liquid chromatography tandem mass spectrometry (LC-MS/MS) (Stevens et al., 2008), two dimensional electrophoresis (2-DE) and mass spectrometry (MS) (De Ceuninck et al., 2005; Hermansson et al., 2004), and more recently by in-solution trypsin digestion of the cartilage explant supernatant followed by LC/MS/MS (Clutterbuck et al., 2011). Quantitative proteomic analyses of cartilage tissue and its secretome are fundamental to the understanding of protein dynamics and for development of diagnostic methods of early disease. Furthermore reliable methods to accurately quantify differentially expressed

proteins allow the identification of potential biomarkers in disease (Simpson et al., 2009). Recently MS approaches to quantification have been applied to cartilage studies using stable isotope labelling with amino acids in cell culture (SILAC )(Polacek et al., 2010a). These relative quantification studies allow the distinction by molecular weight using amino acid labelling during the growth of different cell populations.

In this work we applied relative quantification with a label-free approach to the cartilage secretome. An *in-vitro* model of degradation using stimulation of cartilage explants with IL-1β enabled us to identify and quantify reproducible data sets of defined proteins.

## Materials and Methods

### Cartilage isolation and explant culture

Full thickness equine articular cartilage was harvested from the entire surfaces of the metacarpophalangeal joints of eight skeletally mature horses aged between 8 and 14 years, with grossly normal joints under sterile conditions obtained from an abattoir. Ethical approval was not required for the study as samples were deemed waste products of meat processing. Cartilage was diced into explants approximately $2mm^2$, mixed and placed in complete medium [Dulbecco's modified Eagle's medium (DMEM), supplemented with foetal calf serum (10% v/v), 100U/ml penicillin, 100U/ml streptomycin (Invitrogen, Paisley, UK) 500 ng/ml amphotericin B (BioWhittaker, Lonza, USA). Explants were washed twice with serum-free DMEM (to

deplete serum and synovial proteins) and allowed to equilibrate in complete medium for 24h at 37$^{o}$C in 5% $CO_2$ in 12 well plates (2ml/well). Media was then replaced with serum-free DMEM prior to incubation, supplemented were applicable with human recombinant IL-1β (10ng/ml; R&D Systems, Abingdon, UK) to induce cartilage degradation or DMSO (Il-1β diluent) as a control. The media (with and without IL-1β) was exchanged 48h after initiation of treatment, and cultures harvested after 96h. In this study 48h and 96h supernatant samples were pooled, following the addition of protease inhibitors (Complete protease Inhibitors, EDTA-free, Roche, Lewes, UK) and stored at -80$^{o}$C prior to downstream analysis thus representing the total secretome over 96h. Protein concentrations of supernatants were estimated by Bradford assay (Thermo Scientific, Rockford, USA ). Cartilage explants were lyophilized in order to obtain a tissue dry weight for normalisation.

The viability of the chondrocytes within the explants treated for 96h with or without IL-1β was verified with trypan blue staining in a separate experiment (n=3). Cells were isolated from the explants using collagenase as previously described (Peffers et al., 2010). Isolated chondrocytes were counted and tested for cell viability employing the trypan blue (Gibco, Paisley, UK) exclusion assay using 0.4% trypan blue. Statistically significant differences between chondrocyte viability of control and treated cultures were analysed using paired Students T test following normality testing. The analyses were undertaken using SPSS (IBM, Hampshire, UK).

**Western blot analysis of the secretome**

Western blotting with matrix metalloproteinase 3 (MMP-3) was used as a complimentary methodology to validate the early OA model. Volumes of cartilage explant supernatant from all donors and conditions were adjusted to represent equal dry weights of cartilage. Human recombinant MMP-3 (Merckt, Darmstadt, Germany) was used as a positive control. Samples were heated to 80$^{o}$C for 10 min in NuPAGE® LDS sample buffer (Invitrogen, Paisley, UK) and electrophoresed for 1h at 200V under reducing conditions on Novex 4-12% SDS-PAGE gels (Invitrogen, Paisley, UK). Protein transfer to nitrocellulose was performed using the Invitrogen X Cell Sure Lock apparatus according to standard protocol. Membranes were blocked with TBS (pH 7.4) containing 0.1% Tween-20 (Invitrogen, Paisley, UK) (TBST) and 5% dried skimmed milk for 1 h at room temperature. A goat polyclonal MMP-3 primary antibody (Abcam, Cambridge, UK) was diluted to 1:1000 in milk powder/Tween and added to the membrane for overnight incubation at 4°C. Following washing in TBST, membranes were incubated for 1h at room temperature with the secondary antibody conjugated to horseradish peroxidise (HRP); polyclonal rabbit anti-goat IgG HRP (Abcam, Cambridge, UK) at 1:5000 diluted with TBST containing 5% dried skimmed milk. Chemiluminescence was used to detect the protein bands using Western Lightning™ and Western Lightning Plus Chemiluminescence reagents (Perkin Elmer, Beaconsfield, USA). ImageJ software (http://rsbweb.nih.gov/ij/) was used to quantify bands using densitometry. The relative intensity of IL-1β treated samples were compared to control for each donor.

**1-D SDS PAGE separation and in-gel trypsin digestion**

Cartilage explant supernatants using 10µg protein loading per lane from four individual donors were analyzed by one dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis (1D-SDS-PAGE) followed by in-gel tryptic digestion to assess quantitative/qualitative differences in protein profiles, as previously described (Peffers et al., 2012).

**Protein identification by linear ion trap quadruple (LTQ) mass spectrometry**

Peptides generated from each gel digest were prepared for LC-MS/MS on a Dionex Ultimate 3000 coupled to a Thermo Electron LTQ as previously described (McClean et al., 2007). Raw spectra were converted to Mascot generated files (MGF) using Proteome Discoverer software (Thermo, Hemel Hempstead, UK). The resulting MGF files were searched against the Swiss-Prot databases (version 9, taxonomy; mammalian) sequence databases using an in-house Mascot (Perkins et al., 1999) server (Matrix Sciences, London, UK). Search parameters used were; peptide mass tolerances 1.8Da, fragment mass tolerance of 0.6Da, 1+, 2+ and 3+ ions, missed cleavages; 1, and instrument type ESI-TRAP. Modifications included were; fixed; carbamidomethyl cysteine and variable; oxidation of methionine.

**In-solution tryptic digestion and mass spectrometry using linear ion-trap Orbitrap mass spectrometer (LTQ-Orbitrap Velos)**

Cartilage supernatants for each individual donor for each experimental condition were supplemented with a digestion enhancer; 1% (w/v) Rapigest (Waters,

Manchester, UK) for 10min at 80$^{o}$C in 25mM ammonium bicarbonate. Protein samples were reduced and alkylated with 3mM DTT (60$^{o}$C for 10 minutes) and then 9mM IAA (30 min in the dark at room temperature) with trypsin at a ratio of 1:50 protein: trypsin ratio overnight at 37$^{o}$C. Rapigest was inactivated by incubating for 45min at 37$^{o}$C with trifluoroacetic acid (VWR International) to a final concentration of 0.5% (v/v). The soluble phase was retrieved following centrifugation at 13500rpm for 15 minutes and used for LC/MS/MS.

LC/MS/MS analysis was performed on an aliquots of tryptic peptides from individual digests, in duplicate or triplicate using a nanoAcquity$^{TM}$ ultraperformance LC (Waters, Manchester, UK) in line with a LTQ-Orbitrap Velos (Thermo-Fisher Scientific, Hemel Hempstead) as previously described (Claydon et al., 2012).

**Label-free peptide quantification**

The Thermo raw files of the acquired spectra were analysed by the Progenesis™ LC-MS software (version 3.1.4003, Nonlinear Dynamics) for label-free quantification. Progenesis™ LC-MS takes profile data of the MS scans and transforms them to peak lists. One sample was selected as a reference after checking the 2-D mapping (*m/z* versus retention time), and the retention times of the other samples within the experiment were aligned. Features without the 1+, 2+, 3+ and 4+ charge and isotope peaks of ≤2 were excluded from further analysis. Samples were divided into the appropriate groups using between subject design (between horse variation, between control and IL-1β treatment variation and between technical replicate

variations). Feature raw abundances were normalised which corrects for factors due to experimental variation (such a sample load).

Following feature picking the top 3 spectra for each feature were exported from Progenesis™ LC-MS and utilized for peptide identification with a locally implemented Mascot server; Mascot (version 2.2) in the Ensembl database for horse (*Equus caballus*; EquCab2.56.pep (ftp://ftp.ensembl.org/pub/current_fasta/equus_caballus/pep/) containing 22,640 protein sequences. Search parameters used were: 10ppm peptide mass tolerance and 0.6Da fragment mass tolerance, one missed cleavage allowed, fixed modification; carbamidomethylation, variable modifications; methionine oxidation, hydroxylysine, hydroxyproline. Mascot determined peptides with ion scores of 23 and above and only proteins with at least one unique peptide ranked as top candidate were considered and re-imported into Progenesis™ LC-MS software.

Following the import of the Mascot results for quantification, statistical analysis was performed on all detected features using transformed normalized abundances for one-way analysis of variance (ANOVA). The total cumulative abundance was calculated by summing the abundances of all peptides allocated to the respective protein. All peptides (with Mascot score >23 and $p<0.05$) of an identified protein were included and the protein *p* value (one-way ANOVA) and q-value (*p* value adjusted to false discovery rate (FDR)) were then performed on the sum of the normalized abundances for all runs. Adjusted ANOVA values of $p<0.05$ and additionally regulation of >2-fold or < 0.5-fold were regarded as significant.

**Submission of MS data to PRIDE**

Progenesis™ LC-MS generated files used to search Mascot provided fragmentation spectra, protein and peptide identifications using the *Equus caballus* database are available in the PRIDE database (Jones and Cote, 2008) (http://www.ebi.ac.uk/pride/) at the European Bioinformatics Institute under accession numbers 19447 - 19479. The data was converted into PRIDE using the PRIDE Converter (http:// code.google.com/p/pride-converter).

**Gene Ontology, Pathway Enrichment Analysis and Protein Network Analysis**

The gene symbols for each identified protein were searched in the Ensembl database for horse and converted to the gene symbol of the corresponding human orthologue. The resulting gene list was used for pathway enrichment analysis for up-regulated and down-regulated proteins, were the relative cumulative peptide intensities between IL-1β stimulated and control explant secretomes was ≤0.5 (≥ 2.0) and the corresponding adjusted to FDR p value calculated with the ANOVA test was ≤ 0.05. The enrichment analysis tool of the ConsensusPathDB program (version 12) (Kamburov et al., 2009) acquired the enriched pathways from both gene sets.

Protein network analysis was carried out with the Search Tool for Retrieval of Interacting Genes/Proteins (STRING) tool version 8.2 (http:// string.embl.de) (Jensen et al., 2009)on gene symbols for identified up-regulated and down-regulated proteins ($p < 0.05$ and > 2-fold or < 0.5-fold) as well as unregulated proteins ($p > 0.05$ and regulation between 2 and 0.5-fold). The protein interaction maps were created by allowing for experimental evidence in addition to the

predicted functional links; co-occurrence, co-expression, databases, and text-mining. The minimum required confidence was set for the software at 0.7 (high confidence).

Protein classification with gene ontology (GO) was used to identified proteins with gene ontology (GO) term annotations 'secreted', 'extracellular matrix' and 'glycolysis' based on *homo sapiens* with DAVID (DAVID bioinformatics resources 6.7).

## Results

### Chondrocyte viability and model evaluation using western blotting

Chondrocyte viability was unaffected by IL-1β treatment of explants for 96 hours using Students T-test (Table 1).

| Donor | Control | IL-1β |
|---|---|---|
| 1 | 95 ± 1.4 | 93.5 ± 0.7 |
| 2 | 96 ± 0 | 94 ± 1.4 |
| 3 | 97.5 ± 0.7 | 96.5 ± 0.8 |

**Table 1. Cell viability of cartilage explants was unaffected following IL-1B stimulation.** Trypan blue staining was used to determine the chondrocytes viability following culture for 96 hours. Data are given as mean of two technical replicates ± SD.

In order to relate the results of the study to OA, and validate our inflammatory model, MMP-3 protein in the secretome was evaluated with western blotting (Figure 1). There was an increase in MMP-3 expression following IL-1β stimulation in explant supernatant from all donors.

**Comparative analysis by mass spectrometry of proteins secreted from IL-1β-stimulated cartilage explants**

Catabolic stimulation of equine cartilage explants with IL-1β was used as an *in-vitro* model of early OA. The repertoire of proteins secreted by explants in response to IL-1β treatment in comparison to un-stimulated (control) explants was analysed by LC-MS/MS. Media from control and IL-1β-stimulated cartilage explants were analysed by 1D SDS-PAGE and stained with Coomassie Brilliant Blue to reveal their qualitative differences (Figure 2A). The major bands were cut from the gel, digested with trypsin and the peptides identified using LC-MS (Figure 2B). Proteins included in results had a Mascot score >40 with 2 or more identifying peptides, and a confidence interval of 95%.

Although some slight differences in protein loading were observed, as determined by densitometry (data not shown), the appearance of additional bands was observed by 1-D SDS-PAGE and Coomassie staining in all donors following IL-1β stimulation. Both control and IL-1β treated cartilage explant secretomes expressed a number of cartilage matrix proteins such as COMP. Furthermore, when exposed to IL-1β-stimulation proteins associated with catabolic aspects of cartilage matrix turnover such as MMP-3 were increasingly evident. Interestingly, the expression of proteins involved in glycolysis; lactate dehydrogenase, α-enolase and pyruvate kinase, from chondrocytes particularly within the IL-1β-stimulated explants was revealed. These findings suggested a more comprehensive investigation of the equine cartilage secretome was required.

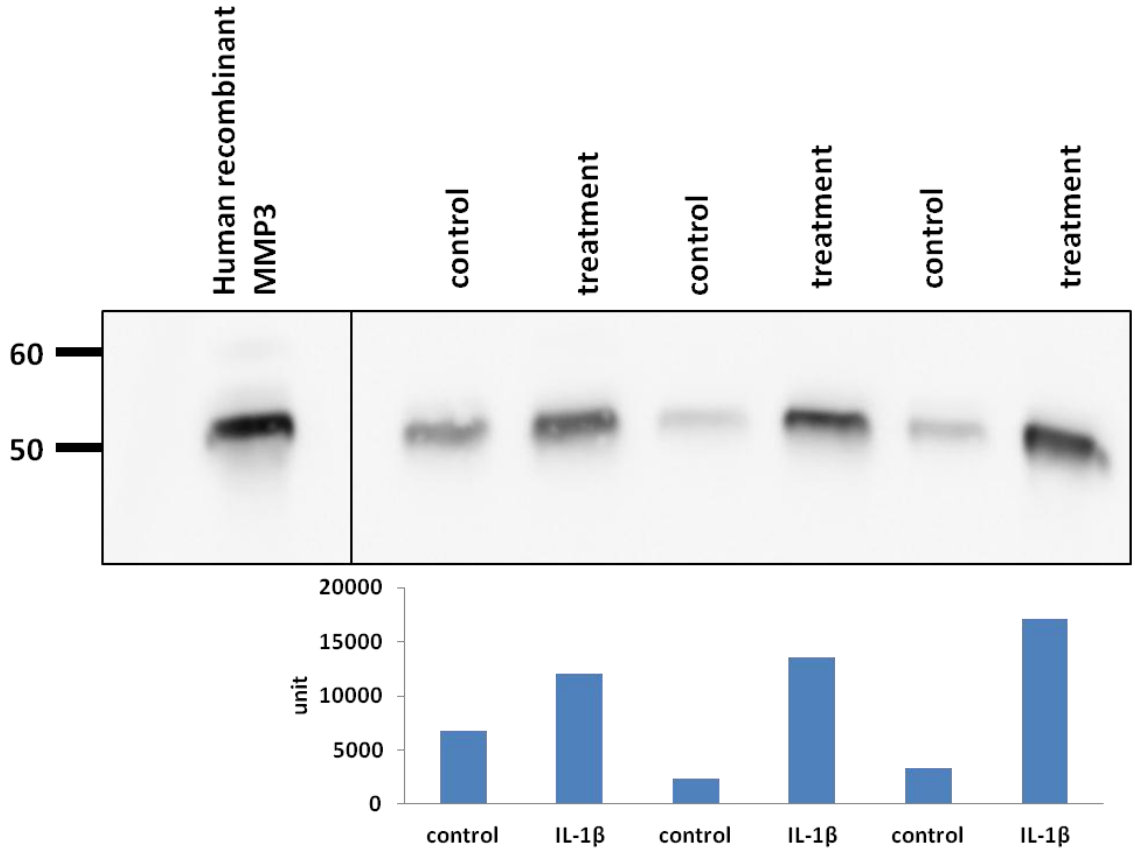



**Figure 1. Cartilage explants stimulated with Il-1β demonstrate increased MMP-3 protein expression**. Western blot analysis using antibodies to MMP-3 of cartilage explant supernatant cultured with and without IL-1β for 96 hours. Representative images of western blots for control and IL-1β treated cartilage explants. Volumes of supernatant loaded were normalised to dry weight of cartilage. Histogram shows relative intensity of each band based on arbitrary units following analysis in ImageJ.

**Label-free relative quantification of the equine OA secretome**

Principal component analysis (PCA) of all identified peptides revealed that the peptides cluster according to the control versus IL-1β treatment with a principal component of 25% (Figure 3A). Quantitative analysis of two technical replicates for each of eight donors with and without IL-1β for each condition was performed using Progenesis™ LC-MS software as described. We identified a total of 248 proteins, of which 224 were identified with 2 or more peptides (Figure 3B). 115 proteins were differentially expressed between IL-1β stimulation and control

(p≤0.05 (adjusted to FDR) and more than 2-fold change), 100 of which had ≥2 peptides. 81 proteins were up-regulated in IL-1β stimulated conditions 68 of these with ≥ 2 peptides, 34 were down-regulated in IL-1β stimulated conditions 32 of these with ≥ 2 peptides (Table 2A, 2B, 2C). Log2 IL-1β treatment/control for some key proteins differentially expressed in this study are depicted in Figure 4. Our IL-1β stimulated explant model was validated by the LC-MS/MS results. Figure 5 illustrates expression profile views of selected proteins produced by Progenesis™ LC-MS. As predicted this demonstrated an increase in cytokine driven proteases and ECM proteins.

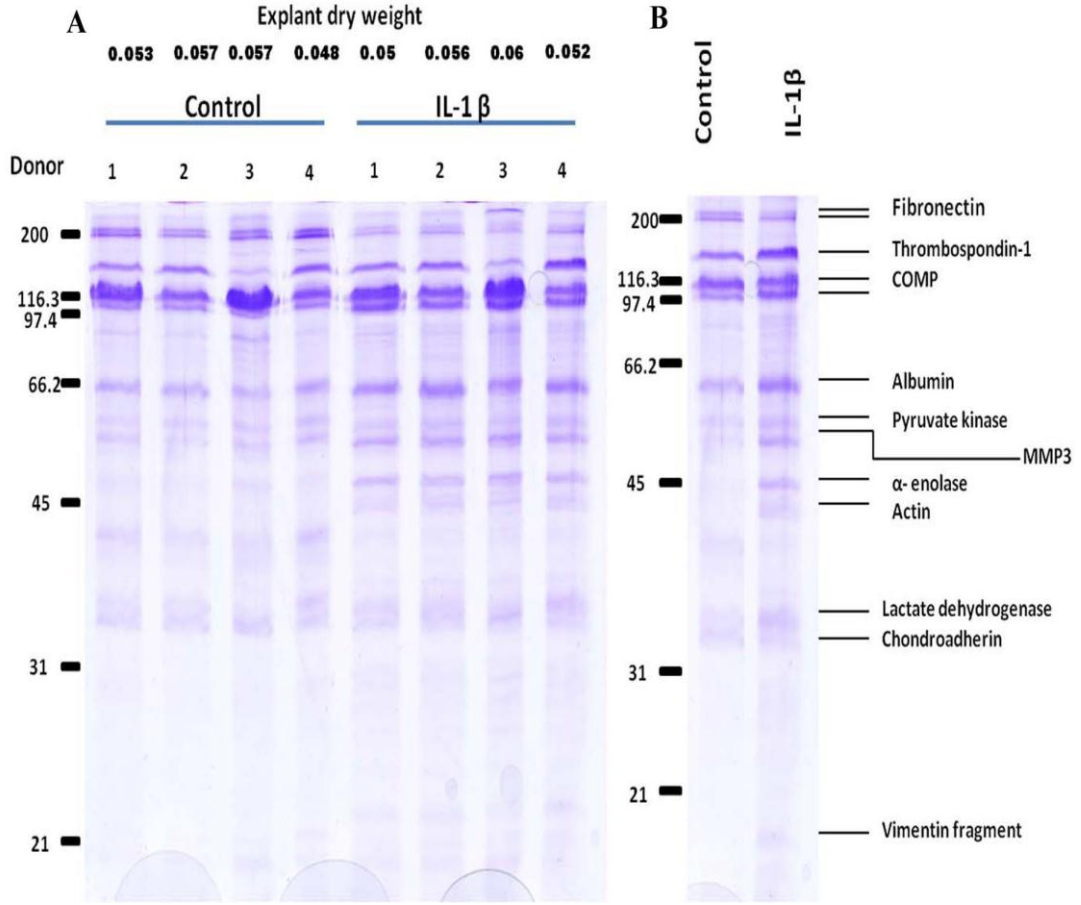



**Figure 2. Analysis of proteins induced in cartilage explants by IL-1β treatment.** A) Equine articular cartilage explants (n=8) were cultured in media supplemented with 10ng/ml IL-1β or un-supplemented media (control). Culture media were collected at 2 and 4 days and pooled for further analysis by SDS-PAGE and staining with Coomassie Brilliant Blue. Equal supernatant protein loading of 10µg of protein per well allowed a qualitative comparison of the secretomes. Dry weight of cartilage explants from each well is indicated on the image which is representative of all donors which produced similar results. B) The differentially abundant proteins in the media marked at the positions of the bands were excised from the gel, trypsin digested, and the protein content of each single band was analysed using peptides identified using LC-MS/MS. Proteins indicated on the gel correlate to the size and are the primary protein identified in that gel analysis.

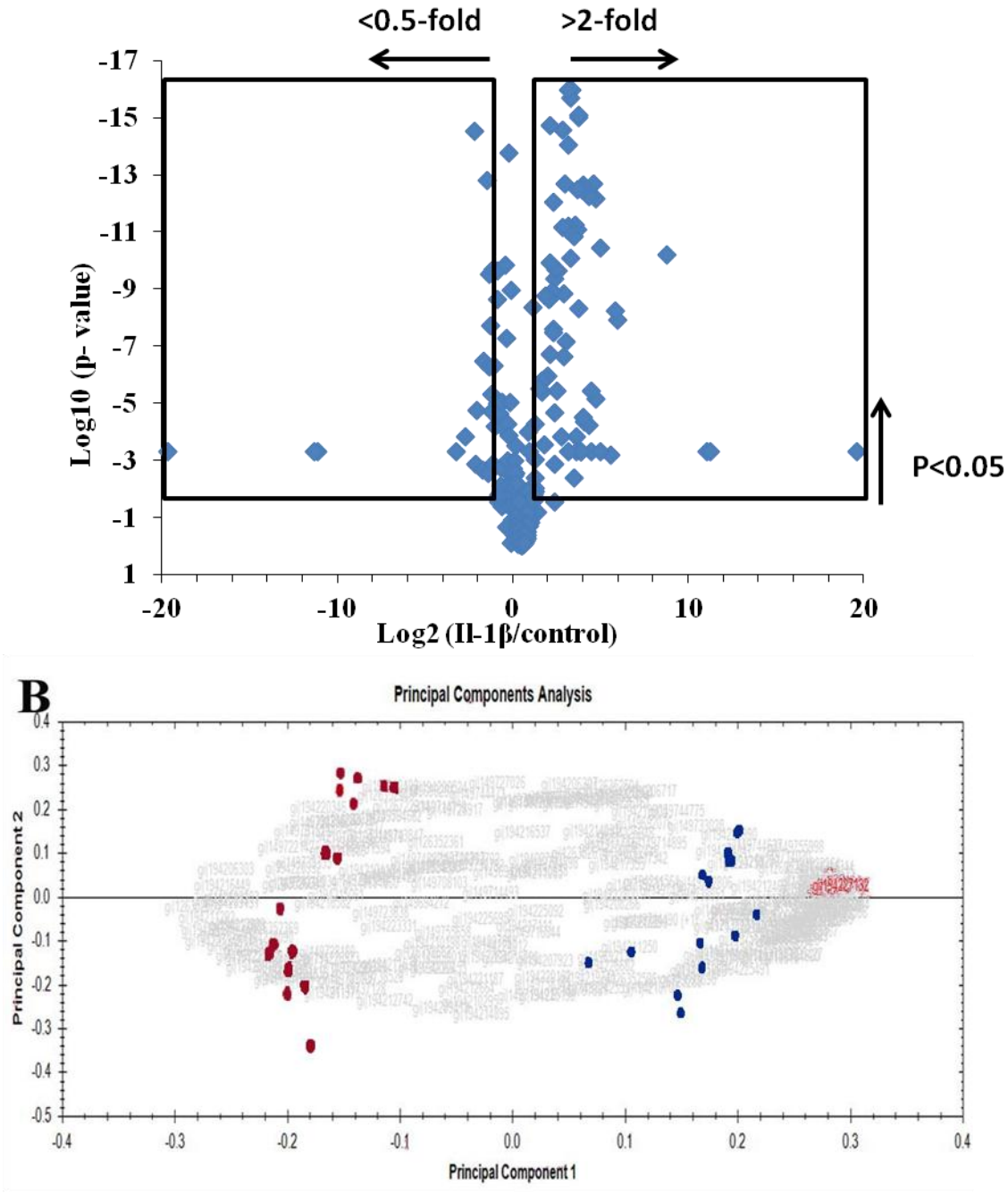



**Figure 3. Label-free quantification of the early OA secretome using Progenesis™.** A) Volcano plot showing differentially abundant proteins (log2 fold change, x axis; log10 p value ANOVA, y axis) from equine secretome control against IL-1β with two or more peptides. Horizontal line indicates p <0.05 and vertical lines indicates between 2 fold and 0.5 fold abundance changes. P-values of 0.000 were set to 0.0005 in order to limit scaling. Black boxes are regulated proteins; between the lines indicate unregulated proteins. B) Principal component analysis clustered the samples into two groups; control (red dots) and IL-1β treatment (blue dots), with a principle component 1 of 25%.

**Table 2A.** Detailed information on the differentially expressed proteins in control and IL-1β treated cartilage explants*

| Accession | Protein description | Peptide count | Anova (p)* | Adjusted p value (q) | Max fold change | Highest mean condition | Log2 T/C | Sub-cellular location | Functional annotation (GO ontology) |
|---|---|---|---|---|---|---|---|---|---|
| gi\|194227196 | Gamma-enolase | 6 | 8.22E-02 | 4.93E-02 | 15.17 | Control | -4.02 | c | glycolysis |
| gi\|149707887 | CSE1 chromosome segregation 1-like | 4 | 2.88E-05 | 3.15E-05 | 9.06 | Control | -3.26 | c,n | export receptor for importin-alpha |
| gi\|149720563 | A-kinase anchor protein | 3 | 1.02E-04 | 1.00E-04 | 8.69 | Control | -3.11 | n | compartmentation protein kinase |
| gi\|149737045 | HHIP-like 2 | 2 | 2.06E-09 | 4.24E-09 | 7.75 | Control | -2.95 | s | carbohydrate metabolism |
| gi\|194214305 | clusterin | 11 | 2.96E-14 | 2.59E-13 | 7.04 | Control | -2.84 | s | apoptosis |
| gi\|149731850 | secreted phosphoprotein 1 isoform 1 | 5 | 9.70E-06 | 1.16E-05 | 5.65 | Control | -2.56 | c,n | chaperone |
| gi\|194225491 | collagen, type X, alpha 1 | 14 | 3.63E-07 | 5.75E-07 | 4.68 | Control | -1.62 | s | extra cellular matrix |
| gi\|194210084 | collagen alpha-1(VI) chain | 13 | 3.78E-07 | 5.75E-07 | 4.35 | Control | -2.02 | s | extra cellular matrix |
| gi\|194223750 | glycogen phosphorylase, brain form | 8 | 2.66E-03 | 2.15E-03 | 3.97 | Control | -1.97 | c | carbohydrate metabolism |
| gi\|194206303 | HLA-B associated transcript 2 | 11 | 5.32E-07 | 4.32E-06 | 3.82 | Control | -1.33 | m | presents antigen to immune system |
| gi\|149743847 | fibulin 1 | 3 | 1.39E-02 | 1.02E-02 | 3.62 | Control | -1.58 | s | extra cellular matrix |
| gi\|126352708 | chitinase 3-like 1 | 10 | 5.34E-06 | 6.77E-06 | 3.60 | Control | -1.87 | s | cellular responce to environment |
| gi\|149704620 | S-methyl-5-thioadenosine phosphorylase | 2 | 7.24E-04 | 6.28E-04 | 3.47 | Control | -1.79 | c | polyamine metabolism |
| gi\|194208128 | angiomodulin, partial | 2 | 1.29E-08 | 2.29E-08 | 3.46 | Control | -1.80 | s | cell growth regulation |
| gi\|194225666 | sperm-membrane associated protein P47 | 4 | 2.58E-10 | 6.28E-10 | 3.45 | Control | -1.83 | m | cell adhesion |
| gi\|149742627 | cofilin-1 | 6 | 3.65E-10 | 8.24E-10 | 3.43 | Control | -1.75 | c | actin polymerisation |
| gi\|149691897 | procollagen C-endopeptidase enhancer 2 precursor | 7 | 3.61E-07 | 5.56E-07 | 3.08 | Control | -1.64 | s | collagen metabolism |
| gi\|126352584 | connective tissue growth factor | 3 | 1.65E-09 | 3.46E-09 | 2.89 | Control | -1.49 | s | Proliferation, differenciation chondrocytes |
| gi\|149722894 | neuroblastoma-amplified protein | 3 | 2.47E-09 | 2.21E-09 | 2.73 | Control | -0.88 | n | unknown |
| gi\|194216449 | Peptidylglycine alpha-amidating monooxygenase | 3 | 9.60E-06 | 1.16E-05 | 2.73 | Control | -1.42 | s,m | protein regulation |
| gi\|194226973 | melanoma inhibitory activity | 2 | 5.54E-06 | 6.93E-06 | 2.66 | Control | -1.44 | s | chemokine |
| gi\|126352530 | AKT interacting protein | 3 | 2.50E-10 | 2.50E-09 | 2.60 | Control | -0.83 | m | K channels |
| gi\|194228715 | exosome complex exonuclease RRP42 | 3 | 1.32E-03 | 1.10E-03 | 2.46 | Control | -1.36 | c,n | exosome complex |
| gi\|126352554 | retinoic acid receptor responder | 2 | 7.96E-05 | 8.11E-05 | 2.44 | Control | -1.33 | m | regulator cell proliferation |
| gi\|126352506 | cullin-3 | 4 | 1.37E-02 | 1.01E-02 | 2.40 | Control | -1.18 | n | scaffold protein |
| gi\|126723289 | tissue inhibitor of metalloproteinase-1 | 5 | 1.49E-02 | 1.08E-02 | 2.33 | Control | -1.30 | s | protease inhibitor |
| gi\|149722142 | cysteine and glycine-rich protein 1 | 4 | 2.15E-02 | 1.52E-02 | 2.33 | Control | -1.29 | n | neuronal development |
| gi\|149705894 | cysteine-rich with EGF-like domains 1 | 3 | 7.73E-06 | 9.54E-06 | 2.18 | Control | -1.08 | s | transport regulator |
| gi\|194213893 | tumor protein p73 | 2 | 8.11E-05 | 8.18E-05 | 2.17 | Control | -1.07 | n | apoptosis |
| gi\|126722914 | immunogobulin gamma 1 heavy chain | 2 | 1.93E-03 | 1.57E-03 | 2.12 | Control | -1.10 | s | innate immune response |
| gi\|126352340 | centromere protein E | 10 | 5.41E-05 | 5.62E-05 | 2.05 | Control | -0.95 | c,n | nucleosome complex |
| gi\|149710093 | leucine-rich repeats and WD repeat domain 1 | 2 | 3.00E-02 | 2.03E-02 | 2.01 | Control | -1.01 | n | chromatin metabolism |
| gi\|149695415 | heat shock 70kDa protein 1-like | 7 | 8.87E-13 | 4.05E-12 | 9.07 | IL-1β | 3.75 | c | anti-apoptotic |
| gi\|126352604 | heat shock protein 90kDa alpha | 3 | 0.00E+00 | 0.00E+00 | 1370.34 | IL-1β | 11.31 | c | anti-apoptotic |

**Table 2B.** Detailed information on the differentially expressed proteins in control and IL-1β treated cartilage explants*

| Accession | Protein description | Peptide count | Anova (p)* | Adjusted p value (q) | Max fold change | Highest mean condition | $Log_2$ T/C | Sub-cellular location | Functional annotation (GO ontology) |
|---|---|---|---|---|---|---|---|---|---|
| gi\|149725588 | heterogeneous nuclear ribonucleoprotein A3 | 2 | 3.31E-13 | 1.66E-12 | 57.34 | IL-1β | 5.75 | c,n | DNA repair |
| gi\|194211344 | high-mobility group box 1 | 2 | 6.54E-04 | 6.50E-03 | 34.04 | IL-1β | 5.62 | n | DNA binding protein |
| gi\|194224640 | coatomer protein complex, subunit gamma | 5 | 0.00E+00 | 0.00E+00 | 20.89 | IL-1β | 4.35 | c | protein transport |
| gi\|194219069 | clathrin, light polypeptide A isoform 1 | 3 | 6.89E-07 | 1.00E-06 | 19.33 | IL-1β | 4.28 | cv,m | protein uptake at the plasma membrane |
| gi\|149712445 | valosin | 3 | 9.20E-06 | 1.12E-05 | 17.61 | IL-1β | 4.25 | c | DNA repair |
| gi\|149733345 | heat shock 70kDa protein 1A | 6 | 8.87E-13 | 4.05E-12 | 14.86 | IL-1β | 3.83 | c | anti-apoptotic |
| gi\|194225092 | phosphoglycerate kinase 1 | 9 | 0.00E+00 | 0.00E+00 | 14.29 | IL-1β | 3.82 | c | glycolysis |
| gi\|149732058 | histone cluster 3, H2bb | 3 | 9.99E-13 | 4.32E-12 | 14.22 | IL-1β | 3.81 | c,n | transcription regulation |
| gi\|149737681 | protein disulfide-isomerase A6 precursor | 2 | 3.86E-14 | 2.90E-13 | 12.99 | IL-1β | 3.68 | c, ER | chaperone |
| gi\|149713810 | endoplasmic reticulum protein 29 | 2 | 2.28E-13 | 1.48E-13 | 10.95 | IL-1β | 4.06 | c, ER | processes secretory proteins |
| gi\|194217031 | thioredoxin | 2 | 4.73E-05 | 4.96E-05 | 10.74 | IL-1β | 3.45 | s,n,c | redox reaction |
| gi\|149727540 | transketolase | 10 | 0.00E+00 | 0.00E+00 | 10.36 | IL-1β | 3.34 | c | transferase |
| gi\|194218400 | xanthine dehydrogenase/oxidase | 4 | 1.03E-12 | 4.32E-12 | 8.98 | IL-1β | 3.16 | s | purine degradation |
| gi\|194208492 | lactate dehydrogenase B isoform 2 | 3 | 6.05E-05 | 6.23E-05 | 8.58 | IL-1β | 3.21 | c | glycolysis |
| gi\|149744840 | solute carrier family 27, member 1 | 2 | 9.92E-05 | 9.82E-05 | 8.21 | IL-1β | 2.94 | m | fluid transport |
| gi\|194205801 | matrix metallopeptidase 1 | 18 | 2.43E-11 | 7.29E-11 | 8.14 | IL-1β | 3.03 | s | protease |
| gi\|149751386 | phosphatidylethanolamine-binding protein 1 | 5 | 6.31E-13 | 3.01E-12 | 7.99 | IL-1β | 2.95 | c | serine protease inhibitor |
| gi\|149705551 | scinderin isoform 1 | 4 | 0.00E+00 | 0.00E+00 | 7.75 | IL-1β | 2.96 | c | cell proliferation and differentiation |
| gi\|194211488 | matrix metallopeptidase 13 | 7 | 8.70E-03 | 6.71E-03 | 7.70 | IL-1β | 2.97 | s | protease |
| gi\|194221214 | beta actin | 6 | 9.99E-16 | 1.31E-14 | 7.26 | IL-1β | 2.83 | c | cell motility |
| gi\|194211629 | phosphoglycerate mutase 1 | 4 | 6.66E-15 | 6.35E-14 | 7.16 | IL-1β | 2.80 | c | glycolysis |
| gi\|194225438 | peptidylprolyl isomerase B | 3 | 4.07E-06 | 5.40E-06 | 7.07 | IL-1β | 2.85 | c, ER | protein folding and transduction |
| gi\|153792484 | alpha-enolase | 19 | 9.51E-12 | 3.22E-11 | 7.03 | IL-1β | 2.76 | c | glycolysis |
| gi\|149728149 | triosephosphate isomerase 1 | 9 | 1.11E-16 | 1.66E-15 | 6.72 | IL-1β | 2.81 | c | glycolysis |
| gi\|194211290 | heterogeneous nuclear ribonucleoprotein D | 2 | 4.58E-11 | 1.30E-10 | 6.59 | IL-1β | 2.64 | c,n | transcription regulation |
| gi\|149716415 | serum amyloid A | 5 | 1.89E-12 | 7.35E-12 | 6.29 | IL-1β | 2.61 | s | acute phase reactant |
| gi\|194208939 | NME1-NME2 protein | 2 | 8.77E-11 | 2.36E-10 | 6.17 | IL-1β | 2.59 | c,n | cell proliferation, differenciation, development |
| gi\|149733335 | vimentin | 23 | 0.00E+00 | 0.00E+00 | 6.06 | IL-1β | 2.58 | cy | intermediate filaments |
| gi\|194221417 | progerin isoform 1 | 12 | 2.55E-15 | 2.80E-14 | 5.99 | IL-1β | 2.60 | c | intermediate filament |
| gi\|149730943 | desmin | 10 | 7.69E-12 | 2.69E-11 | 5.84 | IL-1β | 2.54 | c | intermediate filament |
| gi\|194209022 | eukaryotic translation elongation factor 1 α2 | 2 | 1.34E-11 | 4.40E-11 | 5.74 | IL-1β | 2.50 | c | GDP exchange |
| gi\|194209740 | histone cluster 1, H2ag | 2 | 1.05E-07 | 1.64E-07 | 5.71 | IL-1β | 2.47 | c,n | transcription regulation |
| gi\|194220200 | testis expressed sequence 15 | 5 | 1.68E-12 | 6.76E-12 | 5.65 | IL-1β | 2.54 | n,m | testis specific, several cancers |
| gi\|149709520 | peroxiredoxin 5 | 4 | 3.69E-10 | 8.24E-10 | 5.62 | IL-1β | 2.47 | c | oxidative stress |

**Table 2C.** Detailed information on the differentially expressed proteins in control and IL-1β treated cartilage explants*

| Accession | Protein description | Peptide count | Anova (p)* | Adjusted p value (q) | Max fold change | Highest mean condition | $Log_2$ T/C | Sub-cellular location | Functional annotation (GO ontology) |
|---|---|---|---|---|---|---|---|---|---|
| gi\|126352349 | heterogeneous nuclear ribonucleoproteins A2 | 4 | 1.11E-16 | 1.66E-15 | 5.57 | IL-1β | 2.46 | c,n | RNA processing |
| gi\|149695427 | HNRPA1 protein isoform 1 | 2 | 1.42E-11 | 4.52E-11 | 5.25 | IL-1β | 2.42 | c,n | RNA metabolism |
| gi\|194206498 | annexin A2 isoform 2 | 4 | 3.41E-14 | 2.75E-13 | 5.25 | IL-1β | 2.39 | s | membrane binding protein |
| gi\|194210704 | GDP dissociation inhibitor isoform 2 | 4 | 8.52E-14 | 5.96E-13 | 5.15 | IL-1β | 2.33 | c | regulates GDP/GTP |
| gi\|149701529 | semaphorin III/collapsin-1 | 4 | 4.55E-09 | 8.69E-09 | 4.90 | IL-1β | 2.25 | s | guides axons towards their target, |
| gi\|149712406 | coagulation factor XIII, A1 polypeptide | 3 | 2.42E-09 | 4.88E-09 | 4.82 | IL-1β | 2.31 | s | coagulation |
| gi\|194210816 | fructose-bisphosphate aldolase A | 4 | 1.91E-11 | 5.88E-11 | 4.76 | IL-1β | 2.24 | c | glycolysis |
| gi\|194227132 | catalase | 4 | 9.92E-07 | 1.42E-06 | 4.43 | IL-1β | 2.15 | p | anti-oxidant |
| gi\|149699070 | histone H4 replacement CG3379-PC | 3 | 8.12E-08 | 1.31E-07 | 4.22 | IL-1β | 2.16 | n | nucleosome compnemt |
| gi\|194223331 | zinc finger protein 618 | 2 | 2.44E-05 | 2.70E-05 | 4.21 | IL-1β | 2.43 | n | transcription regulation |
| gi\|194210100 | lamina-associated polypeptide 2 isoform alpha | 6 | 2.95E-02 | 2.01E-02 | 4.10 | IL-1β | 2.03 | n | organisation nuclear envelope |
| gi\|149730340 | TNF receptor-associated factor 3 protein 1 | 4 | 5.75E-12 | 2.15E-11 | 3.76 | IL-1β | 1.97 | c | regulates NFK B and MAPK pathways |
| gi\|194211375 | pyruvate kinase PK-R isoenzyme | 5 | 1.29E-03 | 1.08E-03 | 3.68 | IL-1β | 1.85 | c,n | glycolysis |
| gi\|149713770 | uncharacterized protein C2orf67 isoform 1 | 4 | 1.00E-06 | 1.42E-06 | 3.60 | IL-1β | 1.87 | u | unknown |
| gi\|194226682 | pyruvate kinase 3 | 16 | 2.66E-15 | 2.80E-14 | 3.56 | IL-1β | 1.81 | c,n | glycolysis |
| gi\|194218887 | Parkinson disease protein 7 | 5 | 6.88E-11 | 1.90E-10 | 3.49 | IL-1β | 1.84 | c,n | oxidative stress |
| gi\|194220952 | protein disulfide-isomerase A4 precursor | 4 | 2.26E-04 | 2.16E-04 | 3.48 | IL-1β | 1.81 | ER | chaperone |
| gi\|194217817 | peptidyl-Pro cis trans isomerase isoform 1 | 3 | 1.23E-10 | 3.16E-10 | 3.27 | IL-1β | 1.74 | n | protein folding |
| gi\|194225174 | lipocortin-1 | 5 | 4.30E-08 | 7.16E-08 | 3.26 | IL-1β | 1.67 | c,n | regulates phopholipase A2 |
| gi\|194209131 | ribosomal protein S28 | 2 | 3.08E-09 | 5.99E-09 | 3.20 | IL-1β | 1.63 | c,mi | ribosome |
| gi\|149743864 | non-selenium glutathione phospholipid | 2 | 2.04E-07 | 2.04E-07 | 2.95 | IL-1β | 2.16 | c | anti-oxidant |
| gi\|194222987 | gelsolin | 11 | 1.69E-13 | 1.04E-12 | 2.88 | IL-1β | 1.52 | s | actin modifying protein |
| gi\|194226345 | phosphodiesterase 5A, cGMP-specific | 4 | 3.21E-10 | 7.48E-10 | 2.75 | IL-1β | 1.47 | c | signal tansduction |
| gi\|194221085 | preproalbumin | 6 | 1.03E-06 | 1.43E-06 | 2.70 | IL-1β | 1.43 | s | transporter |
| gi\|194227342 | peroxisomal long-chain acyl-coA thioesterase | 2 | 2.63E-10 | 6.28E-10 | 2.54 | IL-1β | 1.38 | p | fatty acid metabolism |
| gi\|194220937 | matrix metalloproteinase 3 | 16 | 2.83E-04 | 2.68E-04 | 2.45 | IL-1β | 1.26 | s | protease |
| gi\|126723507 | malate dehydrogenase, cytoplasmic | 5 | 4.81E-09 | 9.01E-09 | 2.30 | IL-1β | 1.20 | c | TCA cycle |
| gi\|194221681 | matrilin 3 | 2 | 1.47E-06 | 2.00E-06 | 2.21 | IL-1β | 1.12 | s | extra cellular matrix |
| gi\|149727656 | IMP cyclohydrolase | 7 | 2.00E-05 | 2.28E-05 | 2.06 | IL-1β | 0.98 | c,mi | purine biosynthesis |
| gi\|126352369 | glyceraldehyde-3-phosphate dehydrogenase | 3 | 4.44E-06 | 5.82E-06 | 2.06 | IL-1β | 1.08 | c,n,m | glycolysis |

* These proteins were selected from the 248 proteins differentially identified from the 8 donors. The following criteria were applied; the proteins were identified and quantified with ≥2 unique proteins in all samples,> 2-fold change and p-value anova <0.05 as determined by Progenesis. Log2 T/C indicates Log 2 of Il-1β treatment/control. This data is depicted graphically for some selected proteins in Figure 3. Location: c; cytoplasmic, cy; cytoplasmic vesicle, ER; endoplasmic reticulum, m; membrane, mi; mitochondria, n; nucleus, p; peroxisome, s; secreted.

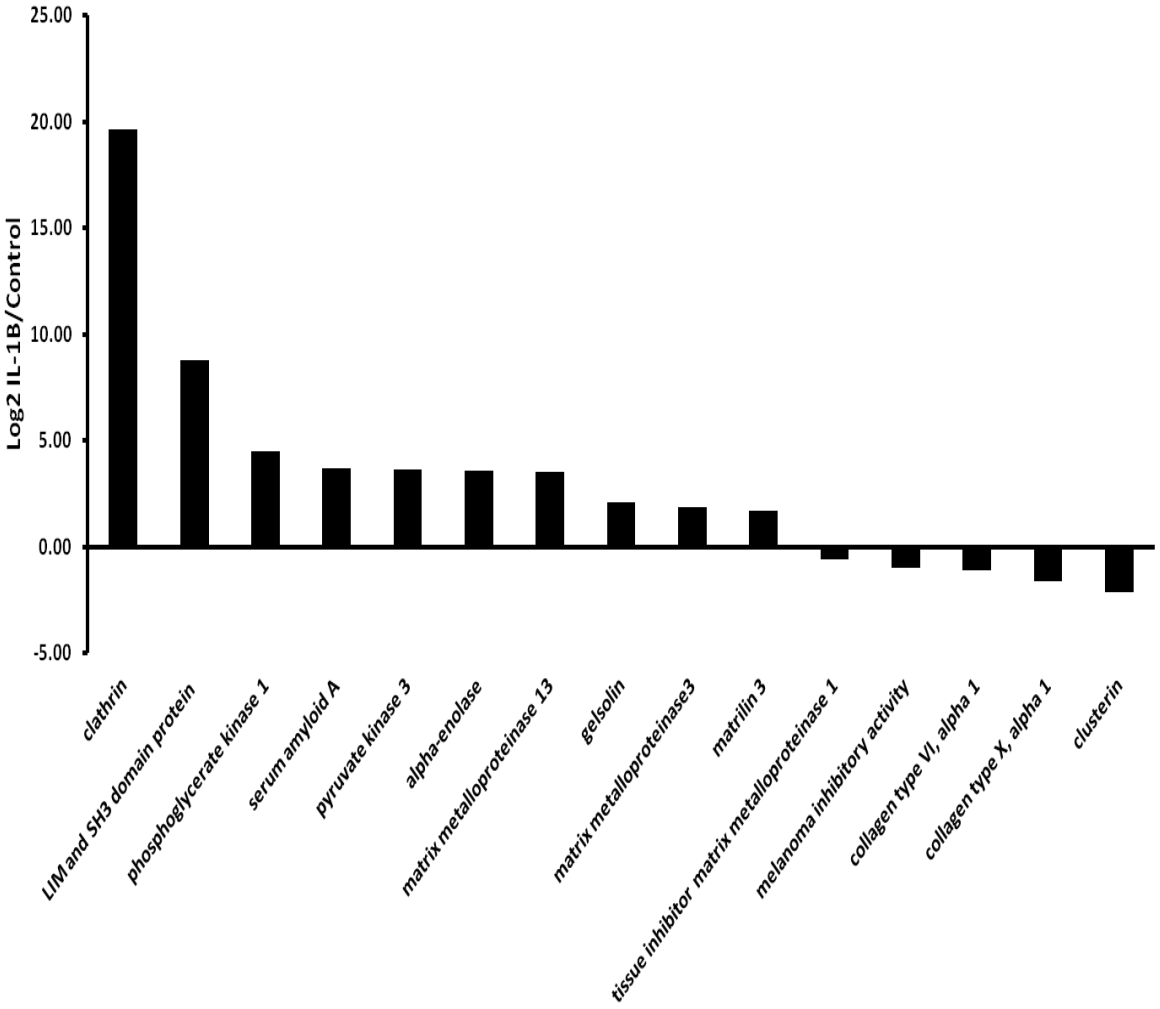



**Figure 4. Differentially expressed proteins in the equine cartilage explant secretome.** Histogram depicts results of label-free relative protein quantification using Progenesis™ software for selected differentially expressed proteins. Changes in protein expression of log2 IL-1β stimulation/control for proteins identified with ≥2 peptides and FDR adjusted p≤0.05 and more than 2-fold regulated are depicted.

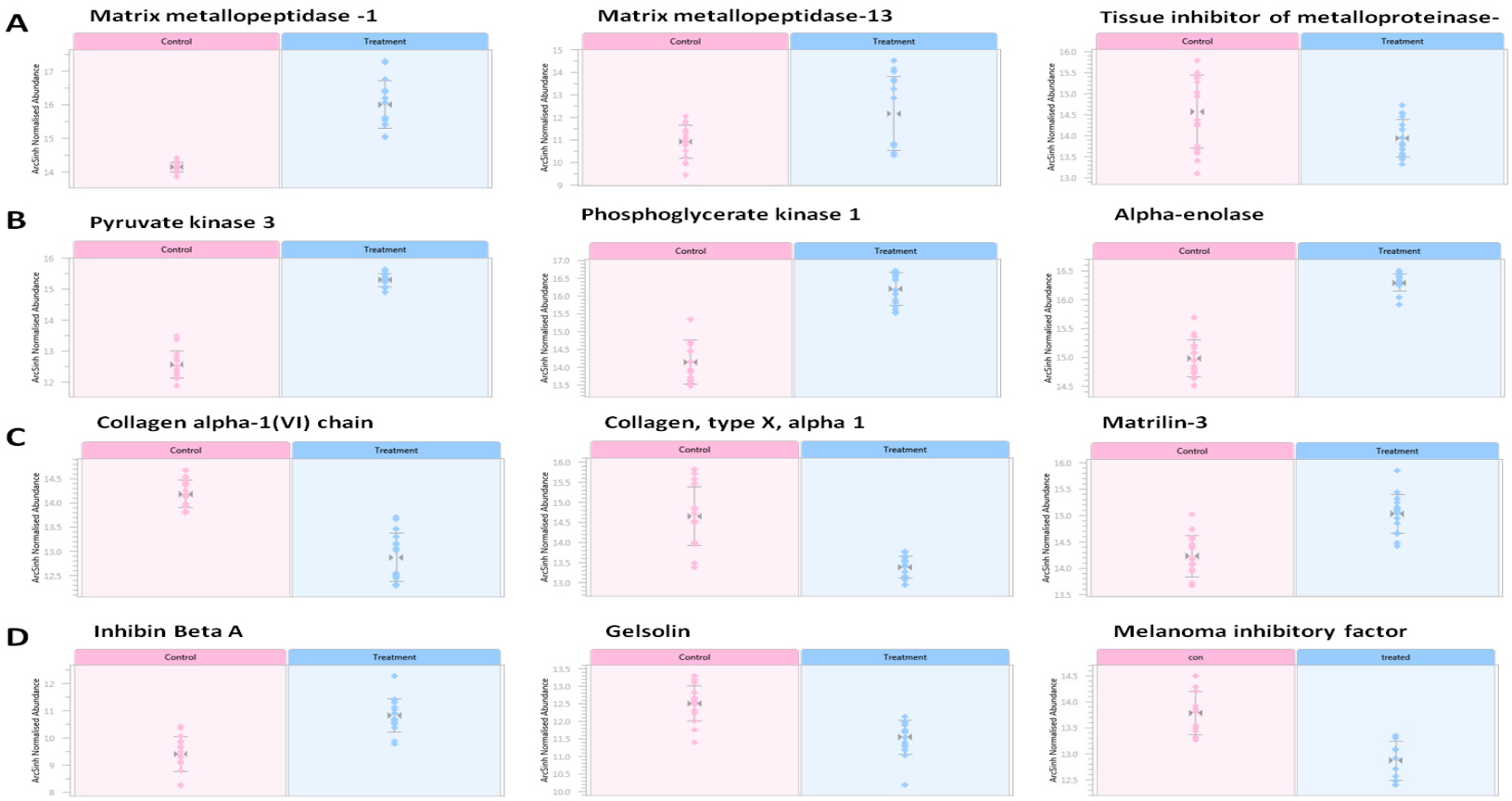



**Figure 5**. Expression profile view of selected proteins produced by Progenesis™ LC-MS. All proteins were identified by 2 or more peptides with greater than 2 fold abundance change and p values <0.01. Control samples are red dots on pink background and IL-1βtreatment blue dots and background. A; cartilage proteases and inhibitors, B; glycolytic proteins, C; extracellular matrix proteins, D; secreted proteins. Plots display the mean arcsinh transformed normalised volume for each group. Error bars demonstrate 3 standard errors within groups.

**Network and pathway enrichment analysis, gene ontology**

Using ConsensusPathDB for pathway enrichment analysis, several significantly ($p \leq 0.01$) up-regulated pathways in the IL-1β stimulated (early OA model) conditions were identified (Table 3). These included 'metabolism of carbohydrates' (Reactome), 'glycolysis and gluconeogenesis' (Wiki pathways, KEGG pathways) and 'glucose metabolism' (NetPath) reflecting altered state of glucose metabolism in our model. No pathways were down-regulated in the IL-1β stimulated secretome.

| Pathway name | Set Size | Measured Genes | p-value | q-value | Pathway Source |
|---|---|---|---|---|---|
| Metabolic pathways - Homo sapiens (human) | **937** | **14** | 0.000122 | 0.00488 | **KEGG** |
| Metabolism of carbohydrates | **129** | **10** | 0.00195 | 0.0391 | **Reactome** |
| Glycolysis and Gluconeogenesis | **47** | **9** | 0.00391 | 0.0391 | **Wikipathways** |
| Glucose metabolism | **62** | **9** | 0.00391 | 0.0391 | **Reactome** |
| EGFR1 | **249** | **8** | 0.00781 | 0.0521 | **NetPath** |
| Glycolysis / Gluconeogenesis - Homo sapiens (human) | **65** | **8** | 0.00781 | 0.0521 | **KEGG** |

**Table 3. Pathway enrichment using ConsensusPathDB pathway enrichment analysis**.

ConsensusPathDB pathway enrichment analysis using proteins which were significantly up-regulated in the IL-1β stimulated explant secretome as identified by a fold change >2 and $p<0.05$ (adjusted to FDR) using Progenesis™. Wilcoxon signed rank test was used to test enrichment analysis.

Protein network analysis followed transformation of proteins identified to non-redundant gene identifier lists of the respective human homologues using STRING. Analyses with the subsets of up-regulated, down-regulated or unregulated proteins in the model were undertaken. Whereas down-regulated proteins, contain only one connectivity between clusterin and TIMP-1 the network produced by STRING

for the subset of up-regulated proteins (Figure 6) contains a connected sub network of several proteins involved in glycolytic pathways such as PKLR (pyruvate kinase LR), TPI1 (triosephosphate isomerise), GAPDH (glutaraldehyde phosphate dehydrogenase), PGAM1 (phosphoglycerate mutase 1), PGK1 (phosphoglycerate kinase 1), ENO1 (alpha1-enolase), ENO2 (alpha2-enolase), LDHA (lactate dehydrogenase), PKM2 (pyruvate kinase, muscle) and MDH1 (malate dehydrogenase 1). STRING analysis of the subset of unregulated proteins resulted in an overall loosely connected network within which two more highly connected sub networks were contained, one containing the structural components of the ECM and the other involving complement pathway proteins.

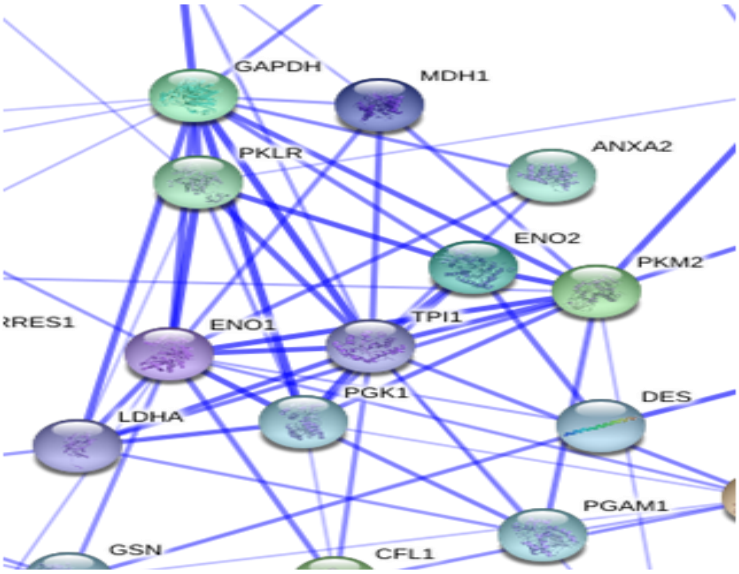



**Figure 6**. STRING generated protein-protein interaction (PPI) map of up-regulated proteins in IL-1β-stimulated cartilage explant secretome ($p < 0.05$ and more than 2 fold up-regulated). Following identification of peptides with Mascot in the Ensembl horse database, identified peptides were merged into proteins and significantly regulated proteins were subjected to protein network analyses using STRING. This allows for experimentally verified and predicted PPIs. High confidence level (0.700) was set in the analysis as the minimum required confidence. The confidence view is shown here. Stronger associations are represented by thicker lines.

Using DAVID gene ontology analysis all genes identified were loaded into the functional annotation chart. We determined that 34% of genes were identified as 'secreted' and 21% were identified as belonging to 'extracellular matrix'. Next we

considered the GO term ‘extracellular matrix’ and ‘glycolysis’ for the sets of genes produced previously by Progenesis™ considered to be differentially regulated (up-regulated, down-regulated) or not changed in regulation. For the term ‘extracellular matrix’ 9% genes were evident in the up-regulated group, 13% of genes in the down-regulated and 28% genes in the unregulated group were defined with this term. When the term GO term ‘glycolysis’ was used 15% of genes from the up-regulated group only were identified.

**Discussion**

The dominance of matrix components within cartilage, where the cell population is less than 1% of tissue volume presents one of many challenges to cartilage proteomic analysis (Schulz et al., 2006). In contrast the secretome provides a more amenable sample for investigation of the pathogenesis of OA and biomarker discovery using mass spectrometry based proteomics. In this study we have explored the quantitative differences found in the equine cartilage media when challenged with IL-1β, resulting from an established model of early OA in both human (Ismaiel et al., 1992) and horse (MacDonald et al., 1992). The horse provided a suitable species for the study of cartilage in health and disease (Arner et al., 1998; Catterall et al., 2006; Clutterbuck et al., 2011; Polacek et al., 2010a) whilst the cytokine stimulated cartilage explant secretome presented an appealing secondary proteome for elucidating pathological pathways involved in matrix degeneration and remodelling. Cartilage proteomics has progressed on from protein identification to quantitation; the protein abundance differences in cartilage under different conditions at a given time (Ong et al., 2003). Many cartilage studies have identified profiles of secreted proteins which have increased

our knowledge of physiological processes in health and disease (Catterall et al., 2006; Clutterbuck et al., 2011; Polacek et al., 2010a; Wilson et al., 2008). These studies have employed relative quantification approaches using platforms including 2-D gel approaches (Catterall et al., 2006; Hermansson et al., 2004), SILAC studies (Polacek et al., 2010a), isobaric tags for relative and absolute quantitation (iTRAQ) (Stevens et al., 2009), and quantitative western blotting (Clutterbuck et al., 2011). Our approach to the determination of quantitation uses label-free techniques, undertaken for the first time to our knowledge, in order to obtain relative quantification for a large set of secreted protein. This approach has the advantages of being economical due to no labelling reagent costs, increased sequence coverage and overall proteome coverage and significantly it enables the use of explants. Biomarker discovery and disease pathways in OA are likely to benefit from this form of global-view quantitation. Indeed sets of proteins with unique expression patterns in response to a particular stimulus, rather than single specific proteins will be involved in pathological processes.

Inflammation has an important role in OA (Fernandes et al., 2002) with inflammatory mediators being secreted by chondrocytes (Gruber et al., 2004). Inflammatory cytokines such as IL-1β stimulate the production of matrix metalloproteinases (MMPs), enzymes that can degrade all components of the extracellular matrix (Goldring and Goldring, 2004). Initial results derived from western blotting and then identified later in the label-free experiments demonstrated that cytokine stimulation indeed provided a model to study differential protein expression in early OA as previously demonstrated. Initial 1D

SDS-PAGE experiments used a qualitative proteomics approach to validate our early OA model. We utilised serum-free media in order to reduce contamination with abundant serum proteins which reduces the identification of lower abundance proteins. In addition pooled, chopped cartilage from skeletally mature horses was mixed in order to minimize potential error caused by the variability in proteoglycan synthesis observed in various sites within joints (Little et al., 2005). Finally as identification of proteins relies on the comprehensiveness of the sequence data for a given species, for our studies, the equine genome was used where possible (Wade et al., 2009). Evidence for our approach is demonstrated by identification of major ECM proteins, plus protease abundance changes following IL-1β stimulation. The data strongly suggested differential protein expression in our model, with induction of some proteins previously identified in proteomic studies in OA identified including alpha-enolase (Ma et al., 2011), MMP-3 (Clutterbuck et al., 2011), lactate dehydrogenase (Ruiz-Romero and Blanco, 2009) and pyruvate kinase (Stevens et al., 2008). Interestingly, in agreement with Haglund *et al.* 2008 (Haglund et al., 2008) results from the 1-D gel indicated that substantial alterations of cartilage matrix components was not part of the initial response to cytokine stimulation.

Next we adopted high-throughput proteomic technologies which create large data sets posing challenges in interpretation. We therefore implemented a number of methodologies to aid in this. The first was the use of a proteomics tool called Progenesis™ LC-MS (Gorman et al., 2009) which allowed us to analyse and compare (in terms of relative quantification), using label-free LC-MS, large sets of data produced by a high mass resolution mass spectrometer. Then we used the

bioinformatic tools DAVID (Huang da et al., 2009), STRING (Jensen et al., 2009), and ConsensusPathDB (Kamburov et al., 2009) in order to interpret the data in relationship to protein functions and identify any central players within our early OA model. As proteins do not act independently within cartilage, the production of interactome networks can provide important information on novel proteins and their networks involved in OA pathogenesis (Iliopoulos et al., 2010). Gene ontology, which uses statistical analysis to validate results has been used previously in proteomic studies of OA (Wu et al., 2007) and was used here to categorise proteins within our early OA model into functional groups.

Data presented following LC-MS/MS of the secretome demonstrated evidence of changing ECM dynamics in response to IL-1β treatment. There was significant up-regulation of important proteolytic enzymes involved in cartilage degradation such as MMP-1, MMP-3 (also validated using western blotting), and MMP-13 (Tetlow et al., 2001) as well as a reduction in TIMP-1 (Martel-Pelletier et al., 1994). Interestingly collagens identified as significantly altering in the media; collagen VI alpha 1 and type X alpha 1 were present in lower levels following IL-1 β treatment. This was in accordance with previous findings in IL-1β stimulated rabbit chondrocytes (Yudoh et al., 2007), human chondrocytes (Shakibaei et al., 2007) and human fibroblast (Nawrat et al., 2005) where a reduction in collagen biosynthesis was evident. In our model we are demonstrating either the cartilage anabolic properties of IL-1β or a lack of release of collagen into the media. Contrastingly, prolonged periods of IL-1β treatment (greater than three weeks) of cartilage

explants cultures have revealed an increase in collagen in the media (Homandberg et al., 1997; Matsuka et al., 1997).

A number of our findings were consistent with previous *in-vitro* studies; Inhibin–Beta A, a member of the transforming growth factor β group, thought to be anabolic for cartilage and previously induced by IL-1 treatment of chondrocytes (Hermansson et al., 2004) was induced in our model. Furthermore, the actin-capping protein gelsolin result concurred with other IL-1 stimulated explant studies were a reduction of the protein in explant media was noted (Wilson et al., 2008). Our findings together with an increase in other cytoskeletal proteins such as β-actin, LIM and SH3 domain protein 1, cofilin-1, desmin and progerin provide additional evidence of cytoskeletal rearrangement previously identified in arthritic cartilage (Lambrecht et al., 2008). Interestingly the non-collagenous oligomeric matrix protein matrilin-3 was also found to be regulated in the media following IL-1β treatment. Our data is in agreement with others findings; matrilin-3 has been demonstrated in advanced OA synovial fluid (Vincourt et al., 2012) and its expression up-regulated in human articular chondrocytes in OA (Pullig et al., 2002a). It has been hypothesised that this is due to a cellular response to the modified microenvironment in OA. In contrast, 2-DE mouse studies found that although retinoic acid (RetA) treatment of mouse cartilage explants increased matrilin-3, Il-1β treatment did not regulate it. Our model used skeletally mature cartilage whereas Wilson *et al.* (Wilson and Bateman, 2008) used 24 day old mouse tissue and the either differences in cartilage maturity, or species differences may affect the response of cartilage to cytokine stimulation.

Data from label-free studies confirmed, using the GO category 'extracellular matrix', that only six genes were differentially expressed according to our criteria, and whilst only matrilin-3 was up regulated, five were down regulated; tissue inhibitor of metalloproteinase (TIMP-1), collagen VI alpha-1 (COL6A1), collagen X alpha-1 (COL10A1), chitinase 3-like 1 and fibulin-1. TIMP-1 is a protease inhibitor capable of inhibiting the activities of matrix metalloproteinases. TIMP-1 protein has been identified as being up-regulated in OA subchondral bone (Hulejova et al., 2007) human articular chondrocytes (HAC), HAC explant culture (Polacek et al., 2010b) and in the IL-1β-stimulated equine explants (Clutterbuck et al., 2011) cultures and HAC secretome in a recent SILAC study (Calamia et al., 2011). However a HAC explant study comparing OA and with normal patients indicated that TIMP-1 synthesis was unchanged in arthritic explants (Martel-Pelletier et al., 1994; Selsted et al., 1996). A further study demonstrated that IL-1β stimulation has a marked inhibitory effect on TIMP-1 expression by chondrocytes (REF?).

COL6A1 is exclusively distributed in the pericellular matrix (Poole, 1997) and is involved in regulating the mechanical environment of chondrocytes, with the pericellular matrix of COL6A1 knockout mice revealing a reduction in mechanical properties (Alexopoulos et al., 2009). Interestingly a HAC SILAC study demonstrated an increase in COL6A1 in the HAC secretomes (Polacek et al., 2010b) and analysis of the OA human articular cartilage proteome revealed an increase in type VI collagen compared to normal (Guo et al., 2008). It has been suggested that type VI may have a protective role for chondrocytes (Peters et al., 2011). Thus this increase may be a

protective response by the chondrocyte to cytokine stimulation. The role of COL10A1 in OA is controversial (Eyre, 2002). COL10A1, produced by hypertrophic cartilage, is normally restricted to a thin layer of calcified cartilage were it is adjacent to bone. To our knowledge it has not previously been identified in the cartilage secretome. Interestingly both chitinase 3-like 1 (a matrix regulating agent) and fibulin-1 (which binds to proteoglycans and is a co-factor of ADAMTS-1) have been previously identified as increased in OA cartilage proteomes (Wu et al., 2007). The mechanisms behind the reductions in these matrix proteins following IL1β stimulation are unclear. The systems involved in the reduced turnover evident by these results may include altered matrix binding, reduced synthesis, or better binding of these matrix components with the cartilage ECM.

Pathway enrichment analysis following label-free relative quantification allowed us to group the observed changes between control and our early OA model samples. Among the significantly up regulated pathways were 'glycolysis, 'metabolism of carbohydrates' and 'glucose metabolism'. Furthermore, protein network analysis using STRING on up-regulated proteins revealed a sub-network of proteins involved in glycolytic pathways; PKLR, TPI1, GAPDH, PGAM1, PGK1, ENO1, ENO2, LDHA, PKM2 and MDH1. The GO term 'glycolysis' was also found in 15% of up-regulated genes whereas no genes in down-regulated or unchanged were matched to this term. Although it could be argued that the increase in glycolytic enzymes is due to necrosis or apoptosis there was no significant difference in cell viability in this study between control and IL-1β stimulation. Indeed it has been previously demonstrated that IL-1 alone does not cause apoptosis in chondrocytes (Blanco et al., 1995). This

suggests that one of the principal features in our model is an altered status of glycolysis and glucose metabolism and is in agreement with finding by Guo et al. 2008 (Guo et al., 2008) in OA cartilage. Chondrocytes survive in a hypoxic environment whilst maintaining tissue integrity via matrix production. They are extremely dependent on glucose metabolism to drive the ECM biosynthetic machinery (Peng et al., 2008) requiring anaerobic glycolysis to generate adenosine tri-phosphate (ATP) indeed glycolysis constitutes 95% of their energy production. Adequate ATP is needed for chondrocytes to respond to stress, such as IL1β stimulation (Fontan et al., 2000) thus a high rate of anaerobic glycolysis is essential for this ATP generation. It is hypothesised that in our model glycolysis is increased in an attempt to increase ECM production in the face of MMP-driven degradation. This leads to an increase in the leakage of the up-regulated enzymes involved in this pathway and so ultimately an increase in these proteins is evident within the early OA secretome. These findings intimate that further research into this field could provide help in the development of therapeutics for early OA.

In our label-free study we identify proteins with a <0.5 or >2 fold change to be differentially expressed. However it is possible that proteins out-with these parameters could pose relevance either as biomarkers of early OA or within the pathogenesis of OA. Previous work has established that threshold parameters applied to data sets can affect the conclusions drawn from studies (Miller et al., 1996; Yao et al., 2008). In data analysis we did assess the validity of conclusions by using a range of threshold parameter and examining the interactions of un-regulated proteins within our early OA secretome.

One of the major findings in this study was the discovery of novel and interesting proteins. There can be a reduction in the identification of less abundant proteins of the ECM when cartilage proteomic studies use whole tissue due to the abundance in cellular proteins. The use of the cartilage secretome allowed a less complex mixture of proteins to be analysed and thus a number of proteins up-regulated following IL-1β stimulation to be identified for the first time. Clathrin light chain (CLH1), an intracellular trafficking protein and the actin-binding protein LIM and SH3 domain protein 1 (LASP1) were found to be over-expressed in our early OA secretome, leading to the possibility of their use as early OA biomarkers. Our study is the first to identify these proteins in spent media of cartilage. Clathrin-coated vesicles (CCVs) form at the plasma membrane where they select protein and lipid cargo for endocytic entry into cells. In addition CCVs form at the trans-Golgi network, where they function in protein transport from the secretory pathway to the endosomal/lysosomal system and secretory granule sorting (Sahota et al.; Yang et al., 1999). In addition clathrin light chain has been implicated in actin organisation (Hyman et al., 2006). Previously the cytokine IL-7, which is produced by OA patients (Long et al., 2008) has been identified as inducing clathrin mediated internalisation (Henriques et al., 2010). It may be that here cytokine stimulated cartilage degradation is increasing intracellular trafficking leading to an up-regulation of clathrin, which is then identified in very high abundance in our secretome. Interestingly as no increase in clathrin heavy chain was identified in the secretome it could be that clathrin light chain is acting independently of the heavy chain. An interesting alternative hypothesis is that there is a secretory pathway stimulated by IL-1β that has not yet been identified.

Interestingly LASP1 was one of the most highly up-regulated proteins identified in the study. A direct link between cytoskeletal remodelling and disease development in OA has been previously identified. In addition the RhoA/ROCK/LIMK/cofilin pathway has been implicated in OA pathogenesis (Appleton et al., 2010; Haudenschild et al., 2010). Taken together the substantial increase in both clathrin light chain and LASP1 points to a change in cytoskeletal modelling in our model.

**Conclusions**

Our aim was to identify and quantify reproducible data sets of defined proteins in an establish model of cartilage degeneration in order to ascertain cellular mechanisms of arthritis. This was accomplished by detecting differentially abundant proteins, some of which have a relationship with the pathogenesis of arthritis from previous works; others were seen for the first time providing exciting new targets for the further investigation of OA pathogenesis. High sensitivity in the proteomics comparison in combination with insightful data mining enabled us to identify some pathways involved in early OA which could be investigated in future research. The evidence presented here increases our knowledge of early OA pathophysiology which could aid in the development of early OA diagnostic markers and therapeutics.

**Acknowledgements**

Mandy Peffers was supported by a Wellcome Veterinary Integrated Research Fellowship.

## APPENDIX TO MANUSCRIPT 1

## A quantitative comparison of label-free and absolute methodologies in the quantification of equine cartilage oligomeric matrix protein and fibromodulin

### Introduction

A QconCAT was developed for quantification of the human osteoarthritic secretome. The QconCAT was used to absolutely quantify cartilage oligomeric matrix protein (COMP) and fibromodulin, in the IL-1β stimulated equine cartilage explant secretome as the q-peptides for these proteins were homologous between human and horse. The aim of this study was to assess the absolute quantification of these proteins compared to the label-free approach using Progenesis™LC-MS (Manuscript 1).

### Materials and methods

#### Cartilage isolation, explant culture, and cytokine stimulation

See Manuscript 1.

#### Determining peptide sequence homology

Using the basic local alignment search tool (BLAST®); (http://blast.ncbi.nlm.nih.gov/Blast) homology between human and *Equus caballus* was identified for the q-peptides for fibromodulin and COMP. SRM methodology was used to quantify these proteins.

#### Protein digestion for absolute quantification

100µg protein for each sample of explant supernatant was detergent-treated, reduced, alkylated and trypsin digested as described above. Due to the poor expression of the QconCAT it was not possible to co-digest it with the analyte sample and so previously trypsin digested QconCAT was spiked into the samples. Samples were run in duplicate technical replicates.

SRM experiments were conducted with 500ng of tryptic analyte peptides spiked in with either 10 fmol, 1 fmol or 0.1 fmol heavy QconCAT, was loaded onto column.

**LC-SRM/MS quantification and data analysis**

Transitions and data analysis for SRM studies have been previously described in the manuscript 3.

**Western blot validation**

Western blotting was used to validate the quantification of COMP and fibromodulin in the control and IL-1β stimulated cartilage explant secretomes for each donor. Volumes of cartilage explant supernatant from all donors and conditions were adjusted to represent equal dry weights of cartilage. Using previously described protocols (Manuscript 1) membranes were probed with primary antibodies against the following; mouse polyclonal to fibromodulin (1:2000 dilution, # 67596 Abcam) and rabbit polyclonal to equine COMP (1:1000, a kind gift from Jay Dudhia, RVC, London, UK). Appropriate secondary antibodies conjugated to horseradish peroxidise (HRP); polyclonal rabbit anti-mouse IgG HRP or polyclonal goat anti-rabbit IgG HRP (both Sigma, Dorset, UK) at 1:2000 were used.

**Statistical analysis**

Statistically significant differences in the absolute quantification of cartilage secretomes between control and treated cultures were identified using mixed effects linear regression to allow for donors with significant biological variation. Statistical analyses were undertaken using S-Plus and Excel software.

**Results**

**Peptide sequence homology**

BLAST indicated that human QconCAT q-peptide amino acid sequences from for each of COMP; SSTGPGEQLR and DTDLDGFPDEK and fibromodulin; IPPVNTNLENLYLQGNR, and LYLDHNNLTR were homologous to the equine

sequence. However as LYLDHNNLTR had been determined a 'type C' peptide (see Manuscript 2), due to poor fragmentation of the parent peptide, this could not be used for quantification purposes.

**Absolute quantification of target proteins using QconCAT**

IL-1β stimulation of equine cartilage explants cultures for 4 days significantly increased COMP release into the media in all donors (mean for SSTGPGEQLR; 3628 to 6282pmol/g dry weight, mean for DTDLDGFPDEK; 3111 TO 6032pmol/g dry weight), $p<0.0001$ (Figure 1).

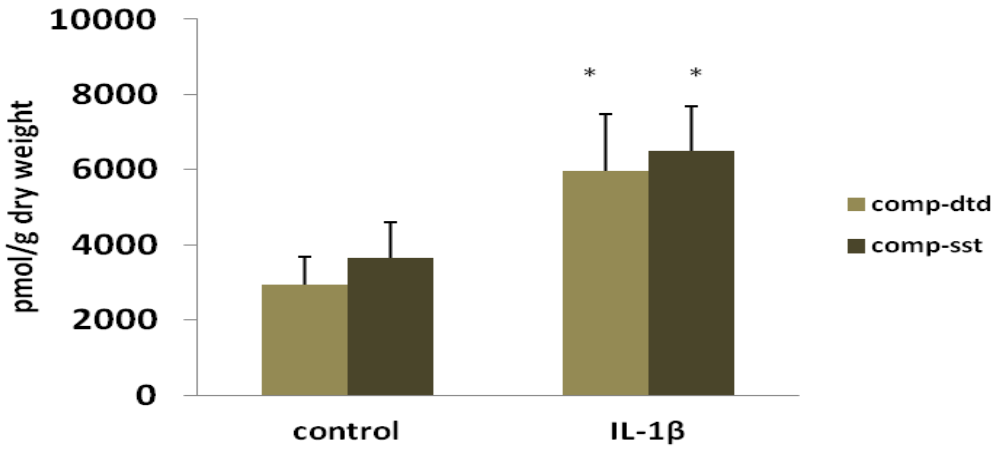



**Figure 1. Effects of 10ng/ml IL-1β treatment for 4 days on COMP protein concentrations in cartilage explant supernatants measured using QconCAT**. Data are represented as pmol/g dry weight. Histograms represent means + SEM. Data evaluated using mixed effects linear regression and * indicates significant difference relative to control. Statistical values for this study were * $p<0.0001$ (n=8).

In the quantification of fibromodulin individual donors produced variable responses to IL-1β stimulation (Figure 2a). Overall there was a non-significant reduction in fibromodulin within the IL-1β stimulated cartilage explant secretome (208 to 186pmol/g dry weight) (Figure 2b).

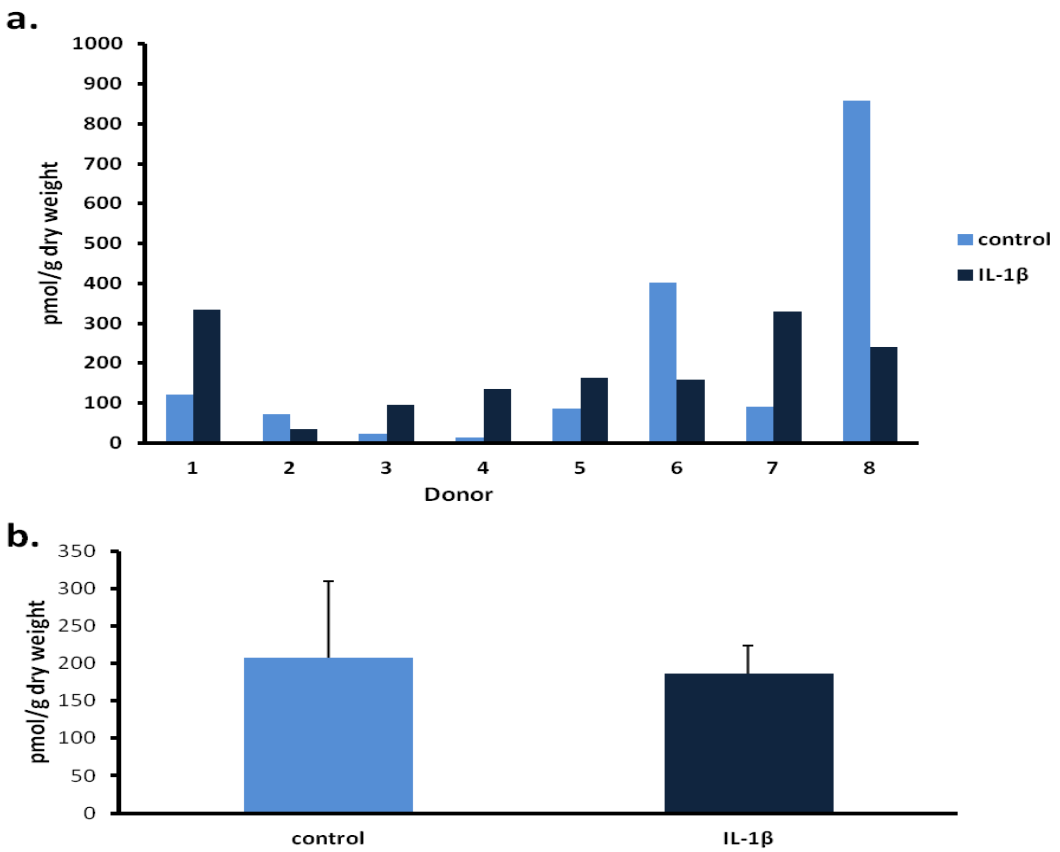



**Figure 2. Effects of 10ng/ml IL-1β treatment for 4 days on absolute fibromodulin protein concentrations measured using QconCAT**. **a**, data represents individual donors for the absolute quantification using the q-peptide IPPVNTNLENLYLQGNR. **b**, data represents the mean absolute quantification (n=8). All data are represented as pmol/g dry weight. Histograms represent means + SEM. Data evaluated using mixed effects linear regression indicated no significant difference for IL-1β stimulated explant cultures relative to control.

**Western blot validation of COMP and fibromodulin**

The absolute quantification of COMP and fibromodulin was validated using western blots probed with equine COMP antibody and human fibromodulin antibody (Figure 3). The position of the bands for each protein probed revealed in the western blots corresponded to the expected size of the full-length proteins. Western blots probed with equine COMP antibody confirmed the findings obtained by QconCAT quantification; there was a general increase in signal in the IL-1β stimulated cartilage explant secretomes compared to control. For western blots probed with antibody to human fibromodulin although results were not as convincing, overall it

was demonstrated that results were similar between western blots and protein concentration as determined using QconCAT.

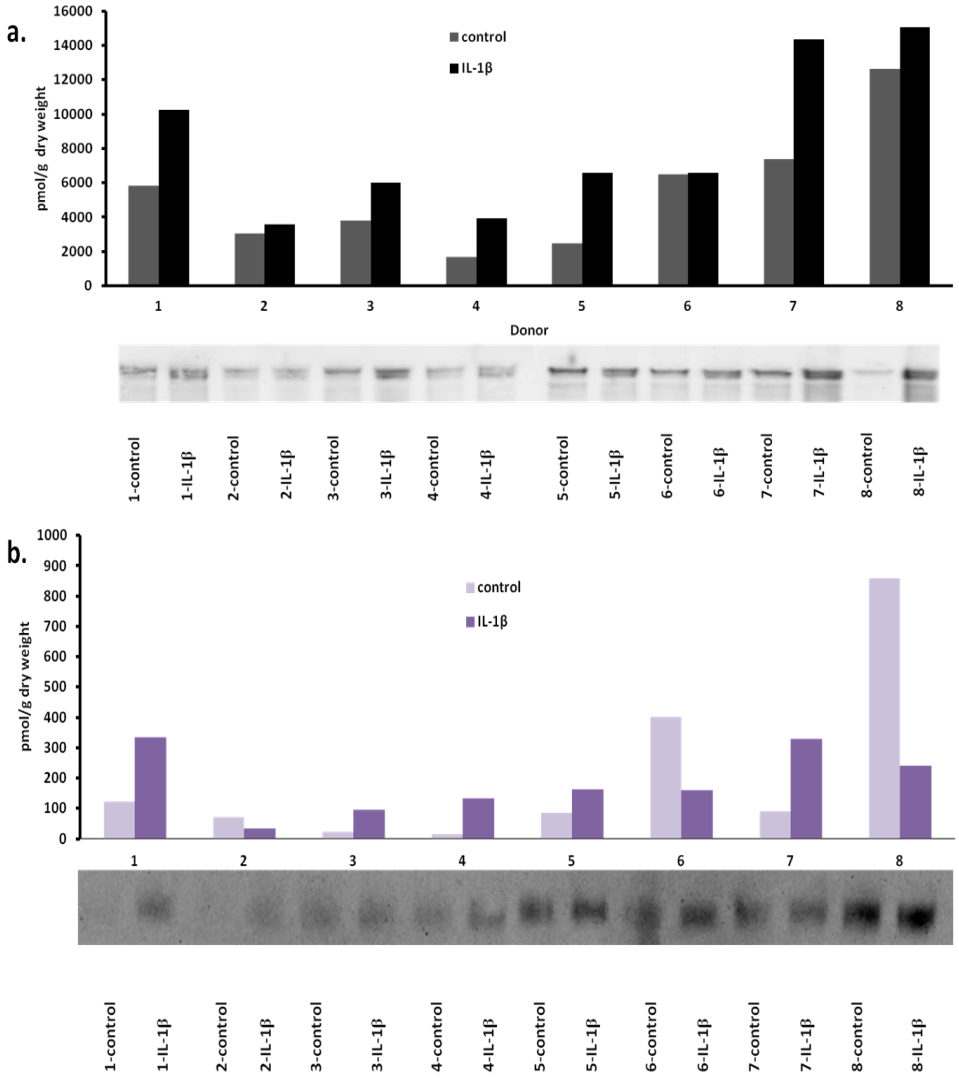



**Figure 3. Western validation of SRM protein quantification results for a. COMP and b. Fibromodulin.** Western blot images of the secretome of for all donors. Protein loading was normalised to dry weight of explants and compared to histograms of absolute amounts using QconCAT quantification.

## Discussion

The determination of fibromodulin and COMP regulation following IL-1β stimulation for 4 days using Progenesis™LC-MS did not identify any significant differences in protein expression (Manuscript 1). Absolute quantification using QconCAT allowed precise amounts of these proteins to be quantified, and was

undertaken in order to compare these results with label-free quantification approach employed by Progenesis™LC-MS.

The QconCAT had been designed using two q-peptides per protein. For COMP both q-peptides (SSTGPGEQLR and DTDLDGFPDEK) were used, however for fibromodulin the physicochemical properties of one of the q-peptides; LYLDHNNLTR were not amenable to quantification using MS. Thus for further analysis only the q-peptide IPPVNTNLENLYLQGNR was used.

Using QconCAT it was determined from results of both q-peptides that there was a significant increase in COMP release into IL1-β stimulated explant media. This was confirmed by probing western blots with equine specific COMP antibodies. The slightly reduced levels quantified by the DTDLDGFPDEK q-peptide could be due to the propensity for incomplete proteolysis caused by an acidic residue aspartic acid (D) adjacent to the site of tryptic cleavage (Siepen et al., 2007). DTDLDGFPDEK consistently producing a lower ratio of light peak area/heavy peak area compared to SSTGPGEQLR. Analysis of Mascot data of analyte digests discovered that occasional miscleavage of DTDLDGFPDEK was evident; this probably led to an underestimation of protein abundance when this peptide was used for absolute quantification. Interestingly label-free quantification for COMP indicated a small (8%) but non-significant increase in COMP in the early OA secretome. Differences in quantification results may be due to the nature of the two methods of quantification. Indeed Progenesis™LC-MS analysis uses all identified peptides for that specific protein in quantification whereas we used only two for absolute quantification. Label-free approaches provide quantitation techniques compatible with whole proteome analysis, a linear dynamic range of 2-3 logs but it gives the lowest accuracy. In contrast stable-isotope labelling provides the greatest accuracy with a dynamic range of 2-5 logs dependant on the MS mode used, where SRM gives the greatest dynamic range (Fries et al., 2007) .

Fibromodulin expression was unaffected by IL-1β stimulation when quantified using both label-free and QconCAT. Western blotting indicated similar trends to our SRM data for most but not all donors. This could be as western blotting analysis may be

subject to more variability than MS based quantification. Results are more subjective and its high sensitivity allows errors in loading to have a substantial effect. Loading cannot be verified with a housekeeping protein as it is difficult to identify a protein with constant expression levels that remain unaltered in secretome experiments.

**Conclusions**

Overall this study identified similar findings in protein quantification of COMP and fibromodulin using a relative and absolute approach to quantification. Western blotting though useful for validation poses more room for experimenter error.

# Manuscript 2

MJ Peffers[1*], RJ Beynon[2], DJ Thornton[3], PD Clegg[1]

**Corresponding Author:** M J Peffers[1]

[1]Dept.of Musculoskeletal Biology, Institute of Ageing and Chronic Disease, University of Liverpool, Leahurst, Chester High Road, Neston, Wirral, CH64 7TE; peffs@liv.ac.uk, pclegg@liv.ac.uk

[2] Proteomics Group, Institute of Integrative Biology, University of Liverpool, Biosciences Building, Crown Street, Liverpool, L69 7ZB; rbeynon@liv.ac.uk

[3] Wellcome Trust Centre for Cell Matrix Research, Michael Smith Building, Faculty of Life Sciences, Oxford Road, Manchester, M13 9PT , UK; dave.thornton@manchester.ac.uk, j.selley@manchester.ac.uk

**Title: Absolute Quantification of Cartilage Extracellular Matrix**

**Running Title**: Quantification of Cartilage Degradation using QconCAT

**Key Words:** absolute quantification, cartilage, osteoarthritis, QconCAT, SRM

## Absolute Quantification of Cartilage Extracellular Matrix


### Abstract

Articular cartilage is composed of a single cell type, the chondrocyte, embedded within an extracellular matrix (ECM). Osteoarthritis (OA) is characterised by the slow degeneration of the ECM. The degradation of matrix proteins is a consequence of proteolytic fragmentation by key proteinases at defined sites. Cleavage of some matrix molecules in cartilage is currently identified semi-quantitatively with specific neoepitope monoclonal antibodies. Although proteomic profiling of cartilage has progressed the absolute quantification of matrix proteins is undefined in cartilage. There is a need for absolute quantification to describe the mechanisms of matrix turnover in health and disease. A QconCAT was designed as a concatenation of tryptic quantotypic peptides to measure absolute amounts of matrix proteins and assess whether degradation of specific matrix proteins at specific sites could be quantified using mass spectrometry. Matrix proteins from normal and OA equine articular cartilage were quantified using selected reaction monitoring (SRM). Absolute quantification values for normal and OA cartilage were identified for aggrecan; first and third globular domains, biglycan, cartilage oligomeric matrix protein, decorin and fibromodulin. A peptide spanning a known biglycan cleavage site was significantly reduced in OA cartilage. However it was not possible to quantify the extent of hydrolysis of targeted matrix proteins due to the variability in peptide abundance between cleaved and non-cleaved peptides. Examination of the relationship between cartilage proteins and mRNA levels revealed that transcripts provided little predictive value with respect to the amount of protein expressed. The equine degradome QconCAT will provide a useful tool in future cartilage research.


### Introduction

Osteoarthritis (OA) is an significant disease of articular cartilage in man and animals caused by multiple overlapping and independent disease mechanisms resulting in

biochemical and biomechanical failure (Brew et al., 2008). Articular cartilage is composed of a single cell type the chondrocyte, embedded at low cell density in a complex extracellular matrix (ECM) that confers its unique load bearing properties. This matrix consists of a dense network of collagen fibres; conferring shape and form, and a high concentration of proteoglycans that provide compressive properties associated with load bearing. There is a constant turnover of these matrix proteins in life with an imbalance between synthesis and degradation being the primary event in OA. The degradation of cartilage matrix is elicited through proteolysis by defined groups of proteinases including matrix metalloproteinases (MMPs) (Burrage et al., 2006) and the 'disintegrin and metalloproteinase with thrombospondin motifs' (ADAMTSs) (Arner, 2002).

There is evidence for increased expression and synthesis of these proteases, in patients with OA in cartilage, synovial tissue and joint fluids (Lohmander et al., 1993b; McCachren, 1991);(Dean et al., 1989). The role these enzymes play in cartilage turnover and disease has not been fully elucidated given that although increased activity of proteases in OA cartilage has been established (Yoshihara et al., 2000), the majority of the enzymes described in cartilage are in the inactive proenzyme form (Dean et al., 1989). Recent studies have suggested that MMPs are principally involved in normal turnover and may have a reduced role in aggrecan breakdown in OA (Struglics and Hansson, 2012). Methods to measure MMP activity *in-situ* remain a challenge (Lombard et al., 2005) and further methods to improve MMP assays are desirable.

The sites of proteinase cleavage for many of the major matrix proteins including aggrecan (Sandy et al., 1991), biglycan (Monfort et al., 2006) cartilage oligomeric matrix protein (COMP) (Dickinson et al., 2003), decorin (Zhen et al., 2008) and type II collagen (Mitchell et al., 1996) are well characterised. The cleavage of aggrecan (Madsen et al., 2010) and type II collagen (Billinghurst et al., 1997) is defined using monoclonal neoepitope antibodies for each hydrolysis site. Whilst neoepitope antibodies have allowed the determination of some of the important sites of proteolytic action in cartilage degradation (Fosang et al., 2010) the results are semi-

quantitative at best and unlike mass spectrometry they do not provide absolute molecular specificity. Western blotting (Heidebrecht et al., 2009) and enzyme-linked immunosorbent assay (ELISA) have historically been the most widely applied methods for targeted semi-quantitative analysis of proteins. Western blotting is excellent at the detection of proteins, as it provides a sensitive and highly specific approach (if good antibodies are available), it is limited when quantitation is needed due to issues of reproducibility and dynamic range. Comparatively, ELISAs are better for quantitation but are less specific as there is no information regarding protein size and they are susceptible to interference from cross-reacting fragments. Mass spectrometry workflows now allow selectivity and sensitivity in protein quantification. Recently there have been major advances in proteomic profiling of cartilage (Garcia et al., 2006; Wilson et al., 2010a). Whilst these techniques have been revealing, they provide a discovery platform to primarily identify cartilage components. There are limitations to these approaches for example in demonstrating low abundance proteins, in particular cleavage products, which are less likely to be identified without refinement in sample preparation and analysis. Indeed, whilst there is value in the discovery of differentially expressed proteins in terms of fold change, a more quantitative data set will allow further interpretation and is needed for predictive biology using a systems biology approach (Otto et al., 2012). Systems biology studies the complex interactions of biological components through the collection, integration and analysis of complex data sets. Mathematical and mechanistic modelling is then used to discover emergent properties in pathways, cells and tissues. One of the keys to this analysis is the input of real numbers into the models for example as absolute amount of proteins (Brownridge et al., 2011). Thus absolute quantification of both intact matrix proteins and the extent of protein cleavage will be required to describe completely the mechanisms of cartilage matrix turnover.

Proteomics has developed into a quantitative science resulting in more refined methods being developed for absolute quantification. This includes the application of selected reaction monitoring (SRM) based experiments using isotope labelled standards (Aye et al., 2012; Gerber et al., 2003). Some of these methods rely on the

principles of surrogacy where, following protease digestion of a protein, the resultant peptides are identified by mass spectrometry (MS) and used to deduce identification of the protein itself. One well described method of absolute quantification is through stable isotope label. There is differential incorporation of stable isotopic nuclei ($^{2}H$, $^{13}C$, $^{15}N$, $^{18}O$) (which acts as a standard) enabling the 'light' and 'heavy' form of the same peptide to be resolved in the MS due to their mass difference. The ratio of signal intensities of the two peptides measures peptide abundance and, correspondingly, the relative abundance of the proteins. A number of methods are available to generate these protein standards (Dupuis et al., 2008). One approach enables the highly accurate parallel absolute quantification of large sets of analyte proteins using an artificial protein QconCAT (Beynon et al., 2005). QconCATs are constructs of a set of mass-tagged internal standard peptide (q-peptides) with sequences unique to the proteins of interest. Multiple peptides are concatenated into a synthetic gene and expressed as a heterologous QconCAT protein in bacterial cultures (Pratt et al., 2006; Rivers et al., 2007), allowing large numbers of biological samples to be analyzed which is cost effective and reliable. The aim of this study was to develop techniques to reveal absolute quantification data of intact matrix proteins, the extent of hydrolysis at specific sites (which could allow OA progression monitoring) in some of these proteins and understand the role of matrix metalloproteinases in normal and OA cartilage. In addition the relationship between transcript and protein abundance was investigated. An innovative platform for the quantification of matrix proteins in equine articular cartilage is described which combines a relatively simple tissue preparation technique, which overcomes some of the technical challenges related to its biochemical properties, with QconCAT technology. This has provided a novel tool for cartilage research through the design and optimisation of SRM assays which was then used in an attempt to define the cartilage degradome in normal and osteoarthritic equine articular cartilage.

## Materials and Methods

### Selection of quantotypic peptides in the QconCAT design

The design for the degradome QconCAT included peptides representative of the matrix proteins; aggrecan, biglycan, cartilage oligomeric matrix protein (COMP), decorin, type II collagen and fibromodulin. A list of peptides from a previous study (Manuscript 1) was used as a library for q-peptide selection. In addition we wished to incorporate additional peptides to quantify proteolytically cleaved molecules and zymogens of specific proteases (Figure 1).

Bioinformatics screening of known aggrecan, collagen-II, cartilage oligomeric matrix protein (COMP), decorin and fibromodulin cleavage sites was undertaken using *in-silico* digests using Protein Prospector version 5.6.2 (http://prospector.ucsf.edu). In addition screening of two important metalloproteinases matrix metalloproteinase 3 (MMP-3) and matrix metalloproteinase 13 (MMP-13) was undertaken as we wished to measure the quantity of zymogen (inactive form of the enzyme including the pro-peptide) compared to active enzyme. A similar approach was used in the quantification of cleavage sites; however the position of cleavage was the junction between the propeptide and the peptide for the enzyme concerned. Finally following further bioinformatics screening of all known markers of collagen synthesis and degradation a marker of collagen synthesis; procollagen type II C-propeptide (PIICP) was identified as a candidate for inclusion. It has been demonstrated that there is an increase in type II collagen synthesis in early OA in man (Otterness et al., 1999). The final QconCAT design contained 28 q-peptides (Table 1a).

With the exception of fibromodulin, which contained a single peptide for quantification, all other intact proteins were quantified with at least two q-peptides. Fibromodulin was quantified with a single peptide as previous studies had demonstrated the IPPVNTNLENLYLQGNR peptide provided an ideal q-peptide for SRM studies (Manuscript 3). Criteria applied in selection of intact protein peptides have been described by Pratt et al. (Pratt et al., 2006) and included their suitability score, physicochemical properties deemed to promote MS detectability, and

uniqueness to a given protein. In addition amino acids which were prone to oxidation (cysteine, methionine or tryptophan) and miscleavage were avoided where possible and all peptides were either lysine or arginine terminated. All peptides where placed in sequence context were possible in order to minimize differences in sequence context and therefore digestion context between QconCAT and analyte.

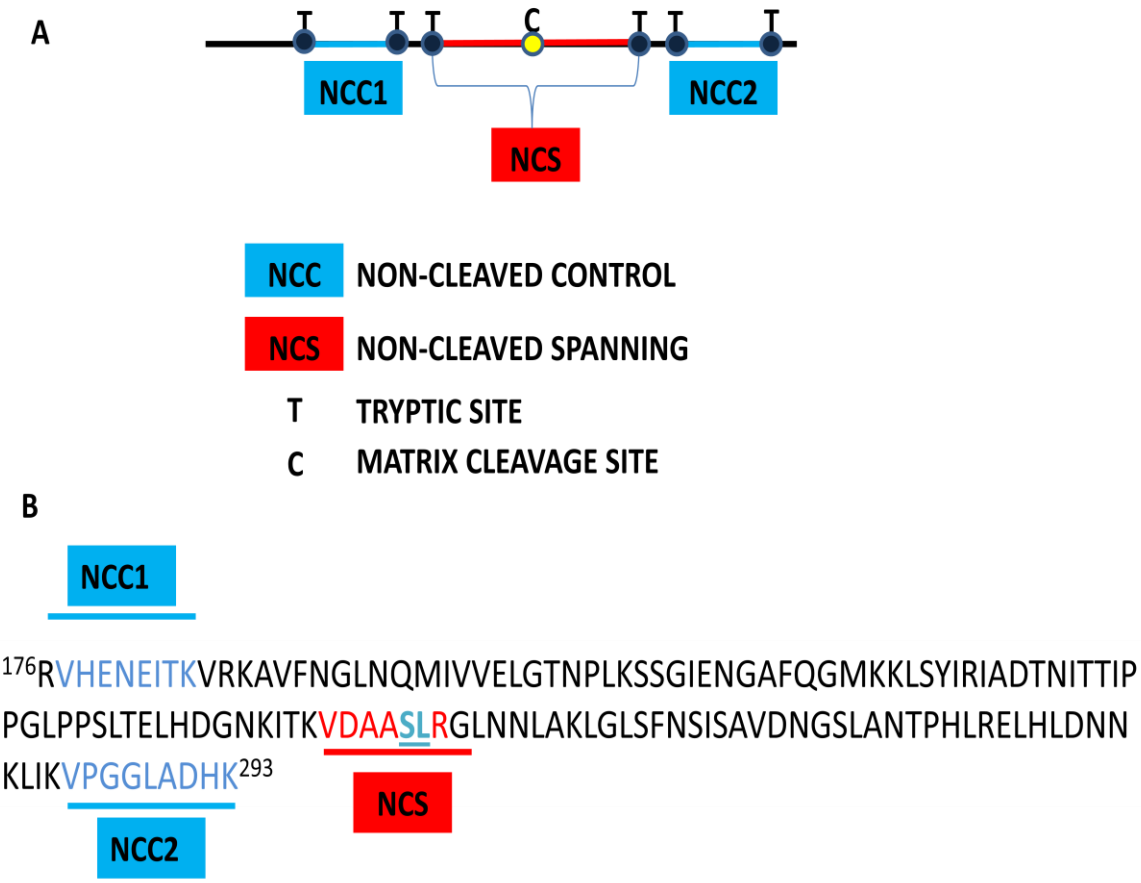



**Figure 1. Principle of the use of QconCAT to quantify cartilage degradation.** A. Tryptic quantotypic Q-peptides distant to the any known cleavage sites were used to quantify intact protein (non-cleaved control peptide; NCC). A tryptic Q-peptide spanning the cleavage site was also incorporated into the QconCAT (non-cleaved spanning; NCS). If no hydrolysis was evident at the given cleavage site our hypothesis was that NCC1=NCC2=NCS. Should cleavage occur at the site the NCS peptide would be less abundant (NCC1>NCS<NCC2). Thus the amount of cleavage of the protein at that site could be quantified. B. Decorin is used as an example; VHENEITK and VPGGLADHK were used as NCC peptides to quantify MMP-3

degradation of VDAASLR between the underlined amino acids S - serine and L – leucine.

| Peptide Order | Protein | Protein Accession | Peptide Type | Q-peptide Amino-acid Sequence |
|---|---|---|---|---|
| 1 | COLLAGEN TYPE II | Q28396 | **NCC** | TGPAGAAGAR |
| 2 | AGGRECAN-G3 | O46556 | **NCC** | YEINSLVR |
| 3 | AGGRECAN- G3 | O46556 | **NCC** | TIEGDFR |
| 4 | AGGRECAN-G3 | O46556 | **NCC** | YQCTEGFVQR |
| 5 | CARTILAGE OLIGOMERIC MATRIX PROTEIN | Q9BG80 | **NCC** | SSTGPGEQLR |
| 6 | CARTILAGE OLIGOMERIC MATRIX PROTEIN | Q9BG80 | **NCC** | DTDLDGFPDEK |
| 7 | MATRIX METALLOPROTEINASE 13 | O18927 | **NCC** | ISELGFPK |
| 8 | BIGLYCAN | Q865A8 | **NCC** | GLQHLYALVLVNNK |
| 9 | DECORIN | Q865A7 | **NCC** | VPGGLADHK |
| 10 | DECORIN | Q865A7 | **NCS** | VDAASLR |
| 11 | MATRIX METALLOPROTEINASE 3 | Q28397 | **NCS** | CGVPDVGHFTTFPGMPK |
| 12 | BIGLYCAN | Q865A8 | **NCS** | GVFSGLR |
| 13 | COLLAGEN TYPE II | Q28396 | **NCC** | GPEGAQGPR |
| 14 | AGGRECAN-G1 | O46556 | **NCC** | YPIVSPR |
| 15 | DECORIN | Q865A7 | **NCC** | ELHLDNNK |
| 16 | FIBROMODULIN | A2Q126 | **NCC** | IPPVNTNLENLYLQGNR |
| 17 | MATRIX METALLOPROTEINASE 13 | O18927 | **NCS** | CGVPDVGEYNVFPR |
| 18 | MATRIX METALLOPROTEINASE 3 | Q28397 | **NCC** | GEILFFK |
| 19 | BIGLYCAN | Q865A8 | **NCC** | NHLVEIPPNLPSSLVELR |
| 20 | BIGLYCAN | Q865A8 | **NCC** | VPAGLPDLK |
| 21 | CARTILAGE OLIGOMERIC MATRIX PROTEIN | Q9BG80 | **NCS** | NTVMECDACGMQPAR |
| 22 | BIGLYCAN | Q865A8 | **NCS** | NMNCIEMGGNPLENSGFQPGAFDGLK |
| 23 | COLLAGEN TYPE II | Q28396 | **PIICP** | SLNNQIESIR |
| 24 | CARTILAGE OLIGOMERIC MATRIX PROTEIN | Q9BG80 | **NCS** | AVAEPGIQLK |
| 25 | MATRIX METALLOPROTEINASE 3 | Q28397 | **NCC** | EHGDFFPFDGPGK |
| 26 | MATRIX METALLOPROTEINASE 13 | O18927 | **NCC** | GETMVFK |
| 27 | BIGLYCAN | Q865A8 | **NCC** | LAIQFGNYK |
| 28 | DECORIN | Q865A7 | **NCC** | VHENEITK |

**Table 1a. Equine cartilage QconCAT signature peptides.** The table includes peptide order within the QconCAT design and peptide type relates to non-cleaved control (NCC), non-cleaved spanning (NCS) or procollagen type II C-propeptide (PIICP). For aggrecan three peptides were located in the 1$^{st}$ globular domain (G1) and a single peptide in the third globular domain (G3).

This was to minimize differences in digestion kinetics. Subsequent to design and construction of the QconCAT a further update of the Ensembl database for horse (*Equus caballus*; (ftp://ftp.ensembl.org/pub/current_fasta/equus_caballus/pep/) indicated that two peptides decorin; ELHLDNNK and COMP; AVAEPGIQLK were no longer identified as unique as they were also present in biglycan and thrombospondin-4 respectively. Subsequently no further analysis is included for these peptides.

Q-peptides that spanned known cleavage sites were selected following prediction of suitable candidates *in-silico* (Table 1b). A similar criterion in selection was adopted as for the used to quantify intact peptides, except we were unable to avoid certain amino acids that are prone to oxidation in some spanning peptides.

| Protein | Q-peptide Amino-acid Sequence | Protease responsible for cleavage |
|---|---|---|
| DECORIN | VDAA**SL**R | MMP-3 (Imai et al., 1997) |
| BIGLYCAN | **GV**FSGLR | MMP-13 (Monfort et al., 2006) |
| BIGLYCAN | NM**NC**IEMGGNPLENSGFQPGAFDGLK | ADAMTS-4 (Melching et al., 2006) |
| COMP | NTVMECDACGMQ**PA**R | ADAMTS-5*(Holden, 2012) |

**Table 1b. Equine cartilage non-cleaved spanning peptides.** The amino acids in bold red indicate the site of protease cleavage. References refer to studies which demonstrated the cleavage of the given protein by the specified protease (Cartilage oligomeric matrix protein; COMP).

The amino acid sequence was sent to PolyQuant GmBH (Enchelon, Germany) for gene synthesis of the QconCAT (Figure 2). The transformation, expression and purification of QconCATs have been described in manuscript 3.

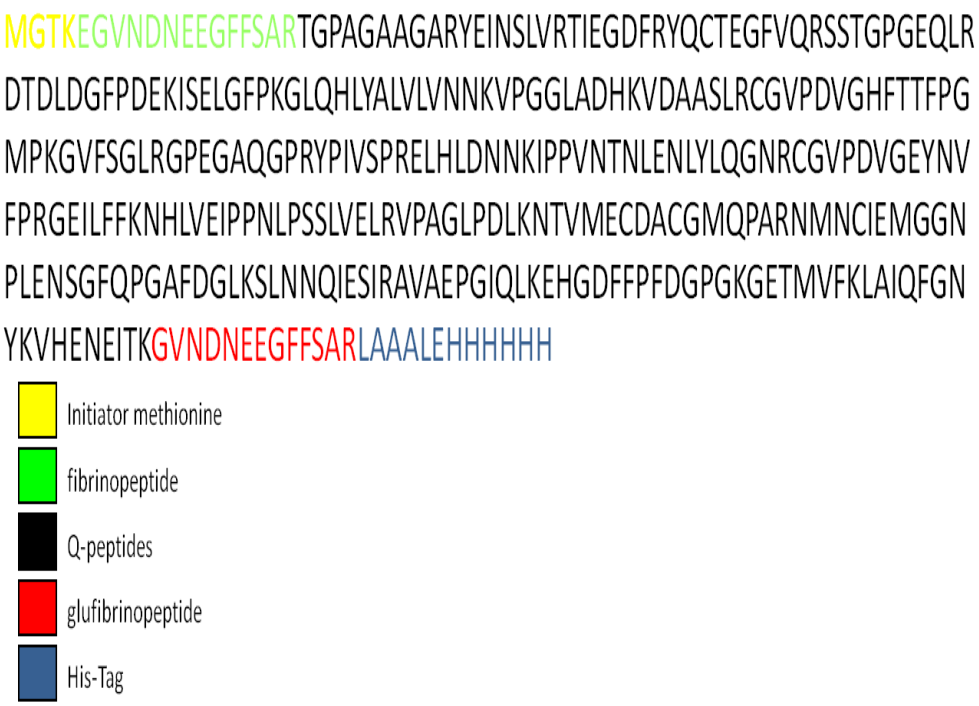



**Figure 2. Amino-acid sequence of equine degradome QconCAT. The QconCAT sequence sent to** PolyQuant GmBH for gene production included an sacrificial peptide containing the initiator methionine (also protects the N terminus of the initial q-peptide from exoproteases), glufibrinopeptide B (Glufib) (for quantification of the QconCAT using an unlabelled Glufib standard), fibrinopeptide (a sequence variant of Glufib, to establish full length expression of the gene), and hexahistidine purification tag (His-Tag).

**Retrospective peptide screening using bioinformatics tools**

Prediction tools available following the production of the QconCAT were utilised to assess peptide detectability and potential for missed cleavage. CONSeQuence (consensus predictor for quantotypic peptide sequence) (Eyers et al., 2011)was applied to the list of peptides by submitting the FASTA files to http://king.smith.man.ac.uk/CONSeQuence. The number of internal miscleaves was set at 0 and prediction type; rank score. A missed cleavage predictor; MCPRED was applied to peptides within the analyte and QconCAT context by submitting FASTA files to (http://king.smith.man.ac.uk/mcpred/). The 'select predictor' SVM (support vector machine) predictor (Webb-Robertson et al., 2008) was selected.

**Characterisation of the QconCAT**

The QconCAT protein was confirmed to be intact and most proteins discernible by the in-gel digestion of a protein band corresponding to the expected molecular mass for the QconCAT of 45KDa using MALDI-TOF as described in Manuscript 3.

**Cartilage sampling and histological assessment**

Equine articular cartilage was obtained from the entire condylar surface of the metacarpophalangeal III bone of the fetlock joints of 11 normal and 9 osteoarthritic horses between 2 and 20 years old, either from an abattoir or with informed consent from thoroughbred racehorses in training prior to euthanasia due to orthopaedic disease. Cartilage for each donor was randomly split and either rinsed and snap frozen in liquid nitrogen for proteomic analysis or stored in RNA Later (Ambion, Applied Biosystems, Warrington, UK) for gene expression analysis. The distal surface of metacarpal III was examined by gross observation immediately after death in the case of 'normal' donors or following post-mortem and storage at -20$^{o}$C for diseased material. Macroscopic scoring of the metacarpophalangeal joint was measured using a macroscopic grading system, recommended for spontaneous OA in Table 2 (Kawcak et al., 2008).

Sections of palmar metacarpal condyles, obtained for histology were placed into 4% paraformaldehyde for 48 hours. Following decalcification in EDTA, they were sectioned at 6µm and stained with haematoxylin and eosin (H and E) and safranin O stains in order to score structural differences and assess glycosaminoglycan (GAG) distribution. Histopathological preparation was undertaken in the Histopathology Department at the University of Liverpool.

Standard sections were examined using an Eclipse 80i (Nikon, Kingston-upon-Thames, UK) light microscope at x4 and x10 magnifications. A DS-U1 camera (Nikon, Kingston-upon-Thames, UK) was used to obtain images. A modified Mankin scoring system was used (as recommended by the OARSI histopathology initiative) for semi-quantitative histological assessment of equine cartilage (Zardi et al., 1985).

| Lesion | Grade | Description |
|---|---|---|
| Wear Lines | 0 | *None* |
| | 1 | *1 or 2 partial-thickness wear lines/joint surface* |
| | 2 | *3 to 5 partial-thickness or 1 to 2 full thickness wear lines/joint surface* |
| | 3 | *>5 partial thickness or >2 full thickness wear lines/joint surface* |
| Erosions | 0 | *None* |
| | 1 | *Partial-thickness erosion,<5mm diameter* |
| | 2 | *Partial-thickness erosion,>5mm diameter* |
| | 3 | *Full thickness erosion* |
| Palmar arthrosis (osteochondral lesions distal palmar MC) | 0 | *None* |
| | 1 | *Partial-thickness erosion, <5mm diameter* |
| | 2 | *Partial-thickness erosion, purple discoloration >5mm diameter* |
| | 3 | *Full-thickness erosion, purple discoloration, >5mm diameter* |

**Table 2. Macroscopic scoring system of the pathology of the distal condyles of MCIII.**

**Cartilage protein extraction and sample preparation for MS**

A method was developed in order to produce a sample that was high pressure liquid chromatography (HPLC) compatible; the addition of impurities or compounds that would irreversibly bind to the pre-column (trap) or column were reduced, whilst washing of both the protein extract and tryptic digests were enabled. Furthermore the removal of a substantial amount of GAGs was believed prudent. As the molecular weight (MW) of the proteins of interest for quantification was 50-250kDa, a filter size of at least twice that was necessary hence a 10kDa molecular weight cut-off (MWCO) was selected. A MWCO spin column enabled the protein sample to be washed and the guanidine to be diluted to an extent that provided optimal conditions for tryptic digestion. As the enzymatic digestion took place above the filter a final spin allowed separation of tryptic peptides to the filtrate. Finally samples were desalted and purified using $C_{18}$ resin in the form of a ZipTip® (Merck Millipore, USA).

Previously frozen cartilage was pulverised and lyophilized in order to obtain a dry weight for normalisation. To 10mg of lyophilised cartilage powder 1ml of guanidine extraction buffer (4M guanidine hydrochloride (GdnHCl), 65mM dithiothreitol (DTT), 50mM sodium acetate) was added and extraction performed with end-over-end mixing for 20h at 4$^{o}$C. 25mM DTT was added 2h prior to 80mM iodoacetamide (IAA) addition for the last 2h in the dark. The soluble fraction was removed following centrifugation for 15min at 13000**g** at 4$^{o}$C. Protein concentrations of aliquots of soluble fraction were estimated by Bradford assay using Coomassie Plus™ protein assay reagent (Thermo Scientific, Rockford, USA) read at 600nm following acetone precipitation. Guanidine-extracted proteins were then washed with ammonium bicarbonate (AMBIC) to give a final concentration of 0.5M GdnHCl on a Vivaspin 2 10kDa MWCO filter. Tryptic digestion (overnight at 37$^{o}$C) was undertaken with the addition of 2μg trypsin (Sigma-Aldrich, Dorset, UK) with a top-up of a further 2μg after 3h. The filtrate of tryptic peptides was obtained following centrifugation for 15min at 13000**g**. 100mM AMBIC was added to the top of the spin column and a further spin for 15min at 13000**g** undertaken. These fractions were pooled and acidified with trifluoroacetic acid (TFA) to 0.1% (v/v).

**1-D SDS PAGE separation and in-gel trypsin digestion**

Cartilage guanidine soluble extracts were analyzed by one dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) to assess intersample consistency and quantitative/qualitative differences between normal and osteoarthritic cartilage. Samples were loaded according to equal volumes following acetone precipitation and resolubilsation in buffer containing 8M urea, 2% (w/v) CHAPS, 0.0002% (v/v) bromophenol blue, plus 0.2% (v/w) DTT.

Aliquots were heated in Laemmli buffer containing 50mM DTT for 5min at 95°C and resolved through 4–12% acrylamide Bis-Tris NuPAGE gels (Invitrogen, Paisley, UK), and proteins were visualized by Coomassie Brilliant Blue. In-gel tryptic digestion of dominant bands was undertaken as previously described (Deckelbaum et al., 1982).

**Protein identification of bands using MALDI-TOF**

Matrix-assisted laser desorption ionisation time of flight time of flight (MALDI-TOF) analysis of protein bands was undertaken as described in manuscript 3. Data was searched against the *Equus caballus* database using the Mascot search engine. Parameters were set to accept one miscleavage, a fixed modification of carbamidomethly cysteine and a variable oxidation of methionine. The peptide mass tolerance for this instrument was set at 0.2Da.

**Monitoring proteolysis of analyte proteins and sample ranging experiments using label-free quantification**

Prior to SRM samples were analysed using label-free quantification on the Synapt G1 Q-TOF instrument, in order to enable a suitable amount of sample to be loaded and to monitor proteolysis of the analyte samples. The peptides were resolved using a NanoAcuity LC system and instrument methods have been previously described (Brownridge and Beynon, 2011). In brief a final on column load of 50fmol of pre-digested protein standard (rabbit muscle glycogen phosphorylase, UniProt Accession number P00489) (Waters, UK) was added to the tryptic peptide preparation (500ng load on column) in order to enable "Hi3," methodology where the intensities of the three most intense unique peptides per protein are compared with those of the protein standard (Ingvarsson et al., 2000). Acquisition was with the instrument in V mode using $MS^E$ data-independent acquisition at a resolution of 20,000.

Data were processed using ProteinLynx Global Server (version 2.4) (PLGS) software (Waters, UK) against a user created database including the pre-digested standard, the QconCAT recombinant protein sequence, and *Equus caballus* databases using an ion accounting algorithm. Peptide and fragment mass tolerance were set to 10ppm and 15ppm respectively. The minimum fragment ion matches per peptide and protein was three and seven respectively, the minimum peptide matches per protein was one. A fixed modification of carbamidomethly cysteine and a variable oxidation of methionine were used. One tryptic miscleavage was specified and the FDR was set at 4%.

**QconCAT digestion**

QconCAT was detergent treated with 1% (w/v) Rapigest (Waters, Manchester, UK) for 10min at 80$^{o}$C in 25mM ammonium bicarbonate. Sequential reduction and alkylation in 3mM DTT (60$^{o}$C for 10 minutes) and then 9mM IAA (30min in the dark at room temperature) was followed by trypsin addition at a ratio of 1:50 protein: trypsin ratio overnight at 37$^{o}$C. Detergent inactivation was then assumed by incubating for 45min at 37$^{o}$C with trifluoroacetic acid (VWR International) to a final concentration of 0.5% (v/v). Following centrifugation for ten minutes at 15000**g** the soluble phase was retrieved and used for LC-MS/MS. Complete proteolysis was monitored on a Synapt G1 Q-TOF by resolving 100fmol on a NanoAcuity LC system using methods previously above. Data was processed as described earlier except with static/fixed modifications carbamidomethyl (C), and variable modifications; oxidation (M), label ([$^{13}C_6$] Lys)/label ([$^{13}C_6$] Arg).

**Selected reaction monitoring optimisation**

SRM assay conditions were optimised using a XEVO TQ (Waters, Manchester) as described in manuscript 3. Trypsin digestion of QconCAT was carried out as described previously (Rivers et al., 2007). SRM experiments were conducted with 1:20 dilution of tryptic analyte peptides as determined by ranging experiments, spiked in with 10fmol previously digested heavy QconCAT, loaded onto column. MS analysis was commenced using methods, parameters and gradients described in manuscript 3.

**Western blotting**

Western blots from spin column supernatants and the protein wash flow through were probed with COMP and MMP-3 antibodies in order to ensure no undigested protein remained and tryptic digestion was complete. Additionally to correlate absolute quantification data with an independent method of protein measurement COMP and MMP- 3 in cartilage extracts was also detected with fluorescence immunoblotting. Following acetone precipitation and resolubilsation in buffer containing 8M urea, 2% (w/v) CHAPS, 0.0002% (v/v) bromophenol blue, plus 0.2%

(v/w) DTT, samples from selected normal and OA donors were heated to 80$^{o}$C for 10min in NuPAGE® LDS sample buffer (Invitrogen, Paisley, UK) and electrophoresed for 1h at 200V under reducing conditions on Novex 4-12% SDS-PAGE gels (Invitrogen, Paisley, UK). Protein transfer to nitrocellulose was performed using the Invitrogen X Cell Sure Lock apparatus according to standard protocol. Membranes were blocked with TBS (pH 7.4) containing 0.1% Tween-20 (Invitrogen, Paisley, UK) (TBST) and 5% dried skimmed milk for 1 h at room temperature. The specificity and cross-reactivity of the non-equine primary antibodies for MMP-3 were confirmed by analysis of Il-1β stimulated equine cartilage explant media (Manuscript 1). The peptide sequences used for GAPDH and β-actin were checked using the NCBI Basic Local Alignment Search Tool; (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins).

Using previously described protocols (Buckland, 2012a) membranes were probed with primary antibodies against the following; rabbit polyclonal to equine COMP (1:1000; a kind gift from Jay Dudhia, RVC, London, UK) (Han et al., 1996) and goat polyclonal to human MMP-3 (1:1000 dilution, # 18898 Abcam, Cambridge, UK), anti GAPDH-horseradish peroxidase (HRP) conjugate (used at 1:10,000; Sigma, Dorset ,UK) and monoclonal mouse anti-β actin (1:1000; Sigma, Dorset, UK). Appropriate secondary antibodies conjugated to horseradish peroxidise (HRP); polyclonal rabbit anti-mouse IgG HRP, polyclonal goat anti-rabbit IgG HRP and polyclonal goat anti mouse HRP (all Sigma, Dorset, UK) at 1:2000 were used. All antibodies were diluted with TBST containing 5% dried skimmed milk. Chemiluminescence was used to detect the protein bands using Western Lightning™ and Western Lightning Plus Chemiluminescence reagents (Perkin Elmer, Beaconsfield, USA). ImageJ software (http://rsbweb.nih.gov/ij/) was used to quantify bands using densitometry.

**RNA Extraction and Reverse Transcription**

Cartilage pieces from donors were dismembranted following freezing in liquid nitrogen, and stored in Tri-reagent (Ambion, Warrington, UK). For RNA extraction the Guanidium-thiocyanate-phenol-chloroform technique, with ethanol extraction was used, followed by the RNeasy (Qiagen, Crawley, UK) column technique incorporating a DNase treatment step. RNA was quantified using a Nanodrop ND-

100 spectrophotometer (Labtech, East Sussex, UK). 16µl RNA was used as the template for reverse transcription. M-MLV reverse transcriptase and random hexamer oligonucleotides (both from Promega, Southampton, UK) were used to synthesize cDNA in a 25µl reaction.

**Quantitative reverse transcriptase PCR design**

Equine specific primers were used that had been described previously (Chen et al., 2010; Radinsky et al., 1990) (Richter and Hormann, 1982). All primers were exon-spanning to allow discrimination of genomic DNA and cDNA. The primers used for target and housekeeping genes are in Table 3.

**Quantitative Real-Time Polymerase Chain Reaction**

Quantitative real-time polymerase chain reaction (RT-PCR) was performed in SYBR® Green PCR master mix (Applied Biosystems, Warrington, UK) in a 25µl reaction with 300nM primer concentration and processed by a 7300 Real Time PCR system (Applied Biosytems, Warrington, UK) using standard amplification conditions. Data was analysed using SDS software system (Applied Biosytems, Warrington, UK). PCR products were measured and relative expression levels normalized to GAPDH and calculated using the $2^{-\triangle Ct}$ method (Livak and Schmittgen, 2001).

**Statistical analysis**

Statistically significant differences between absolutely quantified matrix proteins and gene expression values of normal and osteoarthritic cartilage were analysed using mixed effects linear regression to allow for donors with significant biological variation. Correlation between protein abundance and gene expression was tested using the Pearson's correlation coefficient. The analyses were undertaken using SPLUS 6.1, Minitab 15 and Excel software.

| Gene | Primer Sequence | Accession Code |
|---|---|---|
| **GAPDH** | F: GCATCGTGGAGGGACTCA<br>R: GCCACATCTTCCCAGAGG | AF157626 |
| **ADAMTS4** | F: CAGCCTGGCTCCTTCAAAAA<br>R: CCGCAGGAATAGTGACCACAT | NM_001111299 |
| **ADAMTS5** | F: ACCGATCCTGCAGTGTCACA<br>R: CTGCTCATGGCGAAAAGATTT | EU025851 |
| **Collagen I α1** | F: GACTGGCAACCTCAAGAAGG<br>R: CAATATCCAAGGGAGCCACA | O46388 |
| **Collagen II α1** | F: TCAAGTCCCTCAACAACCAGAT<br>R: GTCAATCCAGTAGTCTCCGCTCTT | NM_001081764 |
| **MMP1** | F: GGTGAAGGAAGGTCAAGTTCTGAT<br>R: AGTCTTCTACTTTGGAAAAGAGCTTCTCT | NM_001081847 |
| **MMP3** | F: TCTTGCCGGTCAGCTTCATATAT<br>R: CCTATGGAAGGTGACTCCATGTG | NM_001082495 |
| **MMP13** | F: CTGGAGCTGGGCACCTACTG<br>R: ATTTGCCTGAGTCATTATGAACAAGAT | NM_001081804 |
| **Aggrecan** | F: GAGGAGCAGGAGTTTGTCAACA<br>R: CCCTTCGATGGTCCTGTCAT | XM_001499504 |
| **Biglycan** | F: TCACCTTCCAGCCCCTAGAGT<br>R: AGAAGCAGCCCCTCCTCAA | NM_001081839 |
| **Decorin** | F: CATCCAGGTTGTCTACCTTCATAACA<br>R: CCAGGTGGGCAGAAGTCATT | NM_001081925 |
| **Fibromodulin** | F: CTTGGCTCCAGACCCTGAAA<br>R: TGCCCCTCGCGTCAGA | NM_001081777 |
| **COMP** | F: GGTGCGGCTGCTATGGAA<br>R: CCAGCTCAGGGCCCTCAT | AF325902 |
| **TIMP 3** | F: CTGCAACTTCGTGGAGAGGT<br>R: ACTCGTTCTTGGAGGTCACG | NM_001081870 |

**Table 3. Primers used for quantitative real-time PCR.** F= Forward primer, R= Reverse primer.

## Results

### Macroscopic grading and histological assessment of samples

Macroscopic grading and a semi-quantitative histological assessment using a modified Mankin scoring system of the metacarpal joint were undertaken. The macroscopic score was used to allocate samples into normal or OA groups, a score of greater than 0 were assigned OA. The average and SEM for the age and Mankin's

score of normal and OA samples were 8.3 ±1.54 years, 3.2±0.73 and 11.2±1.21 years, 4.8±2.14 respectively.

**Comparative analysis of proteins in normal and OA cartilage**

Soluble cartilage protein extracts from normal and OA cartilage analysed by SDS-PAGE revealed no qualitative or quantitative differences using densitometry (data not shown). The major bands were cut from the gel, digested with trypsin and the proteins identified using peptide mass fingerprinting MALDI-TOF (Figure 3). The preliminary protein in the major bands had a Mascot score >62 and a confidence interval of 95%.

**QconCAT protein design, expression and validation**

Bioinformatics screening of aggrecan and collagen-II cleavage sites identified that the QconCAT strategy previously described could not be used at any of the major cleavage sites by using trypsin digestion. This was due to either post-translational modifications such as N-linked glycosylations, (aggrecan) (Figure 4) and hydroxyproline (collagen) around the protein digestion protease sites or problems with amino acids produced at the carboxyl or amino terminals of potential q-peptides for instance glutamic acid residues at the terminals will increase the propensity for miscleavage (Siepen et al., 2007) (Table 4).

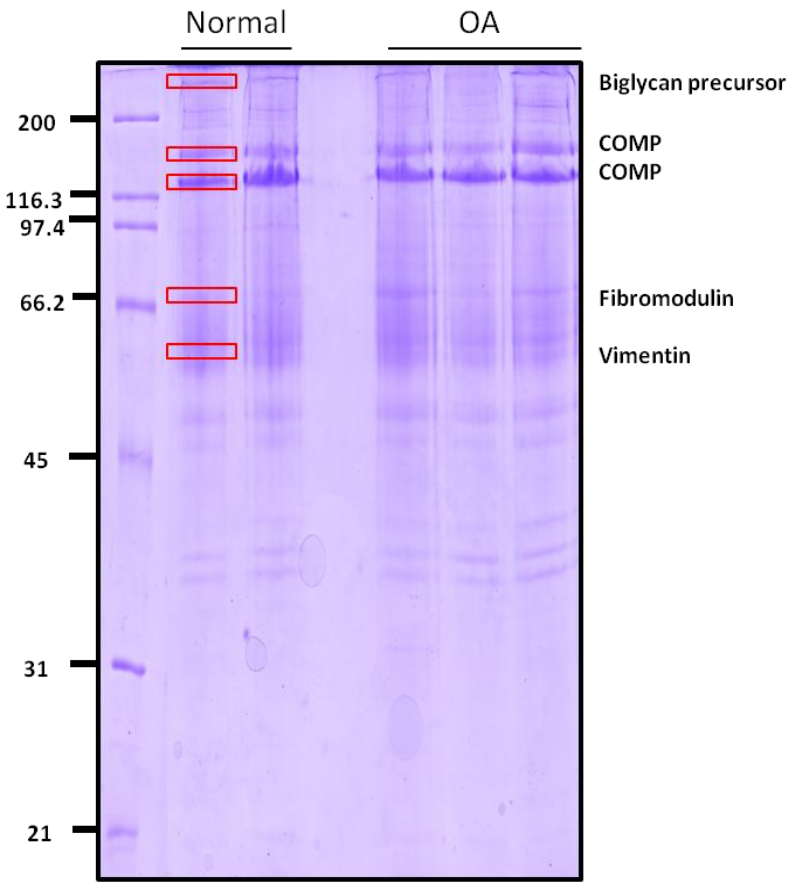



**Figure 3. No qualitative or quantitative differences were evident when 1D SDS-PAGE was used to compare protein components of normal and OA soluble cartilage extracts.** Two normal and three OA samples representative of all samples are shown. Soluble cartilage extract proteins (20µg) from normal and OA donors were resolved by SDS-PAGE and the bands were visualized by Coomassie Brilliant Blue staining. Each lane contains the same proportion of total protein yield in indicate changes in relative amounts. Major bands were excised and analyzed by MALDI-TOF MS after proteolysis.

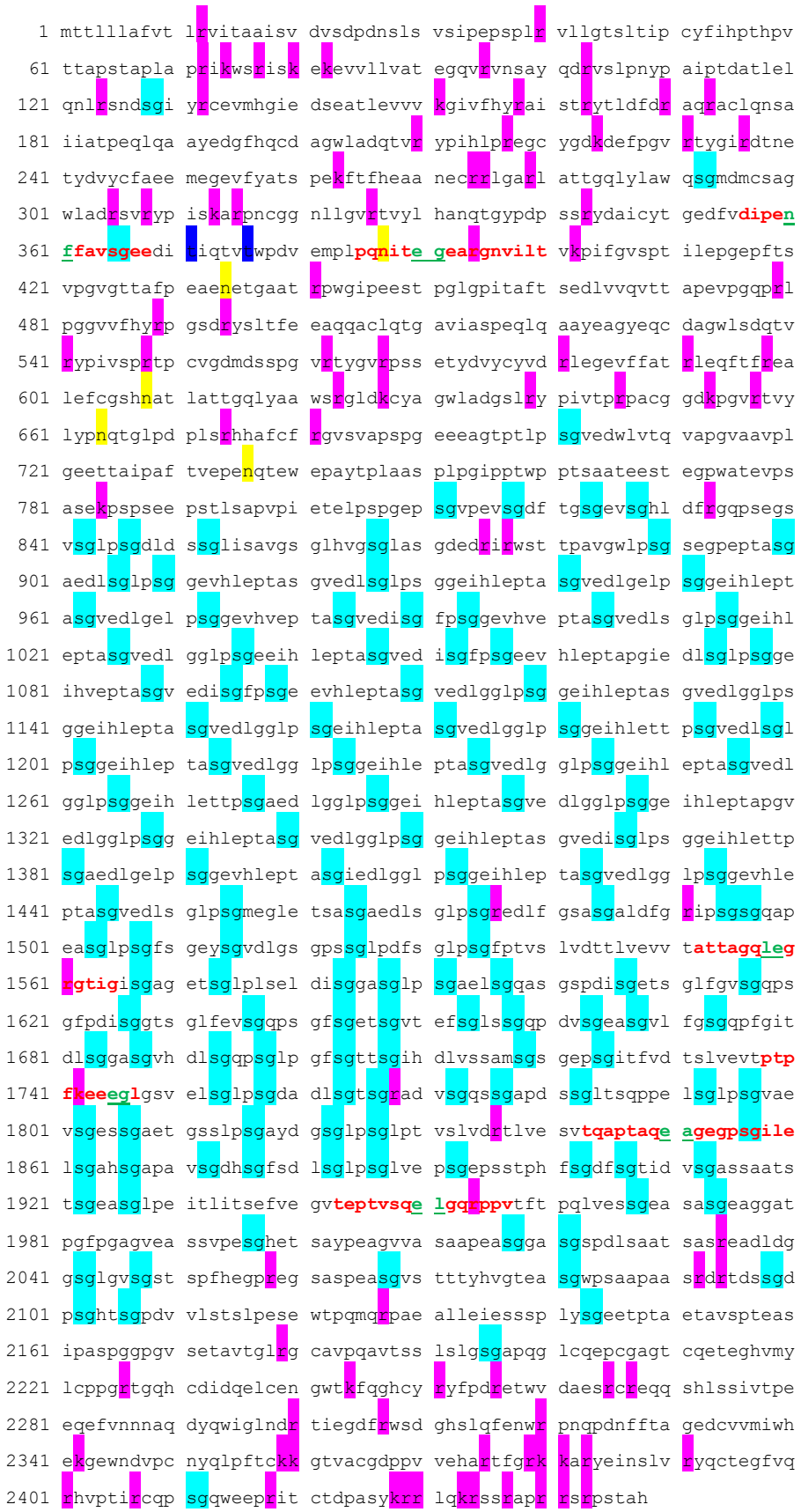


**Figure 4. For aggrecan the abundance of post translational modifications (PTM) surrounding both the tryptic and known cleavage sites meant these sites could not be targeted for q-peptide selection.** Colour coding is as follows; turquoise marked amino acids represent potential chondroitin sulphate attachments, blue marks probable O-linked keratin sulphate glycosylations sites, yellow marks potential N-linked glycosylations, purple represents tryptic cleavage sites which are at the carboxyl side of the amino acids lysine or arginine. Red letters represent the amino acids surrounding the principle aggrecanase and metalloproteinase cleavage sites, green letters are the amino acids between which hydrolysis occurs.

| Cleavage site | Enzyme | | Spanning sequence (NCS) | SPANNING | QconCAT design issue | PTM | Potential |
|---|---|---|---|---|---|---|---|
| DIPE**N** | Trypsin | Carboxyl side of arginine and lysine residues | **>50 amino acids long** | X | | | X |
| NITG**E** | | | | X | | | X |
| ATTAGGL**E** | | | | X | | | X |
| PTPFKEE**E** | | | EE**E**GLGSVELSGLPSGDADLSGTSGR | √ | EE; two glutamic acids* | SG | X |
| TQAPTAQ**E** | | | **>50 amino acids long** | X | | | X |
| TEPTVSQ**E** | | | | X | | | X |
| DIPE**N** | Chymotrypsin | Carboxyl side of tyrosine, phenylalanine, tryptophan and leucine | VDIPE**N**FF | √ | | | √ |
| NITG**E** | | | **>30 amino acids long** | X | | | X |
| ATTAGGL**E** | | | **>30 amino acids long** | X | | | X |
| PTPFKEE**E** | | | KEE**E**GLG | √ | EE; two glutamic acids* | | X |
| TQAPTAQ**E** | | | VESVTQAPTAQ**E**AGEGPSGILE | √ | | SG | X |
| TEPTVSQ**E** | | | VEGVTEPTVSQ**E**L | √ | | | √ |
| DIPE**N** | Glu-C | Carboxyl side of glutamate or glutamate | **N**FFAVSGEE | √ | | SG | X |
| NITG**E** | | | | X | | | X |
| ATTAGGL**E** | | | | X | | | X |
| PTPFKEE**E** | | | **too short** | X | | | X |
| TQAPTAQ**E** | | | | X | | | X |
| TEPTVSQ**E** | | | | X | | | X |
| DIPE**N** | Asp-N | Amino side of aspartate residues | DIPE**N**FFAVSGEED | √ | EE; two glutamic acids* | SG | X |
| NITG**E** | | | **>60 amino acids long** | X | | | X |
| ATTAGGL**E** | | | **>40 amino acids long** | X | | SG | X |
| PTPFKEE**E** | | | **>30 amino acids long** | X | EE; two glutamic acids* | SG | X |
| TQAPTAQ**E** | | | **>40 amino acids long** | X | | SG | X |
| TEPTVSQ**E** | | | **>100 amino acids long** | X | | | X |
| DIPE**N** | Elastase | Carboxyl side of glycine, alanine, proline | PE**N**FFAV | √ | | | √ |
| NITG**E** | | | too short | X | | | X |
| ATTAGGL**E** | | | **E**GRGTI | √ | | | X |
| PTPFKEE**E** | | | PFKEE**E**GLG | √ | EE; two glutamic acids* | | X |
| TQAPTAQ**E** | | | too short | X | | | X |
| TEPTVSQ**E** | | | | X | | | X |

**Table 4. The major aggrecan cleavage sites were inappropriate for designing NCS q-peptides.** The amino acid sequences around of the major cleavage sites for aggrecan were assessed for the potential to produce quantotypic q-peptides for QconCAT design. The size and amino acid composition of the NCS sequences was assessed for suitability and at this point potential sites were nominated in the 'spanning' column. Further analysis of post-translational modifications (PTM) was undertaken; SG denotes probable N-linked glycosylations in the peptides. The final column identified possible cleavage sites for a given protease for QconCAT design.

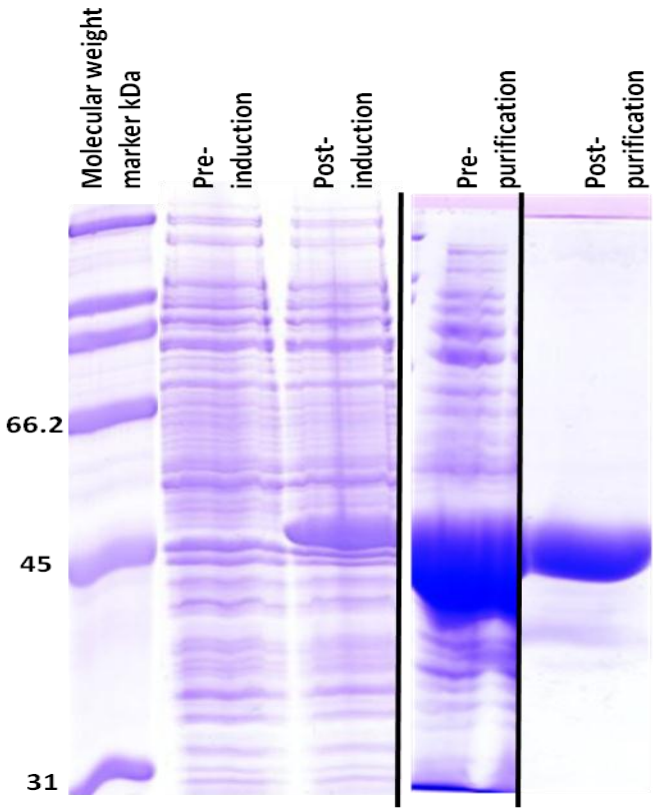



**Figure 5. Electrophoretic profile of the QconCAT at induction and purification.** Whole cell lysate of E.coli BL21DE3 before and after induction of expression of the QconCAT gene supplied in pET21b plasmid. Inclusion bodies were solubilised and purified by nickel affinity chromatography. Following His-tag purification of the QconCAT imidazole-eluted fraction is evident at 45kDa. Black lanes mark none adjacent lanes from the same gel.

The in-gel tryptic digest of this band was analysed with MALDI-TOF mass spectrometry and identified 16 out of 28 peptides from the product ion spectra (Figure 6) in addition to N-terminal fibrinopeptide and C-terminal glufibrinopeptide indicating full length expression of the QconCAT. No miscleaved peptides were identified using label-free quantification indicating complete QconCAT proteolysis.

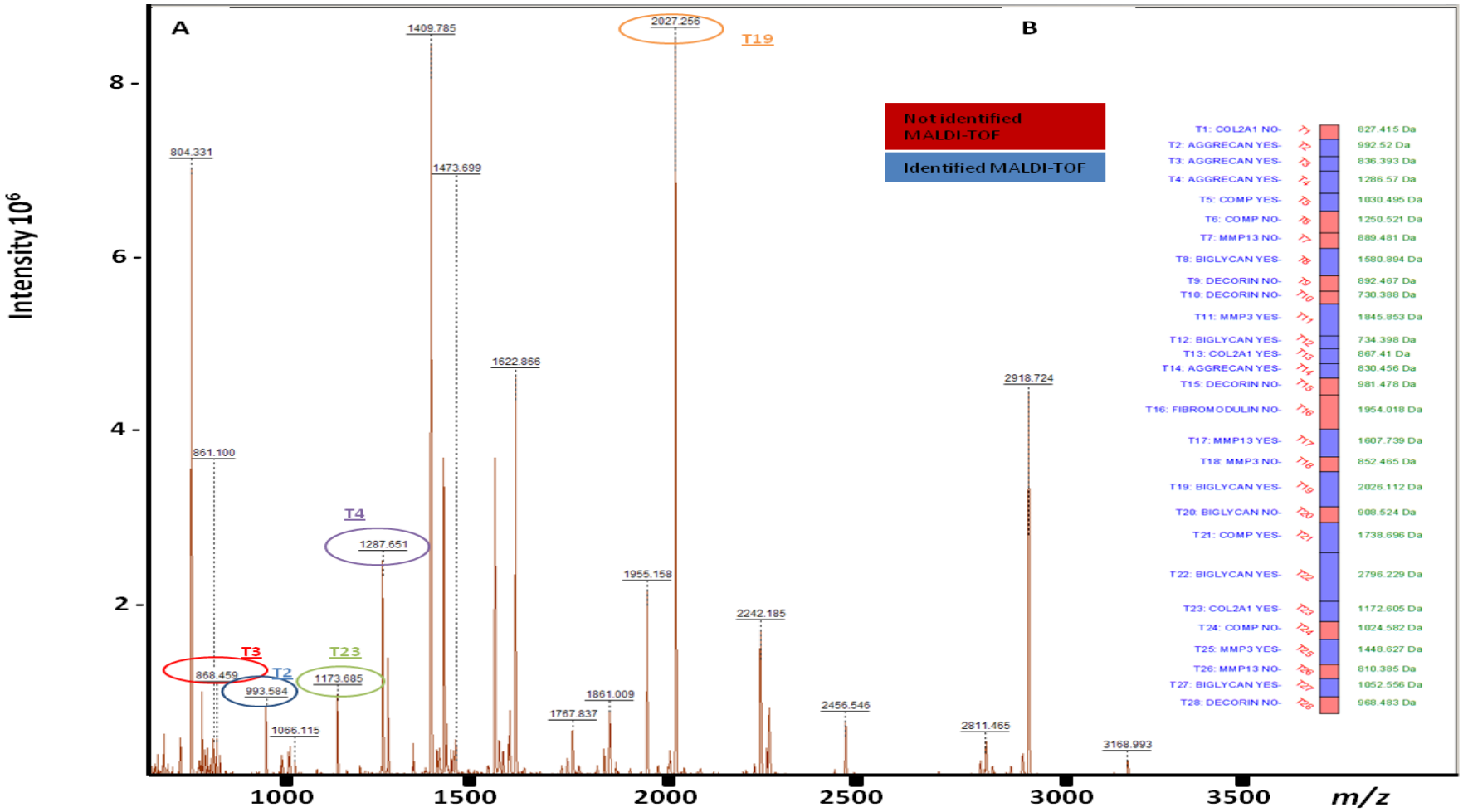



**Figure 6. Validation of the equine degradome QconCAT.** (A) In-gel tryptic digestion of the $^{13}C_6$ lys/ $^{13}C_6$ arg labelled QconCAT was analysed with MALDI-TOF to give a peptide mass fingerprint of derived peptides. The some peptides are indicated as ringed with their respective peptide number. (B) A schematic representation of the disposition of each peptide, those identified (blue) within the QconCAT construct using MALDI-TOF and the respective single charge *m/z*.

**Peptide choice and detectability**

The selection of suitable peptides for quantification was critical. Where possible at least two peptides per protein where chosen in an attempt to reduce the probability of poor peptide ionisation or detectability by the mass spectrometer which would prevent quantification. For proteins we wished to quantify in cartilage utilizing an SRM platform experiment using QconCAT, a simple classification at 10fmol QconCAT spiked into the analyte was applied to peptides as described by Brownridge *et al.* 2011 (Brownridge and Beynon, 2011) and in manuscript 3 (Table 5). This loading was chosen as it was used as the QconCAT loading in quantification experiments. 'Type A' quantifications are defined as were both QconCAT and native peptides are observed. For 'type B' quantifications the peptide is detected in the QconCAT but not in the native peptide, and for these peptides sample protein abundance sets the limit on detection. Neither QconCAT nor native peptides are detected in 'type C' quantification, typically due to poor peptide fragmentation or chromatographic behaviour. Of the 28 peptides composite of the QconCAT we wished to use 26 peptides in order to quantify their constituent proteins within the cartilage soluble extraction; 14 were type A, 6 were type B and 6 were type C. This enabled the quantification of six proteins in normal and OA equine cartilage soluble extracts using the criterion that at least one q-peptide per protein was detected in all samples. Type A peptides were deemed suitable for quantification of the soluble extract of cartilage using an SRM approach in cartilage.

The agreement in quantification between different peptides targeting the same protein was assessed following the quantification of cartilage proteins using SRM. The quantification determined by selected peptides was plotted against each other (Figure 7). There was considerable scatter around the equality line. Quantification with the aggrecan G1 peptide YPIVSPR consistently gave a greater value than the G3 peptides (3.4 fold differences compared to TIEGDFR); which gave similar results. For COMP quantification one of the peptides NTVMECDACGMQPAR gave an 8.2 fold difference in quantification compared to SSTGPGEQLR.

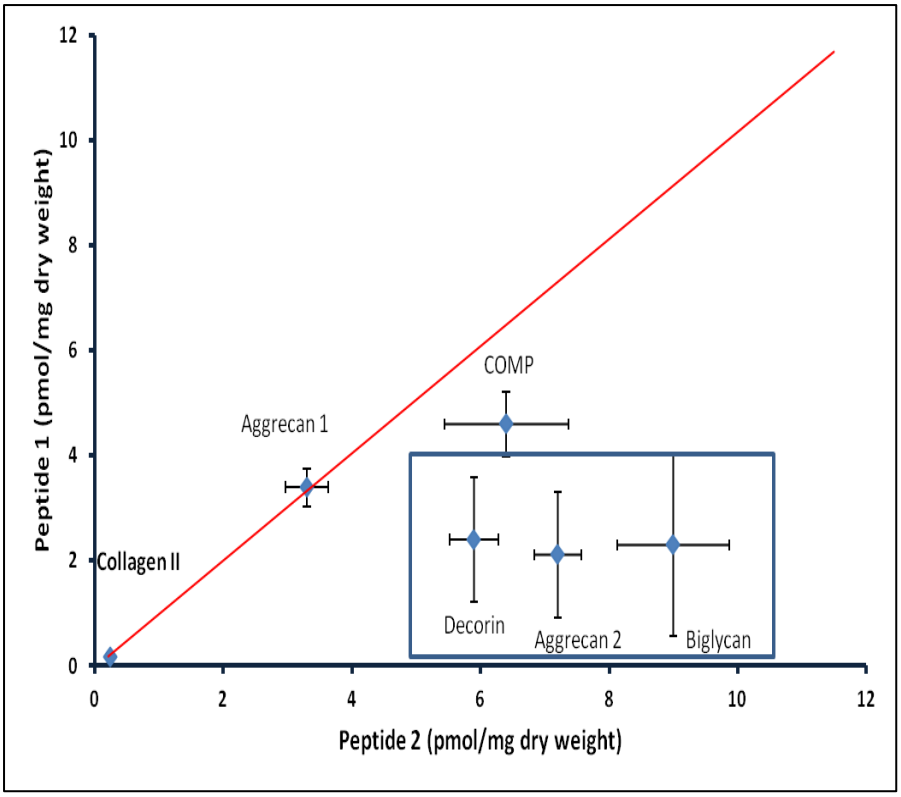



**Figure 7. The agreement in quantification between different sibling peptides targeting the same protein.** The quantitation of intact cartilage proteins determined by each peptide is plotted against each other (siblings). Plots indicate mean ±SEM for each peptide pair (n=20). Quantification values should align on the equality diagonal line (marked in red). Proteins with peptides that vary by two-fold are marked by a blue box. As aggrecan was determined by 4 peptides we compared 2 sets of peptides; peptides the aggrecan G3 peptides YEINSLVR and YQCTEGFVQR (Aggrecan 1) and the G1 peptide; YPIVSPR and remaining G3 peptide; TIEGDFR (Aggrecan 2). For COMP the peptides DTDLDGFPDEK and SSTGPGEQLR were compared. The peptides GVFSGLR and VPAGLPDK were compared for biglycan and the peptides GPEGAQGPR and TGPAGAAGAR for type II collagen.

| Q-peptide | Protein | Evident QconCAT | Evident analyte | Peptide type | Comment |
|---|---|---|---|---|---|
| YPIVSPR | Aggrecan | ✓ | ✓ | A | |
| TIEGDFR | | ✓ | ✓ | A | |
| YEINSLVR | | ✓ | ✓ | A | |
| YQCTEGFVQR | | ✓ | ✓ | A | |
| GVFSGLR | Biglycan | ✓ | ✓ | A | |
| VPAGLPDLK | | ✓ | ✓ | A | |
| GLQHLYALVLVNNK | | X | X | C | poor fragmentation, broad elution profile |
| LAIQFGNYK | | ✓ | X | B | C-terminal peptide |
| NHLVEIPPNLPSSLVELR | | X | X | C | poor fragmentation, broad elution profile |
| NMNCIEMGGNPLENSGFQPGAFDGLK | | X | X | C | split signal; dissimilar oxidation patterns analyte and QconCAT |
| SSTGPGEQLR | COMP | ✓ | ✓ | A | |
| DTDLDGFPDEK | | ✓ | ✓ | A | |
| NTVMECDACGMQPAR | | ✓ | ✓ | A | |
| VDAASLR | Decorin | ✓ | ✓ | A | |
| VHENEITK | | ✓ | ✓ | A | |
| VPGGLADHK | | X | X | C | elution profile, poor fragmentation |
| IPPVNTNLENLYLQGNR | Fibromodulin | ✓ | ✓ | A | |
| TGPAGAAGAR | Collagen Type II | ✓ | ✓ | A | |
| GPEGAQGPR | | ✓ | ✓ | A | |
| SLNNQIESIR | | ✓ | X | B | |
| CGVPDVGHFTTFPGMPK | MMP-3 | X | X | C | split signal; dissimilar oxidation patterns analyte and QconCAT |
| EHGDGDFFPFGPGK | | X | X | C | poor fragmentation |
| GEILFFK | | ✓ | X | B | |
| ISELGFPK | MMP-13 | ✓ | X | B | |
| CGVPDVGEYNVFPR | | ✓ | X | B | |
| GETMVFK | | ✓ | X | B | |

**Table 5. Application of peptide types as determined by SRM experiments.** Q-peptides are classified for quantification purposes as A, B, C. 'Type A' analyte (light) and QconCAT (heavy) peptides (at 10fmol load on column) are detected using a XEVO TQ 'type B' are peptides detected for the QconCAT but not in analyte and when neither QconCAT nor analyte peptides are detected a 'type C' classification is given. Comments relate to possible reasons for non-detection of peptide.

**Q-peptide screening using bioinformatics tools**

A retrospective screening of the q-peptides for peptide detectability and prediction of miscleavage are illustrated in Table 6.

**Quantification of proteins using SRM**

Samples were first analysed using $MS^E$ in order to identify the dilution required of the digest for compatibility with SRM methodologies. This determined that a dilution of 1:20 of the final digest was optimal for all samples. Additionally miscleaves in the target protein were identified consistently for the COMP peptide DTDLGFPDEK at the C terminal as DTDLGFPDEKLR. No other miscleaves in analyte samples were evident.

Where possible, three transitions per peptide were used for quantification. Transitions were defined using Skyline software (O'Connor et al., 1998) and selected after monitoring for the greatest intensity fragments using 50fmol QconCAT digest on the XEVO TQ. y-ions were selected in order to differentiate labelled and unlabelled peptide as the C-terminal residue contained the isotope-labelled amino acid. In addition where possible we selected transitions whose *m/z* was greater than the parent ion *m/z* to maximise specificity (Table 7). Some peptides contained methionines that were subject to variable oxidation, for these transitions of each state of oxidation were utilised and summed. Poor fragmentation of VPGGLADHK and GLQHLYALVLVNNK meant that only a single transition was identified and this was not measurable at a 10fmol load of QconCAT.

The linearity of the response for the QconCAT peptides was established prior to further analyses using a six point standard curve from 0.1fmol to 80fmol load on column labelled QconCAT in an SRM experiment using the XEVO TQ. Data was analysed using a Pearson's correlation coefficient (r). Values of r were between 0.95 and 1.0 which was statistically significant at $p<0.01$ for all peptides.

| | | Analyte context | | QconCAT context | | CONSeQuence | |
|---|---|---|---|---|---|---|---|
| Protein | Q- peptide | N-Terminal Score | C-Terminal Score | N-Terminal Score | C-Terminal Score | Score | Rank |
| Aggrecan | TIEGDFR | 0.49 | 0.47 | 0.41 | 0.42 | 0.267 | 43/59 |
| | YEINSLVR | 0.48 | 0.39 | 0.42 | 0.41 | 0.425 | 22/59 |
| | YPIVSPR | 0.42 | 0.44 | 0.4 | 0.44 | 0.263 | 44/59 |
| | YQCTEGFVQR | 0.39 | 0.39 | 0.42 | 0.42 | 0.458 | 17/59 |
| Biglycan | GLQHLYALVLVNNK | 0.52 | 0.47 | 0.36 | 0.5 | 0.499 | 10/20 |
| | GVFSGLR | 0.51 | 0.43 | 0.36 | 0.49 | 0.323 | 17/20 |
| | LAIQFGNYK | 0.52 | 1 | 0.4 | 0.44 | 0.404 | 13/20 |
| | NHLVEIPPNLPSSLVELR | 0.4 | 0.41 | 0.4 | 0.47 | 0.727 | 2/20 |
| | NMNCIEMGGNPLENSGFQPGAFDGLK | 0.43 | 0.43 | 0.44 | 0.39 | 0.548 | 6/20 |
| | VPAGLPDLK | 0.66 | 0.37 | 0.47 | 0.43 | 0.365 | 15/20 |
| COMP | DTDLDGFPDEK | 0.5 | 0.43 | 0.67 | 0.42 | 0.4 | 20/37 |
| | NTVMECDACGMQPAR | 0.42 | 0.55 | 0.43 | 0.44 | 0.479 | 14/37 |
| | SSTGPGEQLR | 0.49 | 0.43 | 0.42 | 0.67 | 0.359 | 24/37 |
| Type II Collagen | GPEGAQGPR | 0.45 | 0.47 | 0.49 | 0.4 | 0.309 | 54/87 |
| | SLNNQIESIR | 0.35 | 0.52 | 0.39 | 0.46 | 0.481 | 44/87 |
| | TGPAGAAGAR | 0.68 | 0.44 | 0.4 | 0.42 | 0.303 | 56/87 |
| Decorin | VDAASLR | 0.68 | 0.46 | 0.58 | 0.44 | 0.239 | 18/21 |
| | VHENEITK | 0.43 | 0.5 | 0.44 | 0.41 | 0.265 | 17/21 |
| | VPGGLADHK | 0.64 | 0.41 | 0.5 | 0.58 | 0.325 | 15/21 |
| Fibromodulin | IPPVNTNLENLYLQGNR | 0.43 | 0.44 | 0.5 | 0.44 | 0.851 | 1/18 |
| MMP-3 | CGVPDVGHFTTFPGMPK | 0.68 | 0.45 | 0.44 | 0.36 | 0.738 | 2/28 |
| | EHGDFFPFDGPGK | 0.46 | 0.42 | 0.44 | 0.72 | 0.71 | 3/28 |
| | GEILFFK | 0.5 | 0.47 | 0.48 | 0.4 | 0.338 | 16/28 |
| MMP-13 | CGVPDVGEYNVFPR | 0.78 | 0.49 | 0.44 | 0.48 | 0.719 | 6/30 |
| | GETMVFK | 0.56 | 0.49 | 0.72 | 0.4 | 0.22 | 27/30 |
| | ISELGFPK | 0.44 | 0.7 | 0.42 | 0.36 | 0.322 | 23/30 |

**Table 6. Prediction of missed cleavages and peptide detectability.** A prediction tool was used to assess the probability of peptide missed cleavages at N- and C-terminals in analyte and QconCAT context. Using the output score from 0-1, the probability of missed cleavages results in a higher score. For peptide detectability a CONSeQuence score closest to 1 implies good detectability. Each peptide CONSeQuence score is ranked out of the total number of tryptic peptides from its parent protein.

Thus for the quantification of six matrix proteins in our samples the average of three peptides for G3 domain of aggrecan, two peptides for biglycan, two peptides for collagen II, three peptides for COMP and two peptides for decorin, were used. A single peptide was used to quantify fibromodulin and G1 domain of aggrecan. The matrix metalloproteinases MMP-3 and MMP-13 were quantified with one and three peptides respectively.

| Protein | Q-peptide sequence | Precursor m/z | Product m/z | Ion Type |
|---|---|---|---|---|
| Aggrecan | TIEGDFR | 422.20 | 500.24 | y4 |
| Aggrecan | TIEGDFR | 422.20 | 629.28 | y5 |
| Aggrecan | TIEGDFR | 422.20 | 742.36 | y6 |
| Aggrecan | YEINSLVR | 500.20 | 707.43 | y6 |
| Aggrecan | YEINSLVR | 500.20 | 594.35 | y5 |
| Aggrecan | YEINSLVR | 500.20 | 836.47 | y7 |
| Aggrecan | YPIVSPR | 419.23 | 577.36 | y5 |
| Aggrecan | YPIVSPR | 419.23 | 365.20 | y3 |
| Aggrecan | YPIVSPR | 419.23 | 464.27 | y4 |
| Aggrecan | YQCTEGFVQR | 647.29 | 1002.46 | y8 |
| Aggrecan | YQCTEGFVQR | 647.29 | 842.43 | y7 |
| Aggrecan | YQCTEGFVQR | 647.29 | 408.25 | y3 |
| Biglycan | GLQHLYALVLVNNK[+6.0].3 | 529.98 | 593.37 | y5 |
| Biglycan | GVFSGLR | 371.21 | 585.32 | y5 |
| Biglycan | GVFSGLR | 371.21 | 438.26 | y4 |
| Biglycan | LAIQFGNYK[+6.0].2 | 530.30 | 875.47 | y7 |
| Biglycan | LAIQFGNYK[+6.0].2 | 530.30 | 762.39 | y6 |
| Biglycan | LAIQFGNYK[+6.0].2 | 530.30 | 634.33 | y5 |
| Biglycan | NHLVEIPPNLPSSLVELR[+6.0].3 | 678.39 | 906.54 | y8 |
| Biglycan | NHLVEIPPNLPSSLVELR[+6.0].3 | 678.39 | 809.48 | y7 |
| Biglycan | NHLVEIPPNLPSSLVELR[+6.0].3 | 678.39 | 722.45 | y6 |
| Biglycan | NM[+16.0]NC[+57.0]IEM[+16.0]GGNPLENSGFQPGAFDGLK[+6.0].3 | 945.75 | 1229.63 | y12 |
| Biglycan | NM[+16.0]NC[+57.0]IEM[+16.0]GGNPLENSGFQPGAFDGLK[+6.0].3 | 945.75 | 1142.59 | y11 |
| Biglycan | NM[+16.0]NC[+57.0]IEM[+16.0]GGNPLENSGFQPGAFDGLK[+6.0].3 | 945.75 | 1085.57 | y10 |
| Biglycan | NM[+16.0]NC[+57.0]IEMGGNPLENSGFQPGAFDGLK[+6.0].3 | 940.42 | 1229.63 | y12 |
| Biglycan | NM[+16.0]NC[+57.0]IEMGGNPLENSGFQPGAFDGLK[+6.0].3 | 940.42 | 1142.59 | y11 |
| Biglycan | NM[+16.0]NC[+57.0]IEMGGNPLENSGFQPGAFDGLK[+6.0].3 | 940.42 | 1085.57 | y10 |
| Biglycan | NMNC[+57.0]IEMGGNPLENSGFQPGAFDGLK[+6.0].3 | 935.09 | 1142.59 | y11 |
| Biglycan | NMNC[+57.0]IEMGGNPLENSGFQPGAFDGLK[+6.0].3 | 935.09 | 1085.57 | y10 |
| Biglycan | NMNC[+57.0]IEMGGNPLENSGFQPGAFDGLK[+6.0].3 | 935.09 | 810.45 | y8 |
| Biglycan | VPAGLPDLK | 458.28 | 408.75 | y8 |
| Biglycan | VPAGLPDLK | 458.28 | 719.44 | y7 |
| Biglycan | VPAGLPDLK | 458.28 | 816.49 | y8 |
| Collagen II | GPEGAQGPR | 437.72 | 591.33 | y6 |
| Collagen II | GPEGAQGPR | 437.72 | 335.21 | y3 |
| Collagen II | GPEGAQGPR | 437.72 | 463.27 | y4 |
| Collagen II | TGPAGAAGAR | 417.72 | 508.27 | y6 |
| Collagen II | TGPAGAAGAR | 417.72 | 676.36 | y8 |
| Collagen II | TGPAGAAGAR | 417.72 | 579.31 | y7 |
| PIICP | SLNNQIESIR | 590.31 | 510.28 | y4 |
| PIICP | SLNNQIESIR | 590.31 | 623.36 | y5 |
| PIICP | SLNNQIESIR | 590.31 | 979.51 | y8 |
| COMP | DTDLDGFPDEK | 629.27 | 813.35 | y7 |
| COMP | DTDLDGFPDEK | 629.27 | 698.32 | y6 |
| COMP | DTDLDGFPDEK | 629.27 | 1041.46 | y9 |
| COMP | NTVM[+16.0]EC[+57.0]DAC[+57.0]GMQPAR[+6.0].3 | 587.91 | 665.35 | y6 |
| COMP | NTVMEC[+57.0]DAC[+57.0]GM[+16.0]QPAR[+6.0].3 | 587.91 | 681.34 | y6 |
| COMP | NTVMEC[+57.0]DAC[+57.0]GM[+16.0]QPAR[+6.0]_both.3 | 587.91 | 349.23 | y3 |

| Protein | Q-peptide sequence | Precursor m/z | Product m/z | Ion Type |
|---|---|---|---|---|
| COMP | NTVMEC[+57.0]DAC[+57.0]GMQPAR[+6.0].3 | 582.58 | 825.38 | y7 |
| COMP | NTVMEC[+57.0]DAC[+57.0]GMQPAR[+6.0].3 | 582.58 | 665.35 | y6 |
| COMP | SSTGPGEQLR | 519.26 | 705.38 | y6 |
| COMP | SSTGPGEQLR | 519.26 | 762.40 | y7 |
| COMP | SSTGPGEQLR | 519.26 | 863.45 | y8 |
| Decorin | VDAASLR | 369.21 | 523.33 | y5 |
| Decorin | VDAASLR | 369.21 | 381.26 | y3 |
| Decorin | VDAASLR | 369.21 | 638.36 | y6 |
| Decorin | VHENEITK | 488.26 | 438.73 | y7 |
| Decorin | VHENEITK | 488.26 | 739.39 | y6 |
| Decorin | VHENEITK | 488.26 | 610.35 | y5 |
| Decorin | VPGGLADHK[+6.0].2 | 450.26 | 800.44 | y8 |
| Fibromodulin | IPPVNTNLENLYLQGNR | 654.35 | 756.39 | y6 |
| MMP-13 | CGVPDVGEYNVFPR | 807.89 | 1298.65 | y11 |
| MMP-13 | CGVPDVGEYNVFPR | 807.89 | 649.83 | y11 |
| MMP-13 | CGVPDVGEYNVFPR | 807.89 | 987.50 | y8 |
| MMP-13 | GETM[+16.0]VFK[+6.0].2 | 417.21 | 647.35 | y5 |
| MMP-13 | GETMVFK[+6.0].2 | 409.21 | 760.40 | y6 |
| MMP-13 | GETMVFK[+6.0].2 | 409.21 | 631.36 | y5 |
| MMP-13 | GETMVFK[+6.0].2 | 409.21 | 530.31 | y4 |
| MMP-13 | ISELGFPK | 448.76 | 696.40 | y6 |
| MMP-13 | ISELGFPK | 448.76 | 783.43 | y7 |
| MMP-13 | ISELGFPK | 448.76 | 567.36 | y5 |
| MMP-3 | C[+57.0]GVPDVGHFTTFPGM[+16.0]PK[+6.0].3 | 623.63 | 698.36 | y6 |
| MMP-3 | C[+57.0]GVPDVGHFTTFPGM[+16.0]PK[+6.0].3 | 623.63 | 551.30 | y5 |
| MMP-3 | C[+57.0]GVPDVGHFTTFPGMPK[+6.0].3 | 618.30 | 884.46 | y8 |
| MMP-3 | C[+57.0]GVPDVGHFTTFPGMPK[+6.0].3 | 618.30 | 783.42 | y7 |
| MMP-3 | C[+57.0]GVPDVGHFTTFPGMPK[+6.0].3 | 618.30 | 535.30 | y5 |
| MMP-3 | EHGDFFPFDGPGK[+6.0].3 | 485.89 | 723.38 | y7 |
| MMP-3 | EHGDFFPFDGPGK[+6.0].3 | 485.89 | 626.32 | y6 |
| MMP-3 | GEILFFK | 430.25 | 673.44 | y5 |
| MMP-3 | GEILFFK | 430.25 | 560.35 | y4 |
| MMP-3 | GEILFFK | 430.25 | 447.27 | y3 |

**Table 7. Parameters used in SRM assays**. The parent and fragment ion *m/z* and fragment type for all unique peptides. Peptides and transitions used to quantify the soluble cartilage extract in normal and OA cartilages are coloured blue.

For all donors, extracted ion chromatograms were performed for each peptide and the total ion count used to determine the ratio of light peak area/ heavy peak area with 10fmol QconCAT loading. When multiple peptides were employed for quantification the average of these peptides was used for intact quantification.

Results were normalised to dry weight of explants. It was identified that disease status had no significant effect on the concentration of G1 or G3 aggrecan, biglycan, COMP, collagen II, decorin or fibromodulin in cartilage when the either the quantification was compared between mean values of all peptides or by comparing NCC peptides only. However there was a significant reduction in the quantification of the biglycan NCS peptide GVFSGLR between normal (12.8±1.5 pmol/mg and OA cartilage (4.5±1.3pmol/mg) ($p=0.01$) (Figure 8). We were unable to identify and therefore quantify analyte peptides from cartilage extracts for MMP-3 or MMP-13 using QconCAT. However the MMP-3 q-peptide GEILFFK was quantified in the equine secretome (see appendix to manuscript 2).

Although there was no difference using in G1/G3 ratio of concentration between normal and OA cartilage, SRM data indicated that the mean concentration of G3 aggrecan peptides significantly decreased with age relative to the G1 (YPIVSPR) domain with a significant Pearson's correlation coefficient of 0.79 $p<0.05$ (Figure 9).

We were unable to quantify the extent of hydrolysis of targeted matrix proteins due to the variability in peptide abundance between cleaved (NCS) and non-cleaved peptides (NCC). This was because for biglycan, COMP and decorin the concentration of the NCS peptides were greater than the concentration of the NCC peptides (Figure 8).

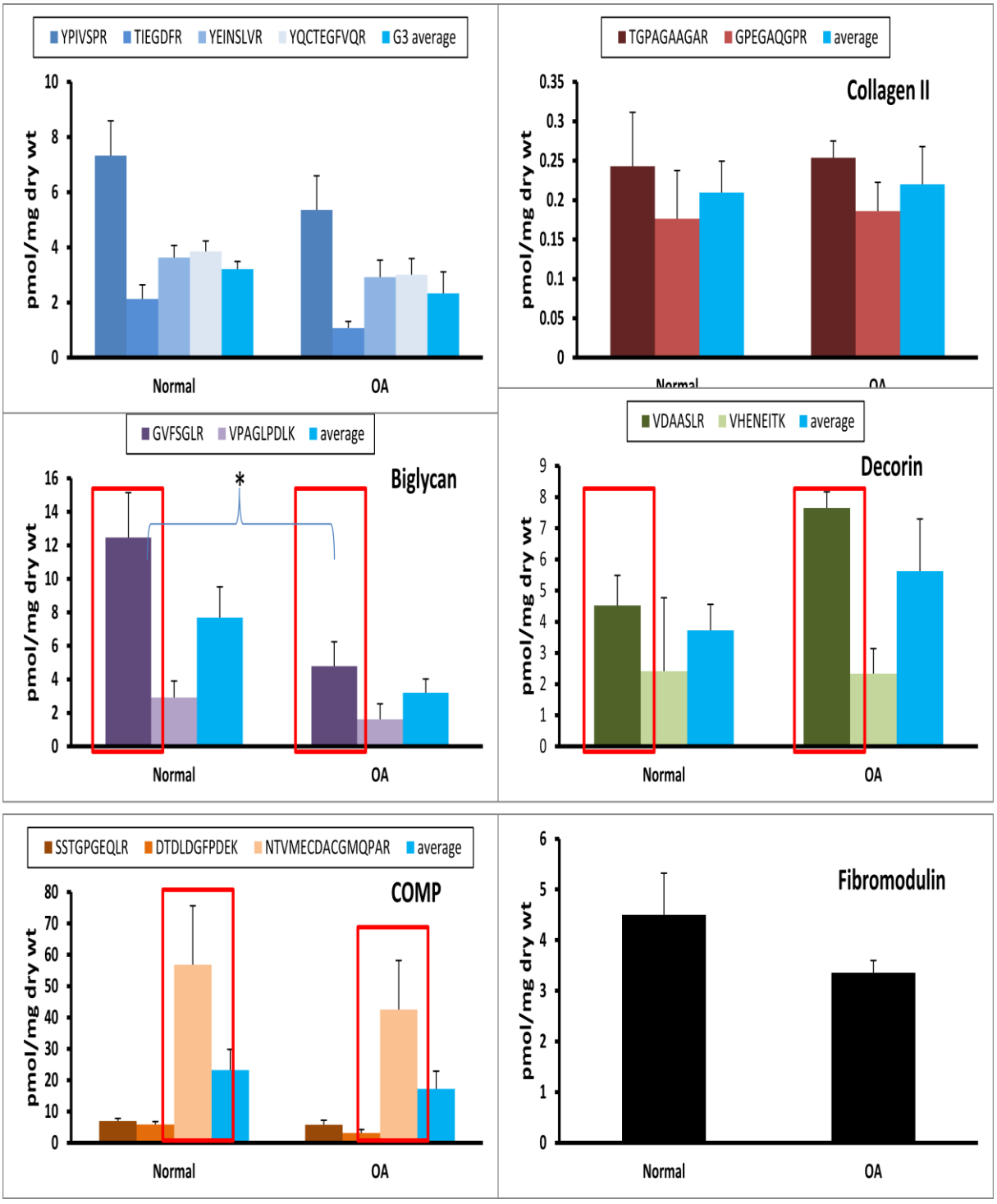



**Figure 8. Quantification of cartilage matrix proteins.** Graph represents individual and mean concentration of peptides ±SEM measured in normal (n=11) and OA (n=9) equine cartilage using QconCAT. Non cleaved spanning peptides are indicted by red boxes. Data was evaluated using mixed effect linear regression. * denotes statistical significance p=0.01.

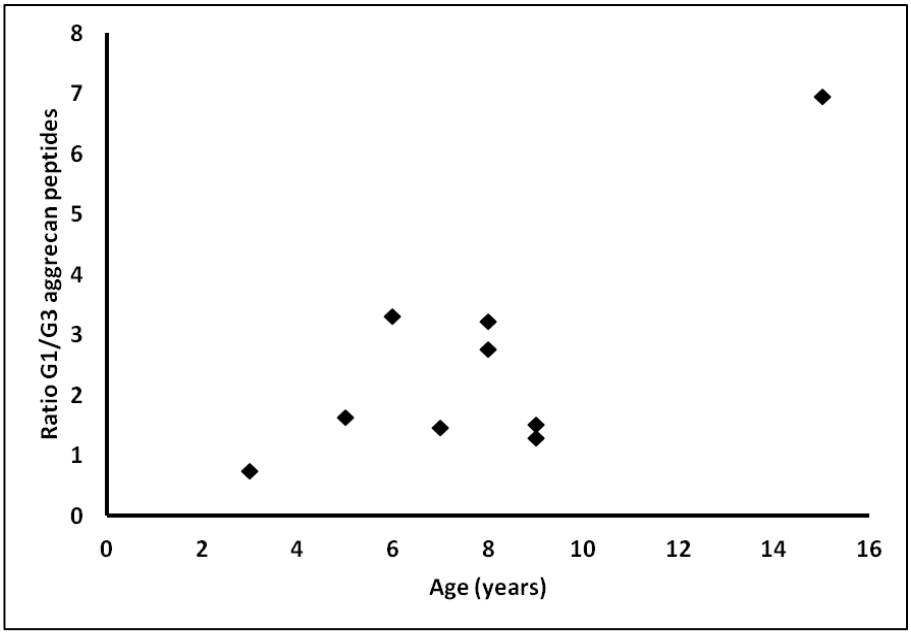



**Figure 9. Relationship of G1 and G3 domains with age in normal equine cartilage**. The scatter plot represents the concentration ratio of G1/G3 determined using SRM analysis plotted against age of horse.

**Validation of SRM data and spin column methodology using immunoblotting**

Immunoblotting revealed a lack of MMP-3 and COMP in the spin column supernatant and protein wash flow through (data not shown). Quantitative immunoblotting was used to confirm the results for the peptides MMP-3 and COMP in normal and OA cartilage. Equal amounts of protein samples were analysed by Western blotting. There was a lack of MMP-3 (data not shown) in the soluble cartilage extract of normal and OA cartilage which agreed with our SRM data. In addition there was no significant alteration in COMP expression between normal and OA cartilage (Figure 10). Despite BLAST predictions neither anti-GAPDH-horseradish peroxidase nor anti-β-actin antibodies worked.

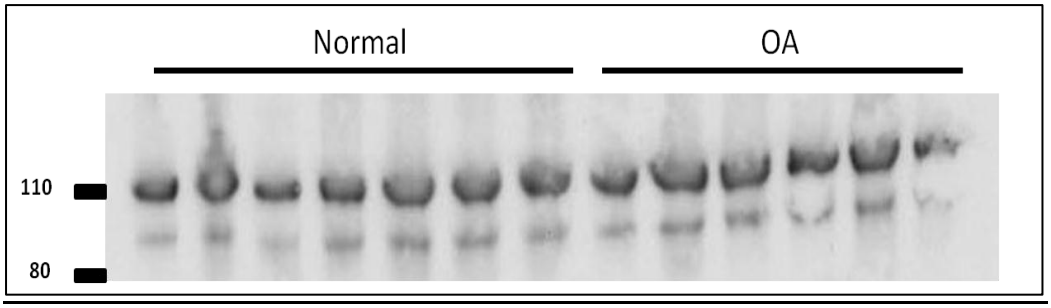



**Figure 10. Effect of disease state on COMP protein levels in equine cartilage.** Western blot analysis, using COMP antibodies of normal and OA soluble cartilage extracts. Densitometry confirmed no significant difference between band intensity between samples (data not shown).

**Gene expression of matrix proteins and proteases in normal and OA cartilage**

To investigate potential patterns of correlation between protein and transcriptional regulation and to determine differential expression of further proteases pertinent to OA pathogenesis comparative gene expression profiles of matrix proteins and proteases were generated for all samples. Significant increase in the expression of MMP-1, MMP-3 and ADAMTS-5 were identified in OA compared to normal cartilage (Figure 11). For matched donors there was only a significant positive correlation of Collagen II gene expression and protein concentration of normal equine cartilage ($r=0.63$, $p<0.02$).

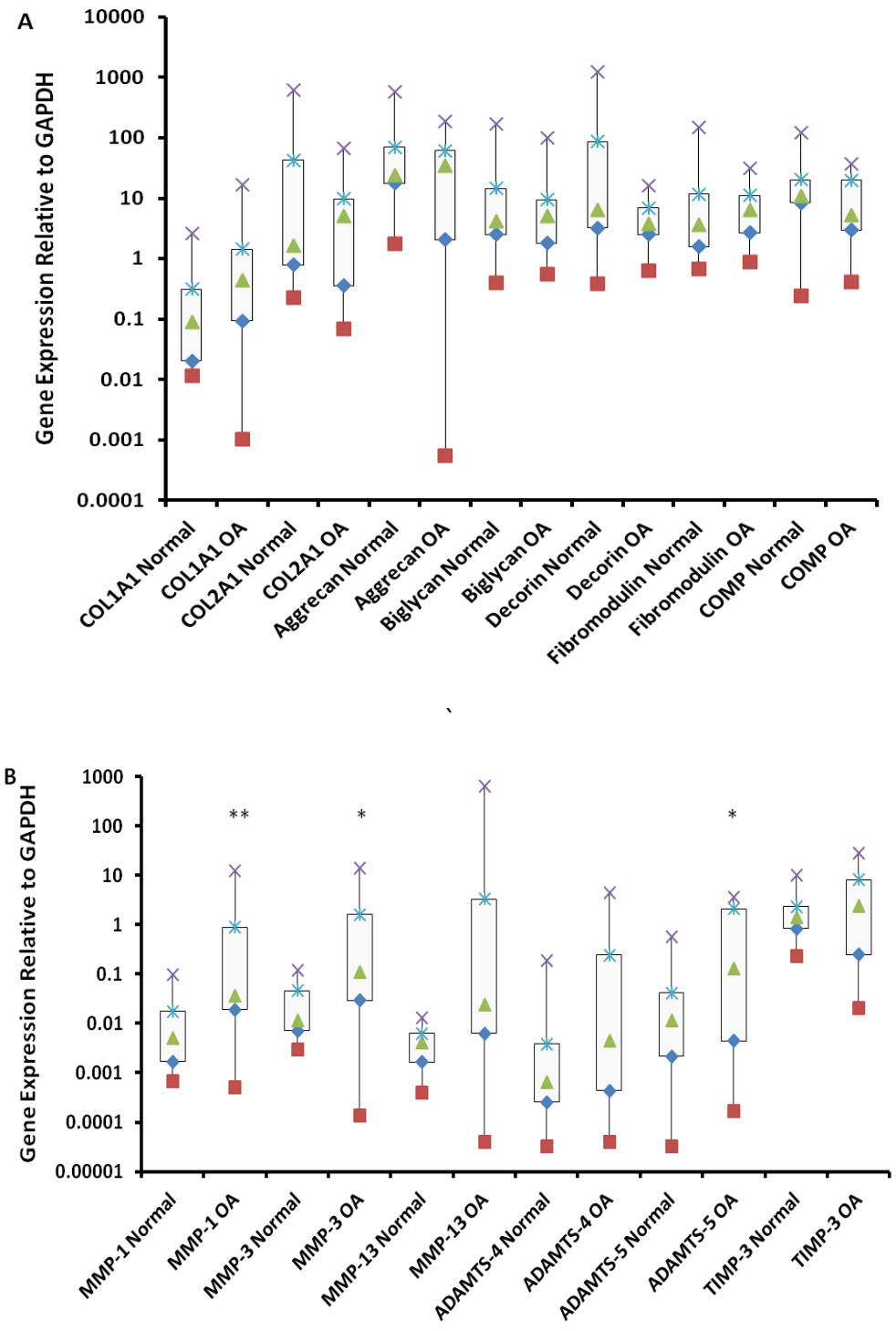



**Figure 11. Mean relative change in gene expression of each gene evaluated in normal (n=11) and OA (n=19) cartilage for A. Matrix genes and B. Proteases.** Statistical significance differences from normal tested using mixed effects linear regression is defined as *($p<0.05$) **, ($p<0.01$). q1 and q3 (blue triangle and blue cross), minimum and maximum values (red square and purple cross), median (green triangle), interquartile range (box) and 95% confidence intervals (whiskers) are shown.

**Discussion**

A prominent feature of OA is the catabolism of the ECM. Many of the proteins as well as the sites and types of proteases involved in cleavage are known. Currently cleavage is identified for some of these sites by neoepitope antibodies. Whilst immunoblotting is sensitive it is poorly quantitative. In this study we wished to develop a novel technique to quantify in absolute terms some sites of cartilage degradation as well as amounts of intact matrix proteins using parallel measurement with QconCAT. We attempted to quantify matrix proteins, the extent of matrix degradation and some of the effectors in normal and OA cartilage using QconCAT.

The lack of published absolute data for intact and hydrolysed matrix proteins is a result of two issues. Firstly the inherent nature of the structure and composition of cartilage, means it is difficult to produce reliable and reproducible data for cartilage. Secondly until relatively recently there has been a lack of a suitable sensitive high throughput proteomic platform. Here we developed a relatively simple technique to safely extract soluble cartilage proteins and in addition the unique capability of triple quadrupole (TQ) mass spectrometry allowed the most sensitive MS platform to be utilised.

Characterisation of samples into normal or OA groups was based on macroscopic grading of the distal condyles of MCIII using the Kawcak *et al.* method (McIlwraith et al., 2010). Some donors that appeared macroscopically normal had histological changes associated with mild OA. Other gross scoring methods containing additional features involving the entire joint (Pritzker et al., 2006) may be better at identifying mild OA, subsequently identified on histological evaluation. Indeed previous studies elucidated microscopic changes in cartilage demonstrating a lack of macroscopic changes (Acebes et al., 2009). Interestingly it has been documented that microscopic changes in the joint, such as in subchondral bone, can precede macroscopic changes associated with OA (Wang et al., 2013). Conversely some samples which on macroscopic examination were classified as OA, on histological assessment produced a relatively low Mankin's score. This was probably due to the

limitations of the microscopic area retrieved, an important consideration given the highly heterogeneous and regional variability in lesions.

As the proteome is complex there is no one method for preparing protein samples for MS. A prerequisite to absolutely quantifying ECM proteins from cartilage is their complete extraction and digestion from the native tissue. It was also necessary to discard, following guanidine extraction, the insoluble collagens and glycosaminoglycans (GAGs) which may interfere with tryptic digestion and can be catastrophic to the fragile capillary chromatography column in the liquid chromatography (LC) system used to resolve the protein digest. Furthermore their abundance can cause dynamic range problems. The insoluble collagen rich fraction was first separated from the soluble guanidine extract. Then we used molecular weight-based fractionation to enable sample clean-up, remove most remaining collagens and GAGs and enable efficient in-solution protein digestion. A MWCO size was chosen for this specific experiment both to prevent the displacement of small intact proteins through the membrane during protein clean-up and to keep high MW GAGs in the final supernatant. The insoluble fraction following tryptic digest was collected and frozen for further analysis. As both biglycan (Schonherr et al., 1995) and decorin (Brown and Vogel, 1989) are known to interact with collagen, analysis of this fraction would be useful to assess the amount of retained proteins in the insoluble fraction.

Complete proteolysis was paramount for the absolute quantification workflow as analyte is compared to standard (QconCAT) at the peptide level. For native proteins a primary factor in the rate and extent of proteolysis is the higher-order structure (Jakoby et al., 2012). Therefore we instigated steps that removed potential impediments for digestion by the reduction of disulphide bonds and the prevention of reformation and by the presence of small amounts of guanidine during tryptic digestion.

The 1-D SDS-PAGE gels of normal and OA cartilage soluble fractions were as predicted, identical in agreement with studies in man (Wu et al., 2007) and mice

(Wilson et al., 2008). Although digested excised bands contained peptides from multiple proteins, as expected the predominant matrix proteins biglycan, COMP and fibromodulin were the abundant identifications. Aggrecan with its high MW was not identified as a major band as its molecular weight is >2500kDa (aggrecan core is 250kDa if all GAG chains are removed).

In this study we adopted SRM of multiple product ions (transitions) for quantification as it currently provides the most sensitive method for protein quantification using MS. The lower detection limit for peptides is enhanced by up to 100 fold compared to full scan MS/MS analysis (Keshishian et al., 2007). The platform was chosen to reduce costs which would have been incurred by running samples on multiple instruments. For instance high abundance intact matrix proteins could have been quantified using accurate mass retention time on a different instrument but this would have doubled instrument time.

The initial goal of this study was to quantify aggrecan cleavage using QconCAT. However this was not possible due to the amino-acid sequence context and positioning of PTMs. The proteolytic enzyme of choice in MS studies trypsin targets the carboxyl-side of arginine and lysine for hydrolysis. Unfortunately these were scarce around the main aggrecan cleavage sites (Madsen et al., 2010). Chymotrypsin would have been the proteolytic enzyme of choice as digestion *in-silico* identified two potential NCS peptides for inclusion into the QconCAT. Indeed recently an assay for detection of two specific cleavage sites on aggrecan used chymotrypsin together with an immunoaffinity based LC-MS/MS method to detect cleavage at the $ARG^{374}$ site and the $AGEG^{1820}$ site (Dufield et al., 2010b). Unfortunately when other proteins of interest for inclusion in the QconCAT were assessed for potential chymotrypsin derived peptides few were identified as suitable. Thus a number of NCS peptides containing cleavage sites of other matrix proteins were included in order to test proof of concept of our approach.

From the initial pool of 26 unique peptides 14 gave reliable signal in analyte and standard and were quantified. For all proteins except fibromodulin we used at least two peptides to quantify intact proteins in the samples. This was owing to either an

inability to detect peptides in both light and heavy forms ('type C' peptides) or an inability to detect analyte peptides. These latter 'type B' peptides (primarily MMPs) may be a result of the presence of analyte peptides below the limits of MS detection or because they were not present in the sample. Using immunoblotting we were unable to identify the presence of MMP-3 in the soluble extraction fraction of normal and OA cartilage in agreement with the SRM data. Previous studies on normal and OA human cartilage using guanidine extraction followed by metalloproteinase assays revealed the presence of acidic and neutral MMPs with elevated levels in OA cartilage (Dean et al., 1989). Therefore we are uncertain why MMP-3 was not identified in our soluble extract. Using immunoblotting with antibodies for MMP-3 (which binds both active and inactive forms) of the spin column supernatant and protein wash flow through also failed to identify the protein. As the majority of MMP-3 present in cartilage is in the proenzyme form (Lombard et al., 2005) it was thought that immunoblotting would identify it, if it was indeed present in the sample. The size exclusion of the membrane within the spin column was selected at 10kDa in order to retain smaller proteins (such as MMP-3 with a MW of 50kDa) we wished to quantify above the membrane prior to digestion. There are possibly two explanations for the failure to identify MMPs in the samples. It is possible that even though the manufacturers try to reduce non-specific protein binding to the membrane (through its composition), that this is occurring and evident only for low abundance proteins such as MMPs. However it was thought that even if this occurred following tryptic digestion the small peptides would be released and so identified in downstream analysis. Alternatively MMPs could be binding to the insoluble collagen fraction. Collagen binding by MMP-1 is known to occur (Tam et al., 2002) but there is no evidence that this occurs in MMP-3 or MMP-13.

A further 'type B' peptide we were unable to detect here was the PIICP peptide. This low abundance peptide may not have been detected due to sample preparation or concentrations present below the limits of MS detection, despite scheduling of transitions to improve detection rate (Picotti et al., 2008). Further work using immunocapture followed by LC-MS/MS, similar to methods used to

detect aggrecan fragments (Dufield et al.), may enable analysis of cartilage synthesis with this peptide.

For the majority of proteins at least two peptides were selected for intact protein quantification to account for redundancy following validation of the peptides in the quantification strategy. Inferring protein abundance from peptide abundance is difficult since peptides for the same protein can disagree. When possible we used proteotypic tryptic peptides identified in our previous studies. Even with this approach, some peptides were still 'type C' and thus redundant for our quantification. There are a number of possible reasons for this. In order to quantify specific areas of a protein, for instance a cleavage site or propeptide junction, it was necessary to use amino acid sequences that would usually be avoided for instance in the peptide CGVPDVGHFTTFPGMPK, a spanning peptide over the propeptide/peptide junction of MMP-3. Although the presence of cysteine (C) is ideally avoided, the reduction and alkylation steps in the methods means inclusion is not problematic (Simpson and Beynon, 2012). This peptide also contains methionine which may present variable oxidation, (a satellite peptide of +16Da). Our study aimed to sum the intensities of all oxidation states in order to quantify this peptide. However only at greater than 100 fmol loading QconCAT on the TQ did we identify predetermined transitions from this peptide possibly due to the split signal and low signal to noise ratio which made it more difficult to distinguish. An additional problem was that for some peptides, including CGVPDVGHFTTFPGMPK, the oxidized form in the QconCAT was much less than for the corresponding analyte peptide making it difficult to produce an adequate transition list for all oxidized forms. Although an attempt was made to force oxidation of the QconCAT using hydrogen peroxide (Corless and Cramer, 2003) this was unsuccessful (data not shown). A further reason for poorly performing peptides is poor fragmentation; leading to a lack of sensitivity in SRM assays. For example the MMP-3 q-peptide EHGDFFPFDGPGK was identified in our previous studies (manuscript 1) in both MS and MS/MS data but when a transition list was produced for this peptide only a single suitable transition was found and this was unusable due to counts in the hundreds at even 100fmol loading. The peptide would have been in greater

abundance in the secretome samples loaded onto the LC in discovery experiments, as well as being in a more 'simple' proteome and for these reasons it was evident in tandem MS but not SRM studies, the latter being undertaken in much diluted samples. This is interesting because CONSeQuence predicted that this MMP-3 peptide was 3/28 in the rankings of MS detectability. Other peptides had poor chromatographic behavior such as GLQHLYALVLVNNK which had a broad elution profile of greater than the retention time window applied in to SRM methods, thus reducing sensitivity.

Subsequent to the design of the QconCAT a new prediction tool for reference peptide selection was released called CONSeQuence (consensus predictor for quantotypic peptide sequence) which predicts peptide detectability with 75% cross-validated accuracy by electrospray ionization MS (Eyers et al., 2011). Interestingly when applied to our QconCAT some 'type C' peptides such as CGVPDVGHFTTFPGMPK, EHGDFFPFDGPGK and NHLVEIPPNLPSSLVELR had a high score indicating detectability, whilst some 'type A' peptides such as YPIVSPR and TIEGDFR had a low score indicating poor detectability. This highlights that a reliance on a single bioinformatics tool is not sufficient for selecting q-peptides.

In order to quantify protein hydrolysis we attempted to develop a technique which quantified a number of NCC sites within the protein and use this, with quantification at a known cleavage site (NCS) to determine the amount of cleavage present. Thus we were assuming that if hydrolysis had occurred at a given site there would be a reduction in quantification of the NCS compared to the NCC peptide for that protein. This approach relied on a good correlation between NCC sibling peptides which are quantified independently. However not only were we unable to use all the chosen NCC peptides for each protein (due to the classification of a number of 'type C' peptides), but following data analysis using the remaining peptides it became evident that inconsistency in peptide quantification would not enable hydrolysis to be quantified with our method. However by analysis of NCS peptides it may be possible to determine the amount of cleavage occurring for at some sites under different conditions. The NCS peptide GVFSGLR which contains an

MMP-13 cleavage site $^{181}$G$^{182}$V in biglycan (Monfort et al., 2006) was quantified as being significantly lower in OA compared to normal samples. MMP-13 can degrade all small leucine-rich repeat proteoglycans (SLRPs) but biglycan and fibromodulin are cleaved preferentially (Monfort et al., 2006). SLRPs interact with collagen influencing fibril formation and interaction (Tchetina et al., 2005). Defective linkage of collagen with subsequent network instability occurs following degradation of the SLRPs. In OA their breakdown may occur prior to major collagen destruction and thus contribute to the process (Heathfield et al., 2004). The biglycan NCS peptide GVFSGLR may provide a useful marker of cartilage degradation as it could be used to determine the role of biglycan breakdown in OA progression.

Quantification performance can be measured by plotting values from sibling peptides against each other. In this study G3 aggrecan and one COMP sibling pair together with type II collagen aligned close to the equality diagonal. Disparity between peptides can be attributed to a number of issues. Interfering signals resulting in low signal to noise will make reliable quantification difficult due to a reduction in the quality of the peak. A reduction in signal can occur due unanticipated post-translational modification (PTM) or to miscleave in the analyte or QconCAT. In addition although protease inhibitors were used in the workflow there is still a possibility of exopeptidases removing terminal amino acids. This would reduce the *m/z* and not allow identification or quantification. Our QconCAT design rejected peptides with known PTMs which would interfere with quantification through signal splitting between modified and unmodified forms. However it may be that although database searching of experimental data was not suggestive of PTMs there were unidentified PTMs, thus underestimating protein quantification.

Reduced quantification values are evident when there is miscleavage in the analyte. Attempts made to identify potential miscleaves using label-free quantification consistently identified a single miscleave related to this study in the analyte as DTDLGFPDEKLR. Indeed compared to one of the other NCC peptides SSTGPGEQLR, there was a reduction in quantification using this peptide. The other NCC peptide;

NTVMECDACGMQPAR produced a much greater quantification than either DTDLGFPDEKLR or SSTGPGEQLR. A higher quantification can be due to QconCAT miscleavage. This is easier to identify than analyte miscleaves as increased column loadings can be achieved in order to recognize low intensity miscleaves. Not only can MS detectors be saturated with more abundant peptides at higher loading, but the LC columns can collapse with protein overloading. As no detection of miscleavage in the QconCAT digest was detected, even at high column loading, this is unlikely to be the reason for NTVMECDACGMQPAR giving a higher quantification of COMP. Previous authors have stated that when miscleaves at high loadings of QconCAT are not identified, the disparity can assumed to be the analyte and so the higher quantification should be used (Simpson and Beynon, 2012). In this study however, when we were unable to identify the cause of the discrepancy we used the mean of all NCC peptides quantified to give an absolute value for intact proteins. For example for intact COMP the mean of all three peptides was utilized. It might be expected that disparity occurs between NCC and NCS peptides where hydrolysis has occurred. For both biglycan (GVFSGLR (NCS) and VPAGLPDLK (NCC)) and decorin (VDAASLR (NCS) and VHENEITK (NCC)) the mean of both peptides were used in quantification. VPAGLPDLK was identified retrospectively as having a high probability of N terminal miscleavage possibly due to a proline residue as the second amino acid (Siepen et al., 2007). Although we were unable to substantiate this, the resulting quantifications with this peptide were lower than with GVFSGLR. However, VDAASLR also identified as having a high probability of N terminal miscleavage using the bioinformatics tools (due to an acidic residue two positions from the cleavage site) but not label-free, produced a quantification value greater than VHENEITK. This is not suggestive of a missed cleavage for VDAASLR in the analyte. Thus peptide disparities meant that the remaining NCS peptides quantifications; GVFSGLR and VDAASLR conferred greater values than the NCC in both normal and OA cartilage. As a result hydrolysis could not be measured using this QconCAT and so we elected to employ the peptides in the quantification of the intact proteins.

The G1 aggrecan peptide; YPIVSPR was more abundant in cartilage than the three G3 aggrecan peptides. Two processes arise in cartilage to explain the differential expression of G1 and G3 in cartilage. Firstly soon after cartilage aggrecan synthesis there is a small loss of G3 (Paulsson et al., 1987). Secondly normal turnover produces a steady decline of G3 (Ratcliffe et al., 1986). This is from cleavages in the chondroitin sulphate (CS) region (Caterson et al., 2000), so that the average size of aggrecan decreases with its age (Dudhia et al., 1996) and hence a large proportion of aggrecan lacks a G3 domain (Paulsson et al., 1987). The QconCAT provides a useful tool to monitor changes in the different regions of aggrecan in cartilage.

Although data did not identify any significant differences in intact matrix proteins between normal and OA cartilage we were able to provide base-line levels for some intact matrix proteins. Our data indicated that of the matrix proteins quantified COMP was the most abundant. Interestingly the concentrations of the small leucine rich proteoglycans (SLRP) were similar to aggrecan. Indeed although in cartilage aggrecan accounts for the greatest proportion of the proteins present by volume (Hardingham et al., 1986) (due to the size of aggrecan and the number of side chains) our results are in agreement with others who hypothesized that SLRPs were present in similar amounts to aggrecan (Hardingham and Bayliss, 1990). Furthermore it has been previously stated of the dermatan sulphate proteoglycans whose members are biglycan and decorin, that decorin is the predominant species in adult human cartilage. This observation was following preliminary $^{35}S$ sulphate labelling studies on a limited number of specimens (Melching and Roughley, 1989). However, the study stated that they could not exclude the possibility that biglycan remained within the cartilage residue following extraction. Although further literature was sought to identify which is the predominant of the two, references were only found that referenced this original paper. This study would suggest that biglycan is the predominant species in normal adult cartilage. However in OA cartilage decorin becomes the predominant.

Quantitative changes in gene expression between normal and OA cartilage were studied to establish whether there was a correlation between protein and transcriptional abundance. In order to understand what is occurring in OA, both

mRNA and protein levels of cartilage are necessary. In cartilage the majority of the protein is within the ECM whilst the mRNA is sourced from chondrocytes. Whilst knowledge of mRNA expression is useful, results are only correlative, rather than causative. As proteins are the molecules that produce events in a cell their quantities need to be studied in association with gene expression. Without collagen Type II we did not identify any correlation. This latter finding should be scrutinised with prudence as collagen II was only quantified in the soluble fraction of the cartilage extract as carryover and does not account for total protein expression. It is not surprising that there was a lack of correlation between mRNA and protein as previously the relationship has been demonstrated to vary in all directions, thus simple deduction of protein from mRNA is insufficient (Greenbaum et al., 2003; Gygi et al., 1999b). There a number of possible reasons for poor correlation. These include post-translational regulation effecting degree of translation, decay differences in mRNA and proteins, and locations or molecular associations of proteins expressed by genes. Our study measured specific proteins in cartilage, which is primarily derived from ECM as cells account for barely 1% of the total volume. Little is known about protein decay of the major ECM proteins, whilst our understanding of mRNA decay is expanding (Tew and Clegg, 2010). More work needs to be undertaken to assess protein turnover in the cartilage ECM in health and disease in order to elucidate what relationship there is between mRNA and protein.

We were interested in gene expression changes between normal and OA cartilage. In the present study despite a large variability in expression levels between donors there was a significant increase in the expression of proteolytic enzymes (MMP-1, MMP-3, and ADAMTS-5) in OA cartilage. These results were consistent with others indicating an increase in catabolic pathways in OA of the metacarpus of the horse (Smith et al., 2006). The role of ADAMTS-5 in OA is controversial (Bondeson et al., 2008). In equine carpi cartilage a recent study revealed a significant increase in ADAMTS-5 in synovial tissue from OA joints, whilst OA cartilage demonstrated a significant increase in ADAMTS-4, (Kamm et al., 2010). It may be that underlying aetiology of OA in the two joints differs. Although there was no significant

difference in MMP-13 expression here, there was an increase in the variability of MMP-13 expression in OA samples with those having a higher Mankin's score having a higher expression of MMP-13 (data not shown). This is in agreement with other work that suggests MMP-13 is detectable in later stage OA (Aigner et al., 2001).

The gene expression of matrix macromolecules was unaltered in this study. Others demonstrate variable results from gene expression analysis of ECM genes in OA; with some indicating differential expression (Brew et al., 2008; Smith et al., 2006) whilst others have not (Martin et al., 2001). We conclude that expression profiles of some genes correlate to OA. However further work using normal and OA cartilage from the same donors but distinct matched joints would be beneficial the further understanding of gene expression changes in OA.

**Conclusion**

The aim of this study was to develop a sensitive, specific and inexpensive assay to quantify ECM proteins and cartilage degradation using QconCAT together with SRM methodology. The study enabled the absolute quantification of matrix proteins for the first time as well as providing a potential tool to measure cartilage degradation at specific sites. This QconCAT will support future studies particularly in measuring the absolute protein abundance in cartilage explant studies.

**Acknowledgements**

Mandy Peffers was supported by a Wellcome Veterinary Integrated Research Fellowship.

## APPENDIX TO MANUSCRIPT 2

### Interrogation of an IL-1β stimulated explant secretome with QconCAT

### Introduction

A QconCAT was developed for quantification of matrix proteins in whole cartilage extracts. However the QconCAT also has applications for absolutely quantifying these proteins in other equine samples. As manuscript 1 had indicated differential expression of MMP-3 protein in the IL-1β stimulated explant secretome by label-free MS and western blotting this model was used to test our QconCAT, particularly in quantifying MMP-3 in a less complex proteome.

The QconCAT was used to absolutely quantify aggrecan, type II collagen, biglycan, decorin and MMP-3. COMP and fibromodulin were not quantified as this had already been previously undertaken in the appendix to manuscript 1.

### Materials and methods

### Cartilage isolation and explant culture

An IL-1β stimulated 96h equine cartilage explant model of early OA was used in the study (n=3). All donor specifications and methods have been previously described in the manuscript; Proteomic characterisation and quantification of an *in-vitro* early equine inflammatory model. Donors used were different to those used in previous studies but were also collected from the entire surfaces of the metacarpophalangeal joints of grossly normal skeletally mature horses aged between 9 and 12 years.

### Protein digestion for absolute quantification

100µg protein for each sample of explant supernatant was detergent-treated, reduced, alkylated and trypsin digested as described in the Manuscript 2. Previously trypsin digested QconCAT was spiked into the diluted samples at 10fmol load on column following analyte ranging experiments using label-free quantification on the Synapt G1 Q-TOF instrument. The peptides were resolved using a NanoAcuity LC

system and Xevo TQ instrument methods have been previously described (Brownridge and Beynon, 2011) and manuscript 2. Samples were run in duplicate technical replicates.

**LC-SRM/MS quantification and data analysis**

Transitions and data analysis for SRM studies have been previously described in the manuscript 2. Results were normalised to wet weight of explants as the explants were to be used in further studies post culture.

**Statistical analysis**

Statistically significant differences in the absolute quantification of cartilage secretomes between control and treated cultures were analysed using mixed effects linear regression to allow for donors with significant biological variation. Statistical analyses were undertaken using S-Plus and Excel software.

**Results**

Two q-peptides were used to quantify each protein except for MMP-3 when only GEILFFK was used as the other q-peptides were 'type C' (not evident in heavy or light channels using SRM). The classification of each q-peptide in terms of non-cleaved spanning (NCC) and non-cleaved control (NCS) classification used in quantification is demonstrated in Table 1. IL-1β stimulation of equine cartilage explants cultures for 4 days significantly increased aggrecan, type II collagen, decorin and MMP-3 in supernatant media (Figure 1). Although quantification values using two q-peptides for the same protein varied, the percentage increase in each when compared to control were not significantly different (Table 1).

| Protein | Peptide amino acid sequence | Peptide type | pmol/mg wet weight Mean concentration control | Mean concentration treatment | % increase |
|---|---|---|---|---|---|
| aggrecan | YEINSLVR | NCC | 4.47 | 20.90 | 368.00 |
| | YQCTEGFVQR | NCC | 8.16 | 29.87 | 265.87 |
| biglycan | GVFSGLR | NCS | 29.22 | 34.38 | 17.65 |
| | VPAGLPDLK | NCC | 2.54 | 3.18 | 25.14 |
| Type II collagen | GPEGAQGPR | NCC | 51.66 | 78.94 | 52.79 |
| | TGPAGAAGAR | NCC | 35.65 | 55.69 | 56.21 |
| decorin | VDAASLR | NCS | 18.07 | 37.96 | 110.08 |
| | VHENEITK | NCC | 17.72 | 43.46 | 145.28 |
| MMP3 | GEILFFK | NCC | 0.00 | 11.84 | Infinite |

**Table 1. IL-1β treatment of normal equine cartilage explants increased the expression of matrix proteins and MMP-3.** The proteins quantified in the study are shown along with the q-peptides used in quantification and their classification (NCC; non-cleaved control, NCS; non-cleaved spanning). Both aggrecan q-peptides were from the G3 domain. Results were normalised to wet weight of cartilage explants. % increase refers to the increase in protein in treated supernatant media compared to control.

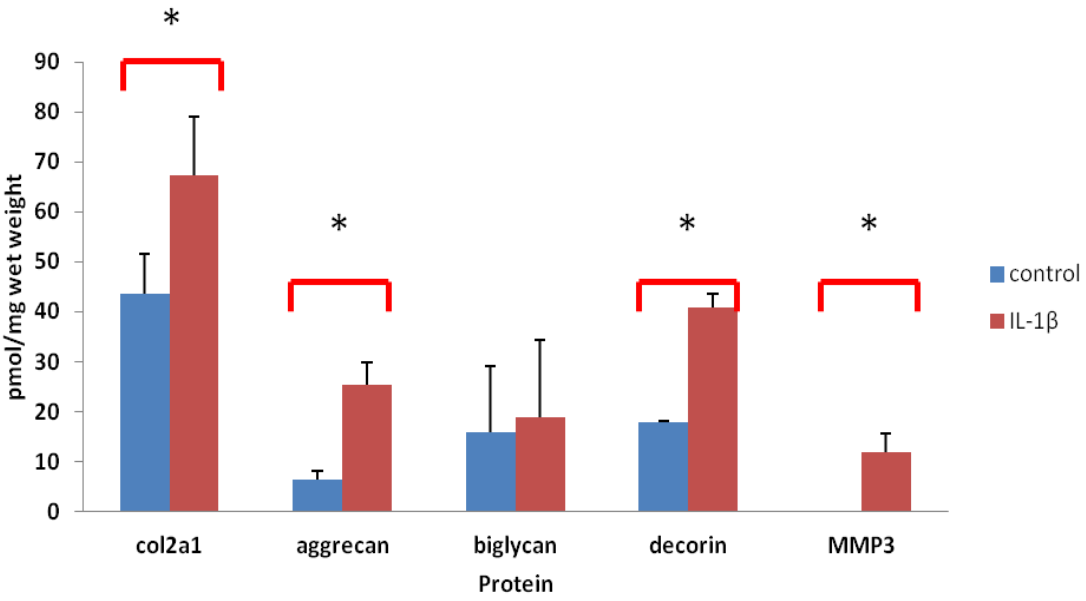



**Figure 1. Quantification of the equine cartilage secretome revealed a significant increase in expression of type II collagen (Col2A1), aggrecan, biglycan, decorin and MMP-3 using QconCAT.** Where two q-peptides were used the mean values from all q-peptide were used for final results. Data are represented as pmol/g wet weight. Histograms represent means + SEM. Data was evaluated using mixed effects linear regression and * indicates significant difference relative to control. Statistical significance is defined for this study as * $P<0.05$ (n=3).

## Discussion

Even in the early stages of disease, the development and progression of OA are thought to involve inflammation. Secreted factors such as proinflammatory cytokines are critical mediators of the imbalance between anabolism and catabolism of the joint tissue involved in OA. IL-1β suppresses type II collagen (Chadjichristos et al., 2002) and aggrecan gene expression (Stove et al., 2000) and stimulates the release of MMP-1, MMP-3 and MMP-13 (Lefebvre et al., 1990). Therefore in the current work the IL-1β stimulated explant culture provides a means of studying early OA changes *in-vitro.*

The QconCAT enabled quantification of aggrecan, type II collagen, decorin, biglycan and MMP-3 in the IL-1β stimulated explant secretome. As demonstrated and

discussed in manuscript 1 the values given by peptides from the same protein varied. Interestingly when NCC and NCS peptides were used to quantify biglycan and decorin there was no apparent change in the percentage increase in the NCS compared to the NCC, following Il-1β stimulation. However there was an reduction in the increase in the amount of the NCC peptides compared to the NCS peptides. This could be due to an IL-1β mediated increase in MMPs resulting in increased degradation of the biglycan peptide GVFSGLR by MMP-3 (Monfort et al., 2006) and decorin VDAASLR by MMP-13 (Imai et al., 1997) as these peptides span the respective cleavage sites.

In manuscript 2 MMP-3 was not quantified. In this experiment undertaken using whole cartilage extract the q-peptide GEILFFK was a 'type B' peptide; signal in the heavy channel (QconCAT) but not the light channel (analyte). Furthermore it was not identified by western blotting either. However, when explant media was probed with anti-MMP-3 antibodies in the manuscript 2 (Figure 1), in agreement with the QconCAT results, it was also demonstrated to have increased expression following IL-1β treatment.

Interestingly this western blot also demonstrated the presence of MMP-3 in the control media. Western blotting is sensitive and highly specific when antibodies are available, and therefore it is excellent for detecting a given protein. In this experiment it highlights that the antibodies used were more sensitive than the SRM approach for the q-peptide used. The SRM experiment did not detect MMP-3 in the control media as levels were apparently below the limits of detection for the assay for this particular q-peptide. Antibodies have limitations concerning reproducibility and dynamic range when it comes to quantification and when it is necessary to determine the concentration of multiple proteins at the same time. In this respect the QconCAT is a useful tool for quantification.

From the matrix proteins quantified the largest change in expression was for aggrecan. This is not surprising as it is one of the first matrix proteins targeted for degradation following cytokine stimulation and in early OA (Fosang et al., 1991;

Madsen et al., 2010). One important finding from this study was the ability to detect type II collagen breakdown early as demonstrated by an increase in Type II collagen in the media after only 4 days. Previous cytokine stimulated explant studies employing an enzyme-linked immunosorbent assay using mouse IgG monoclonal antibody to denatured type II collagen found there was no release of collagen into culture media until 14 days after explant culture. By which time most of the proteoglycans had been lost from the explants. In addition Price et.al 1999 (Price et al., 1999) measured collagen degradation following cartilage explant stimulation with retinoic acid using release of epitope CB11B and found it negligible until 14 days of culture. A further study stimulated equine cartilage explants stimulated with different combinations of IL-1, serine protease activated protein C (APC), TNFα and oncostatin M (OSM) for 96 hours. Collagen release from explants was measured with hydroxyproline assay. There was no significant release of collagen following IL1, TNFα, OSM or APC alone however when IL-1 or TNFα were combined with APC at concentrations greater than 5μg/ml there was a significant increase in collagen release into the media but this required a 60% loss of glycosaminoglycans (Garvican et al., 2010). Hydroxyproline assays are used to quantify collagens given that 14% of amino acids in collagen are hydroxyproline (Weiss and Klein, 1969). The lowest hydroxyproline standard used to make a standard curve is 2μg/ml. Given the average mass of a collagen type II (hypothesised as the predominant collagen released by cartilage explants) is 190kDa then the sensitivity of the lowest standard equates to 1.4μmol/μl. In the SRM assay used the sensitivity was 10fmol/μl. This gives an increase in sensitivity of the SRM to greater than 7 orders of magnitude. Therefore our method may be a useful sensitive tool to monitor early degradation of Type II collagen. Furthermore, whilst the enzymatic pathways of collagen degradation are complex, it is likely that the numerous enzymes involved are activated at varying times and under different circumstances. The numerous pathways leading to OA may therefore result in the formation of differing amounts of the collagen cleavage products, depending on the stage and type of arthritis (Poole et al., 2003). This may explains why significant differential release of type II collagen is evident in this study.

## Conclusion

Whilst antibody-based studies have enabled some questions to be addressed about the sequence of events in cartilage degradation the equine degradome QconCAT could enable further elucidation of mechanisms of cartilage proteolysis as it is a valid method to enables both proteases and matrix proteins to be investigated in a single experiment.

## Acknowledgements

Mandy Peffers was supported by a Wellcome Veterinary Integrated Research Fellowship.

# Manuscript 3

MJ Peffers[1*], RJ Beynon[2], PD Clegg[1]

**Corresponding Author:** M J Peffers[1]

[1]Dept.of Musculoskeletal Biology, Institute of Ageing and Chronic Disease, University of Liverpool, Leahurst, Chester High Road, Neston, Wirral, CH64 7TE; peffs@liv.ac.uk, pclegg@liv.ac.uk

[2] Proteomics Group, Institute of Integrative Biology, University of Liverpool, Biosciences Building, Crown Street, Liverpool, L69 7ZB; rbeynon@liv.ac.uk

**Title: Protein Signatures of the Human Osteoarthritic Secretome**

**Running Title**: **Protein Quantification of the Human Osteoarthritic Secretome**

**Key Words:** cartilage, human, osteoarthritis, QconCAT, secretome

**Protein Signatures of the Human Osteoarthritic Secretome**


## Abstract

Osteoarthritis (OA) is characterized by a loss of extracellular matrix (ECM) which is driven by catabolic cytokines. Proteomic analysis of the OA cartilage secretome allows the global study of secreted proteins. These are an important class of active molecules with roles in numerous pathological mechanisms. Although cartilage studies have identified profiles of secreted proteins, limited quantitative proteomics techniques have been implemented which would enable further biological questions to be addressed. To overcome this limitation we used the secretome from human OA cartilage explants stimulated with IL-1 β and compared proteins released into the media using a label-free LC-MS/MS-based strategy. We then employed QconCAT technology to quantify specific proteins using selected reaction monitoring (SRM). A total of 242 proteins were identified 9 of which identified were differentially expressed by IL-1 β stimulation. Selected protein candidates were quantified in absolute amounts using QconCAT. These findings confirmed a significant reduction in TIMP-1 in the secretome following IL-1β stimulation. Label-free and QconCAT analysis produced equivocal results indicting no effect of cytokine stimulation on aggrecan, cartilage oligomeric matrix protein (COMP), fibromodulin, matrix metalloproteinases 1 and 3 and plasminogen release. Taken together our proteomics approaches to OA research allows comparative protein profiling and absolute quantification affording tools for the detection and quantification of proteins in molecular pathways pertinent to understanding the pathogenesis of OA.


## Introduction

Articular cartilage, an avascular connective tissue provides a nearly frictionless bearing surface for transmitting and distributing mechanical loads between the bones of the skeleton (Mow et al., 1992). The chondrocyte, the sole cell type (Archer and Francis-West, 2003), is embedded within an ECM whose unique load

bearing properties are dependent upon its structural composition and organisation, particularly the interactions between collagens and proteoglycans (Poole et al., 2001). Progressive articular cartilage loss leads to joint pain and dysfunction that is clinically identified as OA. In OA the normal equilibrium between matrix deposition and degradation is disrupted resulting in progressive loss of important ECM components, especially aggrecan and collagens.

Mass spectrometry (MS) has emerged as an important analytical tool for protein analysis with MS-based proteomics enabling proteins within a sample to be identified and quantified. Cartilage proteomic studies have permitted the investigation of cartilage proteins in both the intact cartilage tissue (Iliopoulos et al., 2010; Wilson et al., 2009) and the cartilage secretome (Catterall et al., 2006; Stevens et al., 2008), with a number of studies reporting IL-1 driven protein secretion from cartilage ECM (Ruiz-Romero and Blanco, 2010). The cartilage secretome defined as the proteins identified in the media surrounding the chondrocyte or explants, includes proteins secreted or shed from the cell surface, plus intracellular proteins released into the supernatant due to cell lysis, apoptosis or necrosis (Hathout, 2007). In cartilage explant studies proteins released into media, by chondrocytes and ECM may be similar to proteins released *in-vivo* in cartilage degradation (Wilson et al., 2009) and data from these studies has enabled improved understanding of OA pathogenesis (Iliopoulos et al., 2010). The secretome may be altered by addition of pro-inflammatory cytokines to explants, an accepted method of studying matrix metabolism in experimental investigations of OA *in-vitro* (Arner et al., 1998; Goldring and Goldring, 2004). Indeed cytokine stimulation of normal and OA cartilage explants has been used in numerous studies to initiate a catabolic response (Ismaiel et al., 1992; Song et al., 2007) and assess different facets of the degradative process.

The two types of protein quantification are absolute quantification which determines real amounts of a protein in terms of concentrations for example as copies per cell, and normally uses external or internal standards and relative quantification. The latter determines relative differences in protein abundance.

Within cartilage research there is a need for absolutely quantitative MS in order to define proteins in tangible amounts. This will aid the understanding of and define how protein content of chondrocytes ECM alters in both ageing and disease. Moreover such data will provide necessary information for mathematical modelling of biological systems. Although relative quantification of the cartilage secretome has been undertaken in numerous studies (De Ceuninck et al., 2005; Haglund et al., 2008; Hermansson et al., 2007), these experiments focus on 'discovery' proteomics and the detection of differentially expressed proteins; few studies have attempted to quantify the cartilage secretome in exact amounts. Whist this work has enabled biomarker discovery to progress (Haglund et al., 2008) a more detailed knowledge of the quantities, interactions and dynamics of matrix components and the protease enzymes involved in degradation will increase our understanding of the as yet undefined mechanisms involved in ECM destruction typical of OA. For example a knowledge of the exact nature of protease/ tissue inhibitors of metalloproteinase (TIMP) will further our comprehension of OA pathogenesis which could aid in the discovery of treatments. Metabolic isotope labelling in culture using stable isotope labelling of amino acids in cell culture (SILAC) has been employed in comparative cartilage studies (Polacek et al., 2010a). Whilst SILAC is suitable for the quantification of the same protein under different conditions, it is not suitable for quantification of different proteins under any conditions. Furthermore data is dimensionless leading to difficulties in interpretation.

The two current approaches to absolute quantification are label-free and label-mediated quantification. Label-free methods are based on the direct measurement of the MS acquired signal. Here it is taken that when constituent peptides produced following protein digestion are converted into ions the most abundant proteins will produce the most ions and thus the greatest signal intensities (Zhu et al., 2010). This method provides acceptable quantification for the high abundance components of a sample, but suffers at the low abundance range due to technical variance . Stable isotope labelled quantification includes the use of chemically synthesized peptide standards known as AQUA peptides (Gerber et al., 2003) and QconCAT (Beynon et al., 2005). Both rely on the MS quantification of a known

amount of isotope–labelled peptide standard relative to the otherwise identical non-labelled within an analyte peptide. This approach benefits from a high level of sensitivity with quantification possible to the atomole level.

In addition entire proteins can be used as standards; protein standard absolute quantification (PSAQ). Here a full-length protein with homologous biochemical properties of the target protein are spiked in at the start of the analytical process (Dupuis et al., 2008).

In this study we have used label-free data to give a preliminary assessment of the OA secretome protein profile and protein abundance. Then we used absolute quantification by MS, based on the well established principle of stable isotope dilution, using QconCAT as a multiplexed standard (Beynon et al., 2005) to validate and give absolute baseline values for a number of key matrix proteins. QconCAT are artificial proteins that permit highly accurate parallel absolute quantification of large sets of analyte proteins (Beynon et al., 2005). These constructs of a set of mass-tagged internal standard peptides (q-peptides) contain sequences unique to the proteins of interest. The QconCAT approach uses an artificial gene, cloned into a vector and expressed in bacterial culture containing stable heavy isotope labelled arginine and lysine to produce stable isotope labelled standards. The gene encodes a single ORF consisting of a series of linked peptides, usually tryptic, which each act as individual internal standards for a given protein (Meistermann et al., 2006; Pratt et al., 2006; Rivers et al., 2007). Then either a known amount of QconCAT protein is co-dig ested with analyte proteins or a tryptic digest of the QconCAT is spiked into the analyte tryptic digest at a known amount. MS analysis subsequently allows the quantification of each peptide present in the sample. Relative abundance as well as absolute amount (from known amount of Q-peptide added) of analyte peptide can be measured. A number of papers have used QconCAT technology to quantify multiple proteins including muscle development proteins (Rivers et al., 2007), glycolytic proteins in yeast (Carroll et al., 2011), surface proteins in *Schistosoma mansoni* blood fluke (Castro-Borges et al., 2011), host response to bovine mastitis pathogens (Bislev et al., 2012) and cohesion interactions in human cell lines (Ding et

al., 2011). The advantages of QconCAT include the ability to quantify a large number of proteins simultaneously from a sample, a relatively low cost per peptide, and the ability to synthesize problematic peptides. QconCAT methodology also has its shortcomings. These include the poor solubility of some QconCATs, variable expression, incomplete digestion and post translational modifications during digestion (Aebersold and Mann, 2003).

We hypothesize that there are measurable changes in protein abundance in the human OA secretome following cytokine stimulation which can be absolutely quantified using QconCAT. The aim of this study was to develop and test a targeted quantification method for the multiplexed analysis of proteins involved in the pathogenesis of OA. Here we report the development of a QconCAT in order to quantify the cartilage secretome. A human secretome QconCAT was designed, expressed, validated and then used to quantify cartilage matrix components and proteases in an IL-1β stimulated OA cartilage explant model.

## Materials and Methods

### Peptide selection, preparation and purification of QconCAT

We selected 20 proteins of interest in cartilage degradation. Two proteotypic tryptic peptides per protein were selected from the Global Proteome Machine database (www.thegpm.org) and the Human PeptideAtlas (www.peptideatlas.org) based on criteria described by Pratt et al. (Pratt et al., 2006) including their suitability score, physicochemical properties deemed to promote MS detectability, and uniqueness to a given protein. In addition amino acids that were prone to oxidation and miscleavage were avoided and all peptides were either lysine or arginine terminated (Table 1). Peptides were arranged in sequence context were possible. In the native protein the amino acid sequence prior to the q-peptide were noted and the peptides in the QconCAT ordered where possible to mimic this in order to optimise digestibility. The transformation, expression and purification of QconCAT has been previously described in detail (Rivers et al., 2007). Briefly, following synthesis of the gene by PolyQuant GmBH (Entelechon, Germany) the

QconCAT was ligated into the expression vector pET-21a and expressed in Escherichia coli (E.coli) cultured in minimal media ( 1xM9 salts , 1mM MgSO4, 0.1mM CaCl2, 0.00005% (w/v) thiamine, 0.2% (w/v) glucose, unlabeled amino acids at 0.1mg/ml or 0.2mg/ml histidine, tyrosine, phenylalanine, proline and tryptophan all Sigma-Aldrich) supplemented with $^{13}C_6$ analogues of arginine and lysine containing stable isotope labelled amino acids. Once cells achieved mid log phase ($OD_{660nm}$ 0.6-0.8) expression was induced by addition of 1mM IPTG Isopropyl β-D-1-thiogalactopyranoside (Sigma-Aldrich). After 5h of induction cells were harvested by centrifugation at 1400**g** at 4$^{o}$C for 15min. Cell lysis was undertaken using BugBuster Protein Extraction Reagent (Merck Chemicals, Nottingham, UK).

Inclusion bodies were first re-dissolved in 20mM phosphate buffer, 6M guanidine chloride, 0.5M NaCl, 20mM imidazole, pH 7.4. They were then solubilised using sonication, followed by purification using immobilised metal affinity columns; Ni-MAC (Novagen, Darmstadt, Germany). The purified QconCAT protein was desalted three times by dialysing against 100 volumes 10mM ammonium bicarbonate, pH 8.5, 1mM dithiothreitol (DTT) for 2h changing the buffer each time.

**Characterisation of QconCAT**

The homogeneity of the QconCAT was determined by the in-gel digestion of a protein band corresponding to the expected molecular mass for the QconCAT of 58KDa. Briefly a 5µg aliquot of purified QconCAT protein was run on a 12% SDS-PAGE gel after a 50 min run at 200V, fixed with 40% methanol and 10% acetic acid then stained with Coomassie blue.

In-gel digestion was undertaken as previously described (McIlwraith et al., 2010). 1µl of the digest was mixed with 1µl of α-cyano- 4-hydroxycinamic acid (CHCA; Sigma, Poole, UK) in 50% (v:v) acetonitrile (ACN)/0.1% (v:v) trifluoroacetic acid

| Peptide Order | Protein | Protein Accession | Native N-terminal flanking peptides | Q-peptide amino acid sequence | Native C-terminal flanking peptides |
|---|---|---|---|---|---|
| 1 | MMP16 | ENSP00000286611 | VWK | GIPESPQGAFVHK | ENG |
| 2 | MMP16 | ENSP00000286611 | RVK | EGHSPPDDVDIVIK | LDN |
| 3 | CathepsinD | ENSP00000236671 | QQK | LVDQNIFSFYLSR | DPD |
| 4 | COMP | ENSP00000222271 | CGR | DTDLDGFPDEK | LRC |
| 5 | Col11a2 | ENSP00000372565 | KGK | LGVPGLPGYPGR | QGP |
| 6 | Fibromodulin | ENSP00000347041 | LQK | IPPVNTNLEN LYLQGNR | INE |
| 7 | MMP3 | ENSP00000299855 | TYR | IVNYTPDLPK | DAV |
| 8 | ADAMTS1 | ENSP00000284984 | SDR | DAEHYDTAILFTR | QDL |
| 9 | ADAMTS4 | ENSP00000356975 | LSR | FVETLVVADDK | MAA |
| 10 | CathepsinD | ENSP00000236671 | VSK | YSQAVPAVTEGPIPEVLK | NYM |
| 11 | Link protein | ENSP00000274341 | FLK | GGSDSDASLVITDLTLEDYGR | YKC |
| 12 | MMP3 | ENSP00000299855 | VQK | YLENYYDLK | KDV |
| 13 | ADAMTS5 | ENSP00000284987 | RSK | GLVQNIDQLYSGGGK | VGY |
| 14 | TIMP3 | ENSP00000266085 | VER | WDQLTLSQR | GLN |
| 15 | TIMP4 | ENSP00000287814 | WYR | GHLPLR | KEF |
| 16 | CathepsinK | ENSP00000271651 | HSR | SNDTLYIPEWEGR | APD |
| 17 | Link Protein | ENSP00000274341 | NGR | FYYLIHPTK | LTY |
| 18 | COMP | ENSP00000222271 | AVK | SSTGPGEQLR | NAL |
| 19 | Plasminogen | ENSP00000308938 | PHR | HSIFTPETNPR | AGL |
| 20 | MMP13 | ENSP00000260302 | FWR | LHPQQVDAELFLTK | SFW |
| 21 | MMP13 | ENSP00000260302 | YLR | SYYHPTNLAGILK | ENA |
| 22 | Plasminogen | ENSP00000308938 | LLK | EAQLPVIENK | VCN |
| 23 | Col9a1 | ENSP00000349790 | IQK | VVGSATLQVAYK | LGN |
| 24 | TIMP4 | ENSP00000287814 | GVK | LEANSQK | QYL |
| 25 | MMP1 | ENSP00000322788 | FMK | DGFFYFFHGTR | QYK |
| 26 | ADAMTS1 | ENSP00000284984 | EQK | GPEVTSNAALTLR | NFC |
| 27 | ADAMTS4 | ENSP00000356975 | SIR | NPVSLVVTR | LVI |
| 28 | Aggrecan | ENSP00000268134 | VDR | LEGEVFFATR | LEQ |
| 29 | Fibromodulin | ENSP00000347041 | LER | LYLDHNNLTR | MPG |
| 30 | TIMP1 | ENSP00000218388 | MYK | GFQALGDAADIR | FVT |
| 31 | TIMP1 | ENSP00000218388 | RAK | FVGTPEVNQTTLYQR | YEI |
| 32 | TIMP3 | ENSP00000266085 | VNK | YQYLLTGR | VYD |
| 33 | Col11a2 | ENSP00000372565 | AYR | VARPAQLSAPTR | QLF |
| 34 | ADAMTS5 | ENSP00000284987 | AFR | LPLAAVGPAATPAQDK | AGQ |
| 35 | Aggrecan | ENSP00000268134 | KEK | EVVLLVATEGR | VRV |
| 36 | CathepsinK | ENSP00000271651 | VAR | VGPVSVAIDASLTSFQFYSK | GVY |
| 37 | Col2a1 | ENSP00000338213 | GAR | GAQGPPGATGFPGAAGR | VGP |
| 38 | Col2a1 | ENSP00000338213 | GPK | GPPGPQGAR | GDR |
| 39 | Col9a1 | ENSP00000349790 | GPR | GVQGEQGATGLPGVQGPPGR | APT |
| 40 | MMP1 | ENSP00000322788 | YGR | SQNPVQPIGPQTPK | ACD |

**Table 1 Human cartilage QconCAT signature peptides in QconCAT context order.** The three amino acids found adjacent to the N and C termini of the q-peptide within the native protein are indicated (Matrix metalloproteinase (MMP), collagen (Col), a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS), tissue inhibitor of metalloproteinase (TIMP)).

(TFA) and 1µl spotted on a MALDI plate. Positive-ion MALDI mass spectra (MS) were obtained using an Ultraflex (Bruker, Brenin, Germany) in reflector mode over *m/z* range 900-4500. Monoisotopic masses were collected from centroids of raw unsmoothed data.

**Cartilage isolation and explant culture**

Human articular cartilage (HAC) was obtained following total knee arthroplasty due to OA with informed consent and ethical approval. Full thickness cartilage that appeared macroscopically intact and normal was harvested from the entire surfaces of three male donors aged between 69 and 84 years.

Cartilage was diced into explants of approximately 2mm, mixed and placed in complete medium [Dulbecco's modified Eagle's medium (DMEM), supplemented with foetal calf serum (10% v/v), 100U/ml penicillin, 100U/ml streptomycin (Invitrogen, Paisley, UK) 500ng/ml amphotericin B (BioWhittaker, Lonza, USA). Explants were washed twice with serum-free DMEM (to deplete serum and synovial proteins) and allowed to equilibrate in complete medium for 24h at 37$^{o}$C in 5% $CO_2$ in 12 well plates (2ml/well). Media was then replaced with serum-free DMEM prior to incubation, supplemented with or without human recombinant IL-1β (10ng/ml; R&D Systems) in dimethyl sulfoxide (DMSO) diluent. After 48h, media was removed, centrifuged to remove debris and protease inhibitors (Complete protease Inhibitors, EDTA-free, Roche, Lewes, UK) added. Samples were stored at -80$^{o}$C prior to downstream analysis. Protein concentrations of supernatants were estimated by Bradford assay (Thermo Scientific, Rockford, USA). Cartilage explants were lyophilized to obtain a dry weight for normalisation.

**1-D SDS-PAGE separation and in-gel trypsin digestion**

Cartilage extract secretomes of control and treatment condition for all donors were analyzed by one dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis (1-D SDS-PAGE) to assess quantitative/qualitative differences in protein profiles. 20μg protein was loaded for each lane. In-gel tryptic digests of bands of interest from the 1-D SDS-PAGE was undertaken as previously described (McClean et al., 2007). ImageJ software (http://rsbweb.nih.gov/ij/) was used to quantify bands using densitometry.

**Protein identification of in-gel digests by linear ion trap quadruple (LTQ) Velos mass spectrometry**

Digested samples were analysed by LC-MS/MS using an UltiMate® 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to a LTQ Velos Pro (Thermo Fisher Scientific, Waltham, MA) mass spectrometer. Peptides were concentrated on a pre-column (20mm x 180μm i.d, Waters, Manchester, UK). The peptides were then separated using a gradient from 99% A (0.1% formic acid (FA) in water) and 1% B (0.1% FA in ACN) to 25% B, in 45min at 200nL $min^{-1}$, using a 75mm x 250μm i.d. 1.7μM BEH C18, analytical column (Waters, Manchester, UK). Peptides were selected for fragmentation automatically by data dependant analysis.

Raw spectra were converted to Mascot generated files (mgf) using Proteome Discoverer software (Thermo, Hemel Hempstead, UK). The resulting mgf files were searched against the Human IPI database sequence databases using an in-house Mascot server (Matrix Science, London, UK). Search parameters used were; peptide mass tolerances 10ppm, fragment mass tolerance of 0.6Da, 1+, 2+ and 3+ ions, missed cleavages; 1, and instrument type ESI-TRAP. Modifications included were; fixed; carbamidomethyl cysteine and variable; oxidation of methionine. Data produced were searched using Mascot (Matrix Science UK), against the Human IPI database with taxonomy of Homo sapiens selected. Data were validated using Scaffold (Proteome Software, Portland, USA).

**In-solution tryptic digestion and mass spectrometry using linear ion-trap Orbitrap mass spectrometer (LTQ-Orbitrap Velos)**

Cartilage supernatant fractions or QconCAT were detergent treated with 1% (w/v) Rapigest (Waters, Manchester, UK) for 10min at 80$^{o}$C in 25mM ammonium bicarbonate. In-solution tryptic digestion of protein samples was carried out following sequential reduction and alkylation in 3mM DTT (60$^{o}$C for 10min) and then 9mM iodoacetamide (30min in the dark at room temperature) with trypsin at a ratio of 1:50 protein: trypsin ratio overnight at 37$^{o}$C. Detergent inactivation was then assumed by incubating for 45min at 37$^{o}$C with trifluoroacetic acid (VWR International) to a final concentration of 0.5% (v/v). Following centrifugation for ten minutes at 15000**g** the soluble phase was retrieved and used for LC-MS/MS.

LC-MS/MS analysis was performed using nanoAcquity$^{TM}$ ultraperformance LC (Waters, Manchester, UK) on line to an LTQ-Orbitrap Velos (Thermo-Fisher Scientific, Hemel Hempstead) via a ESI ion source containing a 10um coated Pico-tip emitter (Presearch LTD, Basingstoke, UK). Aliquots of tryptic peptides equivalent to 250ng of cartilage secretome protein or 100fmol of QconCAT (for QconCAT verification and validation) were loaded onto a 180µm x 20mm $C_{18}$ trap column (Waters, Manchester, UK) at 5µl/min in 99% solvent A (water plus 0.1 % FA) and 1% solvent B (acetonitrile plus 1% FA for 5 min and subsequently back-flushed onto a $C_{18}$ pre-equilibrated analytical column (75µm x 15mm Waters, Manchester, UK) using a flow rate of 0.3µl/min. Xcalibur 2.0 software (Thermo -Electron ,Hemel Hempstead, UK) was used to operate the LTQ-Orbitrap Velos in data-dependant acquisition mode. The survey scan was acquired in the Orbitrap with a resolving power set to 30,000 (at 400 *m/z*). MS/MS spectra were concurrently acquired on the 20 most intense ions from the high resolution survey scan in the LTQ. Charge state filtering >1 was used, where unassigned precursor ions were not selected for fragmentation. Fragmentation parameters in the LTQ were: normalized collision energy; 30, activation; 0.250, activation time; 10ms and minimum signal threshold 500 counts with isolation width 2*m/z*.

**Peptide Identification**

Raw spectra were converted to Mascot generated files (mgf) using Proteome Discoverer software (Thermo, Hemel Hempstead, UK). The resulting mgf files were searched against either the Human IPI database, taxonomy; mammalian or QconCAT sequence databases using an in-house Mascot (Yan et al., 2006) server (Matrix Science, London, UK). Search parameters used were; peptide mass tolerances 10ppm, fragment mass tolerance of 0.6Da, 1+, 2+ and 3+ ions, missed cleavages; 1, and instrument type ESI-TRAP. Modifications included were; fixed; carbamidomethyl cysteine and variable; oxidation of methionine.

**Selected reaction monitoring optimisation**

The selected reaction monitoring (SRM) assay conditions were optimised using a XEVO TQ (Waters, Manchester). Trypsin digestion of QconCAT was carried out as described previously (Rivers et al., 2007). The peptides were diluted with 97:3:0.1 ACN: water: FA and 100 fmol of peptides were loaded on column. Optimisation was performed on a XEVO TQ operated with MassLynx 2.4 (Waters, Manchester) coupled to a NanoAcquity$^{TM}$ UPLC (Waters, Manchester, UK). Peptides were loaded using partial loop injection, for three minutes at a flow rate of 5µL/min with 0.1% (v/v) formic acid onto a trapping column (Waters, C18, 180µm x 20mm). Samples were separated by a 30min gradient of 97% A (0.1% (v/v) formic acid), 3% B (99.9% acetonitrile 0.1% (v/v) formic acid) to 60% A 40% B at a flow rate of 300nL/min on a $C_{18}$ analytical column (Waters, NanoAcuity UPLC™ BEH $C_{18}$ 75µm x 150mm 1.7µm column). All transitions were acquired with the following parameters; 3kV ion spray voltage, an 80$^{o}$C interface heater temperature, Q1 and Q3 operating at unit resolution. Cone voltages and collision energies were optimised for each peptide. Dwell times for transitions were determined automatically based on the number of co-eluting peptides but a minimum dwell time of 50msec was maintained.

**Protein digestion and quantification**

10µg protein for each sample was detergent treated, reduced, alkylated and trypsin digested as described above. For SRM experiments, previously trypsin digested QconCAT was spiked into the samples.

SRM experiments were conducted with 500ng of tryptic analyte peptides spiked in with either 10fmol, 1fmol or 0.1fmol heavy QconCAT, loaded onto column. MS analysis was commenced using methods, parameters and gradients described here previously. The ranging ensured that analyte: signal to noise were between 1:10 and 10:1 ratio of a QconCAT loading. MassLynx 2.4 (Waters, Manchester) software was used to produce extracted ion chromatograms of the peptide transitions in order to compare the ratio of analyte to standard. Ratios were converted to fmol and then normalised to dry weight of cartilage explant.

**Label-free peptide quantification**

The Thermo raw files of the acquired spectra were analysed by the Progenesis™ LC-MS software (version 4, Nonlinear Dynamics) for label-free quantification. Progenesis™ LC-MS takes profile data of the MS scans and transforms them to peak lists. One sample was selected as a reference after checking the 2-D mapping (*m/z* versus retention time), and the retention times of the other samples within the experiment were aligned. This was undertaken by studying the chromatogram and aligning on the major peaks. Features without the 1+, 2+, 3+ and 4+ charge and isotope peaks of ≤2 were masked and excluded from further analysis. Samples were then divided into the appropriate groups using between subject design (between donor variation, between control and IL-1β treatment variation and between technical replicate variations). Raw abundances of all features were normalised which corrects for factors due to experimental variation (such a sample load).

Following feature picking we picked the top three spectra for each feature; these were exported from Progenesis™- LC-MS and utilized for peptide identification with a locally implemented Mascot server; Mascot (version 2.3.01) in the Unihuman database. Search parameters used were: 10ppm peptide mass tolerance and 0.6Da

fragment mass tolerance, one missed cleavage allowed, fixed modification; carbamidomethylation, variable modifications; methionine oxidation. Mascot determined peptides with ion scores of 33 and above and only proteins with at least one unique peptide ranked as top candidate were considered and re-imported into Progenesis™ software. Following the import of the Mascot results for quantification, statistical analysis was performed on all detected features using transformed normalized abundances for one-way analysis of variance (ANOVA). The total cumulative abundance was calculated by summing the abundances of all peptides allocated to the respective protein. All peptides (with Mascot score >33 and $p<0.05$) of an identified protein were included and the protein *p* value (one-way ANOVA) were then performed on the sum of the normalized abundances for all runs. ANOVA values of $p<0.05$ and additionally regulation of >2-fold or < 0.5-fold were regarded as significant.

**Gene Ontology**

Using DAVID gene ontology (GO) analysis all genes identified were loaded into the functional annotation chart.

**Statistical analysis**

Statistical analysis for absolutely quantified peptides was undertaken using mixed effects linear regression to allow for donors with significant biological variation with SPLUS 6.1 software.

**Results**

**Comparative analysis by mass spectrometry**

All cartilage was graded as severe OA using a modified Mankin scoring (Bulstra et al., 1989) (original score proposed by Mankin et al 1970) (Mankin and Lippiello, 1970) with scores of between 9 and 13 out of 14. 1D-SDS-PAGE of the cartilage OA secretomes (Figure 1a) demonstrated no significant difference in the profiles of the control and IL-1β treatment using densitometry (data not shown).

The repertoire of proteins secreted by explants was then analysed by MS (Figure 1B). Proteins included in the results had a Mascot score >40 with 2 or more identifying peptides, and a confidence interval of 95%. The cartilage explant secretomes contained a number of cartilage matrix proteins as expected, plus proteins associated with catabolic aspects of cartilage matrix turnover such as MMP-3.

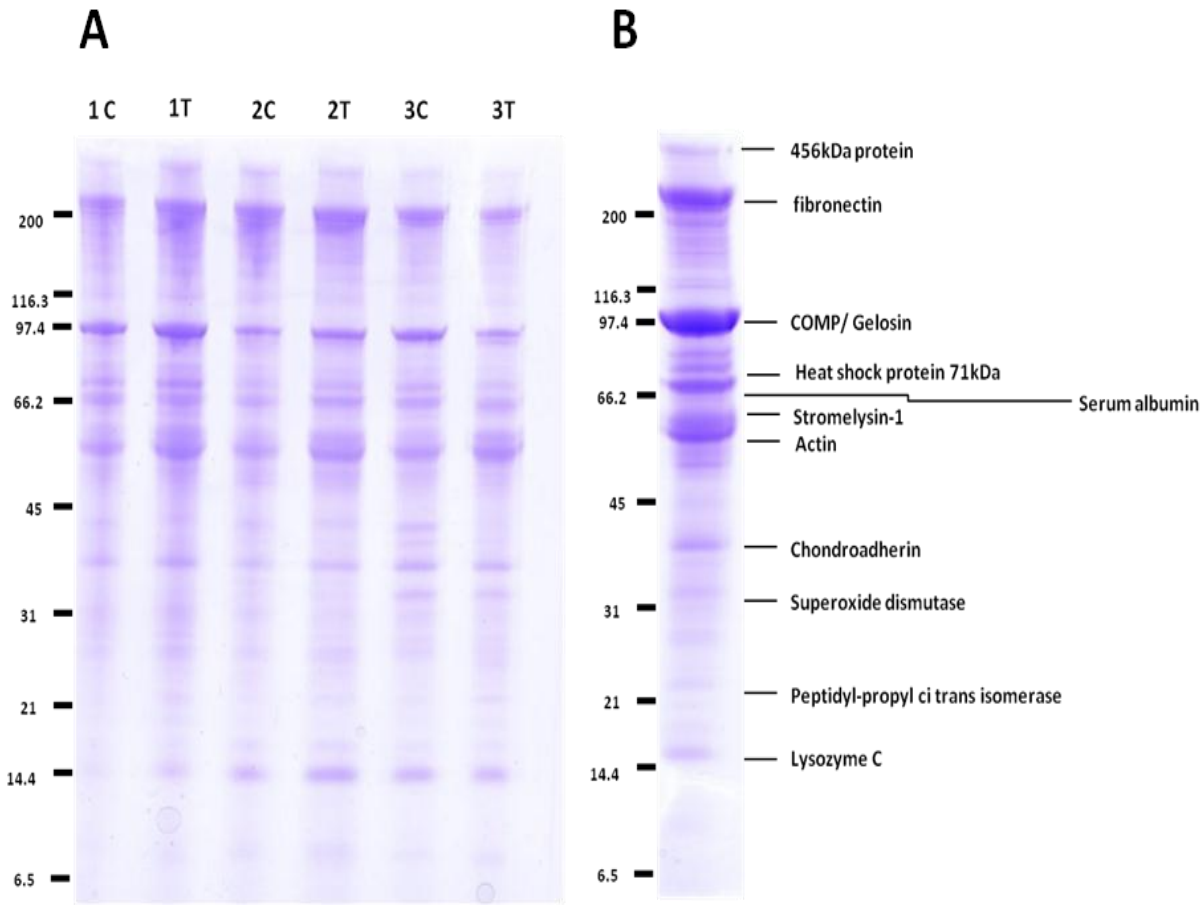



**Figure 1. 1D-SDS-PAGE of the cartilage OA secretomes demonstrated little difference in the profiles following IL-1β treatment.** A. OA human articular cartilage explants (n=3) were cultured in media supplemented with 10ng/ml IL-1β (T) or un-supplemented media (control- C). Culture media were collected at 2 days for further analysis by SDS-PAGE and staining with Coomassie Brilliant Blue. Equal protein loading of 20µg of protein per well allowed a qualitative comparison of the secretomes. B, the most abundant proteins in the media marked at the positions of the bands were excised from the gel, trypsin digested, and the protein content of each single band was analysed using peptides identified using LC-MS/MS. Proteins indicated on the gel correlate to the size and are the primary protein identified in that gel analysis.

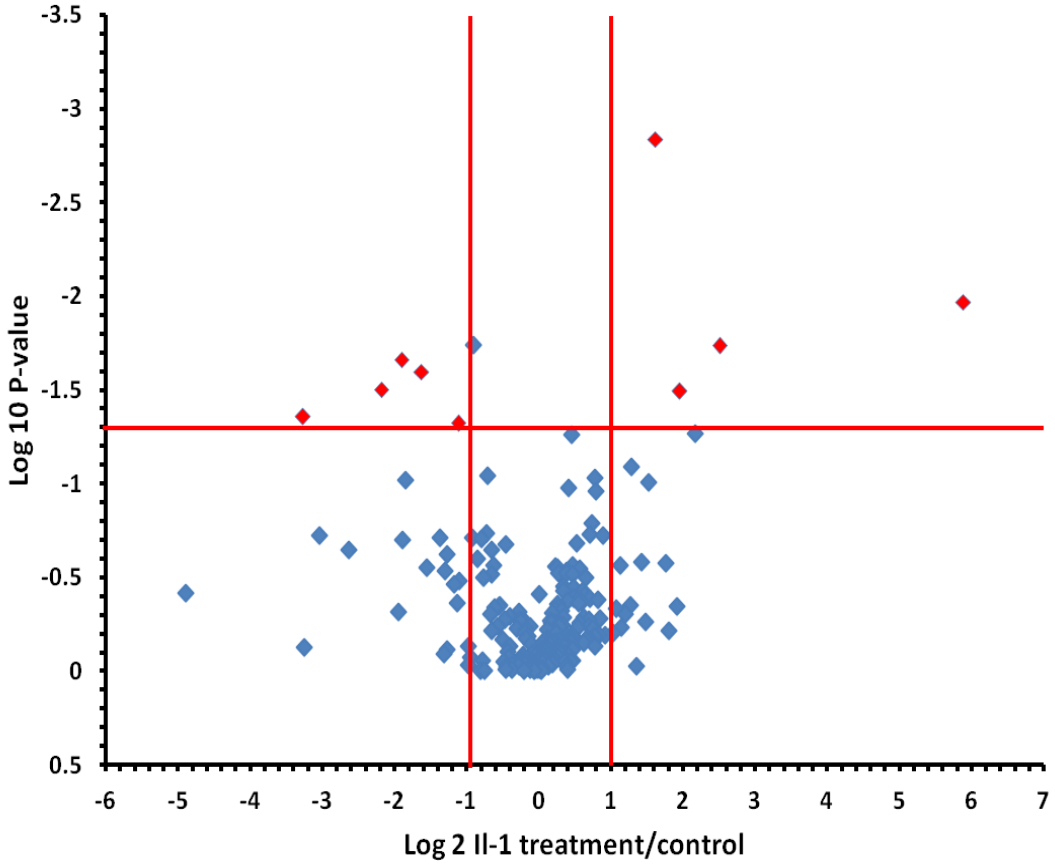



**Figure 2 Comaprison of differential protein abundance presented as a volcano plot.** (log2 fold change, x axis; log10 p value ANOVA, y axis) from control versus treatment quantified with two or more peptides. Horizontal line indicates $p<0.05$ and vertical lines indicate 2 fold and 0.5 fold abundance changes. p-values of 0.000 were set to 0.0005 in order to limit scaling. Red diamonds indicate significant differentially expressed proteins.

Explant experiments were undertaken in duplicate and in-solution tryptic digestion of media was then performed in duplicate. This gave a total of four secretome replicates per donor per condition. We processed these samples for mass spectrometry as described and performed quantitative analysis with Progenesis™ LC-MS software. All identified features in the 2-D maps were aligned between samples, subsequently normalised and assigned to control or IL-1β treatment groups. ANOVA was then performed on normalised peptide intensities. Mascot was then used to identify all features with MS/MS data against the UNIHUMAN

database with search results implemented into the experiment file. Peptide identifications were merged into non redundant protein identifications.

**Label-free protein profiling of the OA secretome**

A total of 278 proteins were identified including aggrecan, fibromodulin, cartilage oligomeric matrix protein, fibromodulin, matrix metalloproteinase 1 and 3, link protein and plasminogen. Gene ontology determined that 53% of genes were identified as 'secreted' and 18% were identified as belonging to ECM. Of these 242 were identified with ≥2 unique peptides (Figure 2).

Nine proteins were differentially expressed between control and IL-1β stimulation ($p<0.05$ and more than 2-fold regulated) (Table 2). Four of these were up regulated following IL-1β stimulation (Figure 3) and five were down regulated (Figure 4). The functions of the proteins along with their cellular location and potential role in OA are indicated in Table 3.

| Highest Mean Condition | Accession | Description | Max fold change | Anova (p) |
|---|---|---|---|---|
| Treatment | P09341 | Growth-regulated alpha protein | 58.99 | 0.01 |
| | P08254 | Stromelysin-1 | 5.70 | 0.02 |
| | Q6IPR1 | LYR motif-containing protein 5 | 3.89 | 0.03 |
| | P01876 | Ig alpha-1 chain C region | 3.06 | 0.00 |
| Control | P36222 | Chitinase-3-like protein 1 | 9.64 | 0.04 |
| | P08571 | Monocyte differentiation antigen CD14 | 4.53 | 0.10 |
| | Q14UF6 | Decay-accelerating factor splicing variant 1 | 4.11 | 0.02 |
| | Q5H9A7 | TIMP metallopeptidase inhibitor 1 | 3.09 | 0.03 |
| | P01034 | Cystatin-C | 2.40 | 0.05 |

**Table 2. A number of differentially expressed proteins were identified by Progenesis™ LC-MS software.** Proteins shown were identified with ≥2 unique peptides and with a >2 fold change in normalised abundance.

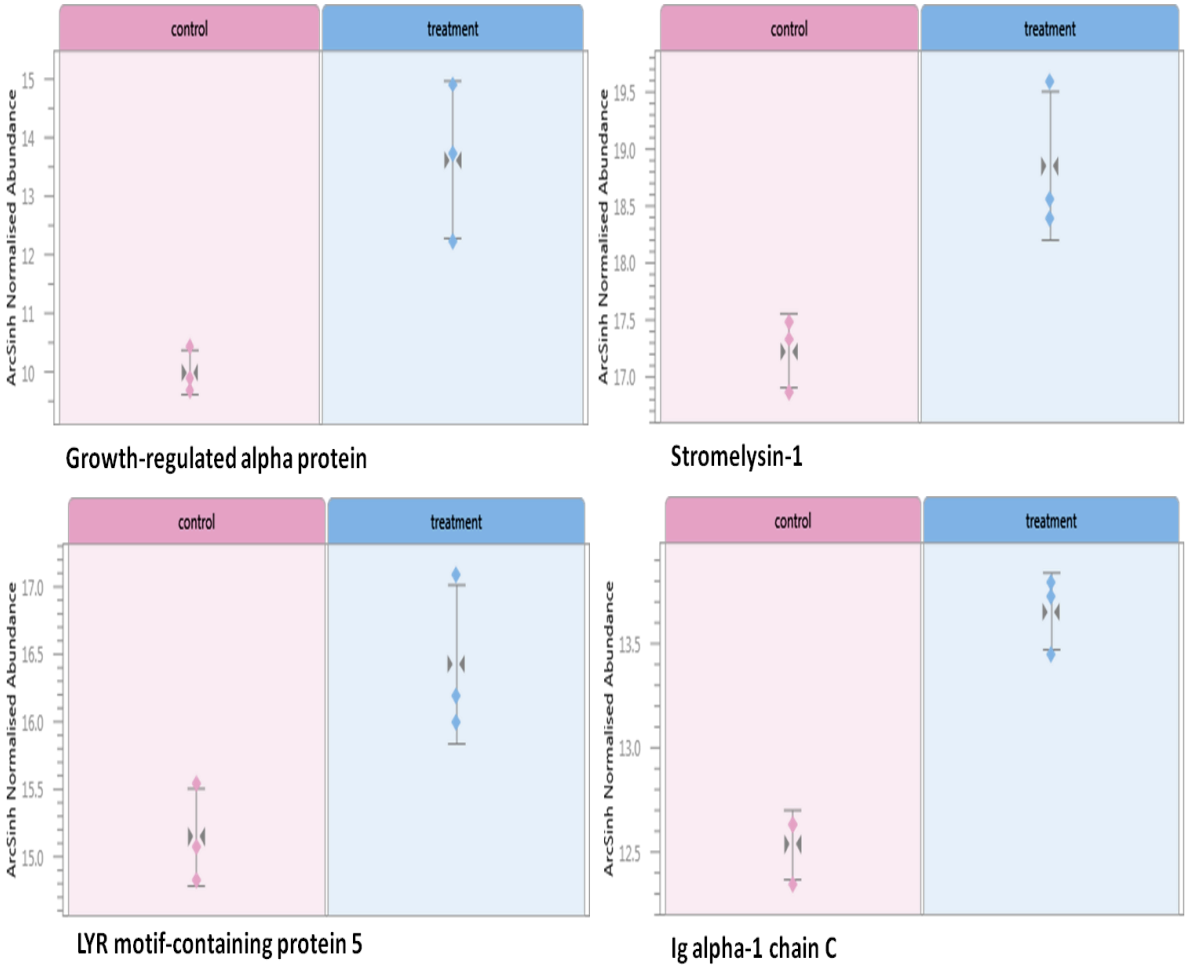



**Figure 3 Expression profile view of proteins significantly higher in IL-1β treated samples produced by Progenesis™ LC-MS.** All proteins were identified by 2 or more peptides with greater than 2-fold abundance change and $p<0.01$. Plots display the mean arcsinh transformed normalised volume for each group. Error bars demonstrate 3 standard errors within groups.

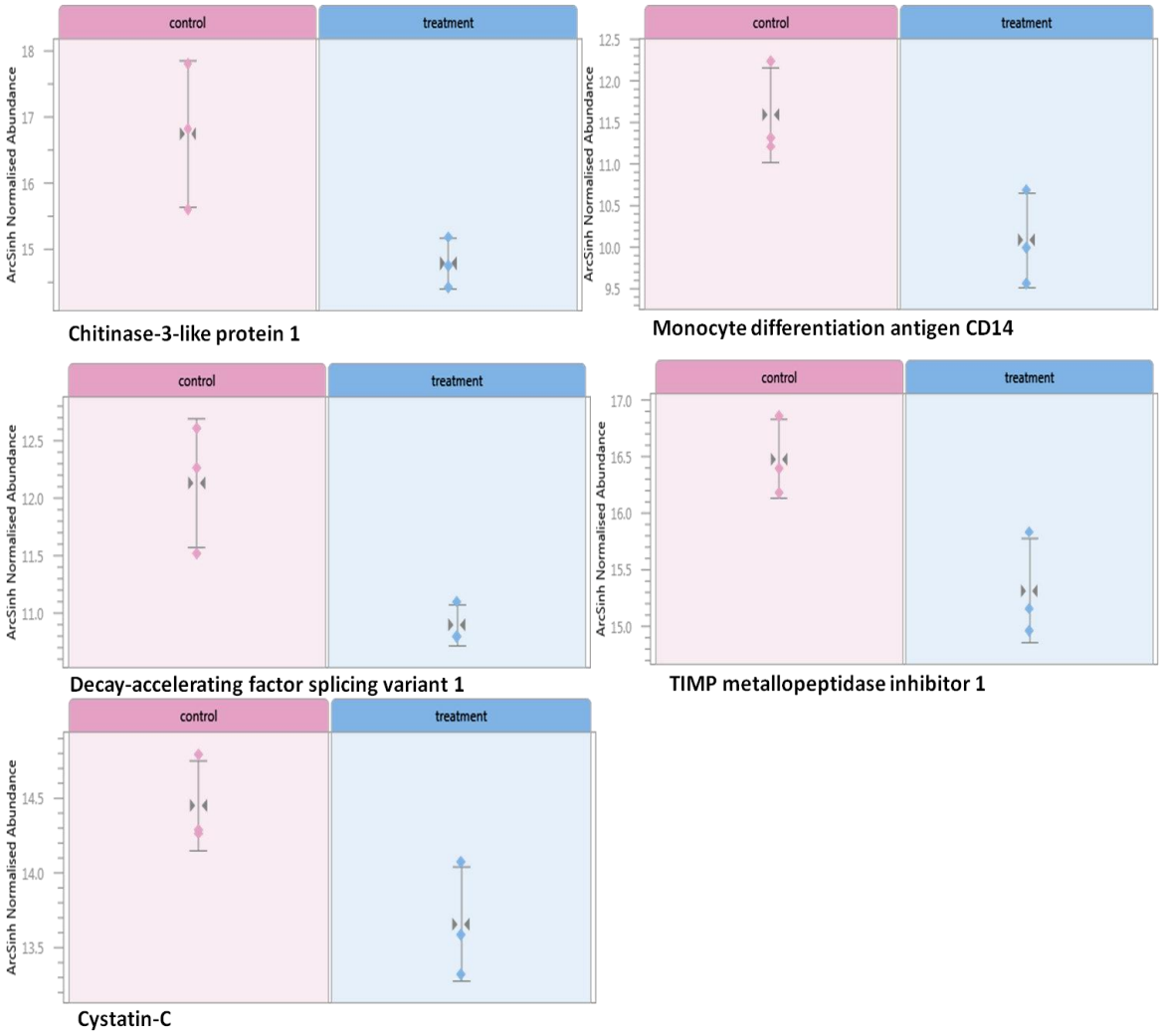



**Figure 4 Expression profile view of proteins significantly reduced in IL-1β treated samples produced by Progenesis™ LC-MS.** All proteins were identified by 2 or more peptides with greater than 2-fold abundance change and $p<0.01$. Plots display the mean arcsinh transformed normalised volume for each group. Error bars demonstrate 3 standard errors within groups.

| Description | Location | Role in arthritis | Function (UNIPROT) |
|---|---|---|---|
| **Growth-regulated alpha protein** | **secreted** | **elevated in RA** | **inflammation** |
| **Stromelysin-1** | **secreted** | **cleaves collagen** | **cleaves collagens of types I, II, III, VII and X** |
| **LYR motif-containing protein 5** | **nucleus, mitochondria** | **unknown** | **protein biosynthesis** |
| **Ig alpha-1 chain C region** | **secreted** | **form complexes in RA** | **immunoglobulin** |
| **Chitinase-3-like protein 1** | **secreted** | **increase in OA** | **tissue remodelling** |
| **Monocyte differentiation antigen CD1** | **cell membrane** | **in synovial fluid early OA** | **innate immune response** |
| **Decay-accelerating factor splicing variant 1** | **cell membrane** | **protect synovial membrane from complement-mediated injury in inflammation** | **inhibitor of complement pathway activation** |
| **TIMP metallopeptidase inhibitor 1** | **secreted** | **control MMPs in cartilage degradation** | **inactivates MMPs** |
| **Cystatin-C** | **secreted** | **mice lacking cystatin C develop OA earlier** | **inhibitor of cysteine proteinases** |

**Table 3. Functions of the differentially expressed proteins along with their cellular location and potential role in arthritis.** The first four proteins represented those with higher expression in Il-1β and the subsequent section those with higher expression in control conditions. (RA; rheumatoid arthritis).

In order to both quantify proteins of interest and validate findings of some of the secreted proteins in our study we employed QconCAT technology.

**QconCAT protein design, expression and validation**

The QconCAT containing proteins of interest to our research studies was designed with two peptides per protein. The QconCAT had an average mass of 58.8kDa which included the N-terminal fibrinopeptide and C-terminal glufibrinopeptide (to allow quantification of the QconCAT) together with a hexahistidine tag for purification. The proteins selected with their respective accession numbers, peptide sequences, *m/z* values of the protonated molecules are presented in Table 1. Following gene synthesis into pET21a vector successful transformation into *E.coli*, induction of expression and purification led to a recombinant protein band which migrated on SDS-PAGE with mobility consistent with an approximate mass of 58kDa implying the QconCAT had been expressed correctly (Figure 5).

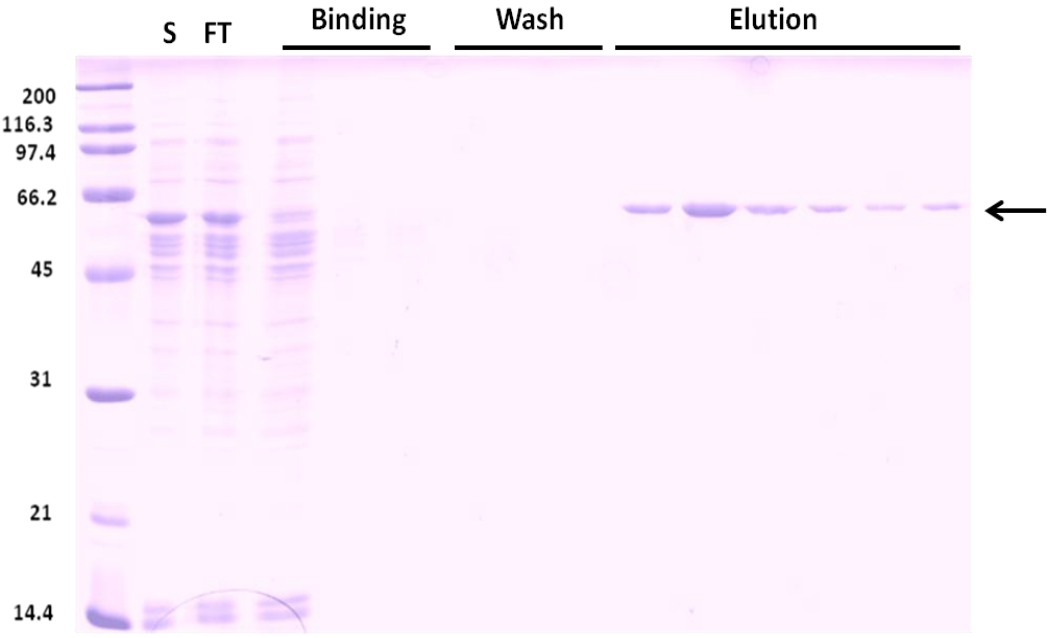



**Figure 5. Purification of the human secretome QconCAT.** The QconCAT was expressed in inclusion bodies which were recovered by centrifugation. Inclusion bodies were dissolved in binding buffer, solubilised using sonication, and then purified using immobilised metal affinity columns. Lanes S and FT, correspond to sequential starting material, flow through fractions. The arrow indicates the QconCAT at approximately 58kDa.

The putative QconCAT protein band from the gel following purification was trypsin digested and analysed with MALDI-TOF mass spectrometry and identified 27 out of 40 peptides from the product ion spectra commensurate with those predicted were observed (Figure 6a and b). Although the peptides were equimolar in proportions within the QconCAT the MALDI spectra demonstrated the expected variability of ionisation between individual peptides. Furthermore the peptides glufibrinopeptide and fibrinopeptide were identified indicating full length expression of the QconCAT, as these peptides are positioned at the beginning and end of the sequence cassette. In order to validate the QconCAT further we identified q-peptides using LC-MS/MS on the LTQ-Orbitrap Velos from an in-solution tryptic digestion, with 100fmol on column QconCAT. Mascot identified 91% sequence coverage of the QconCAT (data not shown). DGFFYFFHGTR and VARPAQLASPTR demonstrated weak fragmentation patterns during the identification and assembly of peptide transitions lists for SRM experiments. No miscleaves were identified indicating complete QconCAT proteolysis.

A 200ml bacterial culture, grown to a cell density of $OD_{A600}$ 0.6–0.8, yielded 33µg of the QconCAT. The identity and chromatographic retention time of the Q-peptides from the unlabeled and labelled QconCAT recombinant proteins were established by preliminary tandem MS analyses of pure QconCATs. The labelling efficiency was high, and the QconCAT peptides were labelled to 98.7%, reflecting the quality of the starting isotopes [$^{13}C_6$]Arg and [$^{13}C_6$]Lys (Figure 6).

The linearity of the response for the QconCAT peptides was established prior to these analyses using labelled and unlabeled QconCAT mixed at different ratios using 50fmol QconCAT loaded on column in an SRM experiment using the XEVO TQ (Figure 7).

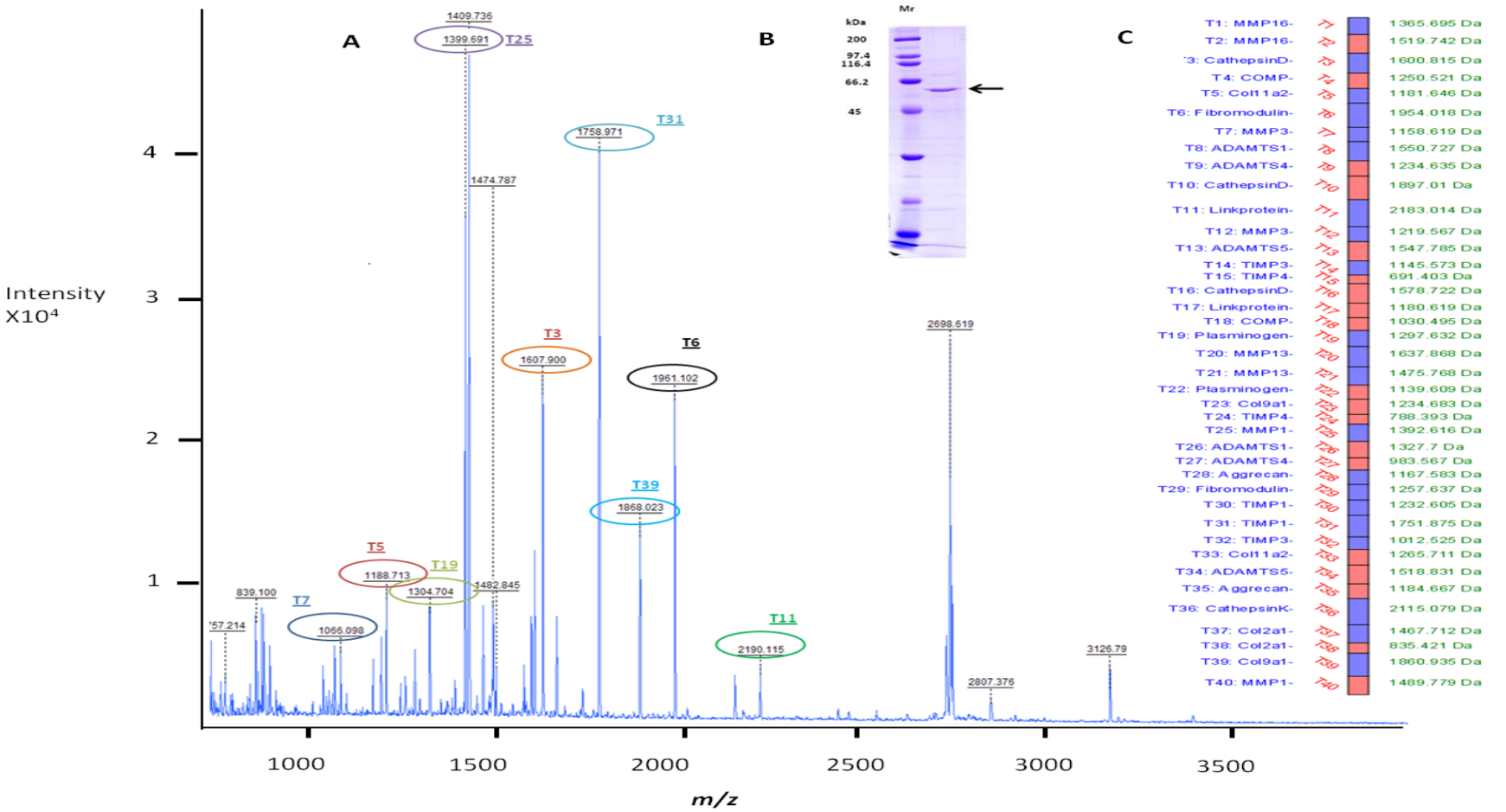



**Figure 5. Validation of the human secretome QconCAT.** (A) In-gel tryptic digestion of the $^{13}C_6$ lys/ $^{13}C_6$ arg labelled QconCAT was analysed with MALDI to give a peptide mass fingerprint of derived peptides. The most intense peptides are indicated as ringed with their respective peptide number. A number of peaks in the lower mass range were identified as matrix peaks. (B) Electrophorectic profile of a 5µg aliquot of purified QconCAT. (C) Schematic representation of the organisation of peptides identified (blue) within the QconCAT.

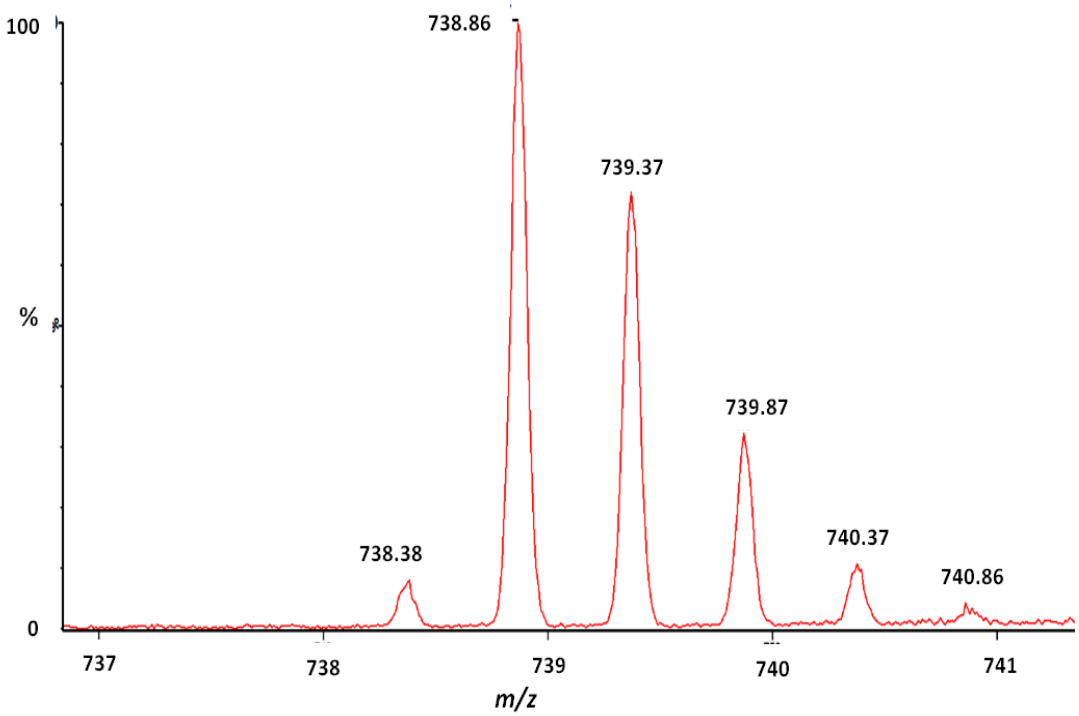



**Figure 6. The labelling efficiency of the QconCAT was high.** An extracted ion chromatogram (EIC) of the q-peptide *m/z* 738.86 ($[M+2H]^{2+}$) was undertaken and the intensity of the minor peak to the left of the major peak was used to calculate the labelling efficiency.

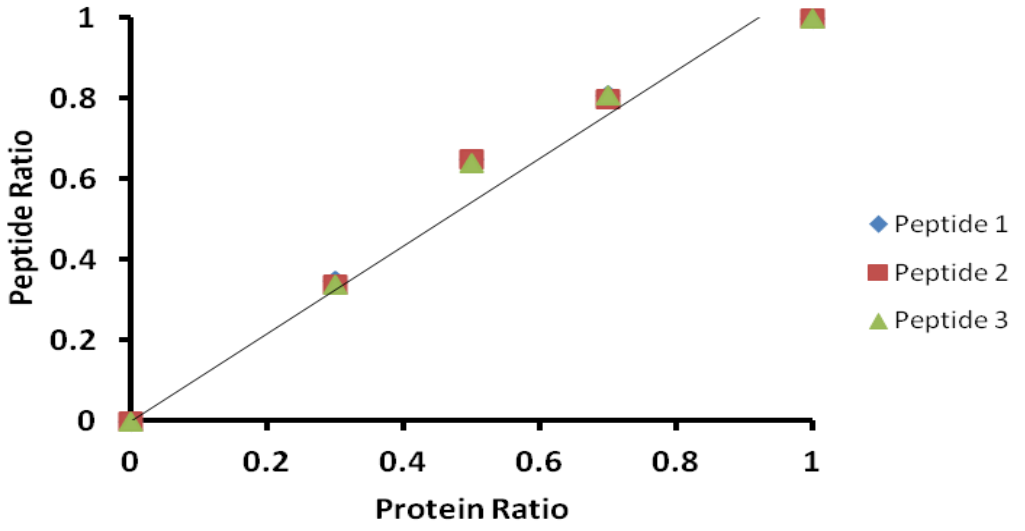



**Figure 7. Linearity of response of using SRM quantification.** Known ratios of heavy and light QconCAT proteins were mixed prior to proteolysis. The digests were then resolved by reversed phase LC-MS and the intensities of the ions corresponding to multiple peptides were measured. Data for three selected peptides, representative of all the peptides are included here.

**Peptide choice and detectability**

The QconCAT was devised to be used in a number of different experiments including the quantification of post translational modifications (PTMs). Thus when designing the QconCAT some peptides were chosen equivalent to were analyte peptides would potentially contain PTM *in-situ* (Huang et al., 1998). Other peptides were chosen from the propeptide sequence of enzymes in order to provide tools for quantifying enzyme activity.

The selection of suitable peptides for quantification was critical. For some proteins we were limited by sequence choices, for example in aggrecan, where the majority of the core protein is heavily modified. Thus we chose two peptides per protein in an attempt to circumvent possible problems such as poor peptide ionisation or detectability by the mass spectrometer which would prevent quantification. For proteins we wished to quantify in the cartilage secretome with a SRM experiment using the QconCAT a simple classification was applied to peptides as described by Brownridge *et al.* (Brownridge et al., 2011). Peptides are classified for quantification purposes as A, B, C for a particular protein loading. 'Type A' quantifications are defined as were both QconCAT and native peptides are observed. For 'type B' quantifications the peptide is detected in the QconCAT but not in the native peptide, and for these peptides sample protein abundance sets the limit on detection. Neither QconCAT nor native peptides are detected in 'type C' quantification, typically due to poor peptide fragmentation or chromatographic behaviour (Table 7). Of the 40 peptides composite of the QconCAT we wished to use 30 peptides in order to quantify their constituent proteins within the cartilage secretome; 12 were type A, 10 were type B, and 8 were type C. This enabled the quantification of seven proteins in our human OA secretome using the criterion that at least one q-peptide per protein was detected in all samples.

| Protein | Q-peptide amino acid sequence | Peptide classification |
|---|---|---|
| Aggrecan | **EVVLLVATEGR** | **A** |
| Cartilage oligomeric matrix protein | **DTDLDGFPDEK** | **A** |
| Cartilage oligomeric matrix protein | **SSTGPGEQLR** | **A** |
| Fibromodulin | **IPPVNTNLENLYLQGNR** | **A** |
| Matrix metalloproteinase-1 | **SQNPVQPIGPQTPK** | **A** |
| Matrix metalloproteinase-3 | **IVNYTPDLPK** | **A** |
| Metalloproteinase inhibitor 1 | **GFQALGDAADIR** | **A** |
| Plasminogen | **EAQLPVIENK** | **A** |
| ADAMTS1 | **DAEHYDTAILFTR** | **B** |
| ADAMTS1 | **GPEVTSNAALTLR** | **B** |
| ADAMTS4 | **FVETLVVADDK** | **B** |
| ADAMTS4 | **NPVSLVVTR** | **B** |
| ADAMTS5 | **LPLAAVGPAATPAQDK** | **B** |
| ADAMTS5 | **GLVQNIDQLYSGGGK** | **B** |
| Aggrecan | **LEGEVFFATR** | **B** |
| Cathepsin D | **LVDQNIFSFYLSR** | **B** |
| Cathepsin D | **YSQAVPAVTEGPIPEVLK** | **B** |
| Cathepsin K | **SNDTLYIPEWEGR** | **B** |
| Link protein | **GGSDSDASLVITDLTLEDYGR** | **B** |
| Metalloproteinase inhibitor 1 | **FVGTPEVNQTTLYQR** | **B** |
| Metalloproteinase inhibitor 3 | **WDQLTLSQR** | **B** |
| Metalloproteinase inhibitor 4 | **GHLPLR** | **B** |
| Cathepsin K | **VGPVSVAIDASLTSFQFYSK** | **C** |
| Fibromodulin | **LYLDHNNLTR** | **C** |
| Link Protein | **FYYLIHPTK** | **C** |
| Matrix metalloproteinase-1 | **DGFFYFFHGTR** | **C** |
| Matrix metalloproteinase-13 | **LHPQQVDAELFLTK** | **C** |
| Metalloproteinase inhibitor 3 | **YQYLLTGR** | **C** |
| Metalloproteinase inhibitor 4 | **LEANSQK** | **C** |
| Plasminogen | **HSIFTPETNPR** | **C** |

**Table 7A. Table represents peptide types as determined by SRM experiments.** Q-peptides are classified for quantification purposes as A, B, C. 'Type A' native and QconCAT peptides are detected. 'Type B' are peptides detected for the QconCAT but not in native form and when neither QconCAT nor native peptides are detected a 'type C' classification is given (ADAMTS; A disintegrin and metalloproteinase with thrombospondin motifs).

**Identification of HAC secretome proteins**

Prior to SRM experiments tryptic digests of all the secretome media samples were analysed using LC-MS/MS on the LTQ-Orbitrap Velos in order to identify the presence of proteins to be quantified within the QconCAT. When these proteins were identified using Mascot we undertook extracted ion chromatograms for the q-peptide *m/z* of the selected QconCAT peptide for confirmation (Table 7b).

| Protein | Peptide | Peptides identified with LC-MS/MS | | | | | | Peptides quantified with SRM | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | control | | | treatment | | | control | | | treatment | | |
| | | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| Aggrecan | EVVLLVATEGR | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Aggrecan | LEGEVFFATR | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | x | x | x | x | ✓ | x |
| COMP | DTDLDGFPDEK | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| COMP | SSTGPGEQLR | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fibromodulin | IPPVNTNLENLYLQGNR | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Link protein | GGSDSDASLVITDLTLE DYGR | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | x | x | x | x | x | x |
| Link Protein | FYYLIHPTK | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | x | x | x | x | x | x |
| MMP-1 | SQNPVQPIGPQTPK | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| MMP-3 | IVNYTPDLPK | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| TIMP-1 | GFQALGDAADIR | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Plasminogen | EAQLPVIENK | x | x | x | x | x | x | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

**Table 7B. Table indicating q-peptides identified from LC-MS/MS studies in human secretome media with and without IL-1β stimulation using UNIHUMAN database with Mascot**. Peptides identified with Mascot were confirmed using extracted ion chromatograms.

**Quantification of proteins using SRM**

SRM of multiple product ions (transitions) were used for quantification as it provided a sensitive method for targeted analyte identification and quantitation. Operating at the MS/MS level, selectivity is provided by incorporating an extra isolation step; monitoring the product ions from a specific set of precursor ions.

Where possible, two transitions per peptide were used for quantification. Transitions were defined using Skyline software (O'Connor et al., 1998) and selected after monitoring for the greatest intensity fragments using 50fmol

QconCAT digest on the XEVO TQ. Fragmented y-ions were selected in order to differentiate labelled and unlabelled peptide as the C-terminal residue contained the isotope-labelled amino acid. In addition where possible we selected transitions whose *m/z* was greater than the parent ion *m/z* to maximise specificity (Table 7c). We were unable to identify any adequate transitions for quantification purposes for two peptides; VGPVSVAIDASLTSFQFYSK (cathepsin K) and DGFFYFFHGTR (MMP-1), emphasizing the need to include at least two peptides within a QconCAT to quantify a protein.

| Protein | Peptide | Precursor *m/z* | Transition *m/z* | Fragment type |
|---|---|---|---|---|
| Aggrecan | **EVVLLVATEGR** | 593.3468 | 632.3362 | y6 |
| Aggrecan | **EVVLLVATEGR** | 593.3468 | 745.4176 | y7 |
| COMP | **DTDLDGFPDEK** | 626.2723 | 807.3456 | y7 |
| COMP | **DTDLDGFPDEK** | 626.2723 | 488.2318 | y4 |
| COMP | **SSTGPGEQLR** | 516.2594 | 699.376 | y6 |
| COMP | **SSTGPGEQLR** | 516.2594 | 756.3991 | y7 |
| Fibromodulin | **IPPVNTNLENLYLQGNR** | 652.353 | 750.3869 | y6 |
| MMP-1 | **SQNPVQPIGPQTPK** | 745.9028 | 837.4781 | y8 |
| MMP-1 | **SQNPVQPIGPQTPK** | 745.9037 | 1161.6524 | y11 |
| MMP-3 | **IVNYTPDLPK** | 580.3217 | 947.475 | y8 |
| MMP-3 | **IVNYTPDLPK** | 580.3217 | 569.3254 | y5 |
| Plasminogen | **EAQLPVIENK** | 570.8157 | 699.3985 | y6 |
| Plasminogen | **EAQLPVIENK** | 570.8168 | 812.4864 | y7 |
| TIMP-1 | **GFQALGDAADIR** | 617.3124 | 717.3489 | y7 |
| TIMP-1 | **GFQALGDAADIR** | 617.3124 | 901.4698 | y9 |

**Table 7C. Parameters used in SRM assays**. The parent and fragment ion *m/z* and fragment type for peptides used to quantify the human OA secretome.

The seven proteins quantified using an SRM approach were aggrecan, COMP, fibromodulin, MMP-1, MMP-3, plasminogen and TIMP-1 (Figure 8a). For COMP we were able to quantify using both peptides; SSTGPGEQLR and DTDLDGFPDEK. DTDLDGFPDEK consistently producing a lower ratio of light peak area/heavy peak area compared to SSTGPGEQLR (Figure 8b). Analysis of Mascot data of analyte digests discovered that occasional miscleavage of DTDLDGFPDEK was evident; this

resulted in an underestimation of protein abundance when this peptide was used for absolute quantification. Therefore we used SSTGPGEQLR for quantification of COMP. Following IL-1β treatment there was an apparent increase in aggrecan, COMP, fibromodulin, MMP-1, MMP-3 and plasminogen and a reduction in TIMP-1. The only protein which achieved statistical significance was TIMP-1, with a reduction in IL-1β treated samples (p=0.0017) although MMP-3 exhibited a trend (p=0.06).

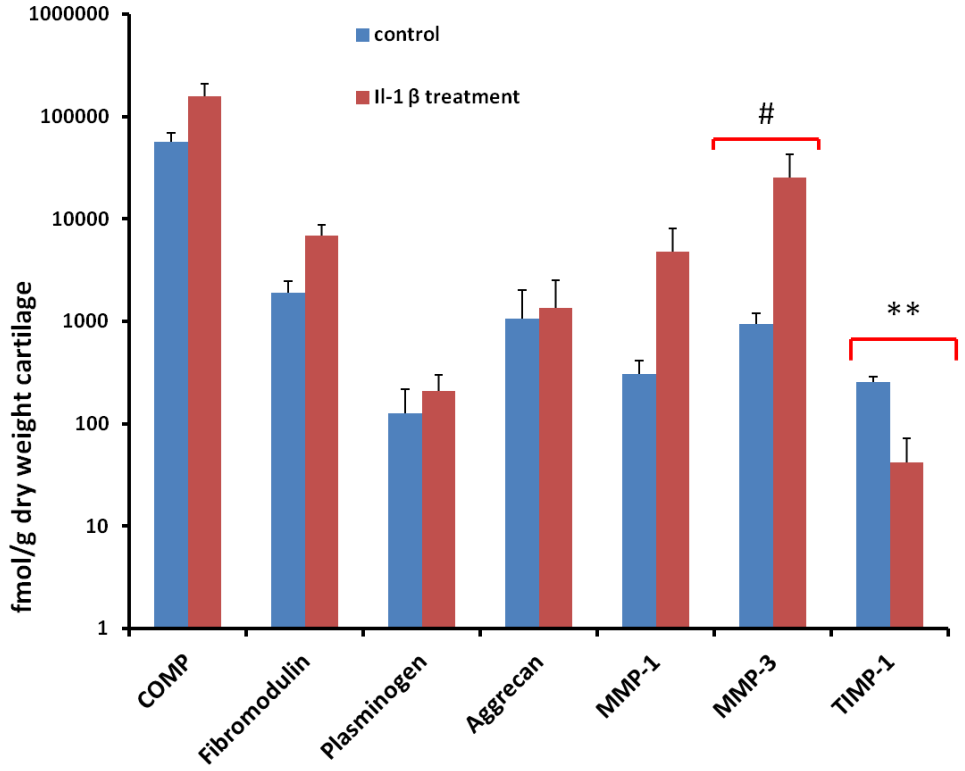



**Figure 8A. Proteins measured in human secretome media using QconCAT.** Extracted ion chromatograms were performed for each peptide and the total ion count used to determine the ratio of light peak area/ heavy peak area at a given QconCAT loading. The protein abundance in the media was then calculated based on the amount of total protein in the media sample. This was then normalised to the dry weight of explants. Mean concentrations and ±SEM (n=3) are indicated. Data were evaluated using mixed effect linear regression and ** indicates significant difference relative to control at the p<0.01 level, # indicates p=0.06.

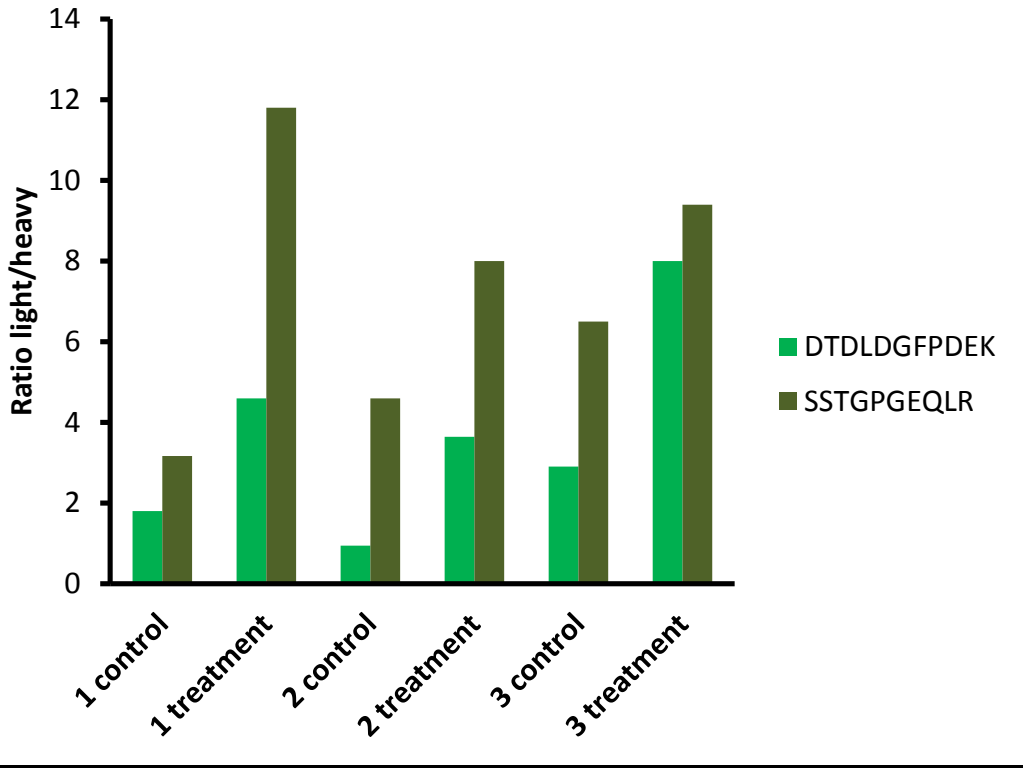



**Figure 8B. Ratios of EIC for peak area light/ peak area heavy peptide for the two q- peptides from the QconCAT used to quantify COMP.** Data shows control and IL-1β treatment values for each of the three human explant donors.

**Discussion**

Cartilage proteomics is developing from simple protein identification through to quantitation. For OA research absolute quantitative proteomics will enable further biological questions to be addressed, by facilitating the experimenter in determining protein abundance in terms of tangible amounts. However, there have been few studies able to absolutely quantify cartilage secreted proteins and experiments have employed relative quantification approaches using platforms including 2-D gel approaches (Catterall et al., 2006; Hermansson et al., 2004), SILAC studies (Polacek et al., 2010b), isobaric tags for relative and absolute quantitation (iTRAQ) (Stevens et al., 2009), and quantitative western blotting (Pullig et al., 2002b). Absolute quantification techniques and its variations have been utilized to measure absolute amounts of a given peptide, allowing quantitative comparisons of different proteins. Whilst AQUA peptides provide synthetic isotope labelled

peptides they remain costly, and usually only allow single proteins to be quantified at a time. In contrast QconCAT technology allows the cost efficient production of heavy isotope labelled standards. The aim of devising this human cartilage QconCAT was to provide an accurate, low cost method able to support large scale protein quantification in cartilage studies.

This study tested the hypothesis that there were measurable changes in protein abundance in the human OA secretome following cytokine stimulation and these could be absolutely quantified using a QconCAT. First the study implemented a comparative proteomic analysis using a label-free LC-MS/MS-based strategy. Specific proteins were then quantified using QconCAT to validate both label-free results and the QconCAT technology itself in the study OA. The model used to assess these methodologies was the human OA cartilage explant secretome following cytokine stimulation. This enabled the absolute quantification of the human cartilage secretome for the first time.

The established cartilage explant model mimics the catabolic events that occur in OA (Ismaiel et al., 1992; MacDonald et al., 1992) . The explants were washed and rested overnight as it has been previously identified that dissection and culture alone activates c-Jun N-terminal kinase (Hermansson et al., 2004). Thus the expression patterns of proteins synthesised by the explants may differ from *in-vivo*. Although the cartilage was already diseased, there is precedence in using such samples as pathological human cartilage has been utilized to examine the role of cytokines and the targeted analysis of protein expression alterations in OA (Goldring et al., 2011). Cartilage explants enables chondrocytes to be retained within their ECM allowing them to retain their phenotypic stability. The ECM also provides native substrates for proteolysis and protein release, equivalent to the shedding the proteins into the synovial fluid (and beyond) during cartilage degeneration. Such explant studies may contribute further understanding of the pathogenesis of OA. Although for this study we were unable to obtain normal mature human articular cartilage this would be beneficial for further interpretation

of our findings and indeed the QconCAT described here could be utilized to this end.

Using 1-D SDS-PAGE we were able to use a qualitative proteomics approach in order to identify the predominate proteins in the HAC OA secretome. These included MMP-3 and COMP. Interestingly, although explants were washed copiously prior to entry to the study serum albumin was identified as a predominant protein in the samples. This may be due to uptake by the cartilage of serum albumin and release into the matrix during the experiment but also highlight the need for proficient explant preparation in proteomic studies as the dominance of a single protein may 'mask' the identity of other less abundant proteins using mass spectrometry. Densitometry of the gel did not identify differential protein expression following 48h stimulation with IL-1β and this may be due to insufficient exposure time to the cytokine or previous findings by others that responsiveness to IL-1$\beta$ is reduced in late stage OA chondrocytes (Forsyth et al., 2005). Indeed other studies using human OA explants determined that at least 14 days of cytokine stimulation was required to identify a significant loss in glycosaminoglycans (GAG) the primary component of which is aggrecan (Ismaiel et al., 1992). This finding is in contrast to our previous work were IL-1β stimulation of normal equine cartilage for 5 days produced a differential expression of cartilage matrix proteins using 1D-SDS-PAGE (manuscript 1). Furthermore, results from the label-free study indicated very few differentially regulated proteins in our model (nine out of 278 proteins identified) in comparison to our other studies in normal equine cartilage stimulated with IL-1β which demonstrated over 100 differentially expressed proteins (Peffers et al., 2012).

High-throughput proteomic technologies created large data sets posing challenges in interpretation. Therefore a proteomics tool called Progenesis™ LC-MS (Zhang et al., 2010) was utilized, enabling the analysis and comparison (in terms of relative quantification) of proteins in our experiment. Following Progenesis™ LC-MS analysis we applied insightful data mining using bioinformatics tool DAVID (Huang et al., 2008) in order to interpret the data in relationship to protein location and function.

As predicted a large number (53%) of proteins identified by GO, which uses statistical analysis to validate results, were secreted. Whilst the dataset of differentially expressed proteins identified by Progenesis™ was small it was interesting that a number of them were involved in inflammation and the innate immune response. Inflammation plays an important role in the pathogenesis of OA (Buckland, 2012b) and following proinflammatory cytokine stimulation of already diseased cartilage it was not surprising that proteins, implicated in pathways associated with rheumatoid arthritis and OA, were identified. It is hypothesised that inflammation might actually be driven by the fragments such as fibronectin (Ding et al., 2009) that are released by cartilage degradation, through activation of the innate immune responses. More recently it has been identified that inflammatory complement cascade has a key role in the pathogenesis of OA (Fernandes et al., 2002). Through proteomic and transcriptomic analyses of synovial fluids and membranes from OA patients, the expression and activation of complement was found to be abnormally high in human OA joints. One interesting finding was the down regulation of decay accelerating factor splice variant (DAF) in IL-1β stimulated explant media. DAF belongs to the complement system and protects cells from complement mediated lysis. Immunohistochemistry studies identified an increased in OA cartilage compared to normal (Davies et al., 1994) and there was an increase in transcript in macroscopically affected OA joint cartilage compared to intact cartilage in the same joint (Geyer et al., 2009). It is possible that aberrant regulation of DAF is occurring due to IL-1β stimulation of already diseased cartilage. Further work into the role of this protein is warranted. Finally as inflammation is an early and persistent event the involvement of joint tissues in OA could be monitored by quantifying levels of a panel of markers such as the inflammatory factors identified as differentially expressed in this study; DAF, growth-regulated alpha protein and IG alpha-1 chain C region.

Many ECM proteins and proteases were identified in the secretome using LC-MS/MS. However some interesting proteins, particularly proteases were not identified such as MMP-1, ADAMTS-4, ADAMTS-1 and the cathepsins. We wondered if a more targeted approach, such as SRM which increases sensitivity

(Hossain et al., 2011), would enable the identification and subsequent quantification of further proteins. Therefore a human cartilage QconCAT was designed. This approach identified and quantified aggrecan, COMP, fibromodulin, link protein, MMP-1, MMP-3 and TIMP-1 from within the secretome.

Peptide repositories available to us on-line together with discovery data were used in the QconCAT design. These data sets include those providing a compendium of targeted proteomics assays (from complex proteome digests) of peptides isolated on a triple quadrupole mass spectrometer as a resource for SRM workflows. The study exemplified the need to use all methods available to select 'proteotypic' peptides prior to inclusion within a QconCAT, although optimal selection of peptides can still have inconsistent success. Once a short list of peptides was highlighted we used a standard QconCAT nomination strategy based on the probability of efficient cleavage and propensity to generate a good MS signal (Brownridge et al., 2012). In addition, we chose peptides both including and without known post-translational modifications (PTMs). Since this study was implemented further informatics tools have become available that may help in the choice of peptides and so increase the likelihood of selecting a detectable peptide. One such tool is CONSeQuence which predicts the probability of a peptide being detected by electrospray ionisation (Eyers et al., 2011).

A strategy was used where two peptides were included per protein in the QconCAT to allow quantification with at least one peptide should one fail. For the seven proteins identified in MS/MS data and then quantified with SRM we were reduced to a single peptide for quantification for all peptides except COMP; for this protein a single peptide was nominated for quantification purposes. The main reason for redundant q-peptides was poor peptide selection, resulting in 'type C' peptides within our QconCAT. The primary reason for type C peptides was poor fragmentation of the parent peptide resulting in inadequate transitions for quantification using an SRM approach. Interestingly both link protein peptides were identified in the label-free experiments (where both MS and MS/MS are used in protein identification), but we were unable to quantify either peptide. One q-

peptide; FYYLIHPTK was identified as a 'type C' peptide. The other peptide; GGSDSDASLVITDLTLEDYGR was readily detected down to 0.1fmol in the standard but was not detected in analyte. It is possible that there may be a previously unidentified PTM on this peptide. In addition this peptide displayed a poor fragmentation pattern as demonstrated by the identification of only a single transition for SRM experiments.

MMP-3 was quantified with a single peptide as the other peptide for this protein was in the pro-peptide region. The nominated q-peptide for the MMP-3 propeptide ('q-propeptide'); WDQLTLSQR was identified as a 'type B' indicating that although the QconCAT heavy peptide was identified the analyte peptide was not. MMP-3 is secreted from chondrocytes as zymogens, which become activated in the extracellular space, a critical level of enzymatic control. Since the propeptide is released on activation it would have been interesting to measure the ratio of q-propeptide to q-peptide in order to gauge MMP activity. Therefore in the design of the QconCAT q-propeptides for MMP-3, -13 and ADAMTS5 were included. Unfortunately apart from the MMP-3 q-propeptide the others were 'type C'. There are a number of possibilities as to why the MMP-3 q-propeptide was not identified in the secretome. It may have been present in lower concentrations than the limits of detection of the SRM experiment. Furthermore whilst final processing to the mature form of the active MMP (where the entire propeptide is removed) requires intermolecular, autoproteolytic cleavage by the target MMP, physiologic activation of MMPs is probably initiated by proteases that cleave specific sites within the propeptide, (Morgunova et al., 1999). Therefore another possibility for not identifying the q-propeptide could be that one of the initial cleavages was within this q-propeptide site.

For aggrecan EVVLLVATEGR was used in quantification experiments as the other peptide LEGEVFFATR was not detected in all samples. The tryptic cleavage site at the N-terminal of this peptide was close to an aspartic acid possibly leading to miscleaves, although no evidence was found for this in MS/MS data. While for COMP we were able to quantify using both peptides, DTDLDGFPDEK consistently

producing a lower ratio of light peak area/heavy peak area possibly resulting from miscleavage occurring due to the acidic residue; aspartic acid adjacent to the cleavage site in the analyte (Liu et al., 1998). Thus SSTGPGEQLR was utilized in the quantification of COMP. For the proteins fibromodulin, MMP-1, TIMP-1 and plasminogen, one peptide per protein was a type 'C' peptide, resulting in a single peptide being used for quantification by SRM.

Il-1β is one of the most significant cytokines in OA (Tetlow et al., 2001) and is assumed to cause damage to OA cartilage through both the induction of protease expression, resulting in cartilage matrix degradation (Arner et al., 1998; Goldring and Goldring, 2004) and reduction in the expression of anabolic genes such as aggrecan and COL2A1 (Aigner et al., 2006). This results in the anabolic-catabolic discrepancy characteristic of OA. Here we examined OA cartilage degeneration in culture using secretome protein identification and both relative and absolute quantification. Relative quantification identified the induction of MMP-3 protein expression following IL-1β stimulation. The absolute quantification also revealed an increase in MMP-3 in all donors in each replicate. However although $p<0.06$ this did not reach statistically significance. Whilst others identified an induction of MMP protein expression along with degradation of matrix constituents in OA HAC explants (Burrage et al., 2006)the induction of MMP-3 demonstrated in this study is not accompanied by an increase in ECM proteins within the secretome media following IL-1β stimulation which might be expected following ECM degradation. This could be due to the short time scale of the experiment or that there is a reduced responsiveness to IL-1β in late stage OA chondrocytes. Statistical significance for MMP-3 may be reached by increasing the power of the significance test (the ability to detect an alternative hypothesis). One way to do this is by increasing the sample size. This provides more information about the mean and offers an improved chance of distinguishing the mean percentage change.

The study quantified the concentration of COMP, fibromodulin and aggrecan, matrix components that have previously been identified in explant secretomes in humans (Wu et al., 2007), mouse (Wilson et al., 2008) and horse (Irving et al., 1990)

with an apparent, though not statistically significant increase in aggrecan, COMP and fibromodulin. One of the most abundant proteins quantified here was COMP, a non-collagenous matrix protein. Its presence in the secretome corresponded to previous studies of cartilage explants (Irving et al., 1990; Wilson et al., 2008; Wu et al., 2007). COMP organises ECM assembly (Halasz et al., 2007) and attaches the chondrocyte to the ECM (Chen et al., 2005). Considered as a marker of cartilage breakdown it has been studied as a biological marker (Hoch et al., 2011; Zhang et al., 2006b). Measurement of intact COMP and fragments thereof in synovial fluid or serum have been shown to correlate to cartilage destruction in OA patient studies (Saxne and Heinegard, 1992) and so it is no surprise that it is abundant in the secretome. It is possible that in the future a QconCAT approach could be used to measure COMP in different tissues and biological fluids.

Fibromodulin, a collagen-binding protein (Heathfield et al., 2004) was also quantified. This protein protects the surface of collagen type I and II fibrils from proteolysis by MMPs (Geng et al., 2006) . Cleavage products of fibromodulin have been identified during IL-1 stimulation of cartilage explant studies *in-vitro* (Heathfield et al., 2004; Monfort et al., 2006) and as such cleavage of fibromodulin may represent an important initial episode that interrupts the collagen fibrillar network leading to more sites for proteases to cleave collagen further. Furthermore it has been suggested that some ECM proteins including fibromodulin become endogenous catabolic factors during joint damage and stimulates innate immune pathways via complement activation (Sofat, 2009). Thus the presence of fibromodulin in relative abundance within the OA secretome presents a means by which ongoing joint damage may be further precipitated.

Plasminogen, a serine protease and important activator of pro-MMPs, has been demonstrated to induce cartilage degradation (Oleksyszyn and Augustine, 1996). Its production is stimulated by Il-1β in cartilage (Collier and Ghosh, 1988). Although it was not evident in initial MS/MS experiments it was quantified using SRM. SRM experiments are unique in their ability for reliable quantification of analyte of low abundance in complex mixtures (Lange et al., 2008). This demonstrates the

advantage of our QconCAT approach to proteomics experiments in the study of low abundance proteins such as plasminogen involved in OA.

MMP activity is regulated by a family of tissue-specific inhibitors including TIMP-1 (English et al., 2006). Presently it is believed that the local balance of MMPs and TIMP activities is crucial for cartilage homeostasis, with disturbances producing higher levels of MMPs over TIMPs resulting in cartilage pathology such as OA. TIMP concentrations generally far exceed the concentration of MMPs in tissue and extracellular fluids, thereby limiting their proteolytic activity to focal pericellular sites by binding to the MMP active sites (Gomez et al., 1997). TIMPs also inhibit cleavage of proteoglycans by aggrecanases (Grad et al., 2006). TIMP-1 specifically has been demonstrated to have inhibitory activity against ADAMTS-4 (Tortorella et al., 1999). In the IL-1β stimulated secretome MMP-3 concentrations exceeded TIMP-1 concentrations as identified using relative and absolute quantification which was probably due to the late stage OA nature of the cartilage explants. Interestingly in the QconCAT study TIMP-1 was the only protein significantly affected by cytokine stimulation agreeing with the label-free study. TIMP-1 mRNA has been previously identified in OA (Kane et al., 2004) and rheumatoid arthritis (Martel-Pelletier et al., 1994). Our results agree with other studies demonstrating that IL-1β stimulation has a marked inhibitory effect on TIMP-1 expression by chondrocytes (Sadowski and Steinmeyer, 2001).

Bringing the results of the different quantification methodologies together it would seem that there is good agreement with the findings from the two studies when proteins were identified as significantly differentially expressed in the secretome. MMP-3 was increased significantly in the label-free study and increased to near significance in the QconCAT study. TIMP-1 expression was identified as significantly reduced using both methodologies. The ECM proteins aggrecan, fibromodulin and COMP were not differentially expressed in the study using either method.

## Conclusions

This study is the first to combine relative and absolute protein quantification in the analysis of the human OA secretome. It enabled the identification of a cohort of proteins expressed by osteoarthritic cartilage with possible roles in its pathogenesis. A human cartilage QconCAT was designed expressed and validated which enabled absolute levels of important proteins in the study of OA to be quantified, demonstrating the usefulness of QconCAT technology in OA research. This is especially useful in quantifying low abundance proteins. Not surprisingly, several of the peptides in this QconCAT did not provide quantotypic performance. However by including two peptides per protein the QconCAT was still robust enough to allow absolute quantification of the majority of proteins and provides a tool for the precise definition of some matrix proteins and proteases within the pathogenesis of OA.

## Acknowledgements

This study was supported by a Wellcome Trust Veterinary Integrated Research Fellowship. We thank the staff at Clatterbridge Hospital, Wirral, Merseyside, for provision of human joint tissue.

# Manuscript 4

MJ Peffers[1*], RJ Beynon[2], DJ Thornton[3], PD Clegg[1]

**Corresponding Author:** M J Peffers[1]

[1]Dept.of Musculoskeletal Biology, Institute of Ageing and Chronic Disease, University of Liverpool, Leahurst, Chester High Road, Neston, Wirral, CH64 7TE; peffs@liv.ac.uk, pclegg@liv.ac.uk

[2] Proteomics Group, Institute of Integrative Biology, University of Liverpool, Biosciences Building, Crown Street, Liverpool, L69 7ZB; rbeynon@liv.ac.uk

[3] Wellcome Trust Centre for Cell Matrix Research, Michael Smith Building, Faculty of Life Sciences, Oxford Road, Manchester, M13 9PT, UK; dave.thornton@manchester.ac.uk, j.selley@manchester.ac.uk

**Title: Characterisation of Neopeptides in Equine Articular Cartilage**

**Running Title**: **Potential Cleavages Sites in Equine Articular Cartilage**

**Key Words:** aggrecan, cartilage, COMP, biglycan, decorin, fibromodulin, neopeptides, osteoarthritis

## Characterisation of Neopeptides in Equine Articular Cartilage


### Abstract

Osteoarthritis (OA) is characterized by a loss of extracellular matrix (ECM) which leads to cartilage degradation and joint space narrowing. Specific proteases including the aggrecanases ADAMTS-4 and matrix metalloproteinase 3 (MMP-3) are important in initiating and promoting cartilage degradation in OA. This study investigated protease specific and disease specific cleavage patterns of some ECM proteins by comparing new peptide fragments; neopeptides, in specific exogenous protease driven digestion of a cartilage extract with an in-vitro model of early OA. A proteoglycan-rich extract was purified from equine articular cartilage, and then digested by the addition of exogenous ADAMTS-4 or MMP-3. In a further experiment equine cartilage explants were treated for 96 hours with interleukin-1 (IL-1β) and the media collected. Proteolysed products following trypsin digestion from all samples were then identified using mass spectrometry. Complete sequences of peptides proteolysed were determined for the major cartilage proteoglycans aggrecan, biglycan, decorin, fibromodulin plus cartilage oligomeric matrix protein (COMP). The generation of peptides varied with enzyme specificity; however some peptides were common to all samples. Specific known cleavage sites were evident for aggrecan, biglycan and COMP as well as a catalogue of potential novel sites following proteolysis. The use of an *in-vitro* model of early OA employing cytokine treated explants enabled novel cleavage sites to be identified under more pathologically relevant conditions. The identification of novel 'neo-terminal' fragments provides a platform for the development of antibodies which could assist in the identification of biomarkers for OA, as well as identifying basic biochemical processes underlying OA.


### Introduction

The unique load bearing properties of articular cartilage are dependent upon its structural composition and organisation, particularly the interactions between collagens and proteoglycans of the ECM (Poole et al., 2001). In normal physiology

these matrix macromolecules are turned over by chondrocytes embedded within the cartilage.

Progressive degeneration of articular cartilage, including proteolysis driven degradation leads to joint pain and dysfunction that is clinically identified as OA. Under normal circumstances, there is equilibrium between matrix deposition and degradation; however this equilibrium is disrupted in OA leading to the excessive degradation of matrix and progressive loss of important matrix components such as collagens and proteoglycans (Martel-Pelletier et al., 1994; Theodoropoulos et al., 2009). Furthermore intact and fragmented ECM peptides, produced following degradation, effect chondrocyte function through integrin receptor signalling (Loeser, 2002).

Studies of the fragmentation of the ECM constituents have elucidated a series of cartilage protein degradations (Fosang et al., 1996a; Milner et al., 2010). *In-vivo* a number of specific proteinases which originate from the chondrocyte or the cells infiltrating inflamed synovium induce cartilage breakdown (Malemud et al., 2003), whilst in cytokine stimulated models of OA the breakdown of cartilage is also by stimulated by a series of proteolytic enzymes through the up regulation of metalloproteinase (Shohani et al., 2010). Initially aggrecan is fragmented and released from cartilage followed by other molecules such as COMP, fibromodulin and collagens.

The two primary enzyme families important in cartilage matrix degradation in OA are the metalloproteinases (MMPs), including matrix metalloproteinase 3 (MMP-3) and the disintegrin with metalloproteinase and thrombospondin motif family (ADAMTS) including ADAMTS-4 (Nagase and Kashiwagi, 2003). MMP-3 is one of the most highly expressed proteases in cartilage capable of degrading proteoglycans including aggrecan (Zhen et al., 2008) as well as activating procollagenases. ADAMTS-4 is an important enzyme in the pathogenesis of OA as demonstrated by its high aggrecanase activity in OA cartilage, and localised expression in areas of aggrecan degradation (Fosang et al., 2008; Tortorella et al., 2000a). It cleaves

aggrecan at five known sites resulting in a loss of the chondroitin sulphate rich regions from cartilage, and thus to a reduction in the ability of cartilage to withstand compressive forces (Struglics et al., 2006). Furthermore it is an important mediator of cytokine stimulated aggrecan loss in normal cartilage (Majumdar et al., 2007). Both these classes of proteases are also involved in the cleavage of other ECM proteins including biglycan and decorin (Monfort et al., 2006), fibromodulin (Kashiwagi et al., 2004) and COMP (Dickinson et al., 2003).

The degradation products of the small leucine-rich repeat protein family of proteoglycans (SLRP) which include fibromodulin, decorin and biglycan have been documented. ADAMTS-4 has been identified as having activity against fibromodulin at the same site of cleavage as MMP-13 (Kashiwagi et al., 2004). Work by Melching *et al.* 2006 (Melching et al., 2006) determined that biglycan was a substrate for ADAMTS-4 and -5 at a site within the fifth leucine-rich region at an asparginine-cysteine bond. In addition biglycan is also susceptible to cleavage by MMP-13 at $Gly^{177}/Val^{178}$ (Monfort et al., 2006). Decorin degradation following MMP induced proteolysis has also identified cleavage sites in cartilage (Monfort et al., 2006) . An important approach to monitoring arthritis is by measuring biological markers of cartilage degradation and repair which reflect changes in joint remodelling. COMP is considered a marker of cartilage breakdown, and has been studied as a biological marker (Mundermann et al., 2005). Measurement of intact COMP and fragments thereof in synovial fluid or serum correlates to cartilage destruction in RA and OA patient studies (Mansson et al., 1995; Saxne and Heinegard, 1992). Numerous sites of degradation in COMP have been postulated for MMPs and aggrecanases (Dickinson et al., 2003).

The progression of OA has been evaluated until recently using imaging techniques such as radiography (Blackburn et al., 1996) and magnetic resonance imaging (Conaghan et al., 2006) or invasive techniques such as arthroscopy to assess gross changes in the joint. However the advent of advanced proteomic techniques has enabled the identification and use of protein biomarkers to become established

(Rousseau and Delmas, 2007) which may aid in not only the monitoring of OA progression but also examine the effects of potential treatments for OA.

A number of peptides derived from cartilage degradation are used as biomarkers of OA. These include markers derived from, bone, cartilage and synovium. Some collagen biomarkers are already utilised in established assays such as the α1 chain of human type II collagen (the HELIX-II neoepitope )(Charni-Ben Tabassi et al., 2008; Eyre and Weis, 2009), Coll 2-1 and Coll 2-1 NO2 (Everts et al., 2003) and type II collagen C-telopeptide fragments (CTX-II) (Christgau et al., 2001). Non-collagenous biomarkers that have been described include COMP (Tseng et al., 2009) and aggrecan (Dufield et al., 2010a) and immunoassays for MMPs are also available (Lohmander et al., 2005).

A number of different approaches have been developed to identify protease substrates in biological samples. These include specific N-terminal tagging of proteins with affinity enrichment and LC liquid chromatography (LC) tandem mass spectrometry (MS/MS) detection (Timmer et al., 2007) an amine-targeted iTRAQ approach (Dean and Overall, 2007), engineered enzymes to selectively biotinylate free protein N-termini for positive enrichment of corresponding N-terminal peptides (Mahrus et al., 2008) and selective recovery of N-terminal peptides (McDonald et al., 2005). Gevaert *et al.* 2003 (Gevaert et al., 2003) developed a peptide isolation procedure based on diagonal electrophoresis and diagonal chromatography; combined fractional diagonal chromatography (COFRADIC) where free amino groups in proteins were first blocked by acetylation and then digested with trypsin. Whilst the process of Edman degradation and sequencing has traditionally been utilised to validate these methods (Edman and Begg, 1967) more recent developments which overcome problems of sequencing multiple closely spaced cleavage fragments on SDS-PAGE using amino-terminal oriented mass spectrometry of substrates (ATOMS) for the N-terminal identification of protein cleavage fragments in solution. In this study we used a novel, simplistic approach employing a number of software tools to mine MS derived data for potential cleavage sites, in order to establish 'proof of concept' for our methodologies. In this

study we define neopeptides as peptides identified by MS which have at least one non-tryptic cleavage site, that is the peptide has been identified without an arginine or lysine fragmented site. We hypothesise that such sites would have been formed by a protease other than trypsin.

Fragmentation patterns of cartilage ECM components were first investigated by western blotting (Fosang et al., 1995) and subsequently by mass spectrometry (Rousseau and Delmas, 2007). Mass spectrometry based proteomic approaches applied to the investigation of human articular cartilage have demonstrated novel potential substrates and cleavage sites for specific enzymes (Rousseau and Delmas, 2007) which may be helpful in the development of biomarkers and clarifying the biology of OA. Peptide ion identification has been significantly enhanced by both improvements in the quality of data acquired, through the use of mass spectrometers with high resolution and high mass/charge *(m/z*) accuracy, and in the data processing, through the use of probabilistic search algorithms such as Mascot. Here identifications are derived from a database search which queries a specific sequence database for the best peptide in order to explain the peaks in the MS/MS spectrum. However recent software developments such as PEAKS® now enable significantly improved sensitivity and accuracy in comparison to existing database searches alone through algorithms which enable sequencing *de novo* (Zhang et al., 2012a). These algorithms derive the peptide sequence directly from the MS/MS spectrum (Ma et al., 2003). Thus a workflow based on LC-MS/MS, de novo sequencing and database searching provides an accurate and convenient method to identify peptide products released by either specific proteases or following cytokine stimulation of equine articular cartilage.

In this study a proteomic analysis of MMP-3 or ADAMTS-4 driven hydrolysis of a crude proteoglycan extract was used to identify neopeptides from aggrecan, biglycan, decorin, fibromodulin plus COMP in equine articular cartilage. In addition we compared these results to IL-1β stimulated equine cartilage extract supernatants in order to identify common neopeptides. This analysis was performed to increase the knowledge of molecular events associated with cartilage

degradation characteristic of OA, identify peptides that may be useful as biomarkers of cartilage disease, as well as to identify novel neopeptides that could be further validated in targeted future studies. A series of conventional and novel proteolytic events have been identified as a result of either specific proteases or the effect of IL-1β. Improved knowledge of specific peptide fragments and the pathways generating these fragments will aid in the identification of markers of joint diseases such as OA.

## Methods

### Articular cartilage isolation and explant culture

Full thickness equine articular cartilage from grossly normal metacarpophalangeal joints of three horses was obtained from the abattoir. Cartilage was diced into explants approximately 2mm x 2mm following sterile dissection, mixed and washed twice with serum-free Dulbeco's Modified Eagles Medium (DMEM) (to deplete serum and synovial proteins). Explants were allowed to equilibrate in complete medium for 24h at 37$^{o}$C in 5% $CO_2$ in 12 well plates (2ml/well).

### Normal equine cartilage IL-1β treated explant studies

After overnight incubation as described explants were washed with serum-free DMEM and the media was then replaced with serum-free DMEM 'complete', supplemented were applicable with either human recombinant IL-1β (10ng/ml; R&D Systems) to induce cartilage degradation or DMSO (Il-1β diluent) as a control. The media (with and without IL-1β) was exchanged 48h after initiation of treatment, and cultures harvested after 96h. The 48h and 96h supernatant samples were pooled, following the addition of protease inhibitors (Complete protease Inhibitors, EDTA-free, Roche, Lewes, UK) and stored at -80$^{o}$C prior to downstream analysis, thus representing the total secretome over 96h.

### Proteoglycan-enriched fraction isolation from cartilage

Equine articular cartilage from the grossly normal metacarpophalangeal joint of three horses obtained from an abattoir was pooled and pulverised with a

dismembranator (Miko, S-Braun, USA). Proteins were extracted with cartilage extraction buffer containing 4M guanidinium chloride, 50mM sodium acetate and proteinase inhibitors (Complete Protease Inhibitors, EDTA-free, Roche, Lewes, UK), pH 6.0 using end-over-end mixing for 20h at 4$^{o}$C. After extraction the soluble fraction (proteoglycan-enriched) was removed following centrifugation for 15min at 13000**g** at 4$^{o}$C. Following dialysis in a 14,000-kD cut-off membrane (Spectrapor, Breda, NL) for 24h at 4$^{o}$C against 0.1M sodium acetate, pH 6.0 in the presence of proteinase inhibitors, the extract was clarified using centrifugation for 15min at 13000**g** at 4$^{o}$C. A crude proteoglycan-enriched extract was isolated by associative caesium chloride density gradient centrifugation (Lark et al., 1997), also in the presence of the proteinase inhibitors. The supernatant was fractionated in an associative caesium chloride (CsCl) density gradient (starting density 1.5g/ml) for 60h at 100,000**g** in an ultracentrifuge (Beckman 50Ti, Gallway, Ireland). The tube was fractionated into quarters; A1-A4 (A1-A2 fraction, identified in previous studies (Heinegard, 1977) as being enriched for aggrecan and small leucine rich proteoglycans was retained for protease digestion and the A1-A2 fraction dialyzed first against 0.1M sodium acetate for 48h at 4$^{o}$C and then against ultrapure water for 36h at 4$^{o}$C, both in the presence of proteinase inhibitors. The samples were then lyophilised. An aliquot of each fraction was assessed for protein content by measuring the protein absorbance at an optical density (OD) of 280nm using a Nanodrop ND-100 spectrophotometer (Labtech, East Sussex, UK). The density of each fraction of the gradient was measured by weighing an aliquot (xml) using a balance. To validate the A1-A2 fractions were enriched for proteoglycans glycosaminoglycan (GAG) analysis of the A1 to A4 fractions was undertaken using a 1,9-dimethyl-methylene blue (DMMB) dye binding microwell spectrophotometric assay. Samples were incubated in 250µl DMMB (16µg/ml 1,9-dimethyl-methylene blue, 2mg/ml sodium formate and 0.2% v/v formic acid pH 3.5) Shark chondroitin sulphate C was used to construct a standard curve (0 to 70µg/ml) and readings were taken at 570nm on a Multiskan EX photometric multiplate absorbance reader (Thermo Electron Corp, Vantaa, Finland). GAG was expressed as µg GAG/ ml crude proteoglycan extract.

**Proteinase digestion of proteoglycan-enriched fraction *in-vitro***

Following initial trial experiments using between 2h and 20h proteinase incubation periods and assessing resultant aggrecan fragments with immunoblotting, the optimum incubations were as stated. Aliquots of A1-A2 extract were digested in proteinase digestion buffer (50mM Tris HCl, 100mM NaCl, 10mM $CaCl_2$, pH 7.5) with either 0.05nmol human recombinant MMP-3 catalytic domain (Calbiochem, La Jolla, USA) for 20h at 37$^{o}$C or with 0.014nmol truncated human recombinant ADAMTS-4 (Calbiochem, La Jolla, USA) for 7h at 37$^{o}$C. Controls for each protease were incubated under the same conditions in the presence of the corresponding recombinant protein formulation buffer. The enzymatic digestion reactions (1ml) were stopped by addition of EDTA (18.3nmol). Molar calculations undertaken to calculate the amount of protease (MMP-3 or ADAMTS-4) to apply were based on a dry molecular mass for equine aggrecan assuming 1000kDa, for MMP-3 the molecular weight (MW) was 22kDa and for ADAMTS-4 was 42kDa.

**Validation of the use of human recombinant proteinases for equine cartilage proteoglycan-enriched fraction digestion**

For MMP-3 and ADAMTS-4 the recombinant human sequence was checked for sequence homology against the equivalent equine protein using NCBI Basic Local Alignment Search Tool; (Protein BLAST; http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). Furthermore, the sequence homology of the human peptides used to generate the neoepitope antibodies to known aggrecan cleavage sites, used in this experiment, were also verified against the equine sequence.

**Deglycosylation of proteoglycan-enriched fraction**

A1-A2 extracts before and after proteinase digestion were digested in deglycosylation buffer (50mM Tris-acetate, 50mM Na-acetate, 10mM EDTA, pH 7.6) for 2h at 37$^{o}$C with chondroitinase AC (0.01U/10µg glycosaminoglycan (GAG)), keratanase I (Seikagaku, Tokyo, Japan) (0.01U/10µg GAG) and keratanase II (Seikagaku, Tokyo, Japan) (0.0001U/10µg GAG). Chondroitinase AC digestion was

verified by measuring the UV absorbance at 232nm using a Nanodrop ND-100 spectrophotometer (Labtech, East Sussex, UK).

**Slot-blot analysis**

Deglycosylated samples of the proteoglycan-enriched fraction, equivalent to 5µg GAG were applied to the nitrocellulose membrane using vacuum manifold device. The membrane was then reduced in 6M guanidinium chloride reduction buffer pH 8 containing 10mM dithiothreitol (DTT) for 15min at room temperature. Following washing twice with 20mM Tris-buffered saline, 0.1% Tween-20, pH 7.5 (TBS-T) the membranes were blocked with 5% milk powder in TBS-T for 1h at room temperature. The membrane was probed overnight at 4$^{o}$C with the following antibodies in TBS-T containing 5% milk; mouse monoclonal to aggrecan ARGxx (BC-3) (Abcam, Cambridge, UK) (1:100 dilution), mouse monoclonal to aggrecan DIPEN (MD Bioproducts, Minneapolis, USA) (1:100 dilution) and rabbit polyclonal to aggrecan (Abcam, Cambridge, UK) (1:1000). The following secondary peroxidise conjugated antibodies were used; goat anti-mouse IgG and goat anti-rabbit IgG both at 1:1000 dilution.

**In-solution tryptic digestion**

Samples of cartilage supernatant from the explant experiments and crude proteoglycan extract from the protease digestion experiments were detergent treated with 1% (w/v) Rapigest (Waters, Manchester, UK) for 10min at 80$^{o}$C in 25mM ammonium bicarbonate. In-solution tryptic digestion of protein samples was carried out following sequential reduction and alkylation in 3mM DTT (60$^{o}$C for 10min) and then 9mM iodoacetamide (30min in the dark at room temperature) with trypsin at a ratio of 1:50 trypsin: protein ratio overnight at 37$^{o}$C. Crude proteoglycan extract samples were desalted and purified using $C_{18}$ resin in the form of a ZipTip® (Merck Millipore, USA). Detergent inactivation was then assumed by incubating for 45min at 37$^{o}$C with trifluoroacetic acid (VWR International) to a final concentration of 0.5% (v/v). Following centrifugation for ten minutes at 15000**g** the soluble phase was retrieved and used for LC-MS/MS.

**LC-MS/MS analysis of IL-1β treated cartilage explant media and proteinase digested cartilage proteoglycan-enriched fraction**

LC-MS/MS analysis was performed using NanoAcquity$^{TM}$ Ultraperformance LC (Waters, Manchester, UK) on line to an LTQ-Orbitrap Velos (Thermo-Fisher Scientific, Hemel Hempstead) via a ESI ion source containing a 10μm coated Pico-tip emitter (Presearch LTD, Basingstoke, UK). Aliquots of tryptic peptides equivalent to 250ng of protein were loaded onto a 180μm x 20mm $C_{18}$ trap column (Waters, Manchester, UK) at 5μl/min in 99% solvent A (water plus 0.1 % formic acid) and 1% solvent B (acetonitrile plus 1% formic acid for 5min and subsequently back-flushed onto a $C_{18}$ pre-equilibrated analytical column (75μm x 15mm Waters, Manchester, UK) using a flow rate of 0.3μl/min. Xcalibur 2.0 software (Thermo -Electron, Hemel Hempstead, UK) was used to operate the LTQ-Orbitrap Velos in data-dependant acquisition mode. The survey scan was acquired in the Orbitrap with a resolving power set to 30,000 (at 400 *m/z*). MS/MS spectra were concurrently acquired on the 20 most intense ions from the high resolution survey scan in the LTQ. Charge state filtering >1 was used, where unassigned precursor ions were not selected for fragmentation. Fragmentation parameters in the LTQ were: normalized collision energy; 30 volts, activation; 0.250, activation time; 10ms and minimum signal threshold 500 counts with isolation width 2*m/z*.

**Data analysis**

For neopeptide identification raw spectra were converted to Mascot generated files (mgf) using Proteome Discoverer software (Thermo, Hemel Hempstead, UK). The resulting mgf files were searched against the Unihorse database using an in-house Mascot (Perkins et al., 1999) server (Matrix Science, London, UK). Search parameters used were; enzyme; none, peptide mass tolerances 10ppm, fragment mass tolerance of 0.6Da, 1+, 2+ and 3+ ions, missed cleavages; 1, and instrument type ESI-TRAP. Modifications included were; fixed; carbamidomethyl cysteine and variable; oxidation of methionine.

For the proteins aggrecan, biglycan, cartilage oligomeric matrix protein, decorin and fibromodulin neopeptides that were present in treated samples exclusively were identified. The probability that a match was correct ($p<0.05$) was determined using the Mascot derived ion score where p was the probability that the observed match was a random event.

In addition raw data files were loaded into PEAKS® Studio 6.0 (Bioinformatics Solutions Inc., Waterloo, Canada) and de novo sequencing and protein identification performed. PEAKS® software employs multiple analytical algorithms and is able to identify the majority of peptides in the data, and can validate database searches using de novo sequencing results. Furthermore it achieves significantly better false discovery rate (FDR) curves than other database software (Zhang et al., 2012a). This means that more peptides can be identified with a lower or equivalent FDR. Estimate FDR function was used in order to create a 'decoy fusion' database based on the Ensembl database for horse (*Equus caballus*; EquCab2.56.pep, (ftp://ftp.ensembl.org/pub/current_fasta/equus_caballus/pep/). Data prerefinement was done by choosing peak centroiding, charge deconvolution, and deisotope options. The quality value was set greater than 0.65. The search parameters for the PEAKS® software included the parent peptide mass accuracy set at 20ppm and the MS/MS fragments 0.6Da mass tolerance. Oxidation was allowed as a variable modification of methionine, carbamidomethylation as a fixed modification of cysteine, and the enzyme specificity was none. The Unihorse database was searched. Results generated using PEAKS® Studio was manually curated against the Mascot search engine results. A 10lgP score of >20 was considered as significant (the score is -10 times the common log of the p value) (Zhang et al., 2012a).

## Results

### Production of a proteoglycan-enriched fraction by caesium chloride density gradient-ultracentrifugation

In this study we were interested in identifying potential cleavage sites in cartilage proteoglycans and COMP. Therefore a proteoglycan-enriched fraction, which was also found to be abundant in COMP, was extracted using the chaotropic agent guanidinium chloride and purified using caesium chloride density gradient centrifugation. Proteoglycans were found, as validated by the GAG assay results at the expected density of 1.3-1.5g/ml (Figure 1). Proteins were predominantly in the A4 fraction as determined by protein absorbance. The proteoglycan-rich A1-A2 fractions were then pooled for further work.

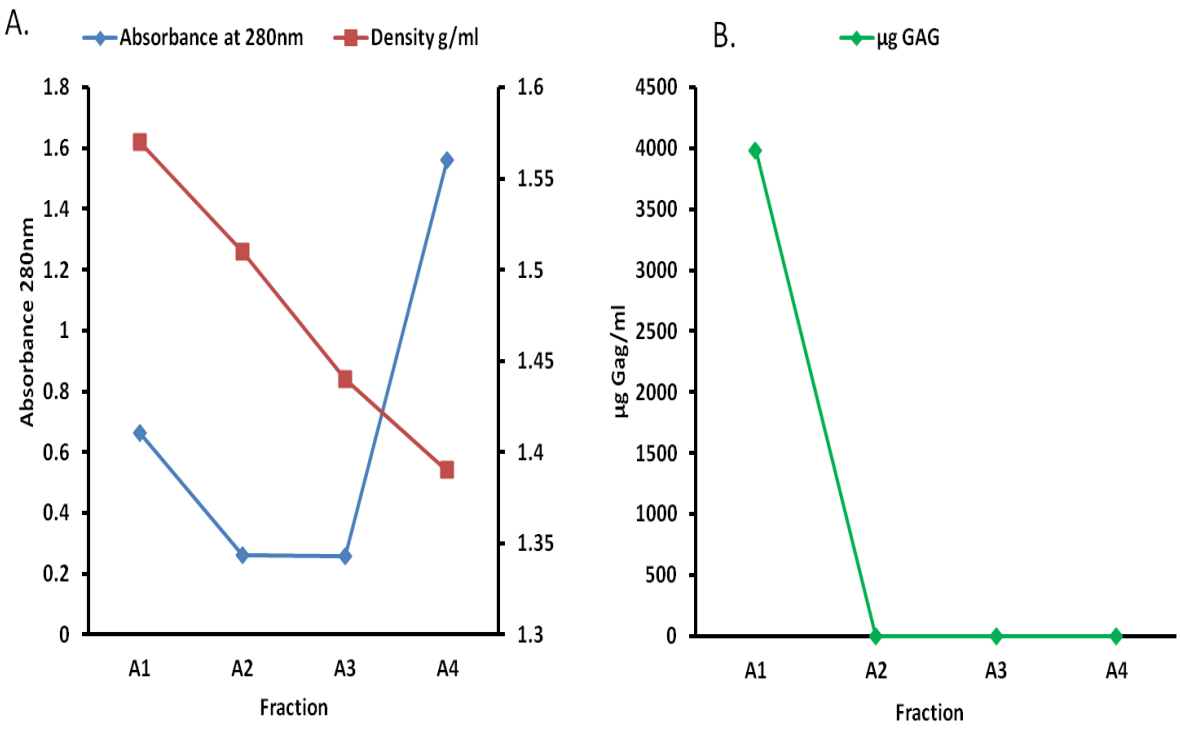



**Figure 1.Guanidinium chloride extraction and CsCl centrifugation produced a proteoglycan-enriched fraction of three pooled samples of equine cartilage.** Solid CsCl was added to extracted soluble cartilage proteins at a starting density of 1.5g/ml. After centrifugation the tubes were fractionated into 4 equal fractions and the density of each fraction measured. Fractions were then measured for A. protein absorbance at 280nm, and assayed to determine B. GAG concentrations of each fraction. Fractions A1 and A2 were then pooled for further work.

**Human recombinant proteinases were validated for the use in equine cartilage proteoglycan-enriched fraction digestion**

As equine recombinant proteins relevant to this study were not freely available human recombinant proteins were used. This approach was first validated by sequence homology studies. The BLAST tool predicted 89% sequence homology between equine and human for the catalytic domain of MMP-3 and 99.6% homology for ADAMTS-4 truncated recombinant protein. Subsequently immunoblotting was used to further confirm our approach. Proteoglycan-enriched fractions extracted from equine cartilage were analysed before and after MMP-3 or ADAMTS-4 digestion. Immunoblotting using anti-ARGxx antibody (ADAMTS-4 derived neoepitope), anti-DIPEN antibody (MMP-3 derived neoepitope) and anti-aggrecan revealed intact aggrecan as well as degradation products consistent with the activity of human recombinant proteinases on equine aggrecan (Figure 2). ADAMTS-4 proteolysis produced a product identified by the aggrecanase derived antibody ARGxx (Figure 2a) and MMP-3 digestion produced a product identified by the DIPEN antibody (Figure 2b).

**Identification of neopeptides formed following MMP-3 or ADAMTS-4 driven cartilage proteoglycan-enriched fraction degradation**

The A1-A2 proteoglycan-enriched fractions from a pool of three individual donors of normal equine articular cartilage were digested by MMP-3 and ADAMTS-4. Data analysis following LC-MS/MS analysis was undertaken initially using a Mascot-driven search of the *Equus caballus* database to identify newly released peptides. Then the raw data files were submitted to PEAKS® software to perform sequencing de novo. This was undertaken in order to confirm the correctness (defined as the probability that an amino-acid sequence derived from the analysis is correct) of the neopeptide identifications as it improves database search performance by combining de novo sequencing and database search peptides (Ma et al., 2003). Searching with the Mascot search engine against the Equus caballus database using a 'no enzyme' search identified 87 and 84 proteins with more than one unique peptide and a FDR of 1.4% and 1.5% for the ADAMTS-4 and MMP-3 digests

respectively. The major proteins identified in the proteoglycan-enriched extracts were primarily ECM proteins (Table 1).

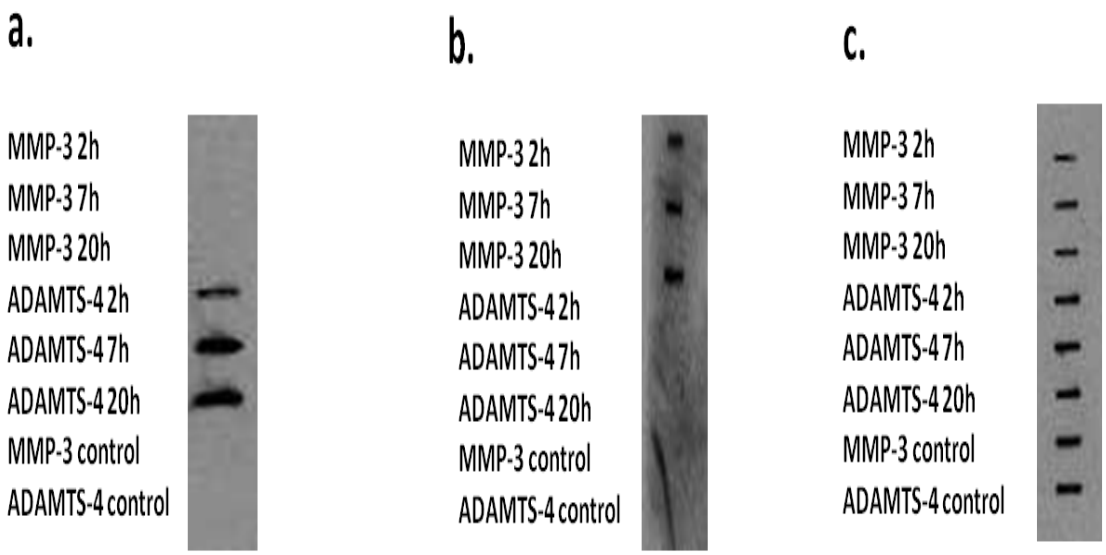



**Figure 2. A slotblot of MMP-3 and ADAMTS-4 digests of cartilage proteoglycan-enriched extracts using anti-neoepitope antibodies reveals human recombinant proteins are active against equine aggrecan.** Proteoglycan-enriched fractions were extracted from pooled equine cartilage of three donors and analyzed both before and after aggrecanase (ADAMTS-4) and MMP-3 digestion using the following antibodies; a. anti-ARGxx, b. anti-DIPEN and c. Anti-aggrecan. Prior to ADAMTS-4/ MMP-3 digestion, only bands identifying intact aggrecan were seen. Following ADAMTS-4 digestion at all incubation times ADAMTS-4 derived degradation products were evident. Similar results were identified following MMP-3 digestion.

The numbers of neopeptides that were generated and positively identified in both protease digests for aggrecan, COMP and fibromodulin but not biglycan or decorin were increased over the levels in control crude proteoglycan extracts incubated under identical conditions, but with the absence of either MMP-3 or ADAMTS-4 (Figure 3). A number of peptides were identified in both protease digested samples and controls Table 2.

| | ADAMTS4 | | | MMP-3 | | |
|---|---|---|---|---|---|---|
| Protein | Significant peptide matches | Significant peptide sequences | emPAI | Significant peptide matches | Significant peptide sequences | emPAI |
| biglycan precursor | 244 | 30 | 87.32 | 118 | 33 | 158.63 |
| decorin precursor | 43 | 21 | 13.29 | 37 | 20 | 16.07 |
| cartilage oligomeric matrix protein | 74 | 33 | 7.07 | 121 | 46 | 24.51 |
| lysozyme C-like isoform 1 | 10 | 7 | 6.56 | 11 | 6 | 6.56 |
| fibromodulin | 42 | 16 | 5.66 | 63 | 21 | 11.87 |
| prolargin-like | 23 | 13 | 4.11 | 23 | 17 | 8.83 |
| ribonuclease 4-like | 4 | 3 | 2.32 | 3 | 2 | 1.72 |
| c-type lectin domain family 3 member A-like | 3 | 3 | 1.56 | 6 | 4 | 0.87 |
| chondroadherin | 9 | 7 | 1.39 | 10 | 8 | 3.81 |
| sushi repeat-containing protein | 10 | 9 | 1.22 | 13 | 11 | 1.37 |
| aggrecan core protein | 89 | 33 | 0.97 | 138 | 49 | 1.93 |
| cartilage intermediate layer protein 1 | 17 | 16 | 0.71 | 21 | 17 | 1.11 |
| lactadherin | 4 | 3 | 0.47 | 9 | 7 | 1 |
| collagen alpha-1(II) chain | 15 | 8 | 0.34 | 8 | 5 | 0.24 |

**Table 1. Major cartilage proteins were identified in crude proteoglycan extract using LC-MS/MS.** The table lists the 14 most prominent proteins identified in this study, based on the emPAI (exponentially modified protein abundance index) (Ishihama et al., 2005) which approximates label-free relative quantification of proteins in a mixture based on protein coverage by peptide matches. Significant peptide matches and sequences are based on Mascot probability based scoring that a match is random set at $p<0.05$

Neopeptides were identified from aggrecan, biglycan, COMP and decorin, and equated to previously identified cleavage sites or novel ones for each protein investigated (appendix to manuscript 4). Results were from the pooled proteoglycan-rich extract from three donors analysed on the instrument in singlicate. Peptides included in the histogram were exclusively identified in the samples assigned and identified using Mascot. The control contained digestion buffer but no exogenous protease.

The number of unique neopeptides identified from each protein following protease digestion was based on 'no enzyme' Mascot searches. Table 3 indicates the number of unique peptides discovered with each protease digestion. In addition PEAKS® was used to interrogate this list of peptides in order to gain confidence in the peptide identifications by demonstrating the peptide-spectrum matching score;

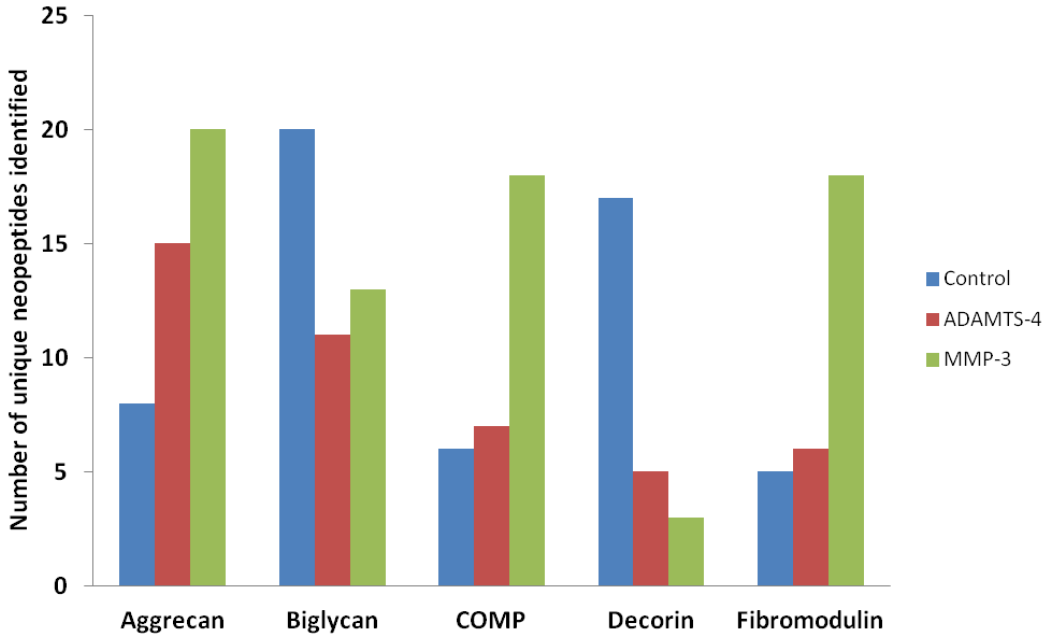



**Figure 3. Digestion of crude equine cartilage proteoglycan with MMP-3 or ADAMTS-4 produced a number of neopeptides.**

| Protein | Control | ADAMTS-4 | MMP-3 | Control + ADAMTS-4 | Control + MMP-3 |
|---|---|---|---|---|---|
| Aggrecan | 8 | 15 | 20 | 1 | 2 |
| Biglycan | 20 | 11 | 13 | 4 | 3 |
| COMP | 6 | 7 | 18 | 5 | 4 |
| Decorin | 17 | 5 | 3 | 5 | 4 |
| Fibromodulin | 5 | 6 | 18 | 10 | 8 |

**Table 2. A number of unique neopeptides were identified following LC-MS/MS.** A number of neopeptides were identified using Mascot following protease digestion with either ADAMTS-4 or MMP-3 or in control conditions of crude equine proteoglycan. In addition some peptides were identified in both the protease digested sample and under control conditions. The number of peptides in these classes is denoted by the columns named 'control +ADAMTS-4' or 'control+ MMP-3'.

PEAKS® Database -10lgP score. The most abundant released neopeptides were derived from aggrecan. Whilst proteases cleaved proteoglycans and COMP, ADAMTS-4 and MMP-3 had various preferences for ECM proteins; ADAMTS-4

generated the most numerous neopeptides from proteolysis of aggrecan and biglycan, whilst except for decorin MMP-3 produced similar numbers of neopeptides for all proteoglycans and COMP.

Although many peptides were generated in the crude proteoglycan digests exclusively by ADAMTS-4 or MMP-3, some neopeptides were generated in both ADAMTS-4 and MMP-3 digestion of crude aggrecan (Table 4a and 4b). The number of times each neopeptide was identified varied from once (for 76% and 74% in ADAMTS-4 and MMP-3 digestion respectively) to 33 times within a single experiment for the biglycan neopeptide $^{152}$NHLVEIPPNLPSS$^{164}$ following MMP-3 digestion. Selected release of some known aggrecan, biglycan and COMP epitopes as peptides was evident.

| | Treatment | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | ADAMTS-4 | | | MMP-3 | | | IL-1β | | |
| Protein | Total No. peptides | #Significant peptides p<0.05 | Significant 10lgP | Total No. peptides | #Significant peptides p<0.05 | Significant 10lgP | Total No. peptides | #Significant peptides p<0.05 | Significant 10lgP |
| Aggrecan | 15 | 1 | 5 | 20 | 5 | 12 | 2 | 0 | 0 |
| Biglycan | 11 | 4 | 8 | 13 | 3 | 10 | 4 | 1 | 1 |
| COMP | 7 | 5 | 5 | 18 | 7 | 12 | 52 | 42 | 45 |
| Decorin | 5 | 1 | 3 | 3 | 1 | 3 | 3 | 1 | 1 |
| Fibromodulin | 6 | 1 | 6 | 18 | 5 | 12 | 3 | 1 | 3 |

**Table 3. Degradation derived potential neopeptides were identified in crude proteoglycan extract and cytokine driven cartilage degradation using LC-MS/MS.** Table indicates unique peptides identified following either crude proteoglycan digestion using MMP-3 or ADAMTS-4 or IL-1β stimulation of cartilage explants. #Significant peptides were identified by Mascot with the probability that a match was correct (p<0.05) derived from the ion score. Significant peptides from PEAKS® analysis had a 10lgP of greater than 20. This equated to a p value of 0.01. For IL-1β stimulated samples the neopeptides were only counted once if identified in multiple samples.

**A**

| Protein | Amino acid preceding peptide residue | Peptide sequence | Amino acid residue after | Significant MASCOT identified peptides; p<0.05 | PEAKS derived 10lgP | Significant PEAKS 10lgP; >20 |
|---|---|---|---|---|---|---|
| Aggrecan | R | LATTGQLY# | L | * | 30.5 | * |
| | R | VSLPNYPAIPTDATLELQN# | L | | 45.09 | * |
| | R | LATTGQLYLAW# | Q | | 31.2 | * |
| | R | EDISGFPSGGE# | V | | 42.38 | * |
| | R | WSDGHSLQFEN# | L | | 43.99 | * |
| | D | WATEVPSASEKPSPSE | V | | No | |
| | Q | VSLPNYPAIPTDATLE | L | | No | |
| | F | YDAICYTGEDFVDIPENFFA | P | | No | |
| | G | LSGTSGRAD | A | | No | |
| | I | EAGEGPSGILE | P | | No | |
| | G | RGQPSEGSVSGL | S | | No | |
| | V | EGPSGILELSGAHSG | V | | No | |
| | R | HDLVSSAMSGSGE | W | | 5.08 | |
| | P | FSGTTSGIHDLVSSAMSG | E | | No | |
| | A | SGVEDLGGLPSGGEIHLEPTAPGV | E | | No | |
| Biglycan | K | NHLVEIPPNLPS# | S | * | 62.49 | * |
| | K | LLQVVYLH# | T | * | 42.42 | * |
| | N | GISLFNNPVPY# | W | * | 59.46 | * |
| | N | *CIEMGGNPLENSGFQPGAFDGLK | L | * | 15.1 | |
| | K | DLPETLNELHLDHN | K | | 30.8 | * |
| | F | NNPVPYWEVQPATF | R | | 30.46 | * |
| | A | IELEDLLR | Y | | 32.15 | * |
| | F | NNPVPYWEVQPA | T | | 28.58 | * |
| | M | CPFGCHC | H | | 16.31 | |
| | D | SLTPTFSAMCPF | G | | 25.7 | * |
| | K | DLPETLNE | L | | 12.59 | |
| COMP | R | AFQTVVLDPEGDAQIDPN# | W | * | 51.82 | * |
| | K | QVCTDIDECETGQHNCVPN# | A | * | 44.46 | * |
| | R | WVVLNQGMEIVQTM# | A | * | 56.86 | * |
| | K | CAPGSCFPGVAC# | S | * | 51.63 | * |
| | N | WFLQHRPQVGY# | N | * | 34.99 | * |
| | Q | NTVMECDACGMQP* | T | | 30.18 | * |
| | R | CEACPPGYSGPTHEGVGMAF* | I | | 20.81 | * |
| Decorin | R | IHEVLDLEPLGPVCPF# | R | * | 53.57 | * |
| | K | NLHALILVN | N | | 19.01 | |
| | K | ASYSGVSLFSNPVQYWEIQPS | T | | 55.77 | * |
| | K | YIQVVYLH | N | | 26.64 | * |
| | Q | MIVVELGTNPLKSSG | I | | No | |
| Fibromodulin | R | DCPQECDCPPNFPTAMY# | C | * | 53.23 | * |
| | R | LSHNSLTNN# | T | | 29.37 | * |
| | R | IPPVNTNLEN# | S | | 29.91 | * |
| | R | INEFSISSFC | Q | | 28.85 | * |
| | R | KVPDGLPSALEQLYLEHNNVY | G | | 28.47 | * |
| | R | ELHLDHN | F | | 29.91 | * |

B

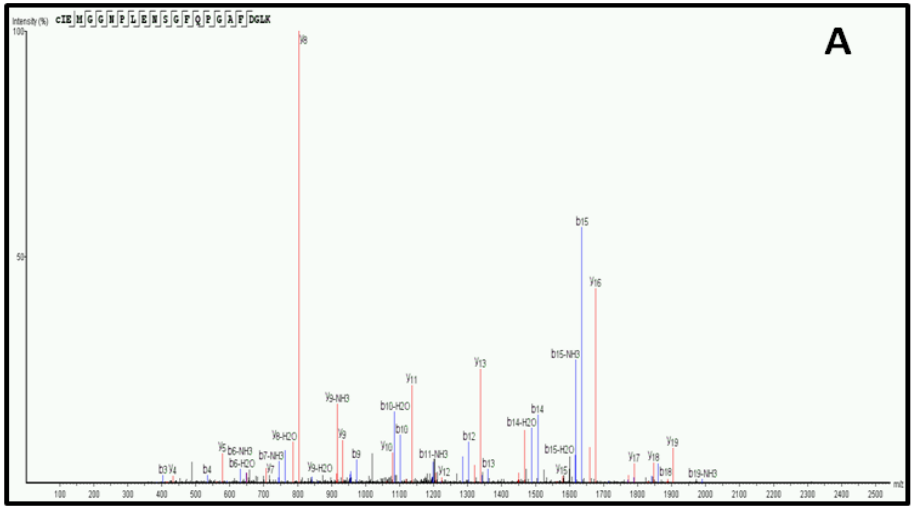
| Protein | Amino acid preceeding peptide residue | Peptide sequence | Amino acid residue after | Significant MASCOT identified peptides; p<0.05 | PEAKS derived 10lgP | Significant PEAKS 10lgP; >20 |
|---|---|---|---|---|---|---|
| Aggrecan | R | VSLPNYPAIPTDATLELQN# | L | * | 53.29 | * |
| | R | LATTGQLY# | L | | 30.54 | * |
| | R | VSLPNYPAIPTDATLE# | L | | 42.07 | * |
| | R | LATTGQLYLAW# | Q | | 29.71 | * |
| | V | EDISGFPSGGE# | V | | No | |
| | R | TYGVRPSSETYDVYC | Y | * | 30.34 | * |
| | R | ACLQNSAIIATPEQL | Q | * | 29.73 | * |
| | R | ITCTDPASY | K | * | 31.6 | * |
| | R | ITCTDPAS | Y | * | No | |
| | K | GTVACGDPPVVEH | A | | 51.9 | * |
| | Y | LAWQSGMDMC | S | | 33.57 | * |
| | R | WSDGHSLQFENWR | P | | No | |
| | R | YDAICYTGEDFVDIPEN* | F | | 53.88 | * |
| | Y | QLPFTCK | K | | 18.47 | |
| | N | SAIIATPEQL | Q | | 13.06 | |
| | R | PSSETYDVYC | Y | | 14.65 | |
| | R | TYGIRDT | N | | No | |
| | K | GEWNDVPCNY | Q | | 21.05 | * |
| | R | VSLPNYPAIPTD | A | | 27.95 | * |
| | M | EGLETSASGAEDLSGL | P | | No | |
| Biglycan | K | LLQVVYLH# | T | * | 28.1 | * |
| | K | NHLVEIPPNLPSS# | L | | 47.95 | * |
| | N | GISLFNNPVPY# | W | | No | |
| | K | VGVNDFCPVGFGV | K | * | 47.03 | * |
| | Y | NGISLFNNPVP | Y | * | 25.76 | * |
| | K | VGVNDFCPVG | F | | 33.22 | * |
| | K | DLPETLNE | L | | 31.84 | * |
| | R | NMNCIEMGGNPLE | N | | 40.5 | * |
| | K | GVFSGLRN* | M | | 28.45 | * |
| | K | NHLVEIPPNLPS | S | | 35.91 | * |
| | K | LLQVVYL | H | | 17.48 | |
| | N | GISLFNNPVPYW | E | | 26.35 | * |
| | P | YWEVQPA | T | | No | |
| COMP | N | WVVLNQGMEIVQTM# | N | * | 39.52 | * |
| | K | QVCTDIDECETGQHNCVPN# | S | * | 54 | * |
| | R | AFQTVVLDPEGDAQIDPN# | W | * | 55.8 | * |
| | Q | CAPGSCFPGVAC# | T | | 33.43 | * |
| | R | WFLQHRPQVGY# | I | | 35.6 | * |
| | R | NALWHTGDTASQ | V | * | 48.64 | * |
| | K | QVCTDIDECETGQHN | C | * | 51.84 | * |
| | K | QMEQTYWQANP | F | * | 48.6 | * |
| | R | AFQTVVLDPEGDAQIDPNWVVLNQGME | I | * | 48.67 | * |
| | R | NALWHTGDTAS* | Q | | 44.81 | * |
| | R | AFQTVVLDPEGDAQIDPNW | V | | 47.75 | * |
| | R | AFQTVVLDPEGDAQIDPNWVVLN | Q | | 57.32 | * |
| | N | WVVLNQGME | I | | 17.7 | |
| | R | AFQTVVLDPEGDAQIDPNWVVLNQGMEIVQTMN | S | | No | |
| | N | SVCINTQ | G | | No | |
| | C | FSQENIIWAN | L | | 9.77 | |
| | I | DPNWVVLNQGMEIVQTMNSDPGLAVGYTA | F | | No | |
| | N | WVVLNQGMEIVQTMN | S | | No | |
| Decorin | R | IHEVLDLEPLGPVCPF# | R | * | 56.99 | * |
| | K | AVFNGLNQ | M | | 26.53 | * |
| | K | ISPGAFTPLV | K | | 20.38 | * |
| Fibromodulin | R | DCPQECDCPPNFPTAMY# | C | * | 60.23 | * |
| | K | IPPVNTNLEN# | L | | 33.69 | * |
| | R | LSHNSLTNN# | G | | 34.82 | * |
| | R | INEFSISSFCTVVDVMN | F | * | 70.96 | * |
| | R | LSHNSLTNNGLAS | N | * | 41.92 | * |
| | R | KVPDGLPSALEQL | Y | * | 42.02 | * |
| | R | SLILLDLSYNH | L | * | 36.48 | * |
| | R | KVPDGLPSALE | Q | | 37.61 | * |
| | R | SAMPADAPLC | L | | 25.38 | * |
| | R | INEFSISSFC | T | | 29.88 | * |
| | K | YVYFQNNQIS | S | | No | |
| | A | LEQLYLEHNNVY | S | | 14.17 | |
| | K | YVYFQNNQISS | I | | No | |
| | R | KVPDGLPSA | L | | 24.32 | * |
| | R | LSHNSLTNNG | L | | 36.18 | * |
| | R | KVPDGLPS | A | | 19.55 | |
| | R | KVPDGLPSALEQLYLEHNNVYSVPDS | Y | | 13.47 | |
| | R | SAMPADAPLCL | R | | 16.32 | |

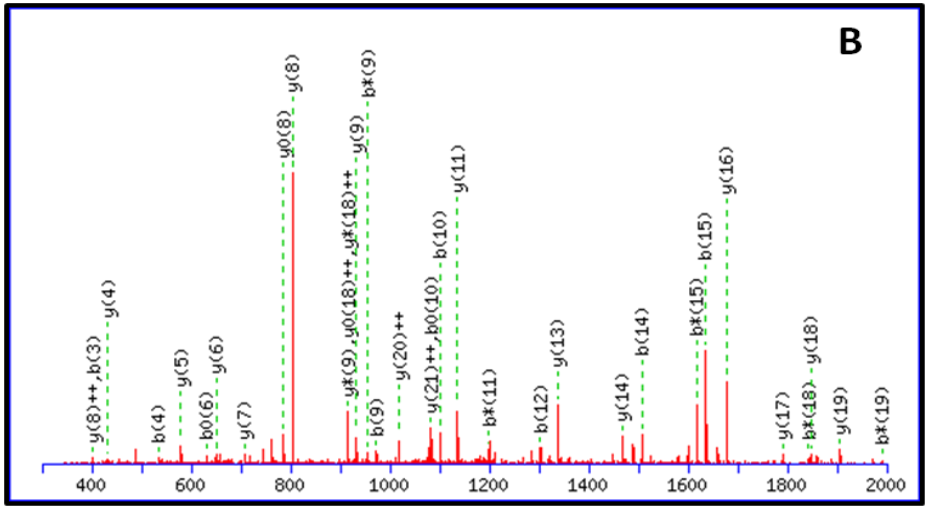
**Table 4. A number of neopeptides unique to either to A. ADAMTS-4 and B. MMP-3 digestion of crude proteoglycan were identified with LC-MS/MS using Mascot**. In the peptide sequence column* denotes neopeptides produced at known cleavages sites to specific proteases and # denotes neopeptide which have been identified following protease digestion with both ADAMTS-4 and MMP-3. 'No' in the 'PEAKS® derived 10lgP' column reflects peptides that were not identified following PEAKS® DB analysis.

Following ADAMTS-4 digestion of crude proteoglycan, the previously identified ADAMTS-4 cleavage site in biglycan $^{190}$N/$^{191}$C (Melching et al., 2006) was evident by the presence of the neopeptide $^{191}$CIEMGGNPLENSGFQPGAFDGLK identified following MS/MS analysis (Figure 4). Furthermore many novel neopeptides were also identified (Table 4a) with a significant probability that the peptide match was correct as determined by Mascot or/and PEAKS®.

In the MMP-3 digest a number of neopeptides representing both novel and known cleavage sites were identified with a significant probability that the peptide match was correct as determined by Mascot or/and PEAKS® including the aggrecan derived semi-tryptic neopeptide $^{344}$YDAICYTGEDFVDIPEN (Table 4B). This correlates to the MMP-derived DIPEN$^{341}$ neoepitope (Lark et al., 1995) (Figure 5). A further semi-tryptic neopeptide at a previously identified cleavage site (the protease that causes this cleavage is unknown (Holden, 2012)) in COMP $^{650}$NALWHTGDTAS was evident.

Interestingly for both these neopeptides PEAKS® analysis significantly identified this peptide whilst, although the Mascot algorithm identified the peptides, the ion scores were not significant. Finally following the identification of the sequence context of all neopeptides within its parent protein the COMP neopeptides produced in both ADAMTS-4 and MMP-3 digests R.AFQTVVLDPEGDAQIDPN.W and N.WVVLNQGMEIVQTM.A were adjacent to each other indicating that both peptides for a cleavage between Asn$^{548}$ – $^{549}$Trp were identified in this study.

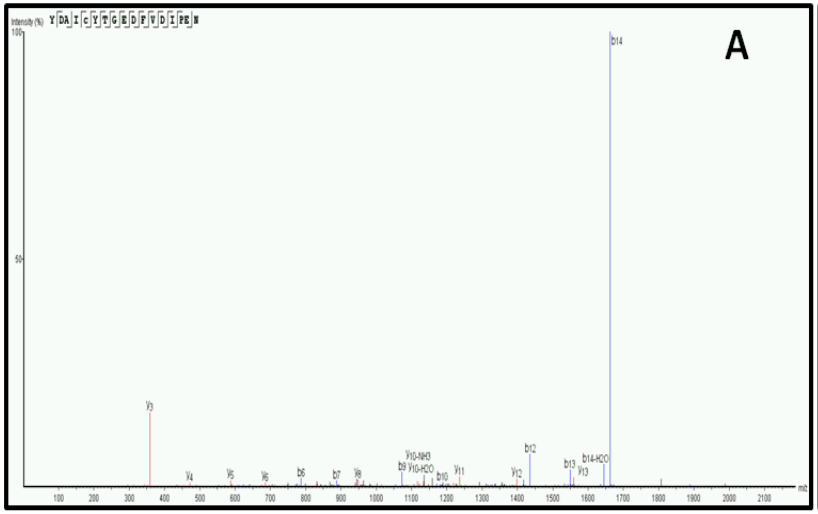


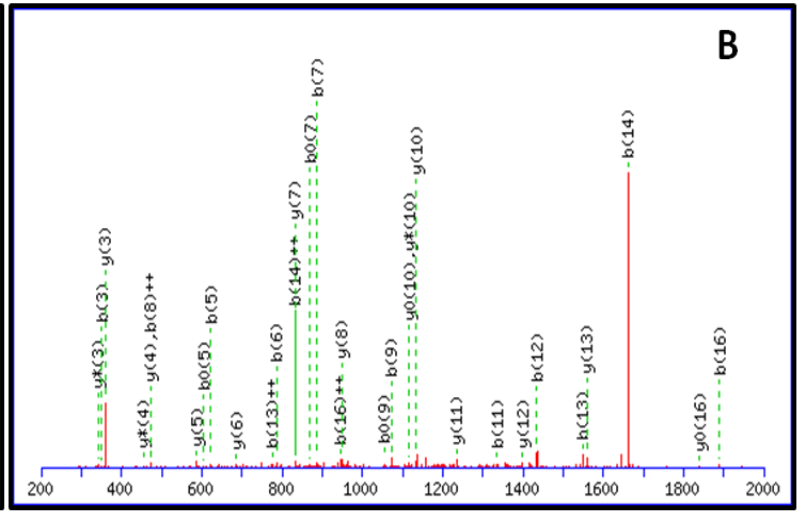




| # | Immoni... | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 133.04 | 161.04 | 143.03 | 144.01 | 81.02 | C(+57.02) | | | | | 23 |
| 2 | 86.10 | 274.12 | 256.11 | 257.10 | 137.56 | I | 2278.08 | 2260.07 | 2261.05 | 1139.54 | 22 |
| 3 | 102.06 | 403.27 | 385.20 | 386.21 | 202.08 | E | 2165.00 | 2146.99 | 2147.97 | 1083.00 | 21 |
| 4 | 104.05 | 534.22 | 516.20 | 517.18 | 267.60 | M | 2035.95 | 2017.94 | 2018.93 | 1018.48 | 20 |
| 5 | 30.03 | 591.37 | 573.22 | 574.10 | 296.11 | G | 1905.16 | 1886.92 | 1888.14 | 952.96 | 19 |
| 6 | 30.03 | 648.41 | 630.43 | 630.43 | 324.62 | G | 1847.92 | 1829.82 | 1830.99 | 924.45 | 18 |
| 7 | 87.06 | 762.43 | 744.36 | 745.39 | 381.65 | N | 1791.06 | 1772.93 | 1773.84 | 895.94 | 17 |
| 8 | 70.07 | 859.50 | 841.33 | 842.37 | 430.17 | P | 1676.85 | 1658.91 | 1659.80 | 838.91 | 16 |
| 9 | 86.10 | 972.55 | 953.70 | 955.55 | 486.71 | L | 1579.80 | 1561.86 | 1562.64 | 790.39 | 15 |
| 10 | 102.06 | 1101.49 | 1083.33 | 1084.44 | 551.24 | E | 1466.68 | 1448.70 | 1449.83 | 733.85 | 14 |
| 11 | 87.06 | 1215.51 | 1197.29 | 1198.44 | 608.26 | N | 1337.63 | 1319.67 | 1320.62 | 669.32 | 13 |
| 12 | 60.04 | 1302.61 | 1284.68 | 1285.66 | 651.77 | S | 1223.62 | 1205.69 | 1206.58 | 612.30 | 12 |
| 13 | 30.03 | 1359.67 | 1341.51 | 1342.54 | 680.28 | G | 1136.65 | 1118.61 | 1119.60 | 568.79 | 11 |
| 14 | 120.08 | 1506.71 | 1488.63 | 1489.61 | 753.82 | F | 1079.62 | 1061.54 | 1062.53 | 540.28 | 10 |
| 15 | 101.07 | 1634.68 | 1616.88 | 1617.66 | 817.85 | Q | 932.59 | 914.57 | 915.58 | 466.74 | 9 |
| 16 | 70.07 | 1731.75 | 1713.93 | 1713.93 | 866.37 | P | 804.54 | 786.51 | 787.40 | 402.71 | 8 |
| 17 | 30.03 | 1788.65 | 1770.76 | 1771.63 | 894.88 | G | 707.51 | 689.36 | 690.35 | 354.19 | 7 |
| 18 | 44.05 | 1859.91 | 1841.87 | 1842.78 | 930.40 | A | 650.48 | 632.34 | 633.32 | 325.68 | 6 |
| 19 | 120.08 | 2006.87 | 1988.99 | 1990.04 | 1003.94 | F | 579.35 | 561.31 | 562.29 | 290.16 | 5 |
| 20 | 88.04 | 2121.90 | 2103.89 | 2104.87 | 1061.45 | D | 432.38 | 414.25 | 415.28 | 216.62 | 4 |
| 21 | 30.03 | 2178.92 | 2160.91 | 2161.90 | 1089.96 | G | 317.22 | 299.21 | 300.19 | 159.11 | 3 |
| 22 | 86.10 | 2292.01 | 2274.00 | 2274.98 | 1146.50 | L | 260.20 | 242.19 | 243.17 | 130.60 | 2 |

| # | b | $b^{++}$ | b^ | $b^{*++}$ | $b^{0}$ | $b^{0++}$ | Seq. | y | $y^{++}$ | y^ | $y^{*++}$ | $y^{0}$ | $y^{0++}$ | # |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161.04 | 81.02 | | | | | C | | | | | | | 23 |
| 2 | 274.12 | 137.56 | | | | | I | 2278.08 | 1139.54 | 2261.05 | 1131.03 | 2260.07 | 1130.54 | 22 |
| 3 | 403.16 | 202.09 | | | 385.15 | 193.08 | E | 2165.00 | 1083.00 | 2147.97 | 1074.49 | 2146.99 | 1074.00 | 21 |
| 4 | 534.21 | 267.61 | | | 516.19 | 258.60 | M | 2035.95 | 1018.48 | 2018.93 | 1009.97 | 2017.94 | 1009.48 | 20 |
| 5 | 591.23 | 296.12 | | | 573.22 | 287.11 | G | 1904.91 | 952.96 | 1887.89 | 944.45 | 1886.90 | 943.96 | 19 |
| 6 | 648.25 | 324.63 | | | 630.24 | 315.62 | G | 1847.89 | 924.45 | 1830.87 | 915.94 | 1829.88 | 915.44 | 18 |
| 7 | 762.29 | 381.65 | 745.26 | 373.14 | 744.28 | 372.64 | N | 1790.87 | 895.94 | 1773.84 | 887.43 | 1772.86 | 886.93 | 17 |
| 8 | 859.34 | 430.18 | 842.32 | 421.66 | 841.33 | 421.17 | P | 1676.83 | 838.92 | 1659.80 | 830.40 | 1658.82 | 829.91 | 16 |
| 9 | 972.43 | 486.72 | 955.40 | 478.20 | 954.42 | 477.71 | L | 1579.78 | 790.39 | 1562.75 | 781.88 | 1561.76 | 781.39 | 15 |
| 10 | 1101.47 | 551.24 | 1084.44 | 542.73 | 1083.46 | 542.23 | E | 1466.69 | 733.85 | 1449.66 | 725.34 | 1448.68 | 724.84 | 14 |
| 11 | 1215.51 | 608.26 | 1198.49 | 599.75 | 1197.50 | 599.25 | N | 1337.65 | 669.33 | 1320.62 | 660.81 | 1319.64 | 660.32 | 13 |
| 12 | 1302.55 | 651.78 | 1285.52 | 643.26 | 1284.53 | 642.77 | S | 1223.61 | 612.31 | 1206.58 | 603.79 | 1205.59 | 603.30 | 12 |
| 13 | 1359.57 | 680.29 | 1342.54 | 671.77 | 1341.56 | 671.28 | G | 1136.57 | 568.79 | 1119.55 | 560.28 | 1118.56 | 559.79 | 11 |
| 14 | 1506.64 | 753.82 | 1489.61 | 745.31 | 1488.62 | 744.82 | F | 1079.55 | 540.28 | 1062.53 | 531.77 | 1061.54 | 531.27 | 10 |
| 15 | 1634.69 | 817.85 | 1617.67 | 809.34 | 1616.68 | 808.85 | Q | 932.48 | 466.75 | 915.46 | 458.23 | 914.47 | 457.74 | 9 |
| 16 | 1731.75 | 866.38 | 1714.72 | 857.86 | 1713.74 | 857.37 | P | 804.43 | 402.72 | 787.40 | 394.20 | 786.41 | 393.71 | 8 |
| 17 | 1788.77 | 894.89 | 1771.74 | 886.37 | 1770.76 | 885.88 | G | 707.37 | 354.19 | 690.35 | 345.68 | 689.36 | 345.18 | 7 |
| 18 | 1859.81 | 930.41 | 1842.78 | 921.89 | 1841.79 | 921.40 | A | 650.35 | 325.68 | 633.32 | 317.17 | 632.34 | 316.67 | 6 |
| 19 | 2006.87 | 1003.94 | 1989.85 | 995.43 | 1988.86 | 994.94 | F | 579.31 | 290.16 | 562.29 | 281.65 | 561.30 | 281.16 | 5 |
| 20 | 2121.90 | 1061.45 | 2104.87 | 1052.94 | 2103.89 | 1052.45 | D | 432.25 | 216.63 | 415.22 | 208.11 | 414.23 | 207.62 | 4 |
| 21 | 2178.92 | 1089.96 | 2161.90 | 1081.45 | 2160.91 | 1080.96 | G | 317.22 | 159.11 | 300.19 | 150.60 | | | 3 |
| 22 | 2292.01 | 1146.51 | 2274.98 | 1137.99 | 2274.00 | 1137.50 | L | 260.20 | 130.60 | 243.17 | 122.09 | | | 2 |
| 23 | | | | | | | K | 147.11 | 74.06 | 130.09 | 65.55 | | | 1 |

**Figure 4. Ion maps for the known ADAMTS-4 derived biglycan cleavage site $^{191}$CIEMG as demonstrated by A. PEAKS® and B. Mascot identify a series of b and y-ions. These charts illustrate the fragment ion sets used in identification using A. sequencing *de novo* and B. spectral library matching.** Top panel for each figure demonstrates the ion chart with intensity of the ions on the y-axis. In the PEAKS® ion match table blue indicates significant b-ions and red significant y-ions. Red signifies significant b and y ions in the Mascot derived ion match table.




| # | Immoni... | b | b-H2O | b-NH3 | b (2+) | Seq | y | y-H2O | y-NH3 | y (2+) | # |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 136.08 | 164.07 | 146.06 | 147.04 | 82.54 | Y | | | | | 17 |
| 2 | 88.04 | 279.10 | 261.09 | 262.07 | 140.05 | D | 1857.78 | 1839.77 | 1840.76 | 929.39 | 16 |
| 3 | 44.05 | 350.13 | 332.12 | 333.11 | 175.57 | A | 1742.76 | 1724.75 | 1725.73 | 871.88 | 15 |
| 4 | 86.10 | 463.22 | 445.21 | 446.19 | 232.11 | I | 1671.72 | 1653.71 | 1654.69 | 836.36 | 14 |
| 5 | 133.04 | 623.41 | 605.21 | 606.22 | 312.12 | C(+57.02) | 1558.88 | 1540.63 | 1541.68 | 779.82 | 13 |
| 6 | 136.08 | 786.38 | 768.30 | 769.29 | 393.66 | Y | 1398.55 | 1380.60 | 1381.58 | 699.80 | 12 |
| 7 | 74.06 | 887.48 | 869.35 | 870.33 | 444.18 | T | 1235.65 | 1217.53 | 1219.19 | 618.27 | 11 |
| 8 | 30.03 | 944.38 | 925.97 | 927.36 | 472.69 | G | 1134.60 | 1116.90 | 1116.90 | 567.75 | 10 |
| 9 | 102.06 | 1073.43 | 1055.51 | 1056.53 | 537.21 | E | 1078.07 | 1059.46 | 1060.76 | 539.24 | 9 |
| 10 | 88.04 | 1188.45 | 1171.15 | 1171.15 | 594.73 | D | 948.66 | 930.42 | 931.45 | 474.72 | 8 |
| 11 | 120.08 | 1335.80 | 1317.56 | 1318.49 | 668.26 | F | 833.49 | 815.46 | 816.38 | 417.20 | 7 |
| 12 | 72.08 | 1434.53 | 1416.69 | 1417.57 | 717.79 | V | 686.53 | 668.32 | 669.31 | 343.67 | 6 |
| 13 | 88.04 | 1549.68 | 1531.61 | 1532.90 | 775.31 | D | 587.33 | 569.26 | 570.31 | 294.13 | 5 |
| 14 | 86.10 | 1662.59 | 1644.70 | 1645.67 | 831.85 | I | 472.39 | 454.23 | 455.16 | 236.62 | 4 |
| 15 | 70.07 | 1759.75 | 1741.74 | 1742.73 | 880.38 | P | 359.19 | 341.15 | 342.26 | 180.08 | 3 |
| 16 | 102.06 | 1888.84 | 1870.78 | 1871.77 | 944.90 | E | 262.10 | 244.09 | 245.08 | 131.55 | 2 |
| 17 | 87.06 | | | | | N | 133.06 | 115.05 | 116.03 | 67.03 | 1 |

| # | b | $b^{++}$ | $b^{0}$ | $b^{0++}$ | Seq. | y | $y^{++}$ | $y^{*}$ | $y^{*++}$ | $y^{0}$ | $y^{0++}$ | # |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 164.07 | 82.54 | | | Y | | | | | | | 17 |
| 2 | 279.10 | 140.05 | 261.09 | 131.05 | D | 1857.78 | 929.40 | 1840.76 | 920.88 | 1839.77 | 920.39 | 16 |
| 3 | 350.13 | 175.57 | 332.12 | 166.57 | A | 1742.76 | 871.88 | 1725.73 | 863.37 | 1724.75 | 862.88 | 15 |
| 4 | 463.22 | 232.11 | 445.21 | 223.11 | I | 1671.72 | 836.36 | 1654.69 | 827.85 | 1653.71 | 827.36 | 14 |
| 5 | 623.25 | 312.13 | 605.24 | 303.12 | C | 1558.64 | 779.82 | 1541.61 | 771.31 | 1540.63 | 770.82 | 13 |
| 6 | 786.31 | 393.66 | 768.30 | 384.65 | Y | 1398.61 | 699.81 | 1381.58 | 691.29 | 1380.60 | 690.80 | 12 |
| 7 | 887.36 | 444.18 | 869.35 | 435.18 | T | 1235.54 | 618.27 | 1218.52 | 609.76 | 1217.53 | 609.27 | 11 |
| 8 | 944.38 | 472.69 | 926.37 | 463.69 | G | 1134.49 | 567.75 | 1117.47 | 559.24 | 1116.48 | 558.75 | 10 |
| 9 | 1073.42 | 537.22 | 1055.41 | 528.21 | E | 1077.47 | 539.24 | 1060.45 | 530.73 | 1059.46 | 530.24 | 9 |
| 10 | 1188.45 | 594.73 | 1170.44 | 585.72 | D | 948.43 | 474.72 | 931.40 | 466.21 | 930.42 | 465.71 | 8 |
| 11 | 1335.52 | 668.26 | 1317.51 | 659.26 | F | 833.40 | 417.21 | 816.38 | 408.69 | 815.39 | 408.20 | 7 |
| 12 | 1434.59 | 717.80 | 1416.58 | 708.79 | V | 686.34 | 343.67 | 669.31 | 335.16 | 668.32 | 334.67 | 6 |
| 13 | 1549.62 | 775.31 | 1531.60 | 766.31 | D | 587.27 | 294.14 | 570.24 | 285.62 | 569.26 | 285.13 | 5 |
| 14 | 1662.70 | 831.85 | 1644.69 | 822.85 | I | 472.24 | 236.62 | 455.21 | 228.11 | 454.23 | 227.62 | 4 |
| 15 | 1759.75 | 880.38 | 1741.74 | 871.37 | P | 359.16 | 180.08 | 342.13 | 171.57 | 341.15 | 171.08 | 3 |
| 16 | 1888.79 | 944.90 | 1870.78 | 935.90 | E | 262.10 | 131.56 | 245.08 | 123.04 | 244.09 | 122.55 | 2 |
| 17 | | | | | N | 133.06 | 67.03 | 116.03 | 58.52 | | | 1 |

**Figure 5. Ion maps for the MMP derived aggrecan cleavage site DIPEN[341] as demonstrated by A. PEAKS® and B. Mascot identify a series of b and y-ions. These charts illustrate the fragment ion sets used for identification of this peptide.** Top panel for each figure demonstrates the ion chart with intensity of the ions on the y-axis. In the PEAKS® ion match table blue indicates significant b-ions and red significant y-ions. Red signifies significant b and y ions in the Mascot derived ion match table.

**Identification of neopeptides following IL-1β stimulation of articular cartilage explants**

In order to identify neopeptides produced in more pathologically relevant conditions media was collected from three donors following IL-1β stimulation of normal cartilage explants. Each donor was analysed separately. The peptides were identified using LC-MS/MS methods. A no enzyme search with the Mascot search engine against the Unihorse database identified between 210 and 245 proteins and had a FDR of between 1.4% and 1.9%. Table 5 lists the top 20 major proteins identified in the media following IL-1β stimulation of cartilage explants based on protein score and emPAI generated from three donors in individual experiments. The ECM proteins COMP, decorin and fibromodulin were in the top 20. The proteins aggrecan and biglycan were also identified in all samples.

| Protein | Mean EmPAI |
|---|---|
| cartilage oligomeric matrix protein | 2632.56 |
| thrombospondin-1 | 3.05 |
| lysozyme C-like isoform 1 | 2.87 |
| pyruvate kinase isozymes M1/M2 isoform M2 | 1.78 |
| decorin precursor | 1.69 |
| alpha-enolase isoform 1 | 1.67 |
| fibronectin | 1.46 |
| phosphatidylethanolamine-binding protein 1 | 0.94 |
| stromelysin-1 precursor | 0.87 |
| thrombospondin-4-like | 0.82 |
| fibromodulin | 0.81 |
| cartilage intermediate layer protein 1 | 0.72 |
| triosephosphate isomerase | 0.66 |
| betaine--homocysteine S-methyltransferase | 0.59 |
| interstitial collagenase precursor | 0.51 |
| beta-defensin-1 | 0.49 |
| actin, cytoplasmic 1 | 0.49 |
| clusterin precursor | 0.47 |
| chondroadherin | 0.46 |

**Table 5. Predominant proteins identified in the media of cartilage explants treated with IL-1β.** The list of the top 20 proteins identified in the media of three cartilage explant donors following treatment with IL-1β based on the mean protein score and mean emPAI.

Next a number of unique neopeptides from each protein of interest, generated following IL-1β explant treatment were identified (Table 6). The most abundant number of neopeptides was derived from COMP.

| Protein | Donor | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Aggrecan | 9 | 4 | 7 |
| Biglycan | 2 | 4 | 0 |
| COMP | 21 | 57 | 25 |
| Decorin | 2 | 1 | 5 |
| Fibromodulin | 4 | 1 | 3 |

**Table 6. The number of unique neopeptides derived from each donor for each protein of interest varied.**

Neopeptides derived from the fragmentation of aggrecan, biglycan, COMP, decorin and fibromodulin were identified. Table 7 indicates neopeptides that were identified in at least two samples.

| Protein | Amino acid preceding peptide residue | Peptide sequence | Amino acid residue after | Generated in ADAMTS-4 digest | Generated in MMP-3 digest | Significant MASCOT identified peptides; p<0.05 | PEAKS derived 10lgP | Significant PEAKS 10lgP; >20 |
|---|---|---|---|---|---|---|---|---|
| Aggrecan | Q | EAGEGPSGILE | L | ✓ | | | No | |
| | V | PIETELPSPGEPSG | V | | | | No | |
| Biglycan | R | LAIQFGNY | K | | | * | 36.23 | * |
| | P | YWEVQPATFRCVTDRLAIQ | F | | | | No | |
| COMP | R | DVDHDFVGDACDSDQDKDGD | G | | | * | 44.09 | * |
| | R | DVDHDFVGDACDSDQD | K | | | * | 48.29 | * |
| | R | CEACPPGYSGPTHEGVGM | A | | | * | 60.66 | * |
| | R | CEACPPGYSGPTHEGVGMAF* | A | ✓ | | * | 61.29 | * |
| | R | FCPDGTPSPCHEKAD | C | | | * | 41.02 | * |
| | K | NTVMECDACGMQP* | A | ✓ | | * | 33.99 | * |
| | R | AFQTVVLDPEGDAQIDPNWVVLNQGMEIVQTMN | S | | ✓ | | 14.25 | |
| | R | SQKNDDQKDTDQDGRGD | A | | | | 39.81 | * |
| | K | QMEQTYWQA | N | | | * | 30.28 | * |
| | R | LVPNPGQEDADRDGVGD | V | | | | 56.21 | * |
| | R | LGVFCFSQENIIWAN | L | | ✓ | | 15.3 | |
| | A | QCAPGSCFPGVAC | T | ✓ | ✓ | | 33.53 | * |
| | R | NALWHTGD | T | | | | 22.78 | * |
| | F | CFSQENIIW | A | | | | 10.63 | |
| | R | SCVCAVGWAGNGL | L | | | | No | |
| | R | NAVDNCPRVPNSDQKDSDG | D | | | | 12.55 | |
| | K | QVCTDIDECETGQHN | C | | ✓ | * | 53.39 | * |
| | R | SCVCAVGWAGNGLLC | G | | | * | 55.88 | * |
| | L | AQCAPGSCFPGVACTQ | T | ✓ | ✓ | * | 42.69 | * |
| | R | KDNCVTVPNSGQEDADRDGIGDACDTD | A | | | * | 50.66 | * |
| | R | KDNCVTVPNSGQEDADRDGIGD | A | | | * | 46.57 | * |
| | R | VSVRPLAQCAPGSCFPGVA | C | | | * | 58.46 | * |
| | R | VSVRPLAQCAPGS | C | | | * | 45.6 | * |
| | Q | CAPGSCFPGVACTQTA | S | | | * | 33.02 | * |
| | R | LGVFCFSQEN | I | | | * | 44.53 | * |
| | R | VSVRPLAQ | C | | | * | 36.83 | * |
| | R | VSVRPLAQC | A | | | * | 43.6 | * |
| | R | AFQTVVLDPEGDAQIDPNWVVLNQ | G | | | * | 59.72 | * |
| | Q | CAPGSCFPGVAC | T | | | * | 42.97 | * |
| | Q | CAPGSCFPGVACTQ | T | | | * | 42.64 | * |
| | K | QVCTDIDECETGQHNCVPN | S | | ✓ | * | 46.28 | * |
| | R | VPNSDQKDSDGD | G | | | * | 42.84 | * |
| | Q | CAPGSCFPGVACT | Q | | | * | 25.94 | * |
| | R | VSVRPLAQCAPGSCFPGVACTQT | A | | | * | 47.6 | * |
| | R | LGVFCFSQE | N | | | * | 23.25 | * |
| | R | KDNCVTVPNSGQED | A | | | * | 40.42 | * |
| | R | SCVCAVGW | A | | | * | 29.85 | * |
| | R | SCVCAVGWAGNGLL | C | | | * | 43.46 | * |
| | R | NALWHTGDTASQ | V | | ✓ | * | 43.02 | * |
| | R | SCVCAVGWAGN | G | | | * | No | |
| | C | AVGWAGNGLLC | G | | | * | 31.5 | * |
| | R | NAVDNCPRVPNSD | Q | | | * | 36.83 | * |
| | R | FYEGPELVADSNVVLD | T | | | * | 38.78 | * |
| | Q | CAPGSCFPGVA | C | | | * | 26.02 | * |
| | R | LVPNPGQEDADRD | G | | | * | 24.18 | * |
| | M | ECDACGMQPAR | T | | | * | 39.5 | * |
| | R | DVDHDFVGDACD | S | | | * | 33.99 | * |
| | R | NTDGDKWGD | A | | | * | 28.11 | * |
| | R | NAVDNCPRVPNSDQKD | S | | | * | 14.29 | |
| | R | VSVRPLAQCAPG | S | | | * | 28.65 | * |
| | R | VSVRPLAQCAPGSCFPGVAC | T | | | | 71.97 | * |
| | A | VGWAGNGLLCGR | D | | | * | 55.97 | * |
| Decorin | K | KASYSGVSLF | S | | | | 31.2 | * |
| | P | PGLPPSLTEL | H | | | * | No | |
| | Q | MIVVELGTNPLKSSG | I | ✓ | | | No | |
| Fibromodulin | R | KVPDGLPSALEQL | Y | | ✓ | | 38.16 | * |
| | R | INEFSISSFCTVVDVM | N | | | * | 53.18 | * |
| | R | KVPDGLPSALEQLYLEHNNVY | S | ✓ | ✓ | | 21.72 | * |

**Table 7. The neopeptides released from cartilage explants following IL-1β treatment.** Table indicates neopeptides identified in at least 2 donor samples. Neopeptides also evident in the ADAMTS-4 and MMP-3 experiments are also indicated with a tick. * indicates neopeptides produced at known cleavages sites to specific proteases. Peptides with a significant ($p<0.05$) match that the sequence was correct based on Mascot ion scores are indicated with a black asterisk. 'No' in the 'PEAKS® derived 10lgP' column reflects peptides that were not identified following PEAKS® DB analysis.

**Discussion**

In OA, the degradation of cartilage is characterised by a loss of ECM caused by secreted proteases; principally matrix metalloproteases (MMPs) (Rivers et al., 2011). The relative over expression of these MMPs in OA appears to be due to a number of factors including mechanical loading (Lucchinetti et al., 2002) and inflammatory cytokines (Martel-Pelletier, 2004). This study was instigated to characterise the peptide products after proteolysis of aggrecan, biglycan, COMP, decorin and fibromodulin using a crude proteoglycan extract following the addition of exogenous MMPs recognised as being upregulated in OA disease progression. These peptides were then compared to those produced by a recognised *in-vitro* model of early OA, using cultured cartilage explants driven by the cytokine IL-1β. This allowed cleavage sites produced under more pathologically relevant conditions to be established and compared with proteoglycan or COMP digests. The proteomics methods utilized in this study enabled 34 aggrecan, 23 biglycan, 23 COMP, 8 decorin and 24 fibromodulin potential novel cleavage sites to be determined from the crude proteoglycan extract by the generation of 'neopeptides', which are non tryptic peptide products produced from either specific MMPs. In addition one aggrecan, one biglycan and two COMP known cleavage sites were identified. Finally a comparison was made between these results and the effect of a cytokine driven production of a cascade of many MMPs on cartilage explants.

The study was interested primarily in identifying peptide products from the major proteoglycans in cartilage. Therefore rather than digesting whole cartilage, a crude proteoglycan extract was prepared. Classic methodologies, previously used to identify cartilage fragments in aggrecan (Lark et al., 1997) were used in order to remove the majority of collagens (as type II collagen is the most abundant protein in cartilage) and other proteins, thus reducing sample complexity for up-stream proteomics workflows. The CsCl A1-A2 fractions containing proteoglycans, including aggrecan as determined by immunoblotting, were then retained for proteoglycan digestion studies. Subsequent proteomics studies demonstrated that the abundant

proteins in the crude proteoglycan extract were the ECM proteins in which we were interested; aggrecan, biglycan, decorin, fibromodulin plus COMP an important non-collagenous ECM protein. Three of these; COMP, decorin and fibromodulin were the most abundant proteins as demonstrated by emPAI after digestion. This was expected as emPAI determines relative abundance of proteins based on the number of peptides identified per protein (Ishihama et al., 2005). The number of unique peptides identified in a protein will depend on factors which can be placed broadly in two categories; whether the peptide is detected and whether the peptide is identified. The latter will be dependent on the database used to search and the fragmentation of the peptide as these spectra are used for searching the spectral libraries. Furthermore within a search peptide identification depends on which PTMs are put in the database search terms. Peptide 'detectability is dependent on amino acid composition and associated physicochemical properties in addition to the separation method in the upstream LC, ionisation and the instrument used for ion detection (Li et al., 2010). Aggrecan was identified as only the eleventh most abundant protein but this is probably due to it being heavily glycosylated resulting in fewer non-modified peptides for identification by mass spectrometry (MS). Post translational (PTM) modifications such as GAGs alter the ion masses used in database searches for peptide identification. Even following treatment with specific enzymes such as chondroitinase ABC to remove specific sugar side chains, short sugar stubs remain attached making identification of the peptide backbone challenging. Though the globular domains contain far fewer PTMs, and therefore peptides which are more likely to be identified in this region, this is a small proportion of the whole protein and the large size of the aggrecan would lead to a low emPAI score.

The use of equine samples in this study were instituted as the horse is an athletic animal and is considered an excellent animal model for human joint diseases, as it enables one to overcome the limitations in joint size and workload that are typical of other small animal models. The activity of the human recombinant proteases against equine tissue was assessed using immunoblotting with neoepitope antibodies of previously identified cleavage sites. When the crude proteoglycan was

digested with ADAMTS-4 ARGXX cleavage sites were evident but no DIPEN cleavages sites (which are derived from MMP activity). In contrast when digestion was undertaken with MMP-3 DIPEN but not ARGXX cleavage sites were seen. It appeared that the human recombinant enzymes were active in equine cartilage digestion.

The concentrations of MMPs used here were derived from previous studies (Zhen et al., 2008) and were used in order to provide adequate peptide generation in the timeframe of the experiment. The time frame of the final study was determined following preliminary work which indicated maximal digestion, as measured by immunoblotting with aggrecan neoepitope antibodies, as 7h and 20h for ADAMTS-4 and MMP-3 respectively. For the IL-1β treated cartilage explant study the concentration of IL-1β was based on many previous works (Chevalier et al., 1996; Ju et al., 2010). Media was collected after 96h based on our previous work on the equine cartilage secretome (manuscript 1).

The proteomics method employed in this study relied upon the search engine Mascot to identify peptides following a no enzyme search. Each peptide is given an ion score and based on a threshold, Mascot then determines how likely it is that the sequence given is true. This is accomplished by the number of fragment ions, both b and y, matched to the given peptide. Subsequently we used a software package; PEAK S® which contains algorithms enabling a number of features, including *de novo* sequencing and PTM identification, following tandem MS (Ma et al., 2003). The *de novo* sequencing feature improves the database search based peptide identifications made by correlating the sequencing de novo sequence and the database sequence, confirming the 'correctness' of neopeptide identifications (were a peptide is defined as 'correct' when it was the best match to the database out of all possible peptides). Peptide identification performance is influenced by accuracy; which can be measured by FDR and sensitivity; measured by the number of peptide spectrum matches and hence peptides. As the aim of our approach was to provide a 'first pass' list of peptides for further investigation cleaved by the ADAMTS-4, MMP-3 or induced following IL-1β stimulation, it was important that peptide identification results were statistically validated to avoid false positives.

Therefore, in the PEAKS® analysis, we employed the decoy-fusion method which joins the decoy and target sequences of the same protein together as a 'fused' sequence (Zhang et al., 2012a). This recently published validation method, overcomes problems previously recognised such as over-confidence, in order to more accurately separate true and false identifications (Brosch et al., 2009). Interestingly the analysis with PEAKS® provided more neopeptides with significant scores than the Mascot driven database search as the similarity between the *de novo* sequences identified in PEAKS® and the database peptide is used in PEAKS® DB's scoring function. A recent study using two data sets to evaluate MASCOT and PEAKS® DB demonstrated the number of peptide spectrum matches identified by PEAKS® DB and Mascot at 1% FDR increased by 42% (Zhang et al., 2012a).

Although protease inhibitors were present throughout the study it is possible and probable that some of the peptides identified were cleaved by endogenous proteases. This is especially probable where peptides are seen with successive amino acids removed such as the exoproteolytic fraying of ITCTDPASY to ITCTDPAS produced following ADAMTS-4 digestion of aggrecan. Most peptides produced by endogenous proteases were eliminated from the data analysis as peptides identified in both control and protease treated conditions were automatically removed. The credibility of our approach was established by a three findings. Firstly the neopeptide DIPEN cleavage site was identified with immunoblotting and a corresponding neopeptide $^{344}$YDAICYTGEDFVDIPEN was demonstrated using LC-MS/MS. Secondly a number of previously identified cleavage sites were evident in our data; two for MMP-3 treatment and three following ADAMTS-4 treatment including the ADAMTS-4 cleavage in biglycan at Asp$^{190}$-$^{191}$Cys (Melching et al., 2006) as suggested by the presence of the neopeptide $^{191}$CIEMGGNPLENSGFQPGAFDGLK in ADAMTS-4 digestion of crude proteoglycan. Finally for COMP following both ADAMTS-4 and MMP-3 digestion we were able to identify peptides on each side of a cleavage site at Asn$^{548}$-$^{549}$Try. It is highly improbable that this would have occurred had there not been a cleavage between these amino acids. In addition these peptides were not evident in control digests. These points would suggest that our approach is valid. Interestingly a previously identified cleavage site in COMP at

$Arg^{649}$-$^{650}Asp$ (Holden, 2012) was identified in the study. However it cannot be stated for certain that this peptide was due to cleavage of the 'unknown protease' since this neopeptide could also have been produced by tryptic cleavage alone since trypsin also cleaves following arginine.

The profile of neopeptides produced was greatest in the media of IL-1β stimulated explants. This may be due to the simpler digest obtained from a media compared to a proteoglycan digest, thus reducing sample complexity and increasing the number of peptides identified, especially medium to low abundance ones. It could also be due to the non-specific nature of the MMPs up-regulated following Il-1β stimulation (Shinmei et al., 1991). Of the two protease digestions MMP-3 was the most active in generating neopeptides from cartilage in agreement with other studies (Zhen et al., 2008). Interestingly both proteases cleaved each proteoglycan assessed within the extracts again in agreement with a study identifying MMP cleavage products in human articular cartilage (Zhen et al., 2008). Of the proteins investigated COMP had the most neopeptides identified following Il-1β stimulation of cartilage explants. This could be due to a number of factors. COMP, a pentomeric protein (each pentomer is 757 amino acids in length) has a relatively even distribution of arginines and lysines within its sequence making it an ideal for tryptic digestion producing tryptic (or semi-tryptic in the case of neopeptides) masses of optimum mass range for detection. In addition it has only two N-linked glycosylations as PTM. From previous studies (manuscript 2) in molar concentration terms it is the most abundant protein. Finally COMP has been investigated as a biomarker of OA in many studies due to its presence in serum and urine (Tseng et al., 2009) and it is hypothesised that in arthritis it must be subject to extensive degradation by proteases thus accounting for the large number of neopeptides in this study.

As a no enzyme search was undertaken on the data this allows tryptic, semi-tryptic and none-tryptic peptides to be identified. The level of confidence in peptide identifications was then tested by Mascot and PEAKS® giving a significance value. The neopeptides that were identified pertaining to known cleavages were all semi-tryptic peptides. That is, they were fragmented by trypsin at only one end of the

peptide. Trypsin is used in MS experiments as it cleaves C-terminal to arginine and lysine resulting in peptides in the preferred mass range for successful fragmentation by tandem MS (MS/MS). In addition it places a highly basic residue at the C-termini and thus is eminently suitable for positive ionisation mass spectrometric analysis. These characteristics aid 'flyability' (how easy it is to observe a peptide) in MS experiments, resulting in informative high mass y-ion series making MS/MS spectra easier to interpret. It is most likely that cleavages in proteins would produce semi-tryptic neopeptides, and these would more easily be identified using MS than none tryptic peptides.

Neopeptides were given significance scores which gives a probability that the sequence established by Mascot or PEAKS® DB was correct and the peptide mass and subsequent fragmentation pattern were not due to another peptide; however this is only a probability and so some neopeptides with a non-significant score could be a true peptide and vice versa.

Significantly more neopeptides were identified in aggrecan than in similar studies (Zhen et al., 2008). Aggrecan is the first matrix component to undergo measurable loss in the progression of OA (Arner et al., 1999), which is principally attributed to ADAMTS-4 and ADAMTS-5 cleavage (Tortorella and Malfait, 2008). In this study only a minority of aggrecan neopeptides were identified in the extended G2 to G3 domain which contains the region for GAG attachments. The reasons for this are two-fold. The extensive PTMs in this region from glycosylations pose a number of challenges to the traditional database search approach. The Mascot algorithm uses a limited number of available PTMs in its searches. Furthermore even in the extensive PTM list in the PEAKS® software GAGs do not feature. Secondly there is a scarcity of arginine and lysine residues in the interglobular domain leading to large tryptic peptides whose masses even without PTMs are often too large to allow identification. Tryptic digestion would produce peptides whose charged forms, especially the lower charged ones, may not be within the mass range of the mass analyzer. Hence it is difficult to identify tryptic peptides in this region.

The DIPEN$^{341}$ cleavage was demonstrated using both neoepitope antibodies and MS following MMP-3 digestion. The other previously identified major aggrecanase derived cleavage sites in equine aggrecan (Fosang et al., 1995; Hughes et al., 1995; Hughes et al., 1992); PQNITEGE$^{373}$ $^{374}$ARGNVILT, ATTAGQL$^{1558}$ $^{1559}$EGRGTIGIS, PTPFKEEE$^{1745}$ $^{1746}$GLGSVELSG, TQAPTAQE$^{1850}$ $^{1851}$AGEGPSGI, TEPTVSQE$^{1950}$ $^{1951}$LGQRPPVT were not identified. This could be due to the size of the semi-tryptic fragments produced being outside the mass range. For instance cleavage at PTPFKEEE$^{1745}$ $^{1746}$GLGSVELSG would produce a small peptide that was less than mass range identifiable; K.EEE and GLGSVELSGLPSGDADLSGTSGR whose mass would be too great. Using *in-silico* digests of aggrecan with chymotrypsin, Glu-C, Asp-N and elastase it was determined that these proteases, commonly used in in-solution digests to produce peptides of a size detectable by MS, did not produce a mass range of peptides any more detectable for aggrecan. Interestingly, apart from decorin, where neopeptides identified were equally distributed throughout the protein, neopeptides generated from COMP, biglycan and fibromodulin were primarily from the C-terminal region. This has previously been described for COMP by Zhen *et al.* 2008 (Zhen et al., 2008).

Many of the neopeptides generated were evident in both MMP-3 and ADAMTS-4 digests and thus generated by multiple proteases, however some unique neopeptides were generated for specific protease digests. Some of these could provide useful pharmacodynamic biomarkers of specific protease activity in articular cartilage. For example for fibromodulin digestion with MMP-3 alone produced the neopeptide $^{325}$INEFSISSFCTVVDVMN$^{341}$, which was identified a number of times and also identified in the *in-vitro* explant culture. This neopeptide and others may function as useful biomarkers of cartilage degradation in articular cartilage. Until recently the progression of OA has been measured almost exclusively by radiography of the joint to assess space narrowing. There is a need to progress to translatable biomarkers that are able to monitor protease activity and OA disease. The peptide fragments demonstrated in this study provide a starting point for further investigations in order to serve as early indicators of cartilage turnover similar those already studied in COMP (Neidhart et al., 1997) and

aggrecan (Dufield et al., 2010a). Indeed COMP has been demonstrated to increase in serum of patients with OA (Clark et al., 1999; Misumi et al., 2001). In this study neopeptides of COMP, a collagen binding thrombospondin matricellular protein family member, were amongst the most numerous identified especially in the *in-vitro* study.

This study has also provided new insights into the possible role of multiple proteases in degrading proteins at the same cleavage sites. Until recently no cleavage site had been identified and attributed to the action of a single specific enzyme for COMP. However ADAMTS-5 cleavage of COMP at $P^{75}/^{76}A$ has now been demonstrated and validated (Holden, 2012). Here the presence of $^{63}$NTVMECDACGMQP$^{75}$ in both ADAMTS-4 digests of crude proteoglycan and following IL-1β stimulation of cartilage explants suggests that ADAMTS-4 is also capable of cleaving COMP at the ADAMTS-5 specific site. Further work is required, for instance by developing neoepitope antibodies that recognizes the cleaved fragment of COMP, to validate this.

**Conclusion**

The study identified many MMP cleavage products of equine cartilage. The catalogue of neopeptides produced provides information about the major proteoglycans susceptible to specific protease activity. In addition the use of an IL-1β stimulated model of early OA enabled a subset of common neopeptides produced by ADAMTS-4 and MMP-3 to be identified, as well as additional neopeptides produced by other MMPs. Several of the neopeptides identified are potential biomarkers of ADAMTS-4, and MMP-3 activity as well as arthritis. An approach has been highlighted to develop a list of potential biomarkers which could be applied to other tissues or diseases of interest.

**Acknowledgements**

Mandy Peffers was supported by a Wellcome Veterinary Integrated Research Fellowship.

**APPENDIX TO MANUSCRIPT 4**

**Diagrams depicting neopeptide locations in aggrecan, biglycan, COMP, decorin and fibromodulin**

**Equine aggrecan (O18832)**

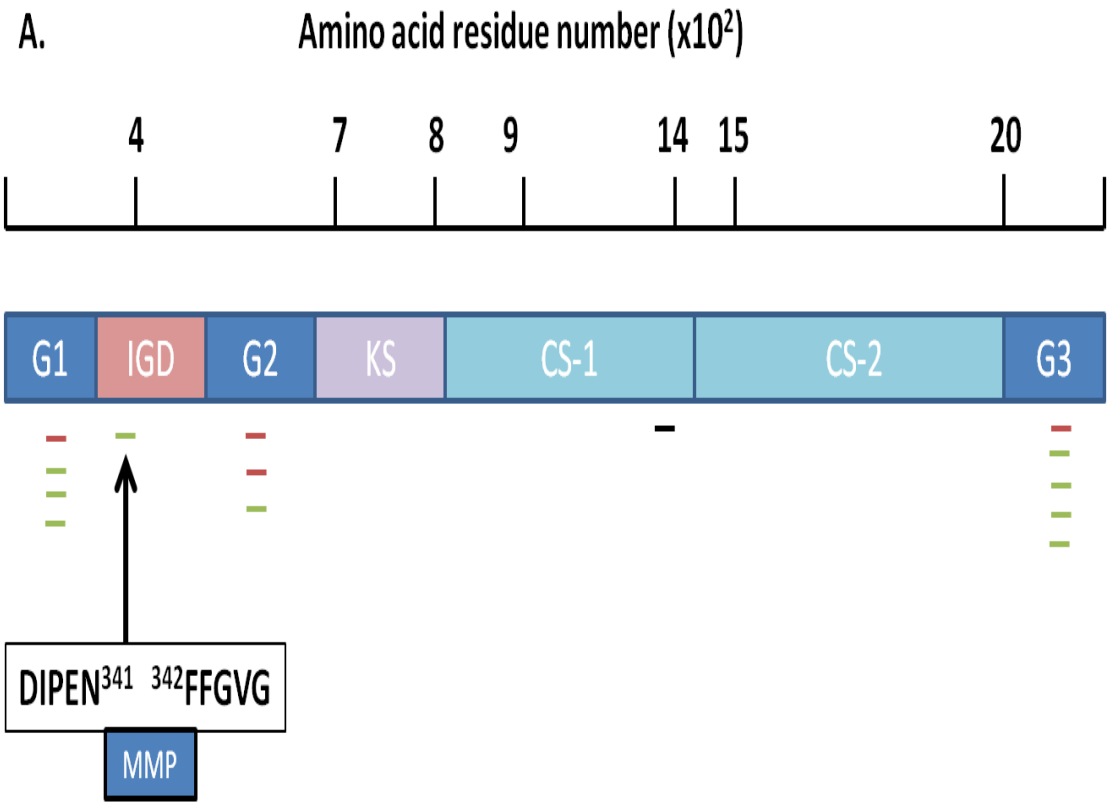
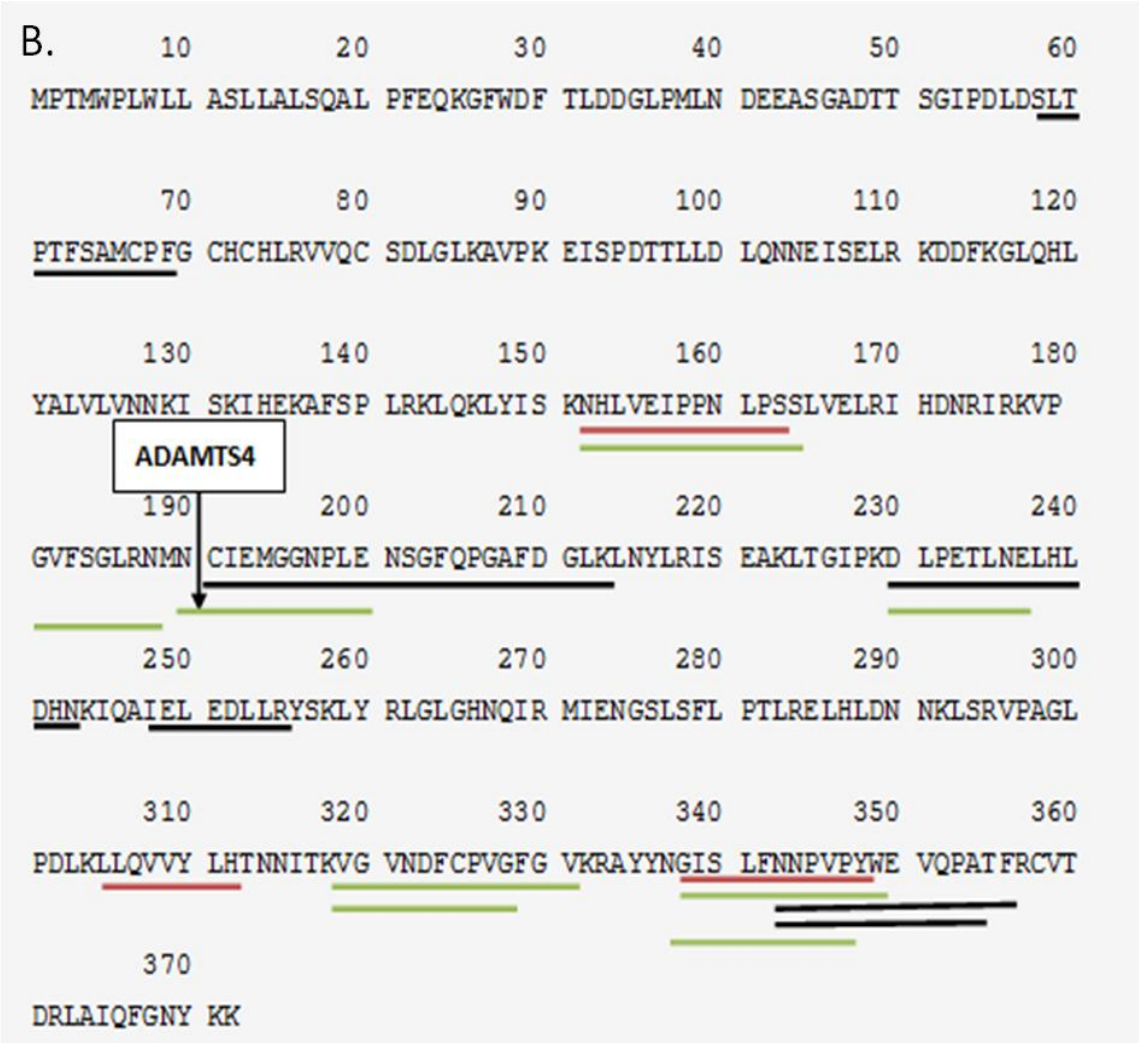
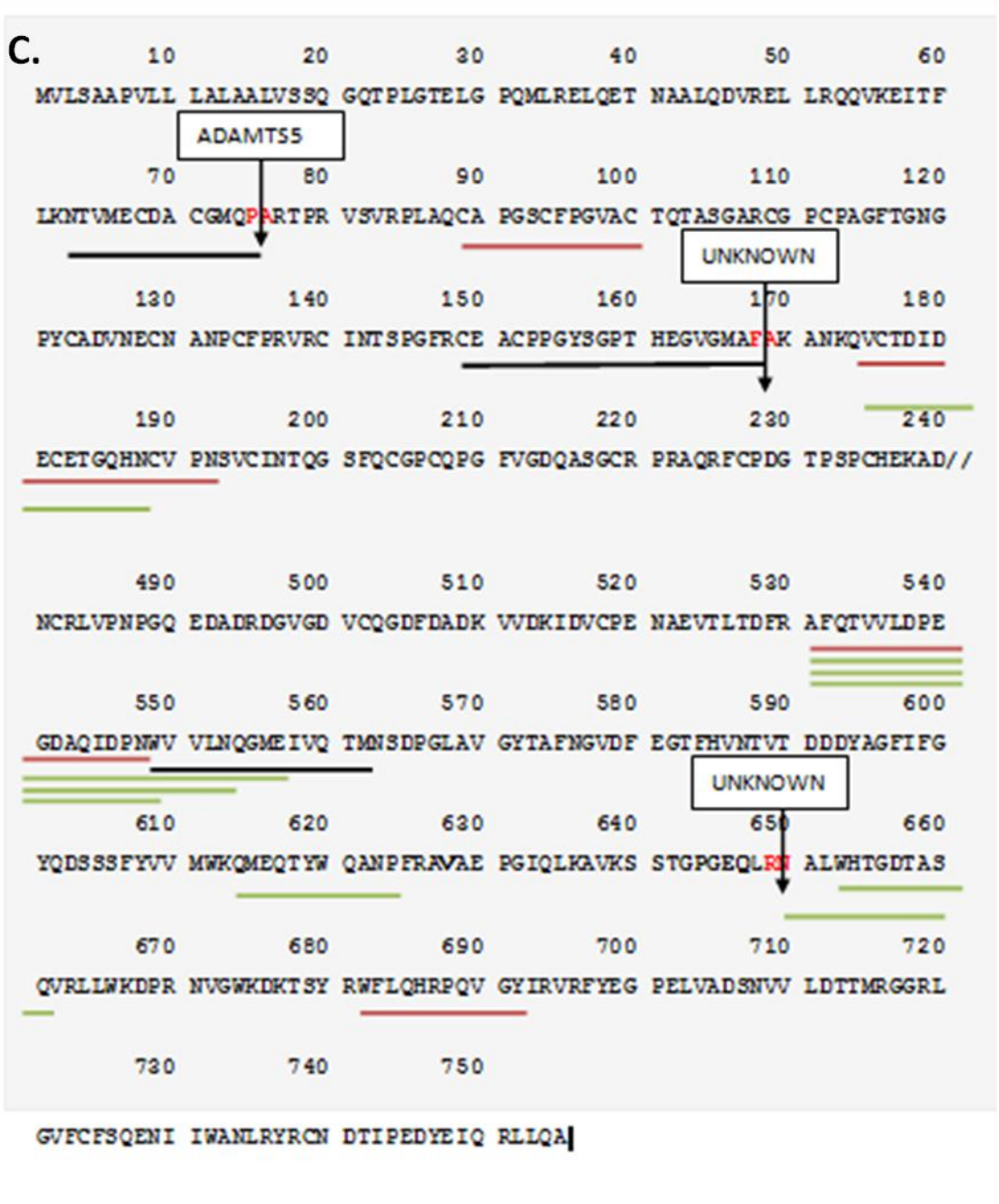



**Diagram A. Schematic representation of the domain organisation and structure of equine aggrecan along with the depiction of the MMP driven cleavage site at $^{341}$N-$^{342}$F.** Black lines represent the location of ADAMTS4 digestion derived neopeptides, green lines MMP-3 digestion derived neopeptides and red represents those identified in both ADMTS4 and MMP3 digestions. All marked neopeptide sequences were significantly identified by Mascot ($p<0.05$)and/or PEAKS® (equivalent $p<0.01$).

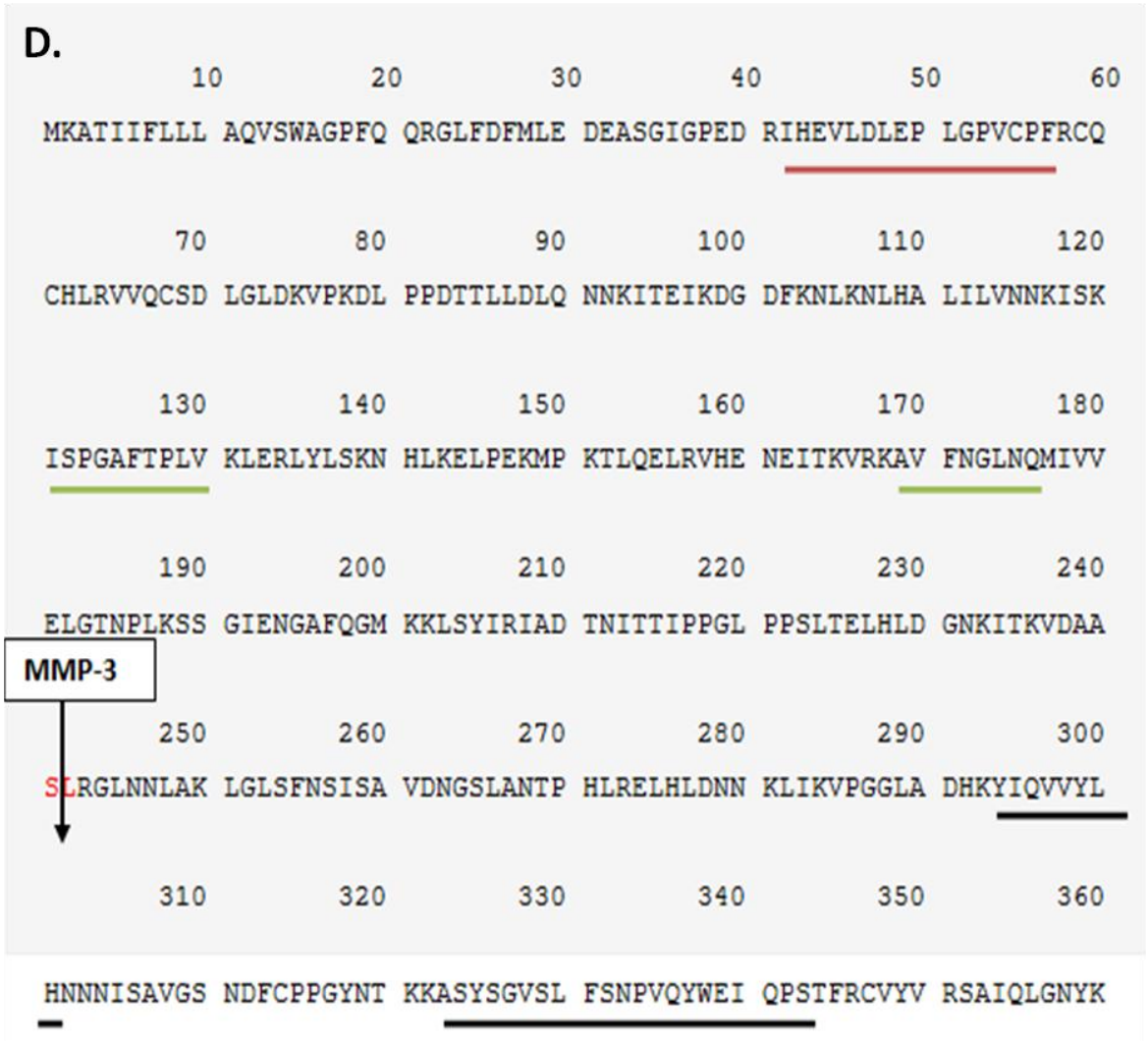
**Equine biglycan (O46403)**



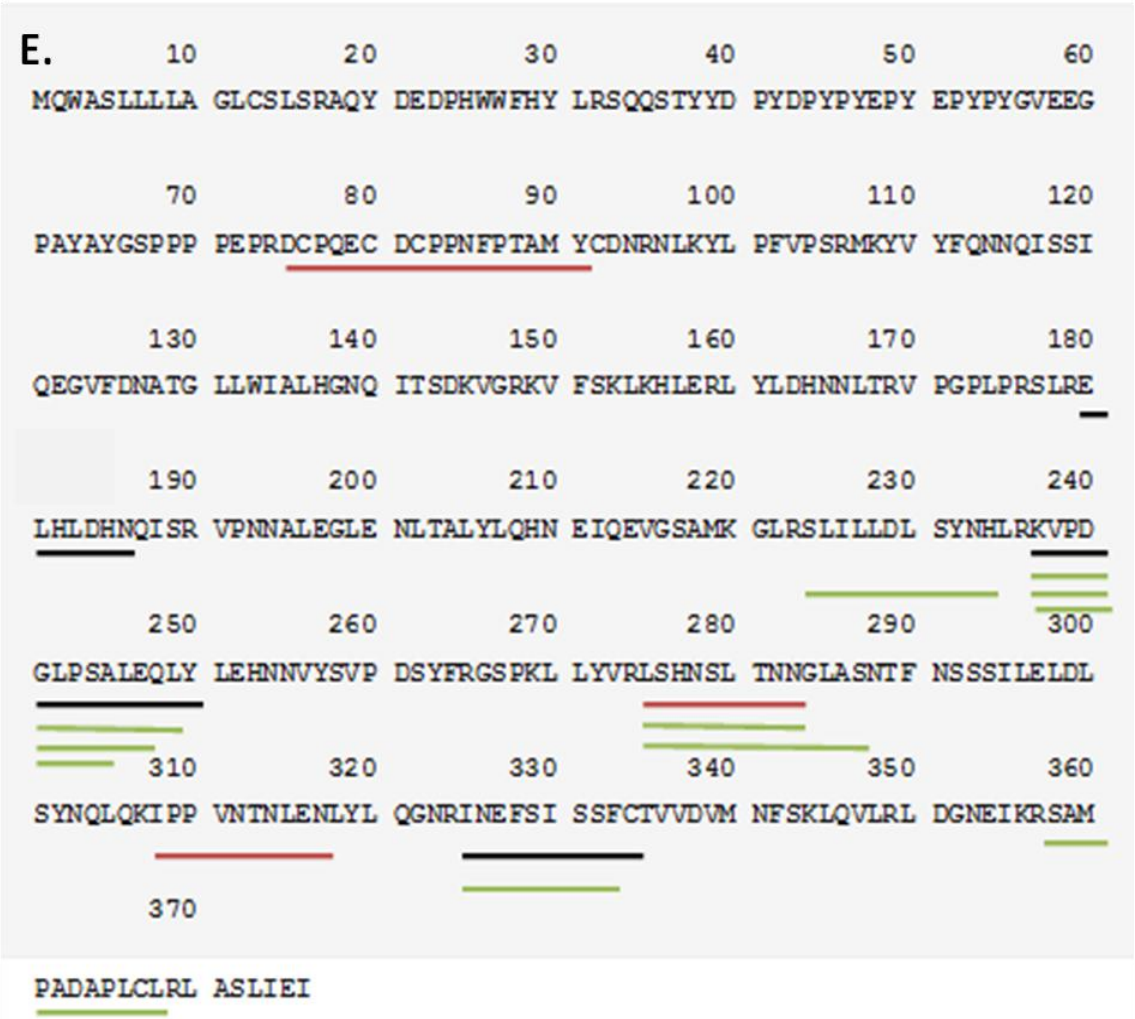
**Equine COMP (Q9BG80)**



**Equine decorin (O46542)**



**Equine fibromodulin (A2Q126)**




**Diagrams B, C, D, E. Equine biglycan, COMP, decorin and fibromodulin amino acid sequences and significant neopeptide locations.** Each protein name is followed by the Uniprot protein accession number the sequence was taken from. For COMP the sequence has been truncated between amino acids 240 and 480. Black lines represent the location of ADAMTS4 digestion derived neopeptides, green lines MMP-3 digestion derived neopeptides and red represents those identified in both ADMTS4 and MMP3 digestions. All marked neopeptide sequences were significantly identified by MASCOT ($p<0.05$) and/or PEAKS® (equivalent $p<0.01$).the location of known protease cleavage sites for ADAMTS4, ADAMTS5, MMP-3 are marked on the relevant protein sequence. In addition two previously identified cleavage sites in COMP are marked (the enzyme responsible for these are not known).

# Manuscript 5

Peffers, M.J[1*]., Cillero-Pastor, B[2]., Eijkel, G[2]., Clegg, P.D[1]., Heeren, R.M.A[2].

**Corresponding Author**: M J Peffers[*]

[1] Institute of Ageing and Chronic Disease, University of Liverpool, Leahurst, Chester High Road, Neston, Wirral, CH64 7TE

[2]Biomolecular Imaging Mass Spectrometry (BIMS), FOM Institute AMOLF, Amsterdam, The Netherlands

**Title: MALDI Imaging Mass Spectrometry Identifies Markers of Ageing and Osteoarthritic Cartilage**

**Running Title: MALDI Imaging Mass Spectrometry Identifies Markers of Ageing and Osteoarthritic Cartilage**

**Key Words:** ageing, cartilage, degradation, MALDI-IMS, osteoarthritis,

## MALDI Imaging Mass Spectrometry Identified Markers of Ageing and Osteoarthritic Cartilage


### Abstract

Cartilage protein distribution and the changes that occur in cartilage ageing and disease are essential in understanding the process of cartilage ageing and age related diseases such as osteoarthritis (OA). This study utilised matrix assisted laser desorption ionization imaging mass spectrometry (MALDI-IMS) to investigate the spatial distribution of different components in ageing and osteoarthritic cartilage sections. The distribution of peptides in young, old and osteoarthritic equine cartilage was compared following tryptic digestion of cartilage slices and MALDI-IMS undertaken with a MALDI SYNAPT™ HDMS system. Following multivariate analysis protein identification was undertaken using database searches. Peptide intensity differences between young, ageing and OA cartilage were imaged with Biomap software. Proteins including aggrecan core protein, fibromodulin, and cartilage oligomeric matrix protein were identified and localised. A number of tentative markers were identified for OA and ageing. In addition a number of potential peptides targeted for degradation in OA were detected. MALDI-IMS provided a novel platform to study cartilage ageing and disease enabling age and disease specific peptides in cartilage to be elucidated and spatially resolved.


### Introduction

Osteoarthritis (OA) is an age related joint disease characterized by a loss of cartilage extracellular matrix (ECM) (Goldring, 2000). Progressive destruction of articular cartilage is a hallmark of OA, leading to chronic pain and lameness. Whilst this heterogeneous condition occurs as a consequence of numerous overlapping independent factors including gender (van Saase et al., 1989), genetics (Spector et al., 1996), and obesity (Carman et al., 1994) age is the most common risk factor for its initiation and progression, with symptomatic OA affecting 10–20% of people aged over 50 years (Lawrence et al., 2008). The explanation for this is an accumulation of 'wear and tear' injuries due to mechanical loading over many years.

Although much work has been undertaken investigating the pathogenesis of OA the molecular mechanisms involved are not fully understood, with few validated markers for disease diagnosis and progression being available. Treatment of OA is limited to pain relief or joint replacement with few disease-modifying therapeutics coming to fruition. Hence methodologies which permit the study in detail of ageing and OA cartilage organisation could improve the knowledge of how ageing increases the risk of OA.

Measurement tools enabling the capture of the dynamic and complex interplay between proteins, lipids, DNA and other molecules are necessary. This interplay involves the appearance, interaction and disappearance of many species on varying time scales. Capturing this spatial and temporal information will help in the quest to understand the pathogenesis of OA and ultimately provide disease modifying treatments.

Mass spectrometry (MS) is an analytical tool that enables the accurate measurement of both mass and charge of molecules. Matrix assisted laser desorption/ionization imaging mass spectrometry (MALDI-IMS) of tissue samples is a powerful technique that allows for spatially resolved, comprehensive and specific characterization of hundreds of unknown molecular species (proteins, peptides, lipids or metabolites) *in-situ* in a single molecular imaging experiment (Seeley and Caprioli, 2008). Furthermore it can spatially resolve below 50μm and it is not necessary to use any kind of labelling (Amstalden van Hove et al., 2010). IMS has been used to identify the molecular distribution of peptides in many diseased tissues including brain (Taban et al., 2007), liver (Lee et al., 2011) and kidney (Meistermann et al., 2006). Additional molecules including lipids (Hankin et al., 2011) and small molecules (Blatherwick et al., 2011) have been localised in numerous tissues. Moreover, specific protein patterns revealed by IMS have been demonstrated as prognostic (Balluff et al., 2011) and diagnostic indicators (Gustafsson et al., 2011). MALDI-IMS techniques have only recently been implemented in cartilage research. One study employed a time of flight secondary ion mass spectrometry (TOF-SIMS) workflow in order to acquire molecular-specific spatial distribution of lipids in normal and OA cartilage (Cillero-Pastor et al., 2012b).

The study revealed a mixture of organic and inorganic distributions that distinguished human normal from OA cartilage. A further study used MALDI-IMS to identify and localise OA specific peptides and proteins in cartilage (Cillero-Pastor et al., 2012a)

The aim of this study was to establish peptide profiles in young, ageing and OA horse cartilage with a high spatial distribution in order to determine changing molecular events distinct between ageing and disease.

**Methods**

**Sample collection, preparation and processing**

Full thickness equine cartilage slices were removed from the mid condyle region of metacarpophalangeal joints collected from an abattoir and snap frozen in liquid nitrogen. All samples were scored macroscopically using Kawcak scoring for pathological grading of the distal condyles of metacarpophalageal III (Kawcak et al., 2008). For age-related studies samples were taken from skeletally mature young (4 years old, n=3) and old (greater than 15 years old, n=3) horses. Equine tissue was readily obtained enabling collection of cartilage samples from macroscopically normal, skeletally mature young and aged horses as well as osteoarthritic cartilage. For young horses, one year is equivalent to about 3.5 years of a human. Hence horses of greater than 15 years old, used in this study equates to humans of older than 52 years. Furthermore, racehorses develop OA at an earlier age due to their strenuous physical activity in a similar way to professional and recreational athletes (Golightly et al., 2009; Neundorf et al., 2010). For OA studies skeletally mature donors greater than 15 years were chosen with mild macroscopic OA changes. Samples were stored at -80$^{o}$C and shipped overnight on dry-ice to the Biomolecular Imaging Mass Spectrometry (BIMS) Molecular Nanophotonics Department, FOM Institute-AMOLF, Amsterdam for experimental analysis.

Cartilage was sectioned at 12µm thicknesses on a cryostat Microm HM 525 (Microm International, Walldorf, Germany) and thaw mounted on glass slides. All samples were dried in a vacuum desiccator for 10min prior to further processing.

**Tissue digestion and matrix deposition for MALDI-IMS**

Young and old cartilage were studied in duplicate whilst normal and OA cartilage were examined in triplicate and then compared by MALDI-IMS. The sections were washed in 70% ethanol and chloroform for 30s each. Trypsin (Sigma-Alrich, Dorset, UK) at 0.05μg/μL was applied using a high-accuracy chemical inkjet printer (CHIP-1000; Shimadzu Biotech, Kyoto, Japan). Each tissue section was spotted with 5nL trypsin (0.05μg/μl) per position in a 150μm spacing raster scheme using cycles of 250pL per droplet. The samples were incubated overnight at 37$^{o}$C. Using a vibrational sprayer (ImagePrep; Bruker, Bremen, Germany) the matrix solution; α-Cyano-4-hydroxycinnamic acid (HCCA) (10mg/mL) in 50% acetonitrile, 50% trifluoroacetic acid (TFA) 0.1%, was applied on top of the tissue section.

**MALDI–IMS**

Tissue sections were optically scanned prior to MALDI–IMS experiments using a 2400 dpi desktop scanner. The resulting images were imported into the MALDI Imaging Pattern Creator software (Waters Corporation, Manchester, UK). A MALDI SYNAPT™HDMS system (Waters Corporation) instrument was calibrated using a standard calibration mixture of polyethylene glycol with a molecular weight of 100-3000 (Sigma-Aldrich). The instrument was operated with a 200 Hz Nd:YAG laser and was configured to acquire data in the positive V-reflectron mode. Data were acquired at a lateral resolution ~ 90μm and at a raster size of 150μm.

**Multivariate analysis and data interpretation following MALDI-IMS**

Within the data analysis workflow a number of AMOLF in-house build MATLAB (The MathWorks, Natick, MA, USA) software tools (Synapt2Tricks, ChemomeTricks, CombiTricks, PEAPI, Synapt_PL) were utilised (Figure 1). Synapt RAW format files were converted to NetCDF format using the DataBridge utility in MassLynx v4.1. The NetCDF files were converted to MATLAB Tricks format using Synapt2Tricks. Selection of the regions of interest (ROIs) to separate on-tissue signal from off-tissue sample was performed using ChemomeTricks. For each ROI an average

spectrum was calculated and these average ROI spectra were combined in one dataset using CombiTricks. Peakpicking was performed on the combined ROI averages dataset using PEAPI. The resulting peaklist was used to integrate the signals of all full ROI files. The integrated full ROI files were combined using CombiTricks. Principal component analysis (PCA) and discriminant analysis (DA) were performed on the resulting file using ChemomeTricks. PCA was used as a data pre-treatment step before DA. This data pre-treatment is necessary to prevent noise related features from being introduced into the DA calculation. The first 20 principal components were used as input for DA to probe spectral similarities and differences between young and old, old normal and old OA or young, old and old OA samples. Some of the resulting discriminant function loadings were used to provide peaks for MS/MS targeting.

Biomap 3.7.5.5 software (Novartis Pharma AG, Basel, Sweden) was used to generate ion images. Normalisation of the intensity of all *m/z* channels was performed using the intensity of the m/z 190 matrix peak. P values for statistical differences found in MALDI-IMS experiments were calculate with one-way ANOVA with a Bonferroni ad-hoc test using SPSS (IBM, Hampshire, UK) following normality testing. Differences were considered to be statistically significant at $p \leq 0.05$. The data were expressed as mean intensity ± standard error mean (SEM).

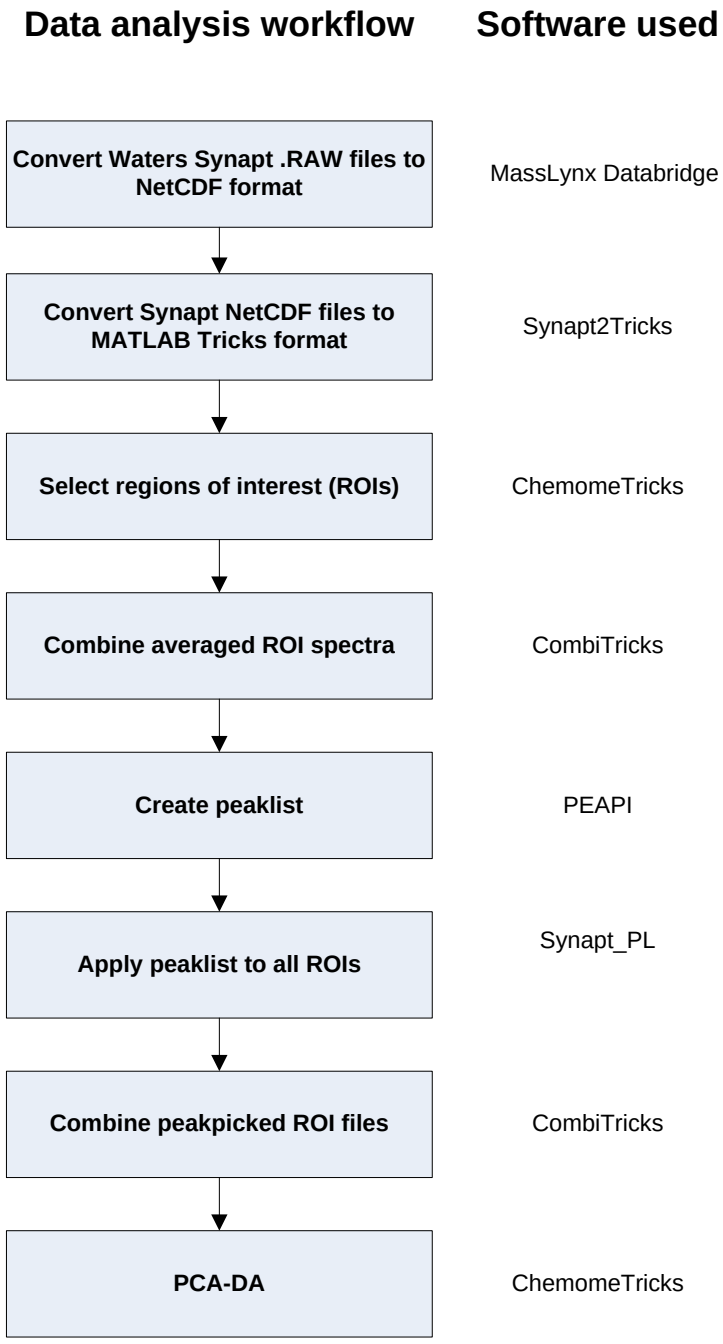



**Figure 1. A number of AMOLF developed software tools were used in the data analysis of MALDI-IMS data.**

### Histological staining and analysis

Following IMS the tissue sections were washed in 70% ethanol for 5min to remove matrix. The slides were immersed in Harris hematoxylin solution (Sigma-Aldrich) for 1min. After washing with water for 15min, they were rinsed in 95% ethanol and counterstained in an eosin solution for 30s. Digital images were acquired with the

Mirax system (Carl Zeiss, Siledrecht, The Netherlands) after dehydrating steps. A modified Mankin scoring system was used for semi-quantitative histological assessment of all equine cartilage samples (McIlwraith et al., 2010).

**Tissue digestion and matrix deposition for profiling experiments**

Protein identification was undertaken with profiling experiments directly from each donor tissue by applying 10µL of trypsin 0.05µg/µL directly and incubating overnight at 37$^{o}$C. HCAA matrix was applied as previously described. Data dependent analysis (DDA) of tryptic peptides was performed with the MALDI SYNAPT™HDMS system. Every MS survey scan was followed by collisional fragmentation of the most intense ions with subsequent collection of MS/MS spectra. Direct MS/MS fragmentations were performed directly from the tissue on the peptides that differentiate either young and old or old normal and old OA tissue. These target peptides were found following Discriminant Analysis (DA). In addition fragmentation was also undertaken on a peptide list generated from bibliographical analysis of previous MALDI-IMS studies in cartilage (Cillero-Pastor et al., 2012c). The resulting data files were submitted to an in-house Mascot (Perkins et al., 1999) server (Matrix Sciences, London, UK) and searched against Unihorse and Swissprot databases. Search parameters used were; peptide mass tolerances 50ppm, fragment mass tolerance of 0.5Da, 1+ ions, missed cleavages; 3, and instrument type MALDI-Q-TOF. Modifications used were variable oxidation of methionine.

**Results**

**Macroscopic grading and morphological evaluation of cartilage samples**

Macroscopic grading and a semi-quantitative histological assessment using a modified Mankin's scoring system of the samples were undertaken. The macroscopic score was used to allocate samples into old normal or old OA groups, a score of greater than 0 were assigned OA. Only macroscopically normal cartilage was used for age related studies. All samples for age related studies had a modified Mankin's grade of 0 (data not shown). The macroscopic grading and Mankin's score

for each sample used in the old normal versus old OA studies are demonstrated in Tables 1A and 1B.

A

| Donor | Structure (/9) | Cellularity (/4) | Cloning (/4) | Total (/17) |
|---|---|---|---|---|
| Normal 1 | 0 | 0 | 0 | 0 |
| Normal 2 | 0 | 0 | 0 | 0 |
| Normal 3 | 0 | 0 | 0 | 0 |
| OA 1 | 3 | 1 | 1 | 5 |
| OA 2 | 2 | 2 | 1 | 5 |
| OA 3 | 1 | 3 | 1 | 5 |

B

| Donor | Wear lines (/3) | Erosions (/3) | Palmar arthrosis (/3) | Total (/9) |
|---|---|---|---|---|
| Normal 1 | 0 | 0 | 0 | 0 |
| Normal 2 | 0 | 0 | 0 | 0 |
| Normal 3 | 0 | 0 | 0 | 0 |
| OA 1 | 0 | 1 | 1 | 1 |
| OA 2 | 0 | 1 | 1 | 1 |
| OA 3 | 0 | 1 | 1 | 1 |

**Table 1. Modified Mankin's grading (A) and macroscopic evaluation (B) and of normal and OA cartilage samples.** Numbers in parenthesis relate to the maximum score for that characteristic with higher scores equating to more severe changes.

**MALDI-IMS of equine cartilage samples**

**a. Multivariate analysis**

**i. Young, old and OA cartilage**

After imaging experiments, data from young, old and OA cartilage groups was analysed together. The spectra of each group revealed different profiles. A

combined spectrum of representative digested equine young, old and OA samples are demonstrated in Figure 2, with peptides specific to each category already evident (Table 2 and 3).

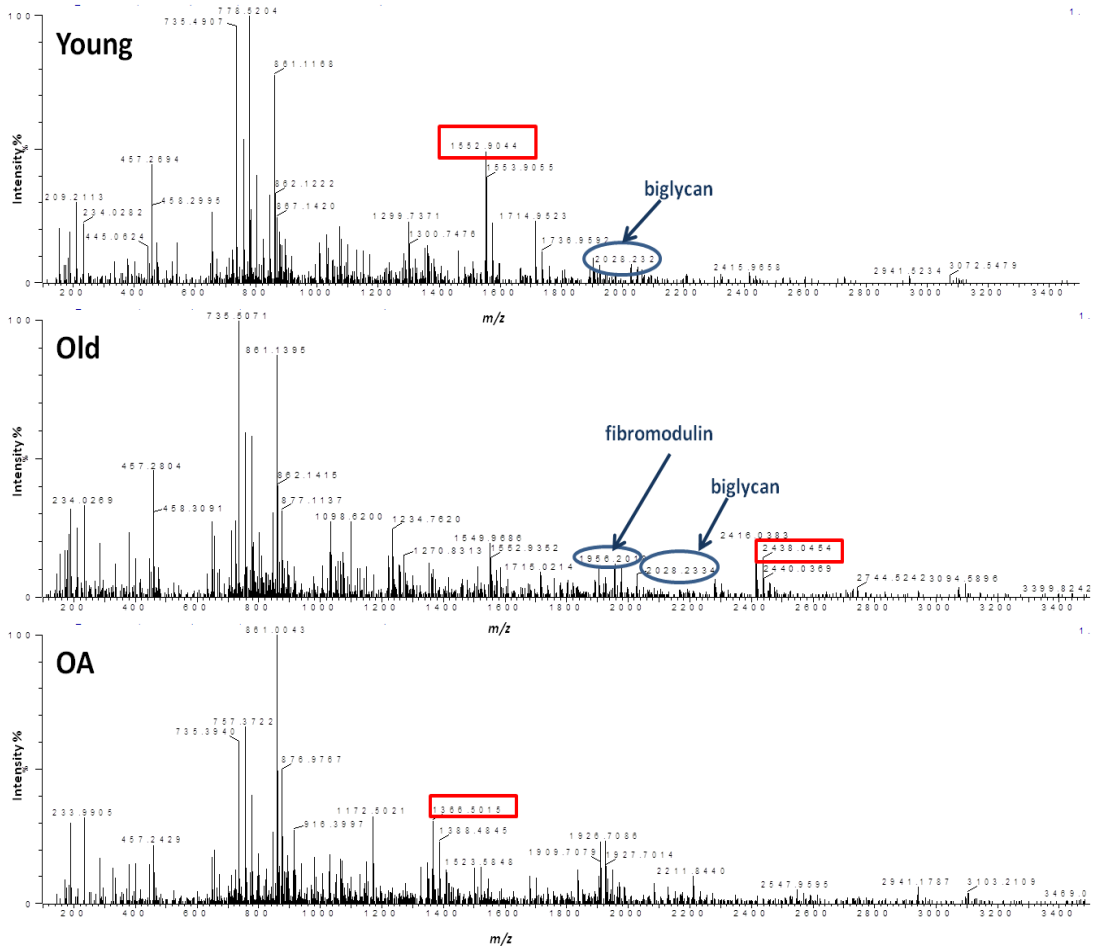



**Figure 2. Combined spectrum of representative digested equine young, old and OA samples reveal different profiles.** A representative spectrum form each group is shown. Examples of peptides specific to each condition are seen in the combined spectra and marked in red boxes.

MALDI-IMS was used for the study of differential peptide distributions in young, old and OA cartilages. PCA analysis, which is unsupervised, revealed differences between each sample type (Figure 3). The first two functions accounted for 13% of the variance.

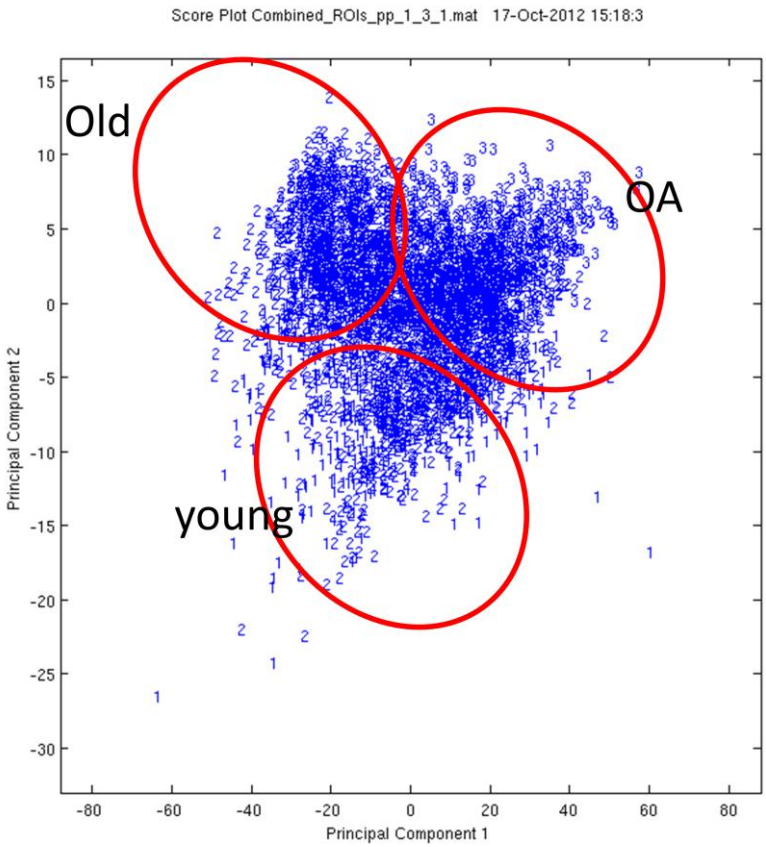



**Figure 3. A PCA scatter plot revealed separation when the first 20 principal components only were taken into account**. Rings denote the different categories; young; 1, old; 2 or OA; 3.

DA was also performed on the first 20 principal components in the dataset in order to remove noise and provide a conservative approach to data analysis.

Following DA the resulting discriminant functions (DFs) classified the data in three groups; young, old and OA. However the first function separated mainly OA from the young samples. Interestingly there was a big contribution of the old samples to the negative part of DF1 indicating that peptides within old samples were also present in OA samples but there were also a contribution of the old samples to the positive part (Figure 4A).

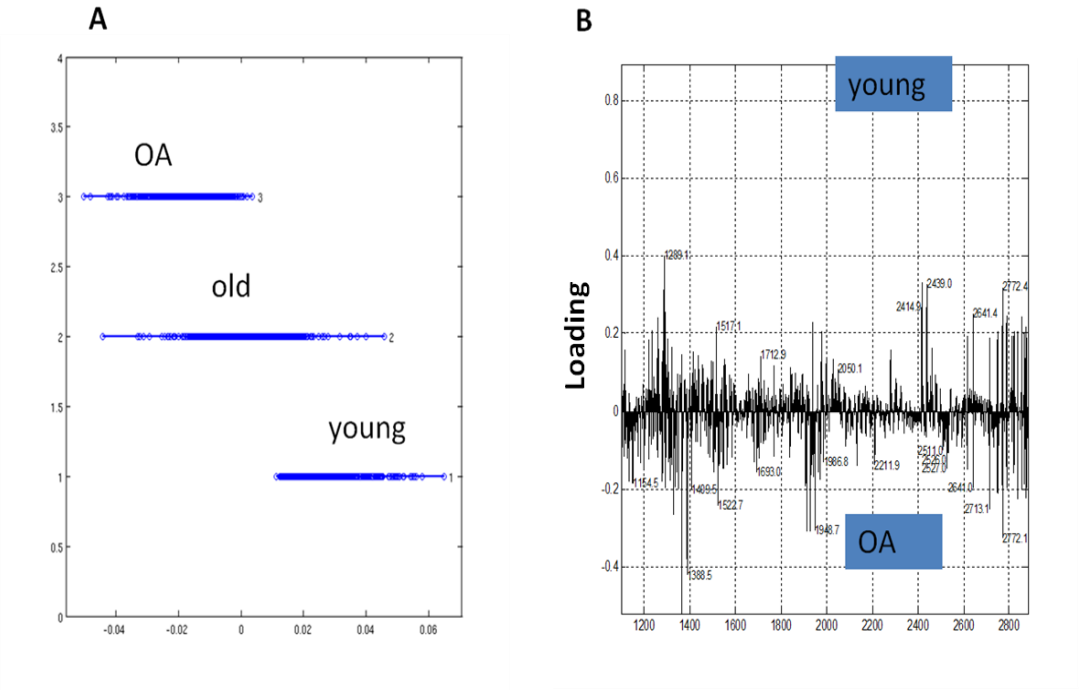



**Figure 4. Young, old and OA samples after MALDI-IMS experiments were analyzed by PCA and DA to classify peptides specific to each group**. A) Data is representative of DF1. The positive part of the graph corresponds mainly to young (and partially old) and the negative to OA and partially to old samples. B) Scaled loading plot of DF1 representing the peptides specific primarily to young and old/OA detected by the MALDI-IMS approach.

Following DA a number of peaks were detected/ identified as specific for young (including 2415.9), old or OA. When the raw data was examined, similar patterns were evident. For instance, using a hierarchical clustering approach to classify the raw spectra we observed differences at 2415.9 *m/z* peak. The intensity of this and other related masses grouped in the same cluster, demonstrated spectral differences even prior to PCA and DA (Figure 5). This peak was also demonstrated to have high loading in young cartilage (Figure 6b). A peptide with mass 2414.9 was subsequently identified as collectin-43 using fragmentation targeting of this mass. It is highly probable that this represents an isotope of 2415.9. The 2415.9 peak has two different distributions in old samples which may be correlated to different areas in a tissue with varying degrees of damage (Figure 5c). This verifies that MALDI-IMS can reveal a peptide distribution profile that could predict and define the type of sample analysed.

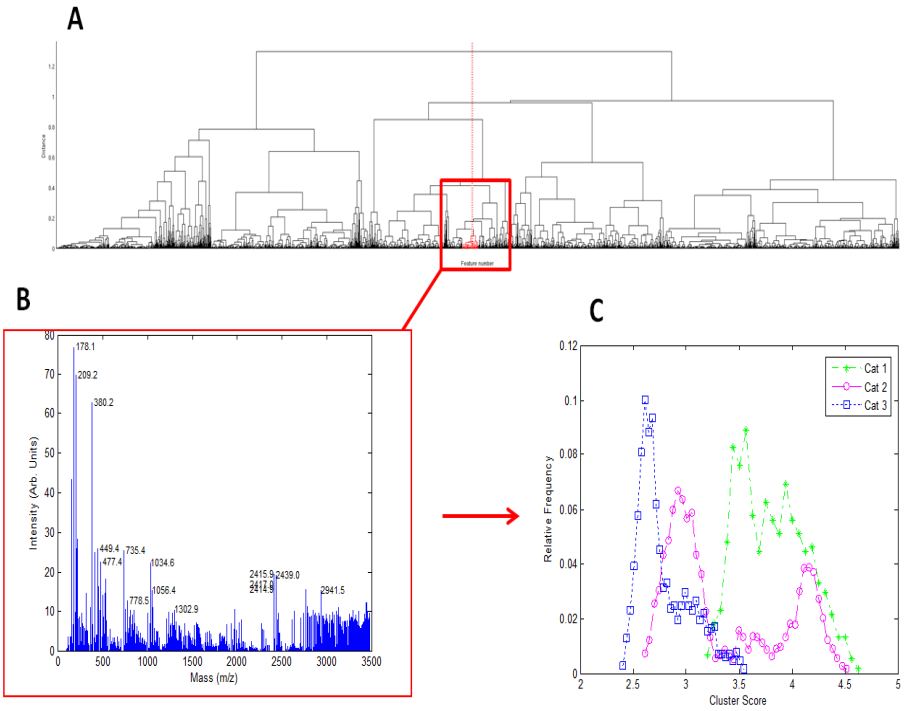



**Figure 5. Analysis of 2415.9 *m/z* cluster, illustrates differences between young, old and OA spectra.** Hierarchical clustering dendrogram (A) showed those peaks that can be correlated and differently distributed among the groups (B). A cluster score histogram (C) demonstrated separation for the 2415.9 m/z cluster between category 1 (young), 2 (old) and 3 (OA), with greatest separation between young and OA.

In addition comparisons were also made using PCA and DA between young and old and old normal and old OA groups separately.

### ii. Young and old cartilage

Differential peptide distributions between young and old cartilage was assessed by MALDI-IMS using three donors for each condition in duplicate. Different spectral profiles of young and old cartilage were evident (Figure 6A and B). Peak picking from within the tissue only was required to reduce the dataset size in order to allow further analysis. The resulting DF1 classified the data into young and old (Figure 6A). The DF1 spectra (Figure 6B) identified the peaks of the positive part of the DF1 were primarily specific from young cartilage samples such as 2415.9 and the peaks of negative part were more abundant in cartilage from old donors.

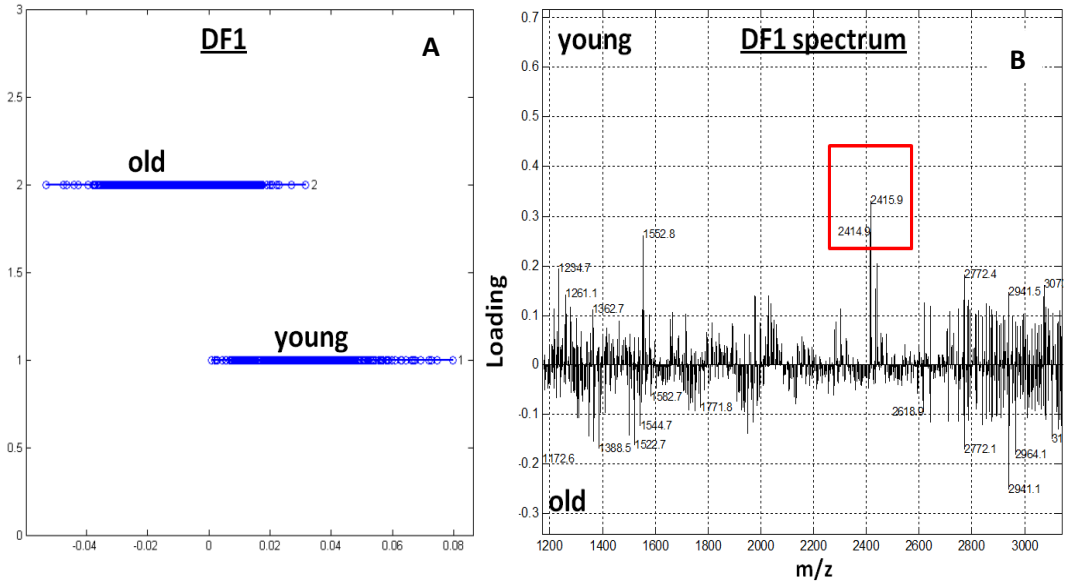



**Figure 6. PCA and DA separated the dataset on age of the sample.** A) The spectra of all the young and old samples after MALDI-IMS experiments were analyzed by PCA and DA to classify peptides specific of each age using a 1 D plot. B) Loading plot of DF1 representing the peptides specific to age of cartilage given by MALDI-IMS. The young specific peaks 2414.9 and 2415.9 are ringed.

### iii. Old Normal and old OA cartilage

Finally normal and OA cartilage from three donors greater than 15 years old in duplicate were assessed for differential peptide distributions using MALDI-IMS. After PCA and DA the resulting DF1 classified the data into normal and OA (Figure 7A). The DF1 spectra (Figure 7B) identified the peaks of the positive part of the DF1 were specific from OA cartilage samples such as 1366.5 (also imaged in Figure 10) and the peaks of negative part were richer in cartilage from normal donors.

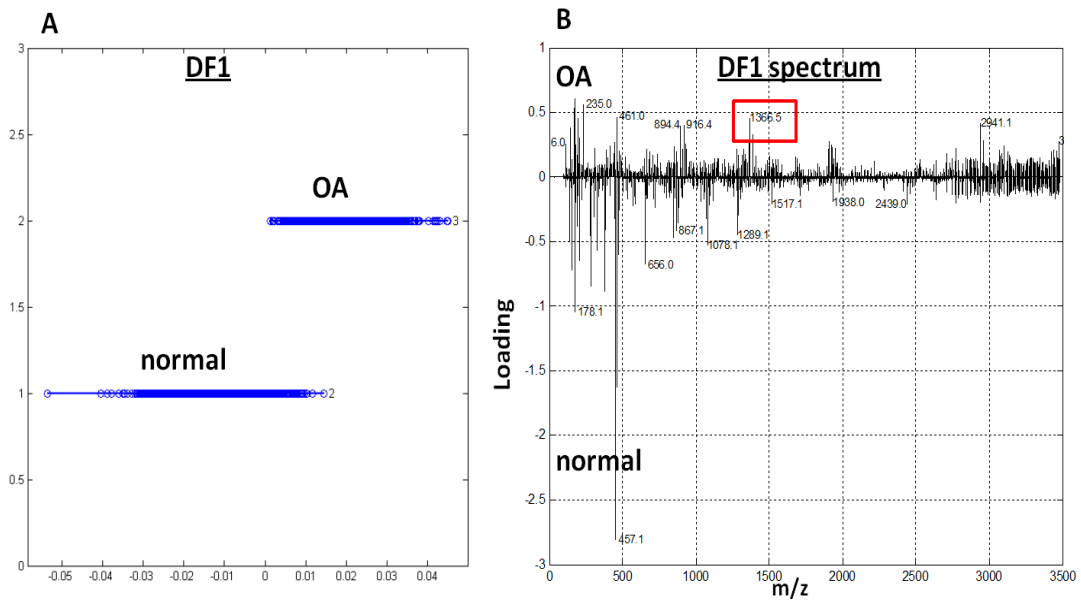



**Figure 7. PCA and DA of old normal and old OA cartilage.** A) The spectra of all the normal and OA samples after MALDI-IMS experiments were analyzed by PCA and DA to classify peptides specific of each category using a 1 D plot. B) Loading plot of DF1 representing the peptides specific to OA given by MALDI-IMS. The OA specific peak 1366.5 is ringed.

### b. Profiling studies of cartilage

MS/MS profiling experiments on cartilage slices were undertaken in order to determine proteins that could be identified directly from the tissue (Table 2). Tryptic peptides of COMP and fibromodulin were common to all groups.

### c. Peptide identification through targeted fragmentation studies

Following PCA and DA of each group; young, old and OA; young and old; old normal and old OA, the peaks with the highest absolute loadings in the DF1 spectra, were targeted directly from tissue slices for MS fragmentation and database searching in order to identify the protein it pertained to. In addition of list of masses were also selected that had been previously identified in human cartilage MALDI-IMS studies (Cillero-Pastor et al., 2012a). For some masses although a good fragmentation spectra was obtained we were unable to identify them (Figure 6).

| Accession | Protein | Observed mass | Score | Peptide |
|---|---|---|---|---|
| FMOD_BOVIN | Fibromodulin | 1557.01 | 67 | R.SLILLDLSYNHLR.K |
| FMOD_BOVIN | Fibromodulin | 1955.23 | 67 | K.IPPVNTNLENLYLQGNR.I |
| PGS1_HORSE | Biglycan | 1312.8 | 41 | K.IQAIELEDLLR.Y |
| CO2A1_MOUSE | Collagen alpha-1(II) chain | 1678.9 | 39 | K.DGETGAAGPPGPSGPAGER.G |
| CO2A1_HORSE | Collagen alpha-1(II) chain | 1427.92 | 23 | K.ALLIQGSNDVEIR.A |
| COMP_HORSE | Cartilage oligomeric matrix protein | 1698.94 | 31 | K.QMEQTYWQANPFR.A |
| MIA3_MOUSE | Melanoma inhibitory activity protein 3 | 1186.74 | 24 | R.EYAPGVLPGKR.D |
| HBBA_HORSE | Hemoglobin subunit beta-A | 1274.85 | 23 | R.LLVVYPWTQR.F |
| MATN3_HORSE | Matrilin-3 | 1618.04 | 20 | K.SRPLDLVFIIDSSR.S |
| CHAD_BOVIN | Chondroadherin | 2198.3 | 29 | R.AGAFDDLTELTYLYLDHNK.V |

**Table 2. Proteins identified from MS/MS profiling experiments of cartilage.** The proteins identified following profiling experiments from cartilage are represented. The table identifies the Swissprot accession number, abbreviation of the protein name, experimental molecular weight of the matched peptides, the score given by the Mascot algorithm, and the sequence of the matched peptides. All Mascot scores were significant ($p<0.05$). All peptide sequences were checked against the equine sequence and were homologous.

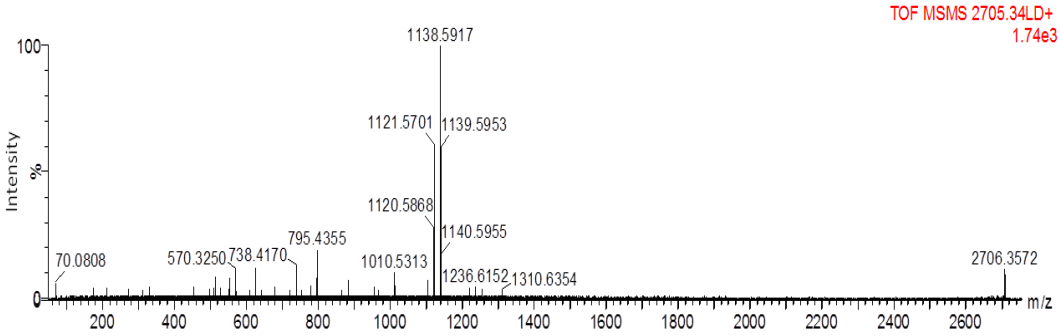



**Figure 6. Database searching of the fragment spectra produced following targeting of the peak 2705.3 was unable to make an identification.** The figure shows an example, representative of a number of peaks targeted for fragmentation, that produced a good series of fragment ions but for which we were unable to obtain a significant match. The spectrum after background subtraction and peak detection illustrates a series of ions which were used to perform a Mascot search against the Swissprot and Unihorse databases. No significant identification was made.

Results from peak lists taken from the DF1 spectra of each comparison targeted identified *m/z* 2414.9 (collectin-43), with a significant Mascot score ($p<0.05$) alone.

Results for targeted fragmentation experiments using masses previously identified in human studies are shown in Table 3.

| Mass | Protein | MASCOT Score | Peptide sequence |
|---|---|---|---|
| 2325.3 | aggrecan | 43 | R.VSLPNYPAIPTDATLELQNLR.S |
| 1312.7 | biglycan | 37 | K.IQAIELEDLLR.Y |
| 2027.2 | biglycan | 47 | K.NHLVEIPPNLPSSLVELR.V |
| 2196.2 | cartilage intermediate layer protein | 67 | R.FLPSEQIQGVVVSAINLEPR.A |
| 2256.1 | cartilage oligomeric matrix protein | 86 | R.FYEGPELVADSNVVLDTTMR.G |
| 1679.8 | collagen alpha-1(II) chain | 75 | K.DGETGAAGPPGPSGPAGER.G |
| 1556.9 | fibromodulin | 42 | R.SLILLDLSYNHLR.K |
| 1955.1 | fibromodulin | 74 | K.IPPVNTNLENLYLQGNR.I |
| 1401.7 | fibronectin | 68 | K.HYQINQQWER.T |

**Table 3. A number of ECM peptides were identified following targeted MS studies on cartilage slices.** Subsequent to MS targeted fragmentation studies and database searching implemented by Mascot using the Unihorse database, the peptides included in this table were identified. All Mascot scores were significant ($p<0.05$). The amino acid sequence of each peptide is included.

**d. Semi quantitative analysis**

The peak distribution intensity differences in young, old and OA cartilage of the identified peptides were semi-quantified using images produced by Biomap software. Whilst there was no difference in the distribution of the majority of the peptides between groups including collagen type II, a significantly different distribution of peptides for COMP; *m/z* 2256.1 and biglycan 2027.2 was observed (Figure 7). As the intensity of the COMP peptide FYEGELVADSNVVLDTTMR was increased in old cartilage it could be a tentative marker of cartilage ageing. Furthermore some of the peaks identified and then imaged with Biomap were demonstrated by DA, to be expressed in the same manner, even at low loadings, for example, the COMP peptide *m/z* 2256.1; FYEGELVADSNVVLDTTMR between old (mean intensity 0.05±0.001) and OA (mean intensity 0.02±0.0002), and old and young (mean intensity 0.03±0.001) plus the biglycan peptide *m/z* 1312.7; IQAIELEDLLR between old (mean intensity 0.075±0.002) and OA (mean intensity 0.038±0.001).

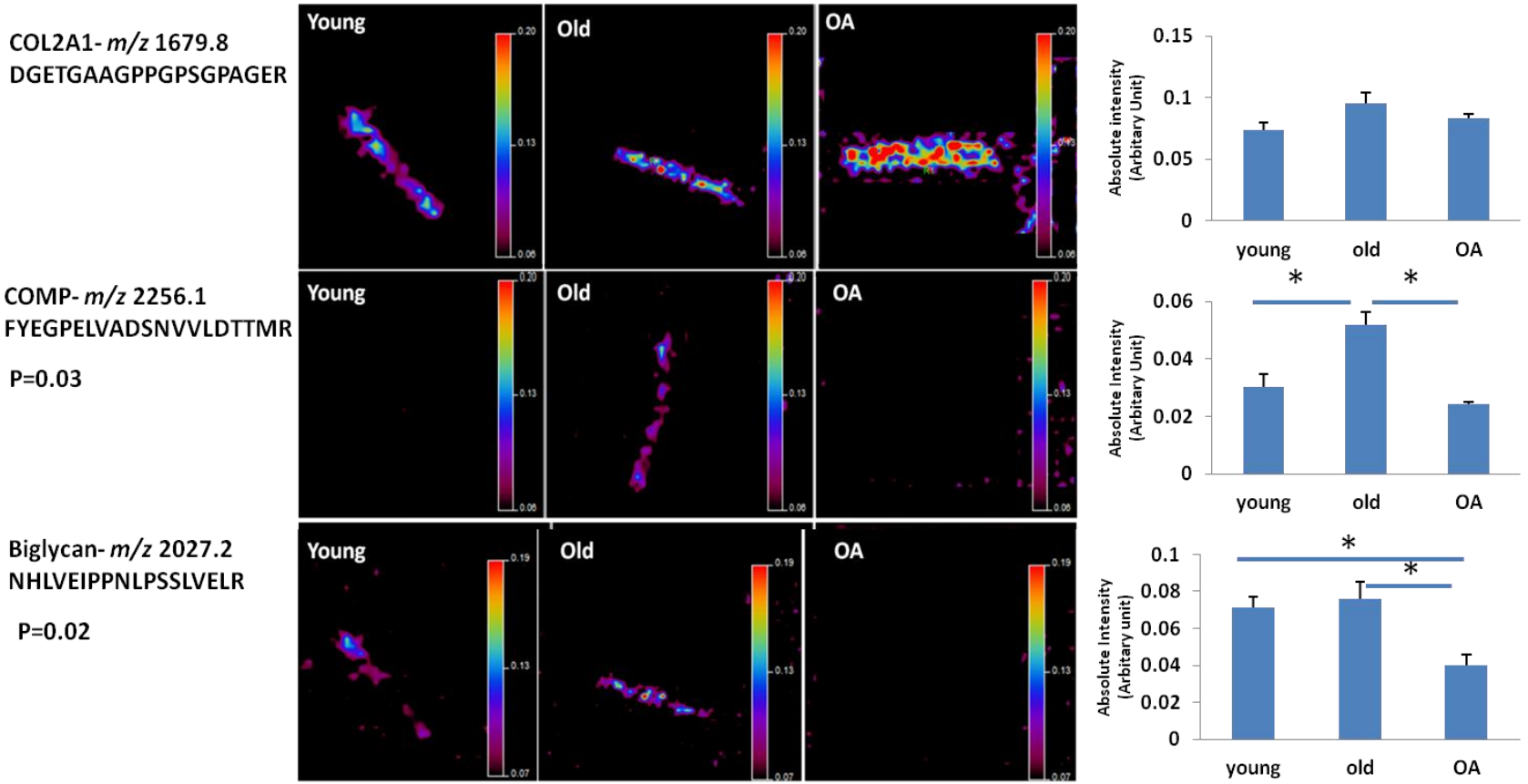



**Figure 7. Significant differential imaging of some ECM peptides.** Biomap was used to quantify the differences in the peptide peak intensity of *m/z* 1679.8, *m/z* 2256.1 and *m/z* 2027.2 between young, old and OA cartilages. Samples representative of each group are illustrated here. Scale bar shows normalised intensities to 190 *m/z* matrix peak. In the scale bar red represents the highest signal. Histograms shows the mean peak intensities ±SEM, n=3 for each peptide. There was no change in the distribution of the type II collagen peak *m/z* 1679.8. However a significant difference was evident for the COMP peptide *m/z* 2256.1 (p=0.03 young versus old and old versus OA) and the biglycan peptide *m/z* 2027.2 (p=0.02, young versus OA and old versus OA).

Further analysis was undertaken comparing the results in this study to those previously undertaken in human cartilage (Cillero-Pastor et al., 2012a). Interestingly distinctive intensity differences were observed for the previously identified fibromodulin peptide ELHLDHNQISR; *m/z* 1361.7 in man (Cillero-Pastor et al., 2012a). This peptide is homologous to the horse and was significantly reduced in OA equine cartilage (Figure 8). However there were no significant differences for the other fibromodulin peptides identified in profiling experiments with *m/z* 1557.0 or 1955.2 between categories identified by either DA or following peak intensity analysis and statistical testing (data not shown).

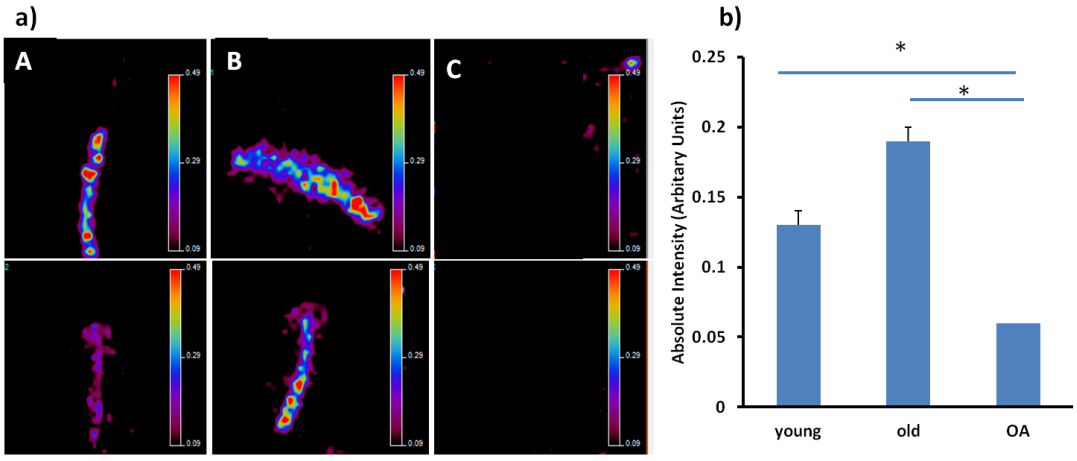



**Figure 8. A significant reduction in peak intensity of peptide m/z 1361.7 in OA cartilage was demonstrated** a) Biomap was used to quantify the differences in the peptide peak intensity of *m/z* 1361.7 between A; young, B; old and C; OA cartilages. Two donors representative of each group are illustrated here. Scale bar shows normalised intensities to 190 *m/z* matrix peak. b) Histogram shows the mean peak intensities ±SEM, n=3. * represents p<0.001, young versus OA, old versus OA and old versus OA.

First the differences found in the distribution of a human OA marker (Cillero-Pastor et al., 2012a); with an *m/z* 1349.6, identified previously as a fibronectin peptide and homologous to the horse sequence was visualised. The highest intensity differences were between young (mean intensity 0.11±0.009) and OA (mean intensity 0.36±0.12) samples (p=0.018, young versus OA). However the intensity difference between old (0.12±0.007) and OA was also significantly different (Figure 9a and b).

Interestingly a peptide *m/z* 1401.7 identified as fibronectin and previously nominated as a human OA marker (Cillero-Pastor et al., 2012a) was visualised in young, old, and OA samples. The highest intensity differences were between young (mean intensity 0.08±0.009) and OA (mean intensity 0.17±0.11) samples (p=0.02, young versus OA) (Figure 9c and d).

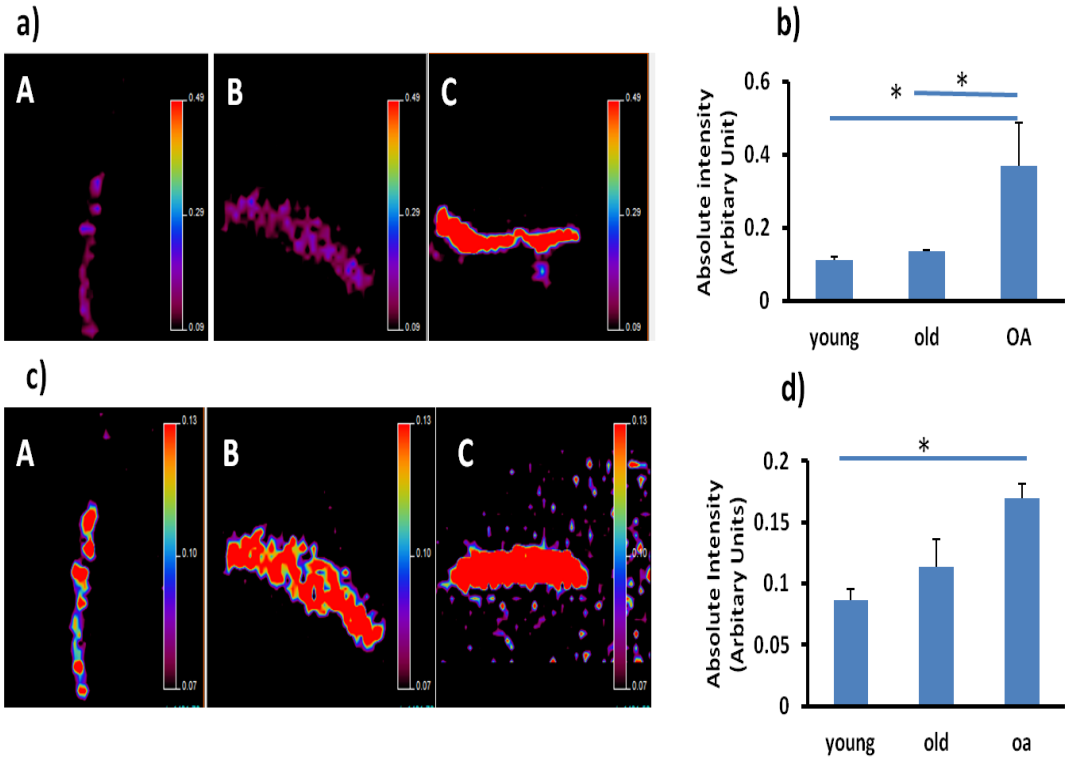



**Figure 9. Significant differential imaging *m/z* 1349.6 and *m/z* 1401.7 in cartilage.** a) Biomap was used to quantify the differences in the peptide peak intensity of a) *m/z* 1349.6 and c) *m/z* 1401.7 between A; young, B; old and C; OA cartilages. Samples representative of each group are illustrated here. Scale bar shows normalised intensities to 190 *m/z* matrix peak. b) Histogram shows the mean peak intensities ±SEM, n=3 for b) *m/z* 1349.6 and d) *m/z* 1401.7 * represents p<0.05; *m/z* 1349.6 young versus OA and old versus OA and *m/z* 1401.7 young versus OA.

Furthermore the peptides with *m/z* 1366.5, 1693.0, 1905.8 were demonstrated by DA as being at greater levels in OA. Of these the peptide *m/z* 1366.5 was found to be significantly imaged at greater intensities in all OA samples (mean intensity 1.00±0.03, p<0.001, young or old versus OA). Mean intensity for young was 0.05±0.003 and old 0.05±0.04 (Figure 10). The mean intensity values for each group

are shown in Figure 10b. In addition the peaks *m/z* 2415.9 and 2414.9 (identified as collectin-43 in fragmentation studies) were differentially expressed in all young samples.

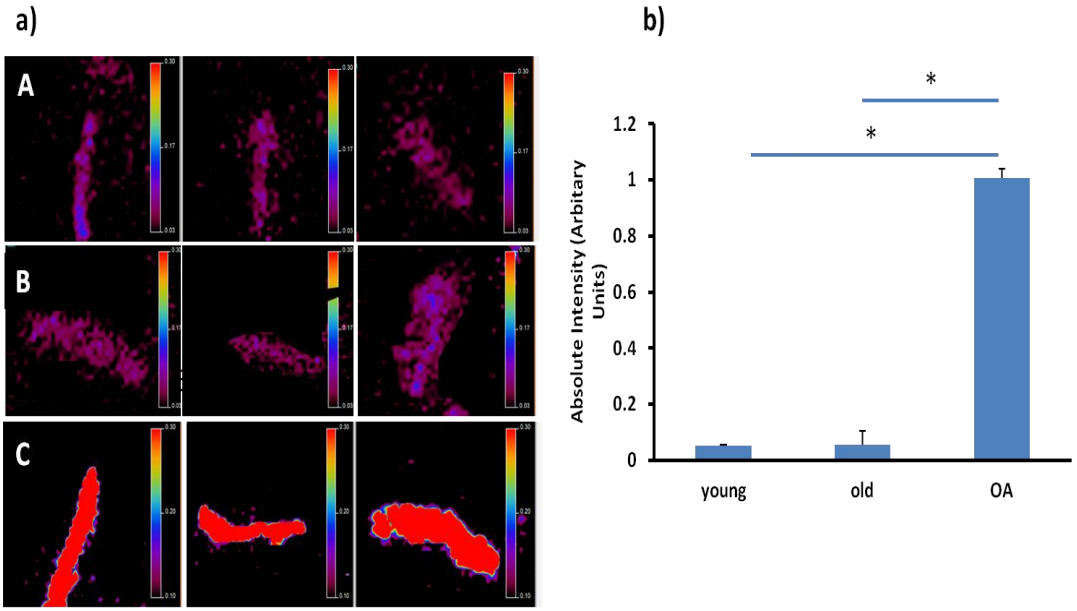



**Figure 10. Distribution and intensity of the tentative OA marker peptide *m/z* 1366.5.** a) The different intensities and distributions of *m/z* 1366.5 peptides in representative samples of A; young, B; old and C; OA samples after MALDI-IMS experiments (n=3). Scale bar shows normalized intensities. Saturation of the images in panel C; OA samples was evident in order to visualise the intensity in young and old samples. Background was evident in panel B but the intensity values calculated were measured creating ROI in order to avoid measurement of other molecules with the same mass that were outside the tissue. b) Histogram represents the mean peak intensities ±SEM, n=3. * represents $p<0.001$, young versus OA and old versus OA.

Finally we measured the intensity of the distribution of peptides in different areas of the cartilage in young, old and OA cartilage in order to assess if the distributions were homogenous. There was an absence of heterogeneity of peptide distribution in all samples.

**Discussion**

MALDI-IMS is a powerful tool to study molecular distributions at tissue surfaces. In this study we utilised recently implemented novel techniques to identify and spatially resolve peptides in ageing and OA cartilage. This technique offers the ability to distinguish peptide patterns within specific locals in cartilage, for instance between the superficial and deep layers (Cillero-Pastor et al., 2012c). The technique has been used to localise molecules such as peptides and lipids in diseases such as cancer (Rauser et al., 2010) in order to provide biomarkers (Stauber et al., 2008) and treatment targets. Whilst a recent study investigated the peptide (Cillero-Pastor et al., 2012a) and lipid (Cillero-Pastor et al., 2012b) distribution in normal and OA cartilage in this study we wished to elucidate whether changes in spatial distribution and presence of peptides occurred in ageing and between ageing and OA.

The samples used in this study were semi-quantitatively assessed macroscopically and microscopically and the OA samples presented with mild arthritic changes associated with early OA as demonstrated by the relatively low Mankin's scores. This was noted as previous studies in human OA cartilage used samples from total knee arthroplasty which would represent severe, late stage OA changes. Despite this, there was a marked difference between the peptides identified in old versus OA group as well as young versus old groups.

Following DA it was demonstrated that the peptide profile of young, old and OA horse cartilages could be distinguished. Interestingly when DA was applied to all three groups (young, old, OA) together there was a large contribution of the old samples to the negative part of DF1 indicating that peptides within old samples were also present in OA samples. This would suggest age-related changes in peptides occur which do not contribute to disease related changes.

Initial profiling experiments were undertaken to determine which proteins could be identified directly from the cartilage. A limited number of identifications were made

when compared to proteomics studies undertaken on equine cartilage extracts or cartilage explant media (Peffers et al., 2012). There were a number of reasons for this. First there was no prior extraction using a chaotropic agent such as guanidinium chloride. Chaotropes disrupt water interactions, promote hydrophobic protein and peptide solubilisation and breakdown the higher order protein structure all of which enable peptide to be more readily analyzed. Separation, depletion, partitioning or enrichment techniques were not used before analysis as this is not practical when MALDI-IMS is used to localize proteins in the tissue. Finally, as cartilage consists primarily of a small number of ECM proteins these can mask the mass spectrum of lower abundance proteins.

Following profiling experiments both Swissprot mammalian and the Unihorse databases were searched using Mascot using peptide mass fingerprints obtained. It was found that the Swissprot database gave the best cross-species matches in terms of scores and number of significant peptides identified, even though these were identified from other mammalian species. When the sequence homology for each peptide was checked with the equine sequence they were found to be the same for all the peptides logged in this study. This is because the peptides have an othologous sequence; their homologous sequences arose from a common ancestral gene during speciation. The equine database is not as well annotated as some others, for instance the human and mouse, and this may account for these differences. Interestingly following targeted fragmentation when the resulting files were searched using Mascot both databases gave similar results. This was probably due to the MS acquisition method differences producing different data, so that subsequent to a targeted fragmentation approach the sequences derived could be matched more definitively to the horse database.

The distribution of peptides in young, old and OA cartilage was examined for the first time. Of the proteins identified in cartilage there were unique peptides representing extracellular matrix proteins such as, COMP, matrilin-3, type II collagen, biglycan and fibromodulin.

The three major findings in this study were the identification of potential degradation sites in OA cartilage in COMP, biglycan and fibromodulin together with a number of potential markers of both age-related and disease-associated changes in cartilage. COMP is a major ECM protein with a role in cartilage structural integrity through its collagen I, II, IX and chondrocyte binding capacity, interaction with other ECM proteins including matrilin-3 (Mann et al., 2004) and role in fibrillogenesis (Halasz et al., 2007). It has been proposed as a biomarker for arthritis (Tseng et al., 2009). Interestingly the peptides identified were from the collagen binding C-terminal of COMP which others have demonstrated is important in intra and extra cellular processes (Briggs and Chapman, 2002). Indeed mutations in the genes encoding COMP and matrilin-3 result in multiple epiphyseal dysplasias (Briggs and Chapman, 2002; Chapman et al., 2001). Measurement of intact COMP and fragments thereof in synovial fluid or serum correlates to cartilage destruction in rheumatoid arthritis (RA) and OA patient studies (Saxne and Heinegard, 1992). Whilst interestingly studies in rats have found the plasma levels of COMP were age dependant (Wester et al., 2003). Using Biomap to determine the peak intensities between different groups there were differences in the distribution of the COMP peptide FYEGELVADSNVVLDTTMR. Its presence exhibited a pronounced decrease in both young and OA compared to old cartilage. This peptide is located in the C-terminal end of COMP which binds collagen I, IX and II, and regulates fibril formation (Halasz et al., 2007). In our previous studies following interleukin -1 stimulation of mature equine cartilage explants we identified a neopeptide which indicated possible degradation of this peptide at $Asn^{712}$ – $^{713}Thr$ (manuscript 4). This would seem to suggest that in OA there is degradation within this peptide resulting in the reduced expression demonstrated when OA cartilage was imaged. Whilst in young cartilage it may represent reduced synthesis or cartilage remodelling. The peptide represents a possible marker of age but not disease related changes in cartilage ECM.

The study also demonstrated a reduction in some peptides for biglycan and fibromodulin, members of the small leucine rich repeat proteoglycan family with important collagen binding properties (Hedbom and Heinegard, 1989; Wiberg et al.,

2003). Structural changes related to ageing are evident in biglycan, were there appears to be a cleavage in the amino terminal domain resulting in a 'no-glycan' biglycan as the terminal peptide containing the glycosaminoglycan chain separates from the protein core (Roughley et al., 1993). Results indicate a reduction in the presence of the mid region biglycan peptide NHLVEIPPNLPSSLVELR in OA cartilage. This either represents a reduction in the synthesis or degradation from disease. However as the other peptide identified in biglycan IQAIELEDLLR (also mid-region) does not reveal reduced intensity in OA it is more likely that it is lost following degradation. Interestingly our previous MS studies (manuscript 4) have identified a potential matrix metalloproteinase 3 (MMP-3) and a disintegrin and metalloproteinase with thrombospondin motifs 4 (ADAMTS-4) driven cleavage site within this peptide at $Ser^{164}$- $^{165}$ Leu. It is hypothesised that in OA there is cleavage at this site by MMPs resulting in a loss of this mass and so a reduction in peak intensity. Fibromodulin peptides were identified and their distribution imaged. One of a number of peptides, identified in MALDI-IMS studies of human cartilage (Cillero-Pastor et al., 2012a), and with the same mass in the horse investigated was the fibromodulin peptide ELHLDHNQISR. This was significantly reduced in OA samples despite the distribution of other fibromodulin peptides identified being unchanged. This would indicate that there is possibly degradation of this peptide as opposed to a reduction in synthesis of fibromodulin. Interestingly again our previous studies (manuscript 4) have identified a potential ADAMTS-4 cleavage site in this peptide at $Asn^{167}$–$Gln^{168}$. ADAMTS-4 is a pertinent enzyme in the pathogenesis of OA and although fibromodulin has been previously identified as a substrate for ADAMTS-4 (Gendron et al., 2007), this was at the Tyr $^{44}$–$Ala^{45}$ bond (Fushimi et al., 2008). Degraded fragments of the core fibromodulin protein have been observed in OA cartilage (Roughley et al., 1996) and with age (Cs-Szabo et al., 1995). Removal of this portion of fibromodulin would result in weaker interactions of collagen fibres to surrounding structures. This illustrates the potential usefulness of MALDI-IMS in identifying and spatially resolving novel cleavage sites with pathological relevance. Indeed this study would indicate that cleavage of the fibromodulin peptide ELHLDHNQISR and the biglycan peptide NHLVEIPPNLPSSLVELR

is disease and not age related. Insights such as this may aid in the understanding of the age-related but not age-distinct disease OA.

A number of tentative OA markers were detected. By means of Biomap software we quantified the intensity of the differences in abundance of peptides with *m/z* 1366.5 and 1349.6 in healthy young, old and OA samples and found their abundance increased in OA as determined by peak imaging. The *m/z* 1349.6 had been identified from human studies (Cillero-Pastor et al., 2012a) as being derived from fibronectin and its sequence homolog confirmed between the horse and man. Furthermore we identified the fibronectin peptide *m/z* 1401.7 as being more abundant in OA cartilage in agreement with others (Cillero-Pastor et al., 2012a). Fibronectin, a ECM glycoprotein, and fibronectin fragments have been associated with enhanced levels of catabolic cytokines and up regulation of MMPs involved in both normal homeostasis and arthritic diseases (Homandberg, 1999). Fibronectin fragments and fibronectin-aggrecan complexes have previously been suggested as biomarkers of OA (Scuderi et al., 2011; Zack et al., 2006). Therefore these peptides may provide promising biomarkers of OA as they are not affected by age-related changes. Fibronectin may provide a key species for potential diagnostic and drug targets. The increase in some fibronectin peptides identified in OA cartilage is most likely to be due to an increase in synthesis. This is because not only are these peptides found along the whole fibronectin sequence but others have demonstrated, in human OA cartilage similar findings which were validated using immunohistochemistry (Cillero-Pastor et al., 2012a).

Whilst a number of studies have interrogated ageing cartilage in order to elucidate the underlying mechanisms in its ageing and how it may contribute to OA none have used MALDI-IMS. In this study we were able to identify peptides and peaks with the potential to differentiate between ageing and OA related changes such as the COMP peptide FYEGPELVADSNVVLDTTMR imaged predominantly in old cartilage and the marker of young cartilage, peak 2414.9 identified as collectin-43. Collectin-43 is a C-type serum lectin with collagenous regions and a member of the collectin family of soluble proteins that are effector molecules in innate immunity

(Holmskov et al., 1994). Biglycan and decorin have been identified as binding collectin-43 and may have an important role in the resolution of C1q-mediated inflammatory processes in cartilage. Biglycan and decorin may down-regulate proinflammatory effects mediated by the collectins (Groeneveld et al., 2005). The identification of this protein in young cartilage is interesting as although there is evidence for ECM components such as fibronectin and hyaluronan to act as toll-like receptors in the innate immune response there was no evidence of disease from histology. It would be interesting to validate this and other findings with immunohistochemistry studies. The presence of collectin-43 requires further investigation to determine what other roles it may perform in normal cartilage physiology.

The protein melanoma inhibitory activity protein 3 (MIA3) was also identified for the first time in MALDI-IMS studies of cartilage. This is a secreted protein expressed by chondrocytes with a fundamental role in the maintaining the chondrocyte phenotype (Bosserhoff et al., 1997). A similar distribution of a peptide with the same mass as MIA3; *m/z* 1186.74 in young, old and OA tissue was evident (data not shown) which was not surprising as even mature chondrocytes secrete MIA3 (Lougheed et al., 2001).

Specific peaks identified from DA loading plots in young, old and OA cartilage slices were targeted for fragmentation in order to obtain the sequence and hence identification of the most differentially expressed peptides between each group. DA loading plots demonstrate a combination of peaks, showing the specific peak profile for each condition. Not all peaks produced fragmentation profiles sufficient for identification with the Mascot algorithm. In some cases this was because parent ion intensities were not very strong. However others peptides produced good fragmentation but were not identified. Therefore we undertook ion mobility studies in an attempt to determine whether when we targeted specific masses for fragmentation we were undertaking tandem MS on more than one peak. Ion mobility MS uses the ion's mass, charge, size and shape (the ion mobility), to determine the migration time to the detector leading to the ability to distinguish

different analyte species (Kanu et al., 2008). The analysis of the ion mobility study data found that for peak 1366 we were indeed targeting a single peak and therefore although we had good MS/MS spectra for many peaks we assume the fragments produced did not correspond to a unique peptide (data not shown). Further studies in the future either using emerging techniques such as desorption electrospray ionization (Wiseman et al., 2006) or laser ablation inductively coupled plasma mass spectrometry (Becker et al., 2010), higher resolution MS instruments or combining separation methodologies prior to MS may help in the identification of these peaks.

There were a couple of key differences between this study and a previous MALDI-IMS study in end stage OA human cartilage (Cillero-Pastor et al., 2012a). Although previous studies enabled differential peptide resolution between superficial and deep layers of cartilage to be identified, using equine cartilage from the metacarpal joint this was not possible. This was due to the thickness of the equine cartilage from the metacarpophalangeal III bone measuring 760±131μm (Brommer et al., 2005), considerably thinner than the human cartilage used (1.7-2.5mm (Shepherd and Seedhom, 1999)). Thus the ∼90μm pixel size used meant that the equine cartilage was only about 6-9 pixels thick. The cartilage from the stifle joint of the horse varies in thickness from 1760-2215μm (Frisbie et al., 2006) may provide an alternative source of cartilage for studies attempting to resolve changes in superficial and deep layers of equine cartilage. Secondly, there was no evidence for heterogeneity of peptide distribution that had been previously identified in OA cartilage. This was probably because the former study used severe OA cartilage whilst here only mild histological changes were evident.

With resolving power of MALDI-IMS certain to increase in the future and improved methods to identify peptides *in-situ* further MALDI-IMS studies using a greater age range of cartilage perhaps including foetal, juvenile together with a series of mature ages may help in understanding why ageing cartilage is more prone to OA.

**Conclusions**

IMS is becoming an established tool for imaging complex biological samples such as cartilage. The *ex-vivo* imaging of aged and diseased cartilage provided 'label-free' and stain-free information about its biomolecular composition. MALDI-IMS on cartilage sections provides a valuable approach for the proteomic investigations of ageing and diseased cartilage.

**Acknowledgements**

Mandy Peffers was supported by a Wellcome Veterinary Integrated Research Fellowship. This work was funded by an EU Farm Animal Proteomics COST initiative (COST Action FA1002).

# Manuscript 6

MJ Peffers[1*], X Liu[2], PD Clegg[1]

**Corresponding Author:** M J Peffers[1]

[1]Dept.of Musculoskeletal Biology, Institute of Ageing and Chronic Disease, University of Liverpool, Leahurst, Chester High Road, Neston, Wirral, CH64 7TE; peffs@liv.ac.uk, pclegg@liv.ac.uk

[2] Centre for Genomic Research, Institute of Integrative Biology, Biosciences Building, Crown Street, University of Liverpool, Liverpool L69 7ZB; xian.liu@liv.ac.uk

**Title: Transcriptomic Signatures in Cartilage Ageing**

**Running Title**: **RNA-Seq Application to Ageing Equine Cartilage**

**Key Words:** ageing, cartilage, ECM, RNA-Seq, Wnt

## Transcriptomic Signatures in Cartilage Ageing


### Abstract

Age is an important factor in the development of osteoarthritis. Whilst some insights into cartilage ageing have been discovered from microarray studies these were not able to elucidate the full range of transcriptomic phenotype of chondrocytes. RNA-Seq is a powerful technique for the interrogation of large numbers of transcripts including non-protein coding RNAs. The aim of the study was to characterise molecular mechanisms associated with age-related changes in gene signatures. RNA sequence libraries were prepared from young and old equine cartilage following ribosomal RNA depletion. Ingenuity Pathway Analysis enabled networks, functional analyses and canonical pathways from differentially expressed genes to be determined. There were 396 genes differentially expressed with the criteria $p<0.05$ and $\pm 1.4$ $\log_2$ fold change; 93 were at higher levels in the older cartilage and 303 were at lower levels in the older cartilage. These represented mRNAs, small non-coding RNAs, pseudogenes, and microRNAs. An over-representation of genes with reduced expression relating to extracellular matrix, degradative proteases, matrix synthetic enzymes, cytokines and growth factors were identified in cartilage derived from older donors compared to young. This points to an age-related failure of matrix, anabolic and catabolic cartilage factors. In addition there was a reduction in Wnt signalling in ageing cartilage. The study has increased our knowledge of transcriptional networks in cartilage ageing by providing a global view of the transcriptome.


### Introduction

Ageing presents huge challenges for society because whilst life span increases, the quality of life faced by individuals in old age is often poor (Beard et al., 2011). The musculoskeletal system in particular is severely affected by the ageing process, with many tissues undergoing changes that lead to loss of function and frailty. Articular

cartilage is susceptible to age related diseases, such as osteoarthritis (OA), although it is not an inevitable result of ageing and is a consequence of a complex inter-relationship between age and further predisposing factors such as obesity (Cooper et al., 2000), injury (Samilson and Prieto, 1983), genetics (Ma et al., 2011) and anatomical configuration (Felson, 2004).

A number of studies have interrogated ageing cartilage in order to elucidate the underlying mechanisms in its ageing and how it may contribute to OA. An age-related reduction in response to insulin-like growth factor in rats resulted in a decline in synthetic activity (Martin et al., 1997). Furthermore, using whole mouse joints, Loeser *et al.* (Loeser et al., 2012) demonstrated that there was a reduction in extracellular matrix gene expression in older sham-operated mice following surgical destabilization of the medial meniscus. A characteristic of ageing articular cartilage is the reduction in the number of chondrocytes within the tissue (Adams and Horton, 1998; Aigner et al., 2004a) and there is evidence of chondrocyte senescence (Lombardi et al., 2005). It is believed that chondrocyte senescence is one of the causes of a decline in the ability of chondrocytes to respond to growth factors; resulting in the anabolic/catabolic imbalance evident in OA (Mueller and Tuan, 2011). One of the consequences of cell senescence is an alteration in cell phenotype (Campisi, 2005) characterised by increased production of cytokines and growth factors. The increase in ageing chondrocytes expressing this phenotype has been proposed to contribute to cartilage ageing and, given the rise in cytokine production in OA, could directly connect ageing to OA development (Loeser, 2010). Furthermore there is evidence for the role of oxidative damage in cartilage ageing from reactive oxygen species (ROS) (Jallali et al., 2005; Loeser et al., 2002), which can result in damage to cartilage DNA (Chen et al., 2008), whilst a link between ROS and development of OA has also been established (Kurz et al., 2002). Hence the outcome of ageing on chondrocyte function is an inability to maintain homeostasis when stressed.

Whilst some insights into cartilage ageing have been learnt from transcriptome profiling studies in ageing joints using microarrays (Loeser et al., 2012), this data did

not identify a specific chondrocyte phenotype associated with ageing alone. In addition microarrays suffer from limitations in coverage and sensitivity meaning that currently a significant part of the chondrocyte ageing transcriptomic phenotype is poorly defined. There is a need to examine and understand the processes and mechanisms involved specifically in cartilage ageing. We were interested in whether age affected gene expression in cartilage. Advances in high-throughput sequencing methodologies are allowing a new approach to studying transcriptomes: massively parallel sequencing of short reads derived from mRNAs; RNA-Seq (Wang et al., 2009). Compared with microarray technologies, RNA-Seq is demonstrated to enable more accurate quantification of gene expression levels (Matkovich et al., 2010). Furthermore RNA-Seq it is an effective approach for gene expression profiling in ageing tissues with a greater dynamic range and the ability to detect non-coding RNAs (de Magalhaes et al., 2010). Therefore we undertook an RNA sequencing experiment on young and old cartilage to characterize molecular mechanisms associated with age-related changes in gene signatures.

## Methods

### Sample collection and preparation

Full thickness equine cartilage from the entire surface of macroscopically normal metacarpophalangeal joints was collected from an abattoir. All samples were scored macroscopically to determine any pathological abnormalities of the distal condyles of the distal third metacarpus (Kawcak et al., 2008). For subsequent RNA-Seq experiments normal cartilage samples from four young horses; aged four years old and four old horses; greater than 15 years old were obtained.

### RNA extraction

Cartilage was prepared by pulverising into a powder with a dismembranator (Miko, S-Braun, USA) following freezing in liquid nitrogen prior to addition of Tri Reagent (Ambion, Warrington, UK). For RNA extraction the Guanidium-thiocyanate-phenol-chloroform technique, with ethanol extraction was used (Chomczynski and Sacchi, 1987), followed by purification using the RNeasy (Qiagen, Crawley, UK) column

technique incorporating a DNase treatment step (Ambion, Warrington, UK) according to the manufacturer's instructions. RNA was quantified using a Nanodrop ND-100 spectrophotometer (Labtech, East Sussex, UK) and assessed for purity by UV absorbance measurements at 260 and 280nm.

**RNA-Seq analysis- cDNA library preparation and sequencing**

Total RNA was analysed by the Centre for Genomic Research, University of Liverpool, for RNA-Seq library preparation and sequencing using the Illumina HiSeq 2000 platform. Total RNA integrity was confirmed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA. Ribosomal RNA (rRNA) was depleted from 8 total RNA samples using the Ribo-Zero™ rRNA Removal Kit (Human/Mouse/Rat EpiCentre, Madison, USA) following the manufacturer's instructions. cDNA libraries were prepared with the ScriptSeq v2 RNA-Seq library preparation kit (Epicentre, Madison, USA) using 50ng rRNA depleted RNA as starting material and following the manufacturer's protocols. Briefly, rRNA-depleted sample was fragmented using an RNA fragmentation solution prior to cDNA synthesis. Fragment size of the final libraries and pooled libraries was confirmed using the Agilent 2100 Bioanalyzer software in smear analysis function. Fragmented RNA was reverse transcribed using random-sequence primers containing a tagging sequence at their 5' ends. 3' tagging was accomplished using the Terminal-Tagging Oligo (TTO) which features a random nucleotide sequence at its 3' end, a tagging sequence at its 5' end and a 3'-blocking group on the 3'terminal nucleotide. The TTO randomly annealed to the cDNA, including to the 3' end of the cDNA. Purification of the di-tagged cDNA was undertaken with AMPure™ XP (Agencourt, Beckmann-Coulter, USA). The di-tagged cDNA underwent 15 cycles of amplification using PCR primer pairs that annealed to the tagging sequences of the di-tagged cDNA. Excess nucleotides and PCR primers were removed from the library using AMPure™ XP (Agencourt, Beckmann-Coulter, USA). The final pooled library was diluted to 8pmol before hybridisation. The dilute library (120µl) was hybridised on each of 3 HiSeq lanes.

**Data processing**

The 100bp paired-end reads obtained by RNA-Seq were compiled using manufacturer-provided pipeline software CASAVA 1.8.2. Reads were then aligned onto the equine chromosomes with TOPHAT 1.3.2 using default settings. Only uniquely mapped reads retained with less than two mismatches were used for analysis. Quality control of the reads in each lane was undertaken using FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc).

The R (version 2.15.1) Bioconductor package edgeR (version 2.13.0) (Robinson et al., 2010) was used to identify differentially expressed genes. EdgeR models data as a negative binomial distribution to account for biological and technical variation using a generalisation of the Poisson distribution model. Prior to assessing differential expression, data were normalised across libraries using the trimmed mean of M-values normalisation method (Robinson and Oshlack, 2010). Genes were deemed differentially expressed with a Benjamini-Hochberg false discovery rate (FDR)-corrected P-value $< 0.05$ and a fold change $\geq 1.4$ (Benjamini and Hochberg, 1995) using a generalised linear model (GLM) likelihood ratio test. Statistical analysis on mapped reads was undertaken with a custom Perl script. All sequence data produced in this study has been submitted to NCBI GEO under Array Express accession number E-MTAB-1386.

**Go ontology and Ingenuity Pathway Analysis (IPA)**

Due to the minimum amount of annotation for the equine genome, equine genes were converted to their human Ensembl orthologs prior to bioinformatics analysis. Functional analysis of age-related differentially expressed genes was undertaken in order to evaluate the differences in gene expression due to age. The functional analysis and clustering tool from the Database for Annotation, Visualisation, and Integrated Discovery (DAVID) (DAVID bioinformatics resources6.7) (Huang da et al., 2009) was used.

Networks, functional analyses, and canonical pathways were generated through the use of IPA (Ingenuity Systems, www.ingenuity.com) to the list of differentially

expressed genes with values adjusted $p<0.05$ and 1.4 $\log_2$ fold regulation. Gene symbols were used as identifiers and the Ingenuity Knowledge Base gene was used as a reference for pathway analysis. For network generation a data set containing gene identifiers and corresponding expression values was uploaded into in the application. Default settings were used to identify molecules whose expression was significantly differentially regulated. These molecules were overlaid onto a global molecular network contained in the Ingenuity Knowledge Base. Networks of 'Network Eligible Molecules' were then algorithmically generated based on their connectivity. The functional analysis identified the biological functions and diseases that were most significant to the data set. Right-tailed Fisher's exact test was used to calculate a p-values. Canonical pathways analysis identified the pathways from the IPA library of canonical pathways that were most significant to the data set.

**Real-time polymerase chain reaction (RT-PCR)**

Samples of RNA from the same pools used for the RNA-Seq analysis were used for RT-PCR. M-MLV reverse transcriptase and random hexamer oligonucleotides were used to synthesize cDNA from 1µg RNA (both from Promega, Southampton, UK) in a 25µl reaction. PCR was performed on 1µl 10x diluted cDNA, employing a final concentration of 3µm of each primer in 20ml reaction volumes on an ABI 7700 Sequence Detector using a SYBR Green PCR mastermix (Applied Biosystems, Warrington, UK). Exon-spanning primer sequences were used that had been validated in previous publications (Peffers et al., 2010; Radinsky et al., 1990; Taylor et al., 2009) or were designed for this study using Primer-Blast; National Centre for Biotechnology Information (NCBI) http://www.ncbi.nlm.nih.gov/tools/primer-blast/. BLAST searches were performed for all sequences to confirm gene specificity. Oligonucleotide primers were supplied by Eurogentec (Seraing, Belgium). Assays for four genes; glyceraldehyde-3-phosphate dehydrogenase (GAPDH), TATA box binding protein (TBP), beta-actin (ACTB) and 18 ribosomal RNS (18S) were selected as potential reference genes as their expression was unaltered in the RNA-Seq analysis. GAPDH was selected as the reference genes from the panel of reference genes by applying a gene stability algorithm (Vandesompele et al., 2002) using genorm$^{PLUS}$ (Biogazelle, Zwijnaarde, Belgium) (Hellemans et al., 2007).

Relative expression levels were normalised to GAPDH and calculated using the $2^{-\Delta Ct}$ method (Livak and Schmittgen, 2001). Standard curves were generated from five serial dilutions for each assay to confirm that all efficiencies were acceptable; within 5% of GAPDH and $R^2$ >0.98. Primers pairs used in this study are listed (Table 1). RT-PCR analysis data was $\log_{10}$ transformed to ensure normal distribution and then analysed using Student's T –test.

**Statistical Analysis**

The analyses were undertaken using S-Plus, SPSS and Excel software.

| Gene | Accession Code | Primer Sequence |
|---|---|---|
| GAPDH* | AF157626 | F: GCATCGTGGAGGGACTCA<br>R: GCCACATCTTCCCAGAGG |
| TBP* | XM_001502211 | F:TGCTGCTGTAATCATGAGGGTAA<br>R:TCCCGTGCACACCATTTTC |
| ACTB* | AF035774 | F:CCAGCACGATGAAGATCAAG<br>R:GTGGACAATGAGGCCAGAAT |
| 18S* | AJ311673 | F:GGCGTCCCCCAACTTCTTA<br>R:GGGCATCACAGACCTGTTATTG |
| RUNX2 | XM_001502519 | F:TCCCTGAACTCTGCACCAAG<br>R:GCCAGGTAGGAGGGGTAAGA |
| IL7R | NM_001081942 | F:GGCTATGCACAGAATGGAGACT<br>R:CAACTGGCTGTAGCACGAGA |
| SRPX | XM_001489643 | F:CTGAGAACAAGGGCG-TTGC<br>R:CCGGAGCGTTGAGTTTGC |
| ACSL5 | XM_001915998 | F:CCTGGGCTCCTATCTCTTGC<br>R:CGGAGATGATCCACTCTGGC |
| DKK | NM_001267802 | F:TAGAACCCTGGGACCTCTGG<br>R:GTGTCACTTTGCAAGCCTGG |
| ADAMTS4# | NM_001111299 | F:CAGCCTGGCTCCTTCAAAAA<br>R:CCGCAGGAATAGTGACCACAT |
| COL1A1* | O46388 | F:GACTGGCAACCTCAAGAAGG<br>R:CAATATCCAAGGGAGCCACA |
| COL2A1* | NM_001081764 | F:TCAAGTCCCTCAACAACCAGAT<br>R:GTCAATCCAGTAGTCTCCGCTCTT |
| COL10A1* | XM_001504101 | F:TGCCCAGTGGACAGGTTTCT<br>R:GTCTTTTCGTTTCTAGTCAGATTTTGAA |
| MMP1# | NM_001081847 | F:GGTGAAGGAAGGTCAAGTTCTGAT<br>R:AGTCTTCTACTTTGGAAAAGAGCTTCTCT |
| MMP13# | NM_001081804 | F:CTGGAGCTGGGCACCTACTG<br>R:ATTTGCCTGAGTCATTATGAACAAGAT |
| IL1β# | NM_001082526 | F: GAGCCCAATCTTCAACATCTATGG<br>R:CAGGCTTGGTAAAAGGACTTGGTAT |
| TNFα# | NM_001081819 | F:GCTCCAGACGGTGCTTGTG<br>R:GCCGATCACCCCAAAGTG |
| TGFβ# | NM_001081849 | F:CCCTGCCCCTACATTTGGA<br>R:CGGGTTGTGCTGGTTGTACA |

**Table 1. Gene primer sequences used in RNA-Seq validation.** *Denotes primer pairs previously published (Taylor et al., 2009) and # denotes primer pairs previously published (Barr, 2011).

## Results

### Preliminary analysis of RNA-Seq data

Eight libraries representing four animals from two groups; young and old (n=4 young and n=4 old) were prepared from cartilage mRNA. For all cartilage samples we performed RNA-Seq using 100bp paired-end sequencing on the Illumina HiSeq 2000 platform and approximately 116-235 million reads were obtained per sample. From these reads low quality reads were eliminated resulting in 7-58 million mapped reads (equal to 6.5-35% of the total reads). In total 3-49 million uniquely mapped read pairs were obtained per sample and aligned to the reference sequence of the equine genome (*Equus caballus*; EquCab2.56.pep, from ftp://ftp.ensembl.org/pub/current_fasta/equus_caballus/pep/).

Identical reads mapped to the same genomic position were retained as duplicates as these were potential real reads. The number of genes per read were normalised to 'reads per kilo base of exon model per million mappable reads' (RPKM); therefore the values were considered the final expression level for each gene (Mortazavi et al., 2008). Using *the Equus caballus* database, analysis demonstrated in total 16,635 genes (from a total of 25,180 genes) were expressed in cartilage, which represented 66% of the equine genome. This data was used for subsequent analysis and is comparable to other recent RNA-Seq studies (O'Loughlin et al., 2012).

### Age-related differential gene expression in cartilage

A multidimensional scaling plot (MDS) (Figure 1) revealed data was clustered tightly in two groups; one for 'older' donors and for 'younger' donors.

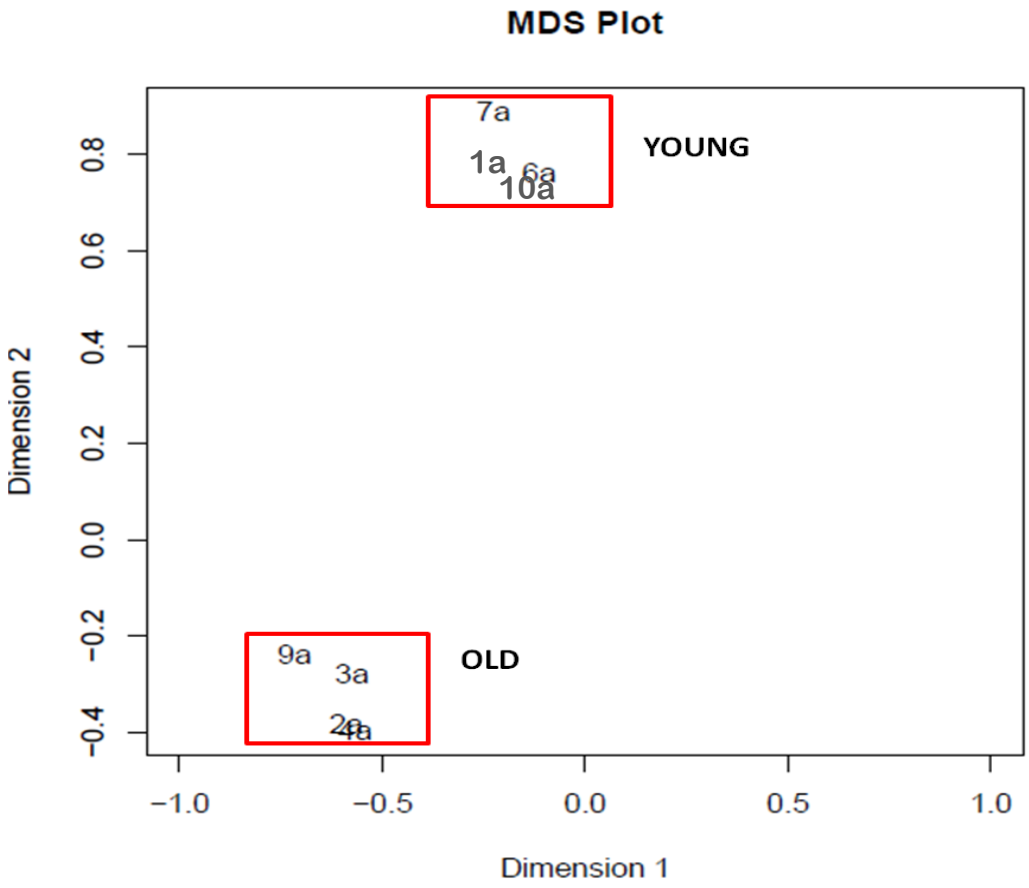



**Figure 1. Principal component analysis reveals the greatest variability in RNA-Seq data is due to age of the donor**.

Alterations in gene expression between young and old cartilage demonstrated significant age-related changes. There were 396 genes differentially expressed with the criteria $p<0.05$ and $\pm 1.4$ $\log_2$ fold change (Figure 2); 93 were at higher levels in the older cartilage and 303 were at lower levels in the older cartilage. Table 1 represents the top 10 genes most differentially expressed up and down in the young compared to the older horses.

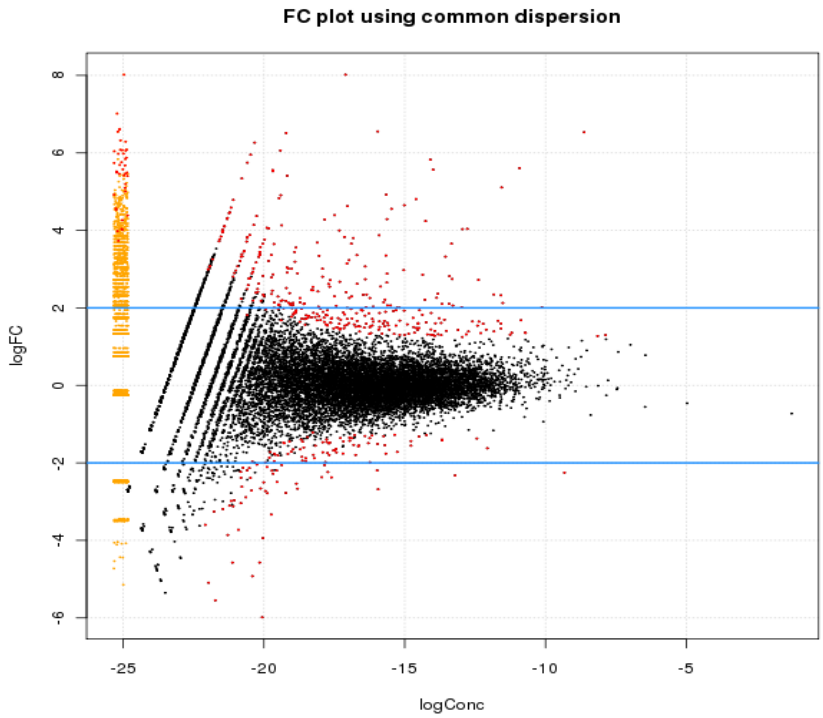



**Figure 2. A set of differentially expressed genes between young and old cartilage were discovered.** Using the common dispersion in edgeR 396 differentially expressed genes were identified with $p<0.05$ (marked in red). In order to enable the expression of all genes to be visualised simultaneously a smear plot was produced. The smear at the left-most edge allows visualisation of genes with zero counts in one of the groups. This was undertaken as if the total counts in one group are zero, the log-fold change is technically infinite, and the log-concentration is negative infinity.

| Gene symbol | Gene name | $log_2$ fold change | Adjusted p value |
|---|---|---|---|
| Genes with increased expression in old cartilage | | | |
| RORA | RAR-related orphan receptor B | -5.98 | 3.26252E-11 |
| SKA1 | spindle and KT associated 1 | -5.55 | 0.015509031 |
| LINGO1 | leucine rich repeat and Ig domain containing 1 | -5.09 | 0.020082666 |
| DKK1 | dickkopf homolog 1 | -4.92 | 0.000622559 |
| COCH | coagulation factor C homolog, cochlin (Limulus polyphemus) | -4.57 | 1.49E-04 |
| RELN | reelin | -4.57 | 1.87E-05 |
| TOX3 | TOX high mobility group box family member 3 | -3.86 | 0.004421009 |
| SLC22A3 | solute carrier family 22 (extraneuronal monoamine transporter), member 3 | -3.73 | 0.000467601 |
| FGF9 | fibroblast growth factor 9 | -3.33 | 4.16E-04 |
| SHCBP1L | SHC SH2-domain binding protein 1-like | -3.26 | 0.026646605 |
| Genes with decreased expression in old cartilage | | | |
| CPZ | carboxypeptidase Z | 32.09 | 1.60E-08 |
| C18orf8 | chromosome 8 open reading frame 4 | 31.09 | 1.40E-04 |
| SRPX | Sushi repeat-containing protein SRPX | 30.69 | 0.000193955 |
| CYP1A1 | cytochrome P450, family 1, subfamily A, polypeptide 1 | 30.68 | 1.19E-03 |
| AQP1 | aquaporin 1 | 30.62 | 1.19E-03 |
| PHEX | Phosphate regulating endopeptidase homolog, X-linked | 30.39 | 0.000354295 |
| EPHA5 | EPH receptor A5 | 30.16 | 0.015509031 |
| CTCFL | CCCTC-binding factor (zinc finger protein)-like | 30.15 | 0.00167332 |
| IL7R | interleukin 7 receptor | 30.14 | 0.006160656 |
| ACSL5 | acyl-CoA synthetase long-chain family member 5 | 30.13 | 0.02104673 |

**Table 1. Genes with the highest and lowest $log_2$ fold change when comparing RNA from young and old cartilage.** $Log_2$ fold change and adjusted p-values were determined in edgeR. The genes demonstrated are the 10 genes with highest and lowest expression in old compared to young cartilage samples.

The top 25 differentially expressed genes are represented in Figure 3. NCBI GEO under accession number E-MTAB-1386 contains a complete list of all genes mapped. Of the subset of 93 genes that were significantly higher in older donors 9 were small nuclear/nucleolar RNAs, 12 pseudogenes, 11 genes which were not identified and a single microRNA (miRNA); miRNA21. Thus 60 known protein coding genes were differentially expressed as higher in the older cartilage. Within the group where gene expression was lower in old compared to young cartilage seven were small nuclear/nucleolar RNAs, one was a pseudogenes and three were not identified. Small nuclear/nucleolar RNAs differentially expressed are featured in Table 2. Thus 352 were used in downstream DAVID and IPA analysis.

| Name | Family | Action | Target | Higher |
|---|---|---|---|---|
| SNORD113 | C/D BOX CLASS | site-specific 2'-O-methylation | not predicted to target rRNA or snRNA. | young |
| SNORA53 | H/ACA box class | pseudouridylation | not identified | young |
| SNORA79 | H/ACA box class | pseudouridylation | not identified | young |
| SNORA48 | H/ACA box class | pseudouridylation | 28SrRNA | young |
| SNORD12/SNORD16 | C/D BOX CLASS | site-specific 2'-O-methylation | 28srRNA, 18SrRNA | young |
| U2 splicesomal RNA | splicesome | complex of snRNA and protein subunits that removes introns from a transcribed pre-mRNA | | young |
| Small nucleolar RNA U89 | H/ACA box class | pseudouridylation | not identified | young |
| SNORA40 | H/ACA box class | pseudouridylation | 28SrRNA | old |
| SNORA5 | H/ACA box class | pseudouridylation | not identified | old |
| SNORA30/SNORA37 | H/ACA box class | pseudouridylation | 28SrRNA | old |
| U6 splicesomal RNA | splicesome | complex of snRNA and protein subunits that removes introns from a transcribed pre-mRNA | | old |

**Table 2. A number of small nucleolar (SNORD) and small nuclear (SNORA and splicesomal RNA) RNAs were identified as being differentially expressed (DGE) in ageing cartilage.** The class action and target of these RNAs are shown with higher DGE in young or old cartilage. U6 splicesomal RNA was identified in the data with two separate accession numbers.

**Age-related changes in important cartilage genes**

There was a reduction in the expression of 42 genes relating to ECM, degradative proteases, matrix synthetic enzymes, cytokines and growth factors in cartilage derived from older donors compared to young. In comparison there was an increase in only three ECM genes; COL10A1, COL25A1 and lubricin together with a single growth factor; fibroblast growth factor 9 in older donors (Table 3).

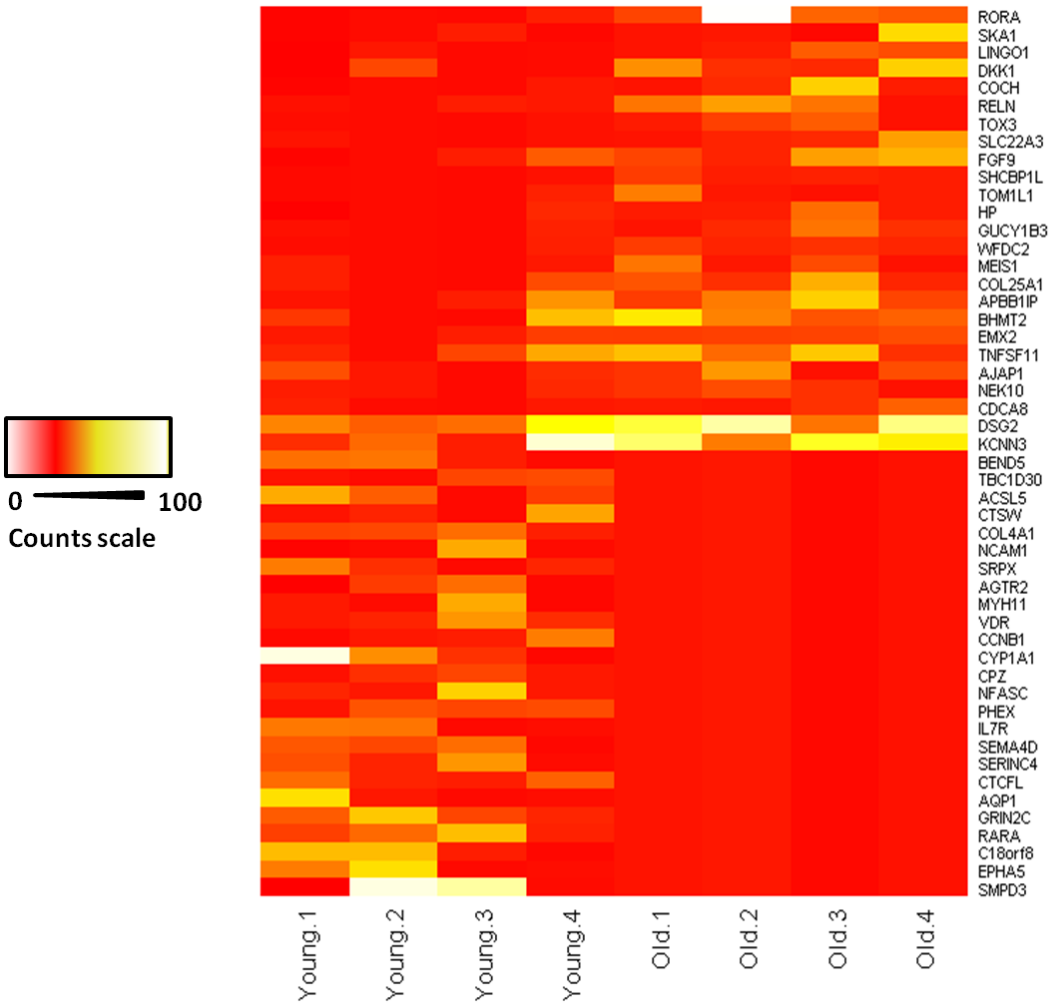



**Figure 3. Most differentially expressed genes in cartilage ageing**. The heat map illustrates the 25 most highly regulated up and down genes in cartilage. The counts represent raw counts for each donor. Significance was set at $p<0.05$ and ± 1.4 $\log_2$ fold change in gene expression based on mapped reads following normalisation and statistical testing in edgeR. Orange represents less counts and white represents a greater number of counts.

| DGE higher in old | | |
|---|---|---|
| **Gene class** | **Gene name** | **Gene symbol** |
| **ECM** | collagen, type X, alpha 1 | COL10A1 |
| | collagen, type XXV, alpha 1 | COL25A1 |
| | lubricin | CSPG4 |
| **Cytokine/GF** | fibroblast growth factor 9 | FGF9 |
| **DGE lower in old** | | |
| **Protease** | A disintegrin and metallopeptidase with thrombospondin 12 | ADAMTS12 |
| | A disintegrin and metallopeptidase with thrombospondin 2 | ADAMTS2 |
| | A disintegrin and metallopeptidase with thrombospondin 4 | ADAMTS4 |
| | matrix metallopeptidase 1 | MMP1 |
| | matrix metallopeptidase 13 | MMP13 |
| | plasminogen activator inhibitor-1 | SERPINE1 |
| | plasminogen activator, tissue | PLAT |
| **Matrix enzyme** | chondroitin sulfate synthase 3 | CHSY3 |
| | hyaluronan synthase 3 | HAS3 |
| | procollagen C-endopeptidase enhancer | PCOLCE |
| **ECM** | asporin | ASPN |
| | biglycan | BGN |
| | cartilage intermediate layer protein 2 | CILP2 |
| | chondroadherin | CHAD |
| | collagen alpha 1(V) chain | COL5A1 |
| | collagen, type I, alpha 1 | COL1A1 |
| | collagen, type I, alpha 2 | COL1A2 |
| | collagen, type II, alpha 1 | COL2A1 |
| | collagen, type III, alpha 1 | COL3A1 |
| | collagen, type IV , alpha 1 | COL4A1 |
| | collagen, type IV, alpha 5 | COL4A5 |
| | collagen, type IX, alpha 1 | COL9A1 |
| | collagen, type IX, alpha 2 | COL9A2 |
| | collagen, type VIII, alpha 1 | COL8A1 |
| | collagen, type XI, alpha 1 | COL11A1 |
| | collagen, type XII, alpha 1 | COL12A1 |
| | collagen, type XIII, alpha 1 | COL13A1 |
| | collagen, type XIV, alpha 1 | COL14A1 |
| | collagen, type XV, alpha 1 | COL15A1 |
| | collagen, type XVI, alpha 1 | COL16A1 |
| | fibulin 1 | FBLN1 |
| | fibulin-7 | FBLN7 |
| | matrilin 2 | MATN2 |
| | matrilin 4 | MATN4 |
| | procollagen V, alpha 2 | COL5A2 |
| | thrombospondin 2 | THBS22 |
| | thrombospondin 3 | THBS3 |
| **Cytokine/ GFs** | fibroblast growth factor 12 | FGF12 |
| | interleukin 11 | IL11 |
| | interleukin 8 | IL8 |
| | interleukin-1b | IL1B |
| | tumor necrosis factor, alpha-induced protein 3 | TNFAIP3 |

**Table 3. Older cartilage demonstrates more quiescent gene expression compared to young cartilage**. The table illustrates significant differential gene expression (DGE) in young and old cartilage of important cartilage extracellular matrix (ECM), cytokines and growth factors (GFs), proteases (causing cartilage degradation) and matrix enzymes (involved in matrix synthesis) (significance was set at $p<0.05$ and $\pm$ 1.4 $\log_2$ fold change in gene expression based on mapped reads following normalisation and statistical testing in edgeR). There was a considerable reduction in all these classes of cartilage genes in old compared to young donors.

**Gene ontology analysis of differentially expressed genes to characterise transcriptomic signatures in cartilage ageing**

DAVID analysis of all differentially expressed genes (DGEs) included annotations for cell adhesion and extracellular matrix (Supplementary table 1 on CD). The genes most differentially expressed, with reduced expression in cartilage from older donors included two involved in Wnt signalling; CPZ (carboxypeptidase Z); and C18orf8 (chromosome 8 open reading frame 4). Furthermore other genes involved in Wnt signalling; secreted frizzled-related protein 2 (SFRP,) Wnt 11 and the Wnt inhibitory factor-1 (WIF1) were also reduced.

Interestingly of the 93 genes expressed in higher levels in older cartilage one of the most highly regulated was a regulator of Wnt signalling; DKK1 (dickkopf homolog 1). DAVID analysis of this group revealed annotations for skeletal and cartilage development, and immune response.

**Differential expressed genes and network analysis**

Both sets of differentially expressed genes associated with ageing were analysed together in IPA with the following criteria; $p<0.05$ and $\log_2$ fold change ± 1.4. Network eligible molecules were overlaid onto molecular networks based on information from the IP knowledge database. Networks were then generated based on connectivity. Supplementary table 2 provided on the CD contains all identified networks and their respective molecules. Interesting age related features were determined from gene networks inferred. According to the top scoring network the differentially expressed genes were from connective tissue disorders, such as collagens; COL12A1, COL16A1, COL1A1, COL25A1 plus leucine rich repeat and Ig domain containing 1 (LINGO), transforming growth factor β-induced 68kDa and coclin (COCH) (Figure 4A).

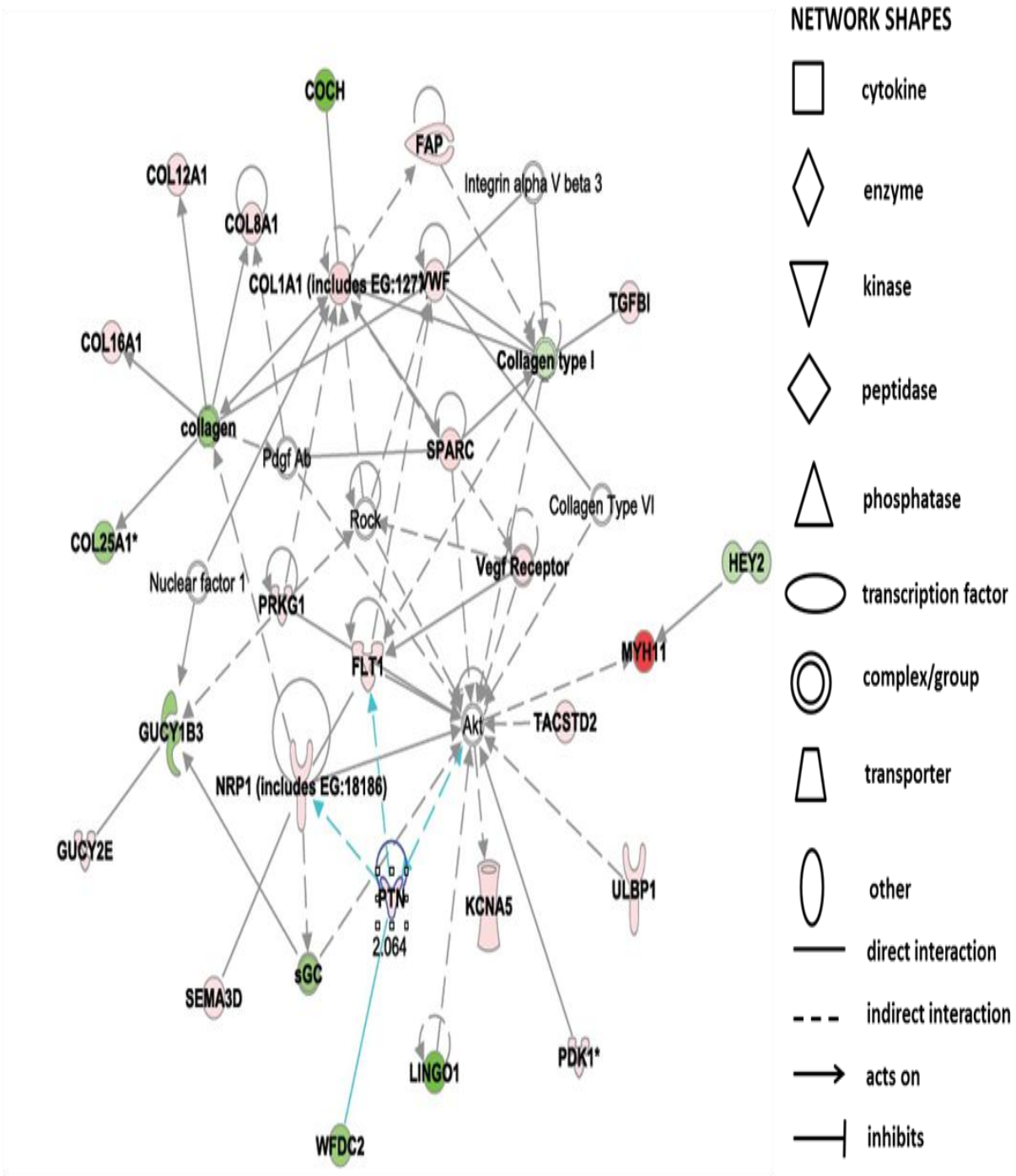



**Figure 4. Top-scoring networks derived from the 352 genes differentially expressed in ageing.** IPA identified connective tissue disorders as the principle associated network functions with scores of 43. The figure is a graphical representation between molecules identified in our data. The green nodes represent up-regulated and red nodes down regulated gene expression in older cartilage. Intensity of colour is related to higher fold change. Legend provides a key to the main features in the network

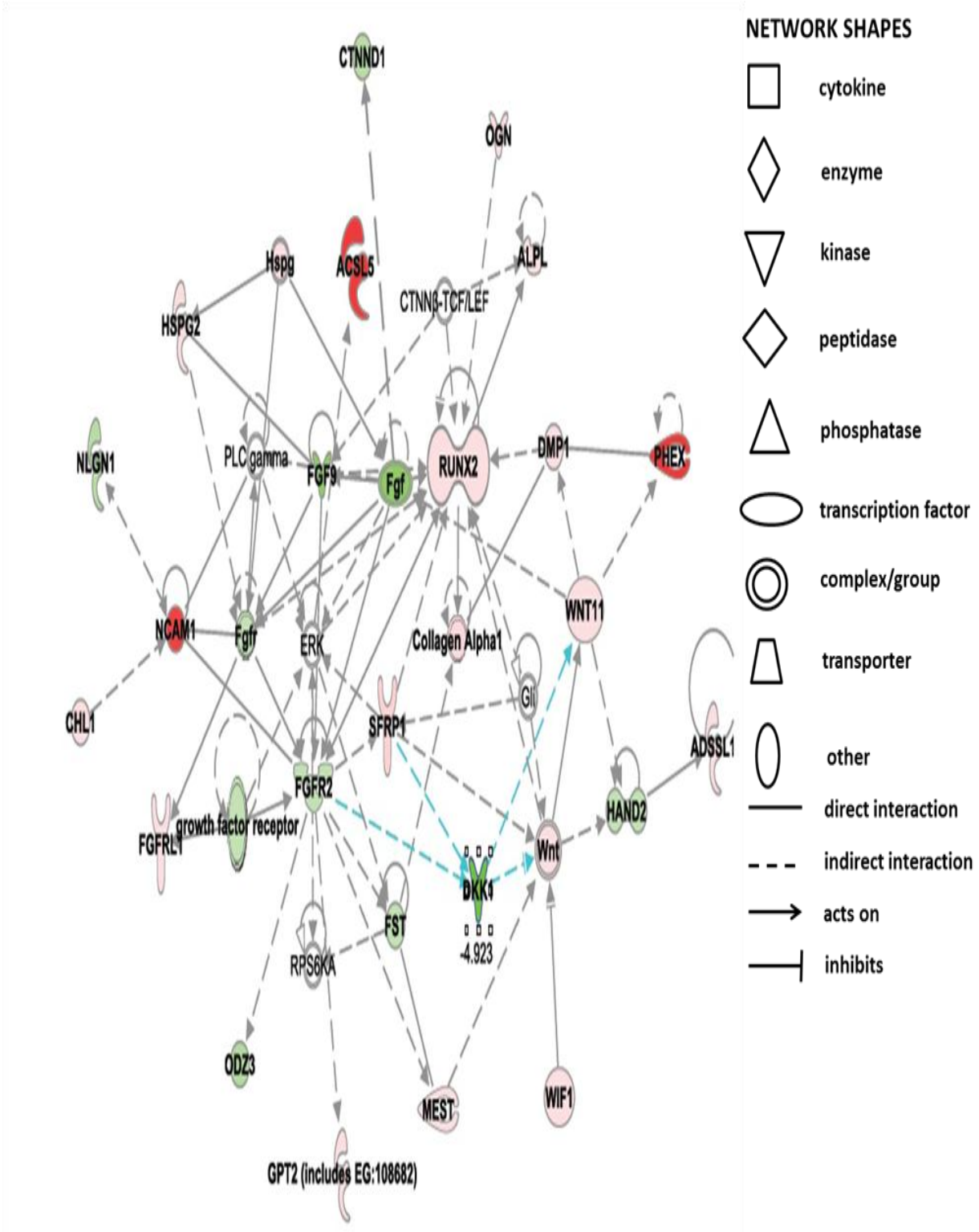



**Figure 5. IPA identified the second top-scoring network derived from the differentially expressed in ageing as** a further connective tissue disorder with scores of 35. The green nodes represent up-regulated and red nodes down regulated gene expression in older cartilage. Intensity of colour is related to higher fold change. Legend provides a key to the main features in the network.

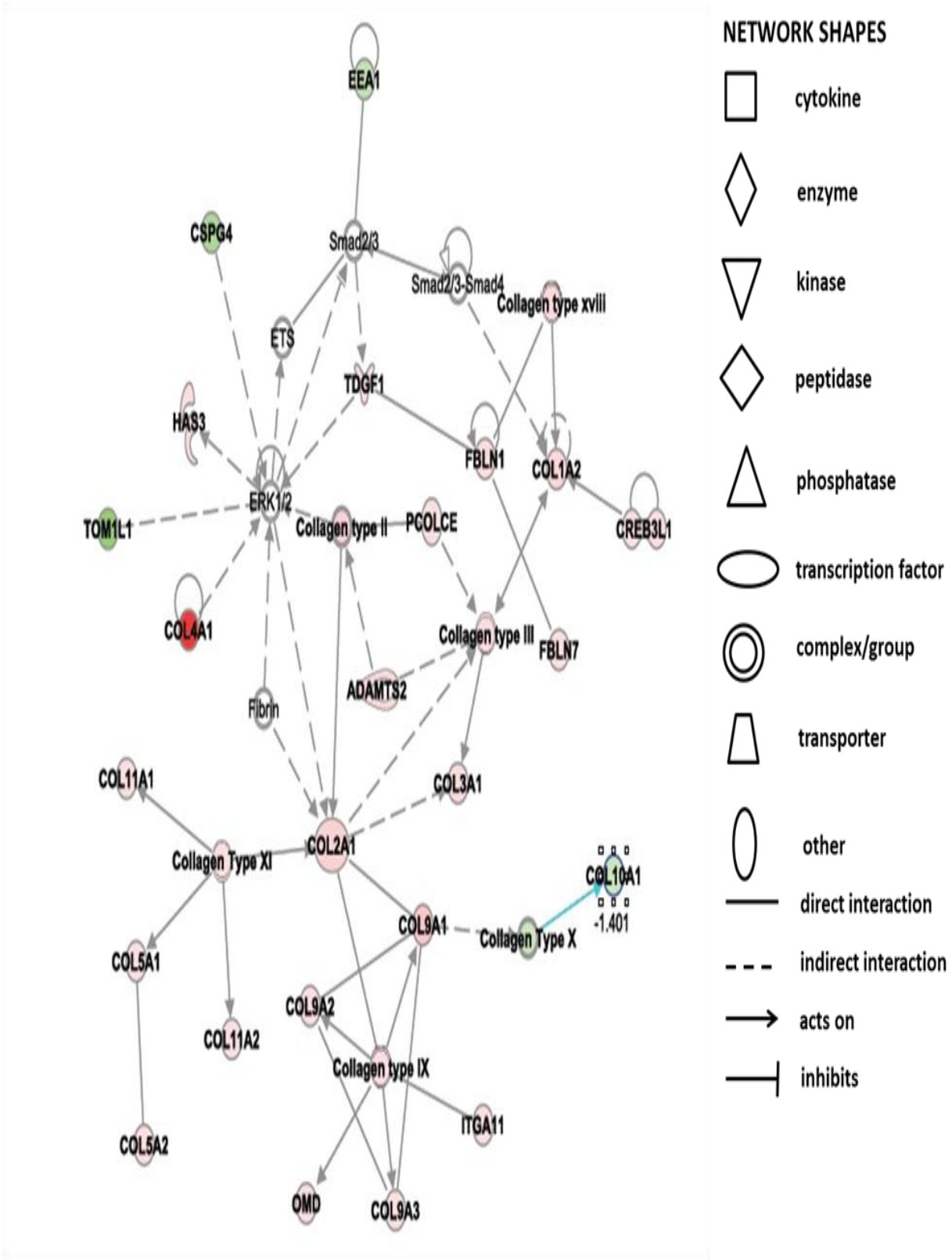



**Figure 6. IPA identified ageing significantly affects the connective tissue development and function network in ageing cartilage.** The green nodes represent up-regulated and red nodes down regulated gene expression in older cartilage. Intensity of colour is related to higher fold change. Legend provides a key to the main features in the network.

Other networks significantly enriched also related to a further network in connective tissue disorders which contained genes including collagens; COL10A1, COL11A1, COL2A1 plus a disintegrin and metalloproteinase with thrombospondin motifs-2 (ADAMTS-2) and fibulin-1 (FBLN1) (Figure 4B). Additionally a connective tissue development network was also significantly affected. The genes most affected in this network included acly-synthetase long chain family member 5 (ACSL5), phosphate-regulating neutral endopeptidase (PHEX) and dickkopf homolog 1 (DKK1) (Figure 4C).

Significant IPA canonical pathways are demonstrated in Table 4 and the associated molecules of the top canonical pathways identified are in Supplementary table 3 provided on CD. These include atherosclerosis signalling, prothrombin activation and rheumatoid arthritis.

| Name of canonical pathway | p-value | Ratio |
|---|---|---|
| **Atherosclerosis signalling** | 3.80E-09 | 15/136 (0.11) |
| **Role of osteoblasts, osteoclasts and chondrocytes in rheumatoid arthritis** | 3.41E-06 | 16/238 (0.067) |
| **Intrinsic prothrombin activation** | 9.82E-06 | 6/35 (0.171) |
| **Hepatic fibrosis and stellate cell activation** | 9.92E-06 | 12/146 (0.082) |
| **Role of macrophages, fibroblasts and endothelial cells in rheumatoid arthritis** | 1.73E-04 | 16/333 (0.048) |

**Table 4. A number of IPA canonical pathways were significantly affected in ageing cartilage.** The significance of the association between the data set and the canonical pathway was measured using a ratio of the number of molecules from the data set that mapped to the pathway divided by the total number of molecules that map to the canonical pathway is displayed. In addition a Fisher's exact test was

used to calculate a p-value determining the probability that the association between the genes in our dataset and the canonical pathway was explained by chance alone.

**Confirmation of DGE using real time PCR measurements of selected genes**

In order to validate the RNA-Seq technology 14 genes were selected to measure using reverse transcription and real time PCR based on differences noted in the arrays and/or their potential importance in the OA process. This was performed on the original RNA from all donors used to perform the RNA-Seq experiment (Table

| | | | Age | | |
|---|---|---|---|---|---|
| **RNA-Seq Results** | **Significant $Log_2FC$** | **Gene name** | **Young** | **Old** | **P value** |
| **higher in old** | -4.9 | **DKK1** | 0.0044±0.006 | 0.0338±0.20 | 0.024 |
| | -1.4 | **COL10** | 0.0006±0.0001 | 0.0013±0.000886 | 0.26 |
| **lower in old** | 2.2 | **RUNX2** | 0.0093±0.009 | 0.0034±0.002 | 0.002 |
| | 1.6 | **SRPX** | 0.0046±0.005 | 0.0006±0.0005 | 0.007 |
| | 30.1 | **ACSL5** | 0.0155±0.0002 | 0.0062±0.005 | 0.09 |
| | 30.1 | **IL7R** | 0.0005±0.0002 | 0.0001±0.0001 | 0.06 |
| | 6.5 | **COL2A1** | 63.3246±54.7 | 1.3165±1.12 | 0.04 |
| | 6.5 | **COL1A1** | 3.4815±1.57 | 0.0278±0.02 | 0.15 |
| | 1.6 | **MMP1** | 0.7093±0.21 | 0.4027±0.14 | 0.055058 |
| | 2 | **MMP-13** | 0.2323±0.15 | 0.044±0.04 | 0.1 |
| | 1.9 | **ADAMTS-4** | 0.5121±0.35 | 0.1345±0.05 | 0.07 |
| | 6.2 | **IL1β** | 0.0057±0.005 | 0.0004±0.0003 | 0.05 |
| **no change** | not significant | **TNFα** | 0.0041±0.004 | 0.001±0.0003 | 0.28 |
| | not significant | **TBFβ** | 1.2865±0.23 | 2.3124±1.24 | 0.15 |

**Table 5. Real-time PCR analysis of 14 selected genes reveals good correlation with RNA-Seq results.** Values are the mean ± SD of relative expression levels normalized to expression of GAPDH. DKK1 = dickkopf homolog 1; COL10 = collagen type X ; RUNX2 = Runt-related transcription factor 2; SRPX = Sushi repeat-containing protein; ACSL5 = acyl-CoA synthetase long-chain family member 5; IL7R= interleukin 7 receptor; COL2A1 = collagen type II, alpha 1; COL1A1 = collagen type I, alpha I; MMP1= matrix metalloproteinase 1; MMP-13 = matrix metalloproteinase 1;, ADAMTS4 = a disintegrin and metalloproteinase with thrombospondin motifs 4; IL-1β = interleukin 1β; TNFα = tumor necrosis factor alpha; TGFβ = transforming growth factor β.

5). Genes were selected based on differences noted in the RNA-Seq results. All genes were found to have comparable results with RNA-Seq data for instance genes identified as having an increase in expression in older samples in the RNA-Seq experiment also gave increased expression relative to GAPDH following real time PCR. Statistical significance was reached for 66% of genes tested using Student's T test. Two genes, whose expressions were not significantly altered in RNA-Seq results; tumor necrosis factor alpha (TNFα) and transforming growth factor β (TGFβ) were also unaltered when assessed with real time PCR.

**Discussion**

Ageing has an important role in the development of OA by making the joint more susceptible to other OA risk factors. In order to provide interventions to prevent age-related changes and reduce the risk of developing OA the underlying mechanisms involved in age-related changes of cartilage require elucidation. Whilst characterising both young and old cartilage at the molecular and systemic levels is essential for identifying the critical signalling pathways. In the present study, we used RNA-Seq technique to undertake deep transcriptome profiling of young and old cartilage for the first time. Furthermore validation studies using real-time PCR demonstrated high correlation between methodologies. This is the first time to our knowledge that this technique has been used to interrogate transcriptional changes in cartilage ageing.

This study built on previous findings that demonstrated that joint ageing causes a reduction in matrix gene expression (Loeser et al., 2012). We took a single tissue,

articular cartilage and undertook RNA-Seq in order to interrogate a greater range of genes for differential expression. Not surprisingly our experiments identified that age of donor accounted for the principal variability in the data. However novel findings of this study included: (1) The age-related gene expression changes identified, were most notably involving reduced DGE in older cartilage; 3.3 fold. (2) An over-representation of genes with reduced expression relating to ECM, degradative proteases, matrix synthetic enzymes, cytokines and growth factors were identified in cartilage derived from older donors compared to young. (3) Cartilage ageing caused a decrease in important Wnt signalling genes. (4) According to IPA pathway analysis the top scoring network for differentially expressed genes were from connective tissue disorders and connective tissue development. (5) IPA also demonstrated significant canonical pathways for atherosclerosis signalling, prothrombin activation and rheumatoid arthritis. (6) There was differential expression of pseudogenes and non-coding RNAs in cartilage ageing; with increased expression of 12 pseudogenes and 10 non-coding RNAs in older and one pseudogene and seven non-coding RNAs in younger cartilage.

Equine tissue was readily obtained enabling collection of cartilage samples from macroscopically normal, skeletally mature and aged horses. Importantly the horse suffers clinical joint diseases similar to man, and the articular cartilage thickness is comparable (Brommer et al., 2005). For young horses one year is equivalent to about 3.5 years of a human (icerydernet/agerelationship, 2011). Hence horses of greater than 15 years old, used in this study equates to humans of older than 52 years. As the cellularity of cartilage is low and a considerable amount of high quality RNA is required for RNA-Seq studies we utilized the entire articular surface of distal metacarpal III bone. Thus load bearing and non-weight bearing cartilage was used. An assessment of macroscopic changes was made, and all samples scored, but a lack of tissue meant microscopic analysis of the samples could not be undertaken. Nevertheless our previous studies indicated a high correlation between gross scoring and Mankin's grading in normal ageing equine cartilage (Manuscript 2).

Annotations of genes at reduced levels in older samples included many relating to ECM, degradative proteases, matrix synthetic enzymes, cytokines and growth

factors. In contrast, within these annotations those at higher levels in older were very small; collagen X, XXV, lubricin and fibroblast growth factor 9. It appears there is an age-related failure of matrix, anabolic and catabolic cartilage factors. Whole mouse joints have demonstrated a reduction in matrix genes with age (Loeser et al., 2012) whilst another study found an age-related decline in matrix production when equine chondrocytes were stimulated with TGFβ1 (Iqbal et al., 2000). Others have provided evidence for a chondrocyte senescence secretory phenotype in ageing, demonstrated by an increase in cytokines (Forsyth et al., 2005; Long et al., 2008) along with MMP production as well as a reduction in growth factors response (Blaney Davidson et al., 2005; Chubinskaya et al., 2002). Here we did not find evidence for an age-related increase in inflammatory environment. Indeed one of the previous studies cited here demonstrated an increase in IL-7 in ageing chondrocytes and in response to fibronectin fragments or IL-1 (Long et al., 2008). Although our experiment did not identify IL-7, interestingly one of the most down regulated genes identified in this study was IL-7 receptor. It has been previously demonstrated that a reduction in IL-7 receptor signalling in ageing β-progenitor cells resulted in ageing-like gene expression profiles (Curtis et al., 2012). Also, whereas others have demonstrated an increase in IL-1 (Forsyth et al., 2005) (where an increase in IL-1 protein was seen in older cultured human chondrocytes) and MMP-13 (Forsyth et al., 2005; Wu et al., 2002) with age in human cartilage, this study identified an age-related decline in their message. In contrast, one MMP-13 study looked at catabolic responsiveness with age whilst another used immunolocalisation of MMP-13 to identify protein. The two are not always related (Greenbaum et al., 2003) for instance reduction of mRNA concurrent with increased protein expression can occur when a protein half-life is increased due to stabilisation components involved in protein turnover. Differences could also be attributed to our age classification of young and old and species distinctions. Alternatively increased matrix enzymes (MMP-1,-13) and cytokines such as IL-1,-8,-11 identified in younger cartilage could be due to increased turnover of cartilage in young. Interestingly a recent study identified low innate capacity to produce IL-1β and IL-6 was associated with the absence of OA in old age (Goekoop et al., 2010).

The reduction in Il-1β evident in older cartilage may represent a protective mechanism against OA.

One of the many novel findings in this study was that in cartilage derived from old donors there was a reduction in the expression of some key Wnt signalling genes in addition to an increase in the Wnt antagonist DKK1 and RunX2, a downstream target of Wnt. Wnt signalling is active in adult cartilage with deregulation being detrimental resulting in age-associated joint pathologies due to excessive remodelling and degradation (Yates et al., 2005). This signalling pathway has also been found to regulate both matrix synthesis in chondrocyte cell lines (Zhu et al., 2008) and stimulate catabolic genes such as MMP-13 and ADAMTS-4 in chondrocytes (Yuasa et al., 2008). A recent study demonstrated that the activation of the Wnt pathway inhibited IL-1-mediated MMP-13 expression in human chondrocytes (Ma et al., 2012b) which was mediated through a direct interaction between NF-κB and β-catenin, identifying a potential protective function of Wnt in aging and OA. One study has linked Wnt signalling with chondrocyte hypertrophy through RunX2 activation (Dong et al., 2006). Whilst elsewhere it was shown that DKK1 is a major player in the cessation of hypertrophic differentiation which can contribute to OA (Buckland, 2012c). Interestingly COL10A1, a marker of chondrocyte hypertrophy was increased in old cartilage. A recent study in mesenchymal stems cells derived from OA patients found that COL10A1 downregulation played a role in the establishment of a defective cartilage matrix in OA (Lamas et al., 2010). It would seem that this increased expression with ageing is not through the Wnt signalling interaction with subsequent RunX2 activation as previously described (Dong et al., 2006). Further credence is given to this hypothesis by our findings that alkaline phosphatase expression, also regulated through, RunX2 was down regulated in old cartilage. Overall Wnt signalling is involved in maintenance of cartilage and the disregulation event here in ageing may be an important event. Interfering with the pathway may contribute to improvements in cartilage regeneration.

Using IPA this study identified age-related changes in pathways and processes including connective tissue disorders and development in which a significant number of genes, regulated both strongly and subtly were enriched. This is not remarkable given the number of matrix genes differentially identified in the study. Whilst some canonical pathways identified as significantly affected by ageing were not surprising either such as the role of osteoblasts and osteoclasts in rheumatoid arthritis, others for instance the pathways for atherosclerosis signalling were. This is a chronic inflammatory process and the DGE of a mixture of proteases along with lipoproteins accounts for this finding. In ageing cartilage further studies to investigate these processes and canonical pathways as well as the molecules involved are clearly required.

One advantage of the use of RNA-Seq to undertake DGE studies is that other sets of RNA molecules from the transcriptome can be identified, such as non-protein coding RNAs ( for example microRNAs (miRNA) and small nucleolar RNA (snoRNA)), a significant part of the transcriptome (Kapranov et al., 2007) and pseudogenes. For example a recent RNA-Seq experiment was conducted in order to identify ribosomal protein pseudogenes (Tonner et al., 2012).

Pseudogenes provide a novel tier of gene regulation through the generation of endogenous silencing (siRNA) or miRNA binding sites which act as decoys for miRNAs (Salmena et al., 2011). Indeed some miRNAs have been demonstrated to target them (Poliseno et al., 2010). It is hypothesised that they act as post-transcriptional regulators of the corresponding parental gene (Muro et al., 2011). Whilst possessing very similar sequences to their counterpart coding genes they are unable to be transcribed due to mutation/ deletion or insertion of nucleotides. A key challenge in RNA-Seq data analysis is to discern reads among multiple potential sources with similar sequences. A recent study provided a specialised pipeline for pseudogene transcription discovery in RNA-Seq (Tonner et al., 2012), which could prove useful in future studies. Transcription of pseudogenes has tissue specificity and can be activated or reduced in disease indicating a possible functional role in cells (Zheng et al., 2007). Some pseudogenes have been identified as increasing

with age, such as pseudogene cyclin D2 in the ovary (Choi et al., 2001). Whilst this study identified the differential expression of pseudogenes in cartilage ageing, it is not known if these are functional or their relevance to cartilage ageing. Recent work by the Encyclopaedia of DNA Elements (ENCODE) Consortium identified that about 8% of the pseudogenes in the human genome are functional (Pei et al., 2012). With the recent publication of GENCODE a reference human genome annotation for The ENCODE Project (Harrow et al., 2012) more light may be shed relating to the role of pseudogenes in cartilage ageing in the near future. Pseudogenes present an interesting area for future research in cartilage ageing and disease.

A single miRNA was identified as differentially expressed in the study. The methodology used here does not enrich for miRNAs. In order to increase the identifications of small miRNAs using RNA-Seq techniques are used to enrich for small RNAs in conjunction with additional miRNA abundance quantification algorithms (Berninger et al., 2008). A single miRNA; miR21 was identified as increased in ageing cartilage. MicroRNAs are short non-coding RNAs which regulate the translation and/or degradation of target message (Ambros, 2004). miR21 has been implicated in inflammation (Sonkoly and Pivarcsi, 2009), cancers including osteosarcomas (Ziyan et al., 2011), and hypomethylation (Pan et al., 2010). Its role in cartilage is not fully elucidated though a study in rats found that miR-21 promoted increased proliferation and matrix synthesis in chondrocytes embedded in atelocollagen gel (Kongcharoensombat et al., 2010). Indeed the finding is interesting as epigenetic changes such as hypomethylation occur with ageing and this is a risk factor contributing to several age-related pathologies (Fraga and Esteller, 2007).

A further set of small RNAs; snoRNAs, a class of small guide RNAs found in the nucleolus were also identified in the study. They direct chemical modification of other RNAs, and like miRNAs are emerging as important regulators of cellular function and disease development. There are two classes including C/D box snoRNAs (SNORDs) and H/ACA box snoRNAs (SNORAs) which are associated with

methylation and pseudouridylation of ribosomal and other RNAs. It has been suggested that snoRNAs fine-tune the ribosome to accommodate changing requirements for protein production during development, normal function and disease (Montanaro et al., 2008). Indeed control of snoRNA expression may play a pivotal role in the regulation of high protein producing cells such as chondrocytes as demonstrated by the phenotypes of ribosomopathies (Narla and Ebert, 2010). There are very few studies into the significance of snoRNAS in cartilage ageing or disease except for a recent study which proposed the use of serum snoRNA U38 and U48 as biomarkers of early cartilage damage. In the study an increase in these snoRNAs was detected in serum following anterior cruciate ligament injury, but was not associated with normal ageing (Zhang et al., 2012b). The snoRNA transcriptome signatures derived from this study in ageing cartilage provides an interesting set of genes for further studies in order to determine if they play a role in ageing.

**Conclusions**

A major strength of this study is that it represents the first application of RNA-Seq technology for transcriptomic studies in cartilage ageing. The study has increased our knowledge of transcriptional networks by providing a global view of the transcriptome. The molecular signatures derived here-of reflect a combination of degenerative processes but also transcriptional responses to the process of ageing. Next generation sequencing provided an ideal quantitative framework to study pathways and networks as an integrated system in order to understand the complex processes of cartilage ageing.

**Acknowledgements**

This study was supported by a Wellcome Trust Veterinary Integrated Research Fellowship.

**General Discussion**

The work in this thesis supports the hypothesis that there are distinct mechanisms involved in cartilage ageing which differ from those in cartilage diseases such as arthritis. A number of distinct alterations in the ageing transcriptome were characterised most importantly reduced expression of extracellular matrix proteins, degradative proteases, matrix synthetic enzymes, cytokines and growth factors as well as alterations in Wnt signalling in old cartilage. This points to an age-related failure of matrix, anabolic and catabolic cartilage factors. Aberrant Wnt signalling may contribute to these changes. By comparison an early OA *in-vitro* model of cartilage degradation determined that although there was an overall reduction in the ECM proteins released into the media following IL-1β treatment there was also an increase in degradative proteases. The thesis achieved its second objective by establishing innovative techniques to identify novel cleavage sites in matrix proteins by using mass spectrometry methodologies LC-MS/MS with insightful data mining and MALDI-IMS. The design and development of cartilage QconCATs enabled absolute quantification of matrix proteins for the first time. Finally the third aim of the thesis was achieved as key proteinases and their inhibitors involved in the pathogenesis of OA were quantified in IL-1β stimulated cartilage explant cultures.

Proteins are the key molecules in a living organism. Proteomics enables the study of the true 'actors' in pathways leading to physiological and pathological changes. Whilst genomics technologies have recently advanced exponentially to enable the transcriptome to be measured in detail this does not always relate directly to what is happening at the protein level and what the 'actors' are doing. More transcripts do not automatically mean more protein as downstream controls on translation, protein folding, and degradation can effect transcription. Indeed a recent study in yeast explored the relationship between transcripts and corresponding protein level variation and the group concluded that underlying transcript levels cannot account for the majority of variation observed in the corresponding protein levels between yeast strains (Foss et al., 2011). Therefore it is important to study both transcript and protein when investigating a tissue or disease. This was the basis for the use of next-generation sequencing technologies and proteomics in the

experimental methodologies used here. Mass spectrometry based proteomics (Aebersold and Mann, 2003) has changed the way in which biological systems are interrogated because it can measure thousands of proteins and PTMs in parallel. It was for this reason that a heavy bias to these techniques was used in this thesis.

Proteomic technologies have been successfully used to understand molecular and cellular mechanisms that contribute to arthritis and its progression as well as identifying potential biomarkers (Ruiz-Romero and Blanco, 2009). However, due to the technical challenges related to their biochemical properties including low cellularity and the ECM composition being abundant in highly anionic macromolecules that interfere with isoelectrofocusing in 2-DE (Wilson et al., 2008), progress has been slow. In order to avoid these problems the majority of cartilage proteomics studies being undertaken in chondrocyte or explant culture (Iliopoulos et al., 2010). In this thesis initial studies undertaken using cartilage explant culture studies in human and horse were utilised to measure and quantify changing protein expression following IL-1β stimulation. This required limited purification or separation techniques which can lead to reduced reproducibility and enabled the differential expression of many proteins to be identified with MS as the analyte was relatively simple. Thus problems produced by the presence of small numbers of high abundance proteins masking the identification of less abundant proteins were avoided. Explants cultures were chosen over monolayer cultures as isolated chondrocytes dedifferentiate in culture leading to a change in phenotypic expression (Holtzer et al., 1960). Additionally the ECM provides native substrates for proteolysis and protein release, equivalent to the shedding of the proteins into the synovial fluid during cartilage degeneration. HAC studies were undertaken in end-stage OA cartilage removed from joints following total knee arthroplasty. Whilst it would have been beneficial to undertake our IL-1β studies in normal human cartilage we were unable to obtain the tissue required. Therefore we chose equine cartilage as an alternative model to study cartilage ageing and arthritis. The horse is a good model as it provides a consistently predictable model of OA that has previously been used to study early pathological events and the horse has been extensively studied with respect to clinical OA (McIlwraith, 2012). In addition a

range of ages of normal cartilage are readily available. The cytokine IL-1β was used as it has been identified as the key cytokine in human OA (Berenbaum, 2007) and levels remain high throughout all stages of OA (Toncheva et al., 2009). The time course used for human OA explant studies was 48 hours whilst that for equine studies was 96 hours. Results from the former study identified few differentially regulated proteins following the shorter time scale of IL-1β stimulation. As previously mentioned this could be due to the time-scale of the study or a lack of responsiveness of end-stage OA cartilage to cytokine stimulation. Further studies using a longer stimulation period would be useful in determining which the case is. In contrast a large number of proteins were identified as differentially regulated in the equine explant model. This model has been described as a model of early OA (Peffers et al., 2012) and was used as such in this thesis. However others may argue that it is more an inflammatory model of cartilage degradation. It is well defined that there is an inflammatory component to OA (Buckland, 2012b; Toncheva et al., 2009). It may be better to describe the model as an explant model of articular cartilage inflammation.

There has been a great improvement in sensitivity and data acquisition speed with the advent of new mass spectrometers based on the triple-quadrupole, quadrupole time-of-flight and Orbitrap technology. These have enabled the identification and subsequent quantification of thousands of proteins in a proteome (Beck et al., 2011) and the development of software tools for identification and quantification. Therefore in our 'discovery' experiments we used an Orbitrap instrument with these capabilities followed by data analysis using a relative quantification software tool. This approach enables the maximum identification of proteins in terms of both numbers of proteins and dynamic expression range. Using this workflow the explant secretome experiments were able to identify both potential pathways involved in the pathogenesis of early cartilage inflammation such as altered glycolysis and cytoskeletal modelling and novel proteins such as clathrin light chain whose role is yet to be determined. This latter class of molecules represent potential diagnostic biomarkers of early cartilage inflammation since there increased expression following IL-1β stimulation represents a response to cytokine

stimulation. These molecules represent interesting molecules for further investigations. In protein biomarker discovery, LC-MS methods are challenging traditional assays (for example ELISAs) because the ability of the mass spectrometer to identify and quantify a protein unambiguously and accurately directly or following enrichment using straightforward-to generate peptide antibodies (Whiteaker et al., 2011). The equine secretome manuscript demonstrated that clathrin light chain was the most significantly elevated protein. Clathrin-coated vesicles (CCVs) form at the plasma membrane where they select protein and lipid cargo for endocytic entry into cells. In addition CCVs form at the trans-Golgi network, where they function in protein transport from the secretory pathway to the endosomal/lysosomal system. This finding raises a number of possible scenarios. It may be that due to cytokine stimulated cartilage degradation there is increased intracellular trafficking leading to an up-regulation of clathrin. Although cell death could be producing the finding there was no obvious cell death demonstrated from trypan blue studies. As there was no increase in clathrin heavy chain in the OA secretome it could be that clathrin light chain is acting independently of the heavy chain. A further alternative hypothesis is that there is a secretory pathway stimulated by IL-1β that has not yet been identified. Further experiments to determine which of these hypothesises are correct would require the use of *in-vitro* models of OA and the identification of patterns of clathrin heavy and light chain gene and protein expression.

The relative quantification experiment data was analysed using a software tool called Progenesis™ which enables label-free quantification. The software outputs a list of proteins identified, together with their fold change compared to a control, a p-value and a q-value (p-value adjusted to FDR). In addition, the number of peptides used in this quantification is identified. During interpretation of the data it is necessary for the experimenter to determine a 'cut-off' within for example q-values, in order to take significant data through for inclusion in results and further bioinformatics analysis. A standard approach, which was used in these studies, would be a fold change of two, q-value of <0.05 and greater than one peptide for identification. It is possible that a single peptide will identify a protein correctly,

although identical sequences could be duplicated in closely related proteins, therefore matching multiple peptide sequences provides greater statistical confidence. In theory some differentially expressed proteins could have been missed. Furthermore some interesting proteins could be missed from the dataset if the fold change was less than two. Therefore it is advisable to look at near significant proteins in the interpretation of the dataset as a whole.

The initial approach to the discovery phase in this thesis combined high performance MS instruments in terms of sensitivity, speed, mass accuracy, and resolution for identification and then relative quantification. The main approaches to protein and peptide quantification using MS from LC-MS data involve either obtaining intensities of peptide precursor ions in a DDA experiment in which peptide precursors are fragmented as they are eluted from the LC system, or measuring peptide fragment ion intensities of peptides, either from DDA experiments, or SRM assays. In the latter a predefined list is used thus rendering the method more sensitive (less noise), and with a greater dynamic range but this, unlike the first method is therefore hypothesis driven. Using SRM assays it is possible to obtain absolute values in terms of protein quantification as copies per cell or normalised to, for example dry weight. Developments in absolute quantification, such as the advent of QconCAT, have been driven by the discipline of systems biology, one goal of which is to delineate protein interaction networks and to measure protein movement within networks. By modelling pathways within protein networks the predicted outcomes of system perturbation can be tested experimentally. Absolute quantification provides these data enabling model parameterisation.

In order to quantify in absolute terms important matrix proteins and proteases in the HAC secretome using SRM assays a QconCAT was designed. This approach was used as a QconCAT is relatively simple to design, allows parallel quantification of tens of proteins in a single experiment, and is extremely cost-effective as once the gene has been manufactured (expressed in heavy labelled media and validated) there is enough for hundreds of experiments. This is because the artificial gene can

be freshly expressed for subsequent studies. In comparison another stable isotope-labelling approach; AQUA peptides (Gygi et al., 1999a) are purchased as a finite amount of a single labelled peptide (cost approximately £1000 per peptide) which once used requires the purchase of more AQUA peptide. Furthermore following identification of sequence homology of the q-peptides the human cartilage QconCAT was also used to quantify some equine cartilage peptides. This demonstrates the flexibility of the QconCAT in quantifying proteins from other species. Further sequence homology identification against mouse for peptides from identical proteins identified 40% of q-peptides were homologous between human and mouse (data not shown). Subsequently an equine cartilage QconCAT was designed in order to provide a tool for measuring matrix proteins, proteases and cartilage degradation at known cleavage sites.

There were differences in the expression of the two QconCATs. The human QconCAT was expressed poorly whilst the equine QconCAT demonstrated good expression. Some QconCATs express better than others, indeed there can be different levels of expression of the same QconCAT from different cultures. Although the reasons for this have not been fully elucidated a number of factors are thought to be important including the sterility of media and culture (due to competition for resources), quantity of *E.Coli* in the culture when the expression is induced and the time the culture is left to grow following induction. In addition level of expression can be increased through increasing the concentration of Isopropyl β-D-1-thiogalactopyranoside, the trigger for transcription (Harman, 2012). Others have used benzyl alcohol in the culture media to slow down the rate of expression and allow the protein time to fold (Jariyachawalid et al., 2012). However as QconCATs do not have secondary structure, methods that claims to help expression through improved folding will not aid QconCAT expression.

Whilst proteins were quantified in the human secretome, equine secretome and equine cartilage extract using QconCAT the QconCAT-LC/SRM approach was not able to quantify all peptides and therefore proteins targeted. In the QconCAT experiments all peptide-levels were classified as type A (good standard and analyte

signals), type B (good standard, missing analyte signal), or type C (neither standard nor analyte signal) for specific loadings on column. Type A quantifications produce actual values; type B quantifications define the upper limit of analyte abundance, type C peptides could not be quantified. In order to quantify the maximum number of peptides in parallel in a single experiment a set loading of QconCAT determined by previous 'ranging' experiments was used. In theory some of the type B peptides could have been quantified by loading more analyte onto the column. Unfortunately this was not possible as column overloading would damage the high pressure liquid chromatography system. However, type B peptides can be used to quantify different analyte samples when the analyte peptide is above the limits of detection with success, as demonstrated by the quantification of MMP-3 in equine secretome media following Il-1β stimulation but not cartilage extract. The importance of using two or more q-peptides to identify a protein is highlighted in this study. A number of peptides were type C and so redundant for quantification. Possible reasons for type C peptides include poor peptide fragmentation (normally determined by a lack of measureable transitions). In addition peptide size, charge, hydrophobicity, and peptide secondary structure, an important factor in determining 'detectability' by electrospray ionisation MS (Eyers et al., 2011) are also important. Whilst the latter could cause a type B peptide it should not produce a type C as there is no secondary structure associated with QconCAT. Furthermore chemical protocols were used to break and prevent reformation of disulphide bonds in the workflows should the QconCAT contain cysteine residues.

An average protein will generate 30-50 tryptic peptides. However some will not be observed in a MS study (Aebersold and Mann, 2003) and not all candidates are suitable for inclusion in a QconCAT. Whilst hydrophilic peptides will not bind to the reversed-phase column used for LC prior to MS analysis very hydrophobic peptides are less likely to elute from the column. The peptides chosen for inclusion in the QconCATs were based on proteotypic peptide databases Global Proteome Machine Database (GPMDB) (Craig et al., 2004) and Peptide Atlas (Desiere et al., 2005) for the human QconCAT and on prior observation in MS/MS studies for the equine QconCAT. These approaches produced 20% and 26% redundancy respectively

indicating no difference in the level of redundant peptides with each approach. Interestingly peptides are included in databases such as GPMDB on the frequency of observation in MS and MS/MS studies and not because they are quantitatively representative of the parent protein. Furthermore in the selection of quantotypic peptides there is no resource to demonstrate completeness of proteolysis, the lack of post-translational modification or uniqueness of the peptide and freedom from isobaric and isomeric peptides derived from other proteins (Simpson and Beynon, 2012). There are useful tools that became available subsequent to the design of the QconCATs in this thesis. One uses four algorithms to assess peptide 'detectability' in an electro-spray ionisation instrument (CONSeQuence) (Eyers et al., 2011). In a yeast test set it improved suitable candidate q-peptide selection. CONSeQuence was used to assess the q-peptides in the equine QconCAT post-design. In theory results should improve the selection of q-peptides by identifying the peptides most likely to be detected, thus reducing type C peptides. Interestingly six type C peptides were identified in the equine QconCAT. Of this half were demonstrated by CONSeQuence to be in the top 10% most detectable peptides whilst some type A peptides were deemed in the bottom 10% (most likely not to be detected). This demonstrates that there are other important factors in determining what makes a good q-peptide. In this QconCAT the major reasons for redundancy were poor fragmentation patterns producing inadequate transitions for detection and broad elution profiles. The latter is a problem as for some of these peptides the elution time was greater than a minute which constituted most of the dwell time for the method (the time the specific transition is searched for by the third quadrupole in the triple quadrupole). For quantification narrow sharp peaks are optimal.

In the quantification experiments there was disparity between some peptides used to quantify the same protein for example in the equine QconCAT the GVFSGLR and VPAGLPDK for biglycan. It has been demonstrated that when this occurs the most common reason is incomplete digestion of the standard or analyte or a failure to detect the analyte due to an unanticipated PTM. A tool is now available that can aid in the identification of possible miscleavages between amino acids (Siepen et al., 2007). When this was used to interrogate the peptides in the analyte and QconCAT

there was a high probability for a miscleave at the N-terminal of analyte peptide VPAGLPDK. This lead to a lower quantification of the protein using this peptide as less completely proteolysed analyte peptide was present for quantification. Indeed VPAGLPDK gave the lower quantification value compared to GVFSGLR. It is difficult to identify miscleaves in MS data of analyte, for example in discovery LC-MS/MS as they are in low amounts and would only be identified if larger amounts of protein were loaded on column. This would be detrimental to the LC as increased total protein would need to be loaded. It is easier to identify miscleaves in QconCAT as the standard is simpler, containing few peptides and so increased loadings (which are less of a problem to the LC) may identify missed cleaves. The potential for previously unidentified PTMs in q-peptides was assessed through data mining of MS/MS data using PEAKS software (data not shown). This program includes advanced PEAKS PTM algorithm (Han et al., 2011) to identify 650 possible PTMs and mutations since traditional database search software can only specify a limited number of possible variable PTMs. There were no possible PTMs identified in the data searched for any of the q-peptides included in the QconCATs. All these points are important for the future development and utilization of QconCAT/ SRM technology in the field of cartilage research.

The equine QconCAT contained a number of tryptic cleavage-site spanning peptides (NCS) in an attempt to quantify cartilage degradation at specific sites using SRM assays. The hypothesis was that non-cleaved tryptic peptides (NCC) would give greater quantification values than NCS peptides if degradation had occurred at the cleavage site. However for most of the quantifications the NCS peptides gave greater values than the NCC peptides. This was probably due to 'normal' variability in quantification between peptides (which has been identified by others as 20% variation being acceptable (Chen et al., 2012)) rather than unidentified sites of degradation in the NCC peptides. Whilst this cannot be ruled out completely data analysis of both IL-1 stimulated equine explant culture media and crude proteoglycan extract hydrolysed with MMP-3 or ADAMTS-4 (manuscript 4) did not identify potential cleavage sites in any of the NCC peptides used in quantification. Interestingly the NCS biglycan peptide GVFSGLR was significantly reduced in

cartilage from OA compared to normal samples indicating degradation at the MMP-13 cleavage site $^{181}$G$^{182}$V in biglycan. Unfortunately there are no commercial anti-epitope antibodies to this site which could be used to validate this result.

The QconCATs were used successfully to measure levels of key proteases and their inhibitors in human and equine cartilage secretomes. The equine QconCAT and human QconCAT offer useful analytical tools for cartilage studies. They could be used in future work to identify and characterize cartilage matrix fragments and proteases in explant degradation studies as an alternative to western blotting. Media samples taken at various time points following IL-1 stimulation for instance could identify the time course of molecular events for the degradation of ECM proteins included in the QconCATs. In addition the QconCATs could be used in different species and tissues.

The knowledge of protease cleavage sites is important for a number of reasons. In numerous cases *in-vivo* cleavage events that are catalyzed by a particular protease reflect its *in-vitro* specificity. Therefore knowledge of *in-vitro* specificity corroborates *in-vivo* cleavage events. Furthermore the position of protease cleavage can influence the biological consequences, especially in cases where several potential cleavage sites are closely related. Finally determination of preferred cleavage sites in known substrates enables protease activity prediction directed toward novel substrates. The thesis developed novel techniques which were tested and validated for the identification of known and novel cleavage sites in matrix proteins by using mass spectrometry methodologies LC-MS/MS or MALDI-IMS. A relatively simple peptide-centric approach technique was developed to identify potential cleavage sites in cartilage ECM proteins. Following extraction of the soluble cartilage components a crude proteoglycan extract was used as a substrate for two important proteases MMP-3 and ADAMTS-4. The use of this analyte and not whole cartilage was for a number of reasons. Firstly we wanted start with a simpler proteome. For downstream MS/MS analysis a more complex mixture will deliver more peptides to the mass spectrometer than can be analysed as the liquid stream flows through the electrospray source. Additionally the data

system would direct MS analysis to more abundant peptides, therefore limiting dynamic range. This is especially important when attempting to identify low abundance peptides such as those produced by degradation. There are a number of published protocols for identifying protease cleavage sites using peptide-centric approach (reviewed by Tholey and Bart 2012 (van den Berg and Tholey, 2012)). These methods are only able to experimentally confirm the internal N-terminal. Many of these methods protect the N-termini by chemical modification (McDonald and Beynon, 2006; Schilling and Overall, 2008). The N-termini formed by proteolysis with the protease are therefore blocked from further proteolysis. These methods require multiple steps and some specialised equipment. The successful approach in this thesis provided a simple method using global MS methodology to produce a library of potential cleavage sites for others to interrogate further. This approach could be applied to other tissues and use other proteases as there is evidence for a role of other classes of enzymes in the cartilage degradation for instance serine proteases, that can directly degrade the ECM or could be involved in the activation of proMMPs (Milner et al., 2001). Further evaluation of many of the neopeptides identified in this thesis as potential biomarkers would be achieved by quantitation in synovial fluid, blood, or urine.

A number of peptides containing potential MMP-3, ADAMTS-4 or IL-1β –driven novel cleavage sites identified in this study were also imaged in MALDI-IMS studies with reduced expression in OA cartilage compared to normal. The intensity of the COMP peptide FYEGELVADSNVVLDTTMR was reduced in OA compared to normal cartilage using MALDI-IMS. In manuscript 4 a potential cleavage at the $Asp^{706}$-$^{707}$Tyr was demonstrated in the media following IL-1β stimulation of cartilage explants. The biglycan peptide NHLVEIPPNLPSSLVELR intensity was reduced in OA cartilage using MALDI-IMS. Manuscript 4 identified potential cleavage sites at $Ser^{164}$-$^{165}$Ser following MMP-3 or ADAMTS-4 treatment and $Ser^{165}$-$^{166}$Leu following MMP-3 treatment. Interestingly the full length tryptic peptide NHLVEIPPNLPSSLVELR was used as a NCC peptide in the equine QconCAT but was not used in quantification due to classification as a type C peptide. This was because it had a broad elution profile and poor fragmentation resulting in an

inability to identify robust transitions for the SRM assays. The peptide was identified by MASCOT with a high score (MASCOT identifies using spectral libraries and not through specific fragments). The differences in peptide detectability are due to the different techniques used. MALDI is laser desorption method which effectively ionises the peptide which is singly charged. ESI used in discovery experiments is a different ionisation technique. It is more challenging to ionise this particular peptide using ESI and in addition to fragment it. A further peptide for fibromodulin ELHLDHNQISR, with a potential cleavage site at $Asn^{186}$-$^{187}Glu$ was imaged with reduced intensity in OA cartilage. Taken together findings from these two studies place more confidence in the conclusions that disease-related degradation can be identified using MS techniques and demonstrate that the two techniques can be used complimentarily to identify sites of cleavage.

The MALDI-IMS technique has advantages over extraction and digestion of cartilage to identify proteins as fundamentally it allows the site of the proteins within the cartilage structure to be identified. As previously discussed the cartilage used in this study was too thin to enable differences in peptide distribution to be visualised in the superficial and deep layers of cartilage unlike those demonstrated in human cartilage (Cillero-Pastor et al., 2012a). Interestingly TOF-SIMS of cartilage was able to resolve the two layers and differences in the the molecular distribution of ions and lipids were demonstrated. This is accounted for by the greater spatial resolution of TOF-SIMS (to the cellular level) compared to MALDI-IMS. The MALDI-IMS technique could be used on thicker equine cartilage such as that from the femerotibial joint as well as in other musculoskeletal tissues such as tendon and ligament as it is an excellent method to identify and localise disease specific peptides and protein. Furthermore it would be interesting to repeat the TOF-SIMS experiments with some methodology modifications to aid the adherence of cartilage to ITO slides. A further modification of the MALDI-IMS technique used in this thesis could be beneficial in further increasing our knowledge of cartilage degradation in ageing and disease. It should be possible to digest the cartilage slices directly with a given protease (such as MMP-3) followed by tryptic digestion in order to discover if different areas of the cartilage are more susceptible to

degradation than others and the nature the degradation takes in terms of peptide cleavage sites.

One of the aims of this thesis was to understand age-related changes in cartilage, the role of these changes in causing OA (since age is the greatest risk factor in OA pathogenesis (Hugle et al., 2012)) and the differences in expression of proteins and transcript between ageing and OA. An attempt was made to characterise age-related changes in cartilage protein signatures using MS of cartilage. The number of samples used (too few; 9 normal in the absolute quantification of equine cartilage ECM) from normal cartilage and the age spread (too few donors over the age range; 2-20 years) of the donors was not compatible with determining this. It would have been better to use larger groups of horses of similar ages (for example a group of 4 year olds and a group of 15 year olds) rather than a large variation in ages as for regression modelling purposes the sample size was too small. Alternatively more donors were required for recruitment into the study in order to determine if age related changes in matrix composition and their levels existed. However the data provides important preliminary results for designing future studies with optimal power and appropriate study design. MALDI-IMS was used to visualise the age-related and disease related changes in peptides and their localisation within cartilage. PCA identified changes in the peptides between young, old and OA samples; peptides in young cartilage were distinct from OA cartilage whilst many peptides were shared between old and OA cartilage. Whilst many of these peptides were not identified the rapid advances in MALDI-IMS (Heeren, 2012) means that in the near future these will be identified enabling pathways distinct between ageing and disease to be elucidated. Furthermore an *in-vitro* model of early OA was used to determine differential expression of proteins in the secretome. IL-1$\beta$ stimulation caused changes in the ECM dynamics with an overall reduction in ECM proteins released into the media but an increase in the proteases MMP-1, MMP-3 and MMP-13. It would be interesting to repeat the experiments and in young and old cartilage to determine age-related effects as donors used in this study were 'middle-aged'. In contrast the RNA-Seq experiments identified an age-related

failure of matrix, anabolic and catabolic cartilage factors and aberrant Wnt signalling.

The RNA-Seq results demonstrated clear age-related transcriptional differences. In ageing and joint disease there is a disruption in the cartilage equilibrium and the synthesis of new matrix components being exceeded by the loss of collagens and proteoglycans (Goldring and Marcu, 2009). The imbalances between anabolic and catabolic processes results in progressive cartilage degeneration. In cartilage ageing there is evidence of senescence resulting in the inability of chondrocytes to maintain matrix turnover (Aigner et al., 2004a). Others have pointed to a 'senescent secretory phenotype' in ageing cells (Campisi, 2005). These cells secrete increased amounts of MMPs and cytokines. Chondrocytes exhibit many changes typical of a senescent cell (Loeser, 2009). Some groups have demonstrated the increased production of MMPs (Forsyth et al., 2005) and cytokine (Long et al., 2008) in older chondrocytes. Both these studies were undertaken in isolated human articular chondrocytes and measured protein levels. This thesis used RNA extracted from chondrocytes *in-situ* and measured gene transcripts. Results demonstrated a reduction in not only matrix proteins (as predicted from previous literature) but MMPs and cytokines (not previously identified to our knowledge). Thus at a transcript level the chondrocytes do not conform to all the characteristics of a 'senescent secretory phenotype'. In order to evaluate these interesting results further studies should be undertaken to measure the protein in ageing chondrocytes to determine if protein for MMPs and cytokines is related to the changes identified in this thesis in transcript. As previously discussed transcript and protein do not always correlate. One way to do this would be to take freshly isolated chondrocytes from cartilage of young and old donors and analys,e using mass spectrometry, the protein expression following cell lysis, trypsin digestion and LC-MS/MS coupled with data analysis using Progenesis™ software.

A further interesting feature of the RNA-Seq results was the disregulation of Wnt signalling and in particular the reduction in WIF and increase in DKK1 identified in cartilage from older donors. WIF promotes chondrocyte differentiation. DKK1 promotes GAG synthesis, SOX9 and type II collagen expression (Shimazaki et al.,

2006) and inhibits chondrocyte hypertrophy (Leijten et al., 2012). It is decreased in OA (Ma et al., 2012a). Furthermore its overexpression leads to amelioration of cartilage destruction in animal models chondrocytes attempt results in this thesis demonstrate a possible mechanism by which the ageing chondrocytes attempt to increase matrix production though unsuccessfully.

For direct comparison of the transcriptome data to disease related changes additional samples from OA equine cartilage would need to be used in similar deep sequencing experiments. Taken together these results indicate different underlying mechanisms involved in cartilage ageing and disease which require further elucidation.

This thesis developed novel proteomic based methodologies which identified and quantified important distinct differences between cartilage ageing and disease. Several proteins not previously described in cartilage were identified. In addition many novel cartilage degradation products were identified and age-related peptides were visualised in cartilage for the first time.

## References


Abiola, P.O., J.V. Parry, and P.P. Mortimer. 1990. Sensitivity for anti-HIV-2 of combined HIV antibody kits. *Lancet*. 336:1386-1387.

Acebes, C., J.A. Roman-Blas, E. Delgado-Baeza, I. Palacios, and G. Herrero-Beaumont. 2009. Correlation between arthroscopic and histopathological grading systems of articular cartilage lesions in knee osteoarthritis. *Osteoarthritis Cartilage*. 17:205-212.

Adams, C.S., and W.E. Horton, Jr. 1998. Chondrocyte apoptosis increases with age in the articular cartilage of adult animals. *Anat Rec*. 250:418-425.

Aebersold, R., and M. Mann. 2003. Mass spectrometry-based proteomics. *Nature*. 422:198-207.

Aigner, T., and J. Dudhia. 2003. Genomics of osteoarthritis. *Curr Opin Rheumatol*. 15:634-640.

Aigner, T., H.A. Kim, and H.I. Roach. 2004a. Apoptosis in osteoarthritis. *Rheum Dis Clin North Am*. 30:639-653, xi.

Aigner, T., B. Kurz, N. Fukui, and L. Sandell. 2002. Roles of chondrocytes in the pathogenesis of osteoarthritis. *Curr Opin Rheumatol*. 14:578-584.

Aigner, T., E. Reichenberger, W. Bertling, T. Kirsch, H. Stoss, and K. von der Mark. 1993. Type X collagen expression in osteoarthritic and rheumatoid articular cartilage. *Virchows Arch B Cell Pathol Incl Mol Pathol*. 63:205-211.

Aigner, T., J. Rose, J. Martin, and J. Buckwalter. 2004b. Aging theories of primary osteoarthritis: from epidemiology to molecular biology. *Rejuvenation Res*. 7:134-145.

Aigner, T., S. Soeder, and J. Haag. 2006. IL-1beta and BMPs--interactive players of cartilage matrix degradation and regeneration. *Eur Cell Mater*. 12:49-56; discussion 56.

Aigner, T., A. Zien, A. Gehrsitz, P.M. Gebhard, and L. McKenna. 2001. Anabolic and catabolic gene expression pattern analysis in normal versus osteoarthritic cartilage using complementary DNA-array technology. *Arthritis Rheum*. 44:2777-2789.

Aimes, R.T., and J.P. Quigley. 1995. Matrix metalloproteinase-2 is an interstitial collagenase. Inhibitor-free enzyme catalyzes the cleavage of collagen fibrils and soluble native type I collagen generating the specific 3/4- and 1/4-length fragments. *J Biol Chem*. 270:5872-5876.

Alexopoulos, L., I. Youn, P. Bonaldo, and F. Guilak. 2009. Developmental and osteoarthritic changes in Col6a1-knockout mice: biomechanics of type VI collagen in the cartilage pericellular matrix. . *Arthritis Rheum*. 60: 771-779.

Alsalameh, S., R. Amin, T. Gemba, and M. Lotz. 2004. Identification of mesenchymal progenitor cells in normal and osteoarthritic human articular cartilage. *Arthritis Rheum*. 50:1522-1532.

Altman, R., E. Asch, D. Bloch, G. Bole, D. Borenstein, K. Brandt, W. Christy, T.D. Cooke, R. Greenwald, M. Hochberg, and et al. 1986. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. *Arthritis Rheum*. 29:1039-1049.

Ambros, V. 2004. The functions of animal microRNAs. *Nature*. 431:350-355.

Amstalden van Hove, E.R., D.F. Smith, and R.M. Heeren. 2010. A concise review of mass spectrometry imaging. *J Chromatogr A*. 1217:3946-3954.

Antipova, O., and J.P. Orgel. 2012. Non-enzymatic decomposition of collagen fibers by a biglycan antibody and a plausible mechanism for rheumatoid arthritis. *PLoS One*. 7:e32241.

Antonsson, P., D. Heinegard, and A. Oldberg. 1991. Posttranslational modifications of fibromodulin. *J Biol Chem*. 266:16859-16861.

Appleton, C.T., S.E. Usmani, J.S. Mort, and F. Beier. 2010. Rho/ROCK and MEK/ERK activation by transforming growth factor-alpha induces articular cartilage degradation. *Lab Invest*. 90:20-30.

Archer, C.W., and P. Francis-West. 2003. The chondrocyte. *Int J Biochem Cell Biol*. 35:401-404.

Arner, E.C. 2002. Aggrecanase-mediated cartilage degradation. *Curr Opin Pharmacol*. 2:322-329.

Arner, E.C., C.E. Hughes, C.P. Decicco, B. Caterson, and M.D. Tortorella. 1998. Cytokine-induced cartilage proteoglycan degradation is mediated by aggrecanase. *Osteoarthritis Cartilage*. 6:214-228.

Arner, E.C., M.A. Pratta, C.P. Decicco, C.B. Xue, R.C. Newton, J.M. Trzaskos, R.L. Magolda, and M.D. Tortorella. 1999. Aggrecanase. A target for the design of inhibitors of cartilage degradation. *Ann N Y Acad Sci*. 878:92-107.

Aye, T.T., T.Y. Low, Y. Bjorlykke, H. Barsnes, A.J. Heck, and F.S. Berven. 2012. Use of stable isotope dimethyl labeling coupled to selected reaction monitoring to enhance throughput by multiplexing relative quantitation of targeted proteins. *Anal Chem*. 84:4999-5006.

Baek, J.Y., S.M. Seo, K.B. Seung, H.J. Park, P.J. Kim, M.W. Park, Y.S. Koh, K.Y. Chang, M.H. Jeong, and S.J. Park. 2011. Clinical outcomes and predictors of unprotected left main stem culprit lesions in patients with acute ST segment elevation myocardial infarction. *Catheter Cardiovasc Interv*.

Baker, A.H., D.R. Edwards, and G. Murphy. 2002. Metalloproteinase inhibitors: biological actions and therapeutic opportunities. *J Cell Sci*. 115:3719-3727.

Balluff, B., S. Rauser, S. Meding, M. Elsner, C. Schone, A. Feuchtinger, C. Schuhmacher, A. Novotny, U. Jutting, G. Maccarrone, H. Sarioglu, M. Ueffing, H. Braselmann, H. Zitzelsberger, R.M. Schmid, H. Hofler, M.P. Ebert, and A. Walch. 2011. MALDI imaging identifies prognostic seven-protein signature of novel tissue markers in intestinal-type gastric cancer. *Am J Pathol*. 179:2720-2729.

Barr, E. 2011. *In* Veterinary School. Vol. PhD. Liverpool, Liverpool.

Barr, E.D., G.L. Pinchbeck, P.D. Clegg, A. Boyde, and C.M. Riggs. 2009. Post mortem evaluation of palmar osteochondral disease (traumatic osteochondrosis) of the metacarpo/metatarsophalangeal joint in Thoroughbred racehorses. *Equine Vet J*. 41:366-371.

Barrett, A.J. 1978. The possible role of neutrophil proteinases in damage to articular cartilage. *Agents Actions*. 8:11-18.

Bau, B., P.M. Gebhard, J. Haag, T. Knorr, E. Bartnik, and T. Aigner. 2002. Relative messenger RNA expression profiling of collagenases and aggrecanases in human articular chondrocytes in vivo and in vitro. *Arthritis Rheum*. 46:2648-2657.

Beard, J.R., S. Biggs, B.D. Bloom, L.P. Fried, L. Hogan, A. Kalache, and S.J. Olshansky. 2011. Global Population Ageing: Peril or Promise. W.E. Forum, editor, Geneva.

Beck, M., A. Schmidt, J. Malmstroem, M. Claassen, A. Ori, A. Szymborska, F. Herzog, O. Rinner, J. Ellenberg, and R. Aebersold. 2011. The quantitative proteome of a human cell line. *Mol Syst Biol*. 7:549.

Becker, J.S., A. Matusch, C. Palm, D. Salber, and K.A. Morton. 2010. Bioimaging of metals in brain tissue by laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) and metallomics. *Metallomics*. 2:104-111.

Belcher, C., F. Fawthrop, R. Bunning, and M. Doherty. 1996. Plasminogen activators and their inhibitors in synovial fluids from normal, osteoarthritis, and rheumatoid arthritis knees. *Ann Rheum Dis*. 55:230-236.

Belluoccio, D., R. Wilson, D.J. Thornton, T.P. Wallis, J.J. Gorman, and J.F. Bateman. 2006. Proteomic analysis of mouse growth plate cartilage. *Proteomics*. 6:6549-6553.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Methodology*. 57:289-300.

Berenbaum, F. 2007. The quest for the Holy Grail: a disease-modifying osteoarthritis drug. *Arthritis Res Ther*. 9:111.

Berninger, P., D. Gaidatzis, E. van Nimwegen, and M. Zavolan. 2008. Computational analysis of small RNA cloning data. *Methods*. 44:13-21.

Beynon, R.J., M.K. Doherty, J.M. Pratt, and S.J. Gaskell. 2005. Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nat Methods*. 2:587-589.

Billinghurst, R.C., L. Dahlberg, M. Ionescu, A. Reiner, R. Bourne, C. Rorabeck, P. Mitchell, J. Hambor, O. Diekmann, H. Tschesche, J. Chen, H. Van Wart, and A.R. Poole. 1997. Enhanced cleavage of type II collagen by collagenases in osteoarthritic articular cartilage. *J Clin Invest*. 99:1534-1545.

Bislev, S.L., E.W. Deutsch, Z. Sun, T. Farrah, R. Aebersold, R.L. Moritz, E. Bendixen, and M.C. Codrea. 2012. A Bovine PeptideAtlas of milk and mammary gland proteomes. *Proteomics*. 12:2895-2899.

Blackburn, W.D., Jr., S. Chivers, and W. Bernreuter. 1996. Cartilage imaging in osteoarthritis. *Semin Arthritis Rheum*. 25:273-281.

Blanco, F.J., R.L. Ochs, H. Schwarz, and M. Lotz. 1995. Chondrocyte apoptosis induced by nitric oxide. *Am J Pathol*. 146:75-85.

Blaney Davidson, E.N., A. Scharstuhl, E.L. Vitters, P.M. van der Kraan, and W.B. van den Berg. 2005. Reduced transforming growth factor-beta signaling in cartilage of old mice: role in impaired repair capacity. *Arthritis Res Ther*. 7:R1338-1347.

Blatherwick, E.Q., G.J. Van Berkel, K. Pickup, M.K. Johansson, M.E. Beaudoin, R.O. Cole, J.M. Day, S. Iverson, I.D. Wilson, J.H. Scrivens, and D.J. Weston. 2011. Utility of spatially-resolved atmospheric pressure surface sampling and ionization techniques as alternatives to mass spectrometric imaging (MSI) in drug metabolism. *Xenobiotica*. 41:720-734.

Bondeson, J., S. Wainwright, C. Hughes, and B. Caterson. 2008. The regulation of the ADAMTS4 and ADAMTS5 aggrecanases in osteoarthritis: a review. *Clin Exp Rheumatol*. 26:139-145.

Bosserhoff, A.K., S. Kondo, M. Moser, U.H. Dietz, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, R. Buettner, and L.J. Sandell. 1997. Mouse CD-RAP/MIA gene: structure, chromosomal localization, and expression in cartilage and chondrosarcoma. *Dev Dyn*. 208:516-525.

Brew, C.J., R.P. Clegg, J.G. Boot-Handford, and T. Hardingham. 2008. Gene expression in human chondrocytes in late osteoarthritis is changed in both fibrillated and intact cartilage without evidence of generalised chondrocyte hypertrophy. *Ann Rheum Dis*. 69:234-240.

Briggs, M.D., and K.L. Chapman. 2002. Pseudoachondroplasia and multiple epiphyseal dysplasia: mutation review, molecular interactions, and genotype to phenotype correlations. *Hum Mutat*. 19:465-478.

Brommer, H., M.S. Laasanen, P.A. Brama, P.R. van Weeren, H.J. Helminen, and J.S. Jurvelin. 2005. Functional consequences of cartilage degeneration in the equine metacarpophalangeal joint: quantitative assessment of cartilage stiffness. *Equine Vet J*. 37:462-467.

Brosch, M., L. Yu, T. Hubbard, and J. Choudhary. 2009. Accurate and sensitive peptide identification with Mascot Percolator. *J Proteome Res*. 8:3176-3181.

Brown, D.C., and K.G. Vogel. 1989. Characteristics of the in vitro interaction of a small proteoglycan (PG II) of bovine tendon with type I collagen. *Matrix*. 9:468-478.

Brownridge, P., and R.J. Beynon. 2011. The importance of the digest: proteolysis and absolute quantification in proteomics. *Methods*. 54:351-360.

Brownridge, P., S.W. Holman, S.J. Gaskell, C.M. Grant, V.M. Harman, S.J. Hubbard, K. Lanthaler, C. Lawless, R. O'Cualain, P. Sims, R. Watkins, and R.J. Beynon. 2011. Global absolute quantification of a proteome: Challenges in the deployment of a QconCAT strategy. *Proteomics*. 11:2957-2970.

Brownridge, P.J., V.M. Harman, D.M. Simpson, and R.J. Beynon. 2012. Absolute multiplexed protein quantification using QconCAT technology. *Methods Mol Biol*. 893:267-293.

Buckland, J. 2012a. Experimental arthritis: RANKL from cartilage linked to subchondral bone loss. *Nat Rev Rheumatol*.

Buckland, J. 2012b. Osteoarthritis: Complement-mediated inflammation in OA progression. *Nature Reviews Rheumatology*. 8.

Buckland, J. 2012c. Osteoarthritis: Control of human cartilage hypertrophic differentiation. *Nat Rev Rheumatol*. 8:368.

Buckwalter, J.A., and H.J. Mankin. 1998. Articular cartilage: degeneration and osteoarthritis, repair, regeneration, and transplantation. *Instr Course Lect*. 47:487-504.

Buckwalter, J.A., P.J. Roughley, and L.C. Rosenberg. 1994. Age-related changes in cartilage proteoglycans: quantitative electron microscopic studies. *Microsc Res Tech*. 28:398-408.

Bulstra, S.K., W.A. Buurman, G.H. Walenkamp, and A.J. Van der Linden. 1989. Metabolic characteristics of in vitro cultured human chondrocytes in relation to the histopathologic grade of osteoarthritis. *Clin Orthop Relat Res*:294-302.

Burrage, P.S., K.S. Mix, and C.E. Brinckerhoff. 2006. Matrix metalloproteinases: role in arthritis. *Front Biosci*. 11:529-543.

Calamia, V., B. Rocha, J. Mateos, P. Fernandez-Puente, C. Ruiz-Romero, and F.J. Blanco. 2011. Metabolic Labeling of Chondrocytes for the Quantitative Analysis of the Interleukin-1-beta-mediated Modulation of Their Intracellular and Extracellular Proteomes. *J Proteome Res.*

Campisi, J. 2005. Senescent cells, tumor suppression, and organismal aging: good citizens, bad neighbors. *Cell*. 120:513-522.

Cantley, C.E., E.C. Firth, J.W. Delahunt, D.U. Pfeiffer, and K.G. Thompson. 1999. Naturally occurring osteoarthritis in the metacarpophalangeal joints of wild horses. *Equine Vet J*. 31:73-81.

Caprioli, R.M., T.B. Farmer, and J. Gile. 1997. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem*. 69:4751-4760.

Carlo, M.D., Jr., and R.F. Loeser. 2003. Increased oxidative stress with aging reduces chondrocyte survival: correlation with intracellular glutathione levels. *Arthritis Rheum*. 48:3419-3430.

Carman, W.J., M. Sowers, V.M. Hawthorne, and L.A. Weissfeld. 1994. Obesity as a risk factor for osteoarthritis of the hand and wrist: a prospective study. *Am J Epidemiol*. 139:119-129.

Carroll, K.M., D.M. Simpson, C.E. Eyers, C.G. Knight, P. Brownridge, W.B. Dunn, C.L. Winder, K. Lanthaler, P. Pir, N. Malys, D.B. Kell, S.G. Oliver, S.J. Gaskell, and R.J. Beynon. 2011. Absolute quantification of the glycolytic pathway in yeast: deployment of a complete QconCAT approach. *Mol Cell Proteomics*. 10:M111 007633.

Castro-Borges, W., D.M. Simpson, A. Dowle, R.S. Curwen, J. Thomas-Oates, R.J. Beynon, and R.A. Wilson. 2011. Abundance of tegument surface proteins in the human blood

fluke Schistosoma mansoni determined by QconCAT proteomics. *J Proteomics*. 74:1519-1533.

Caterson, B., C.R. Flannery, C.E. Hughes, and C.B. Little. 2000. Mechanisms involved in cartilage proteoglycan catabolism. *Matrix Biol*. 19:333-344.

Catterall, J.B., A.D. Rowan, S. Sarsfield, J. Saklatvala, R. Wait, and T.E. Cawston. 2006. Development of a novel 2D proteomics approach for the identification of proteins secreted by primary chondrocytes after stimulation by IL-1 and oncostatin M. *Rheumatology (Oxford)*. 45:1101-1109.

Cawston, T., C. Billington, C. Cleaver, S. Elliott, W. Hui, P. Koshy, B. Shingleton, and A. Rowan. 1999. The regulation of MMPs and TIMPs in cartilage turnover. *Ann N Y Acad Sci*. 878:120-129.

Chadjichristos, C., C. Ghayor, J.F. Herrouin, L. Ala-Kokko, G. Suske, J.P. Pujol, and P. Galera. 2002. Down-regulation of human type II collagen gene expression by transforming growth factor-beta 1 (TGF-beta 1) in articular chondrocytes involves SP3/SP1 ratio. *J Biol Chem*. 277:43903-43917.

Chang, I.S., K.B. Park, Y.S. Do, H.S. Park, S.W. Shin, S.K. Cho, S.W. Choo, I.W. Choo, D.I. Kim, and Y.W. Kim. 2011. Heavily calcified occlusive lesions of the iliac artery: long-term patency and CT findings after stent placement. *J Vasc Interv Radiol*. 22:1131-1137 e1131.

Chapman, K.L., G.R. Mortier, K. Chapman, J. Loughlin, M.E. Grant, and M.D. Briggs. 2001. Mutations in the region encoding the von Willebrand factor A domain of matrilin-3 are associated with multiple epiphyseal dysplasia. *Nat Genet*. 28:393-396.

Charni-Ben Tabassi, N., S. Desmarais, A.C. Bay-Jensen, J.M. Delaisse, M.D. Percival, and P. Garnero. 2008. The type II collagen fragments Helix-II and CTX-II reveal different enzymatic pathways of human cartilage collagen degradation. *Osteoarthritis Cartilage*. 16:1183-1191.

Chen, A.F., C.M. Davies, M. De Lin, and B. Fermor. 2008. Oxidative DNA damage in osteoarthritic porcine articular cartilage. *J Cell Physiol*. 217:828-833.

Chen, F.H., M.E. Herndon, N. Patel, J.T. Hecht, R.S. Tuan, and J. Lawler. 2007. Interaction of cartilage oligomeric matrix protein/thrombospondin 5 with aggrecan. *J Biol Chem*. 282:24591-24598.

Chen, F.H., A.O. Thomas, J.T. Hecht, M.B. Goldring, and J. Lawler. 2005. Cartilage oligomeric matrix protein/thrombospondin 5 supports chondrocyte attachment through interaction with integrins. *J Biol Chem*. 280:32655-32661.

Chen, J., M. Wang, and I.V. Turko. 2012. Mass spectrometry quantification of clusterin in the human brain. *Mol Neurodegener*. 7:41.

Chen, Y.K., C.H. Kao, S.S. Sun, J.A. Liang, C.H. Wang, Y.C. Wu, Y.Y. Lin, K.Y. Yen, and T.C. Hsieh. 2010. Exposing the evil in the dark: the usefulness of delayed-phase FDG PET scan to enhance the detectability of tiny residual skull base osteosarcoma initially concealed by adjacent high physiological brain activity. *Clin Nucl Med*. 35:630-632.

Chevalier, X., P. Claudepierre, N. Groult, and G.J. Godeau. 1996. Influence of interleukin 1 beta on tenascin distribution in human normal and osteoarthritic cartilage: a quantitative immunohistochemical study. *Ann Rheum Dis*. 55:772-775.

Choi, D., S. Yoon, E. Lee, S. Hwang, B. Yoon, and J. Lee. 2001. The expression of pseudogene cyclin D2 mRNA in the human ovary may be a novel marker for decreased ovarian function associated with the aging process. *J Assist Reprod Genet*. 18:110-113.

Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*. 162:156-159.

Christgau, S., P. Garnero, C. Fledelius, C. Moniz, M. Ensig, E. Gineyts, C. Rosenquist, and P. Qvist. 2001. Collagen type II C-telopeptide fragments as an index of cartilage degradation. *Bone*. 29:209-215.

Chubinskaya, S., K. Huch, K. Mikecz, G. Cs-Szabo, K.A. Hasty, K.E. Kuettner, and A.A. Cole. 1996. Chondrocyte matrix metalloproteinase-8: up-regulation of neutrophil collagenase by interleukin-1 beta in human cartilage from knee and ankle joints. *Lab Invest*. 74:232-240.

Chubinskaya, S., B. Kumar, C. Merrihew, K. Heretis, D.C. Rueger, and K.E. Kuettner. 2002. Age-related changes in cartilage endogenous osteogenic protein-1 (OP-1). *Biochim Biophys Acta*. 1588:126-134.

Chughtai, K., and R.M. Heeren. 2010. Mass spectrometric imaging for biomedical tissue analysis. *Chem Rev*. 110:3237-3277.

Cillero-Pastor, B., B. Eijkel, A. Kiss, R.J. Blanco, and R.M.A. Heeren. 2012a. Matrix assisted laser desorption ionization imaging mass spectrometry: A new technique to study human osteoarthritic cartilage *Arthritis and Rhematism*. in press.

Cillero-Pastor, B., G. Eijkel, A. Kiss, F.J. Blanco Garcia, and R.M. Heeren. 2012b. A time of flight secondary ion mass spectrometry based molecular distribution distinguishes healthy and osteoarthritic human cartilage. *Anal Chem*.

Cillero-Pastor, B., G.B. Eijkel, A. Kiss, F.J. Blanco, and R.M. Heeren. 2012c. Matrix assisted laser desorption ionization imaging mass spectrometry: A new methodology to study human osteoarthritic cartilage. *Arthritis Rheum*.

Clark, A.G., J.M. Jordan, V. Vilim, J.B. Renner, A.D. Dragomir, G. Luta, and V.B. Kraus. 1999. Serum cartilage oligomeric matrix protein reflects osteoarthritis presence and severity: the Johnston County Osteoarthritis Project. *Arthritis Rheum*. 42:2356-2364.

Clark, I.M., L.K. Powell, S. Ramsey, B.L. Hazleman, and T.E. Cawston. 1993. The measurement of collagenase, tissue inhibitor of metalloproteinases (TIMP), and collagenase-TIMP complex in synovial fluids from patients with osteoarthritis and rheumatoid arthritis. *Arthritis Rheum*. 36:372-379.

Clark, T.A., C.W. Sugnet, and M. Ares, Jr. 2002. Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science*. 296:907-910.

Claydon, A.J., S.A. Ramm, A. Pennington, J.L. Hurst, P. Stockley, and R. Beynon. 2012. Heterogenous turnover of sperm and seminal vesicle proteins in the mouse revealed by dynamic metabolic labeling. *Mol Cell Proteomics*. 11:M111 014993.

Clutterbuck, A.L., J.R. Smith, D. Allaway, P. Harris, S. Liddell, and A. Mobasheri. 2011. High throughput proteomic analysis of the secretome in an explant model of articular cartilage inflammation. *J Proteomics*. 74:704-715.

Cobon, G.S., N. Verrills, P. Papakostopoulos, H. Eastwood, and A.W. Linnane. 2002. The proteomics of ageing. *Biogerontology*. 3:133-136.

Collier, S., and P. Ghosh. 1988. The role of plasminogen in interleukin-1 mediated cartilage degradation. *J Rheumatol*. 15:1129-1137.

Conaghan, P.G., D. Felson, G. Gold, S. Lohmander, S. Totterman, and R. Altman. 2006. MRI and non-cartilaginous structures in knee osteoarthritis. *Osteoarthritis Cartilage*. 14 Suppl A:A87-94.

Connie, Y., M.D. Chang, A. J., and M.D. Huang. 2011. MR of articular cartilage lesions of the knee

*Applied Radiology*. 40.

Cooper, C., S. Snow, T.E. McAlindon, S. Kellingray, B. Stuart, D. Coggon, and P.A. Dieppe. 2000. Risk factors for the incidence and progression of radiographic knee osteoarthritis. *Arthritis Rheum*. 43:995-1000.

Corless, S., and R. Cramer. 2003. On-target oxidation of methionine residues using hydrogen peroxide for composition-restricted matrix-assisted laser desorption/ionisation peptide mass mapping. *Rapid Commun Mass Spectrom*. 17:1212-1215.

Craig, R., J.P. Cortens, and R.C. Beavis. 2004. Open source system for analyzing, validating, and storing protein identification data. *J Proteome Res*. 3:1234-1242.

Cs-Szabo, G., P.J. Roughley, A.H. Plaas, and T.T. Glant. 1995. Large and small proteoglycans of osteoarthritic and rheumatoid articular cartilage. *Arthritis Rheum*. 38:660-668.

Cummings, J., T.H. Ward, A. Greystoke, M. Ranson, and C. Dive. 2008. Biomarker method validation in anticancer drug development. *Br J Pharmacol*. 153:646-656.

Curtis, H., A. Marusyk, V. Zaberezhnyy, M. Casas, B. Adane, A. Merz, N. Serkova, and J. DeGregori. 2012. Aging-associated alterations in IL-7 receptor signaling and inflammation promote declining B-lymphopoiesis and increased leukemogenesis *The Journal of Immunology, 2012, 188, 109.9*. 188:109.109.

Davidson, R.K., J.G. Waters, L. Kevorkian, C. Darrah, A. Cooper, S.T. Donell, and I.M. Clark. 2006. Expression profiling of metalloproteinases and their inhibitors in synovium and cartilage. *Arthritis Res Ther*. 8:R124.

Davies, M.E., A. Horner, B.E. Loveland, and I.F. McKenzie. 1994. Upregulation of complement regulators MCP (CD46), DAF (CD55) and protectin (CD59) in arthritic joint disease. *Scand J Rheumatol*. 23:316-321.

Day, J.M., A.I. Olin, A.D. Murdoch, A. Canfield, T. Sasaki, R. Timpl, T.E. Hardingham, and A. Aspberg. 2004. Alternative splicing in the aggrecan G3 domain influences binding interactions with tenascin-C and other extracellular matrix proteins. *J Biol Chem*. 279:12511-12518.

De Ceuninck, F., L. Dassencourt, and P. Anract. 2004. The inflammatory side of human chondrocytes unveiled by antibody microarrays. *Biochem Biophys Res Commun*. 323:960-969.

De Ceuninck, F., E. Marcheteau, S. Berger, A. Caliez, V. Dumont, M. Raes, P. Anract, G. Leclerc, J.A. Boutin, and G. Ferry. 2005. Assessment of some tools for the characterization of the human osteoarthritic cartilage proteome. *J Biomol Tech*. 16:256-265.

de Hoog, C.L., and M. Mann. 2004. Proteomics. *Annu Rev Genomics Hum Genet*. 5:267-293.

de Magalhaes, J.P., C.E. Finch, and G. Janssens. 2010. Next-generation sequencing in aging research: emerging applications, problems, pitfalls and possible solutions. *Ageing Res Rev*. 9:315-323.

Dean, D.D., J. Martel-Pelletier, J.P. Pelletier, D.S. Howell, and J.F. Woessner, Jr. 1989. Evidence for metalloproteinase and metalloproteinase inhibitor imbalance in human osteoarthritic cartilage. *J Clin Invest*. 84:678-685.

Dean, R.A., and C.M. Overall. 2007. Proteomics discovery of metalloproteinase substrates in the cellular context by iTRAQ labeling reveals a diverse MMP-2 substrate degradome. *Mol Cell Proteomics*. 6:611-623.

Deckelbaum, R.J., S. Eisenberg, Y. Oschry, E. Butbul, I. Sharon, and T. Olivecrona. 1982. Reversible modification of human plasma low density lipoproteins toward triglyceride-rich precursors. A mechanism for losing excess cholesterol esters. *J Biol Chem*. 257:6509-6517.

Dejica, V.M., J.S. Mort, S. Laverty, M.D. Percival, J. Antoniou, D.J. Zukor, and A.R. Poole. 2008. Cleavage of type II collagen by cathepsin K in human osteoarthritic cartilage. *Am J Pathol*. 173:161-169.

Desiere, F., E.W. Deutsch, A.I. Nesvizhskii, P. Mallick, N.L. King, J.K. Eng, A. Aderem, R. Boyle, E. Brunner, S. Donohoe, N. Fausto, E. Hafen, L. Hood, M.G. Katze, K.A. Kennedy, F. Kregenow, H. Lee, B. Lin, D. Martin, J.A. Ranish, D.J. Rawlings, L.E.

Samelson, Y. Shiio, J.D. Watts, B. Wollscheid, M.E. Wright, W. Yan, L. Yang, E.C. Yi, H. Zhang, and R. Aebersold. 2005. Integration with the human genome of peptide sequences obtained by high-throughput mass spectrometry. *Genome Biol*. 6:R9.

Diaz-Prado, S., C. Cicione, E. Muinos-Lopez, T. Hermida-Gomez, N. Oreiro, C. Fernandez-Lopez, and F.J. Blanco. 2012. Characterization of microRNA expression profiles in normal and osteoarthritic human chondrocytes. *BMC Musculoskelet Disord*. 13:144.

Dickinson, S.C., M.N. Vankemmelbeke, D.J. Buttle, K. Rosenberg, D. Heinegard, and A.P. Hollander. 2003. Cleavage of cartilage oligomeric matrix protein (thrombospondin-5) by matrix metalloproteinases and a disintegrin and metalloproteinase with thrombospondin motifs. *Matrix Biol*. 22:267-278.

Dieppe, P. 1995. Osteoarthritis and molecular markers. A rheumatologist's perspective. *Acta Orthop Scand Suppl*. 266:1-5.

Ding, C., Y. Li, B.J. Kim, A. Malovannaya, S.Y. Jung, Y. Wang, and J. Qin. 2011. Quantitative analysis of cohesin complex stoichiometry and SMC3 modification-dependent protein interactions. *J Proteome Res*. 10:3652-3659.

Ding, L., D. Guo, and G.A. Homandberg. 2009. Fibronectin fragments mediate matrix metalloproteinase upregulation and cartilage damage through proline rich tyrosine kinase 2, c-src, NF-kappaB and protein kinase Cdelta. *Osteoarthritis Cartilage*. 17:1385-1392.

Dong, Y.F., Y. Soung do, E.M. Schwarz, R.J. O'Keefe, and H. Drissi. 2006. Wnt induction of chondrocyte hypertrophy through the Runx2 transcription factor. *J Cell Physiol*. 208:77-86.

Du, J., X. Wu, H. Zhang, S. Wang, W. Tan, and X. Guo. 2010. Mass spectrometry-based proteomic analysis of Kashin-Beck disease. *Mol Med Report*. 3:821-824.

Dudhia, J., C.M. Davidson, T.M. Wells, T.E. Hardingham, and M.T. Bayliss. 1996. Studies on the G3 domain of aggrecan from human cartilage. *Ann N Y Acad Sci*. 785:245-247.

Dufield, D.R., O.V. Nemirovskiy, M.G. Jennings, M.D. Tortorella, A.M. Malfait, and W.R. Mathews. An immunoaffinity liquid chromatography-tandem mass spectrometry assay for detection of endogenous aggrecan fragments in biological fluids: Use as a biomarker for aggrecanase activity and cartilage degradation. *Anal Biochem*. 406:113-123.

Dufield, D.R., O.V. Nemirovskiy, M.G. Jennings, M.D. Tortorella, A.M. Malfait, and W.R. Mathews. 2010a. An Immunoaffinity LC/MS/MS Assay for Detection of Endogenous Aggrecan Fragments in Biological Fluids. Use as a Biomarker for Aggrecanase Activity and Cartilage Degradation. *Anal Biochem*.

Dufield, D.R., O.V. Nemirovskiy, M.G. Jennings, M.D. Tortorella, A.M. Malfait, and W.R. Mathews. 2010b. An immunoaffinity liquid chromatography-tandem mass spectrometry assay for detection of endogenous aggrecan fragments in biological fluids: Use as a biomarker for aggrecanase activity and cartilage degradation. *Anal Biochem*. 406:113-123.

Dupuis, A., J.A. Hennekinne, J. Garin, and V. Brun. 2008. Protein Standard Absolute Quantification (PSAQ) for improved investigation of staphylococcal food poisoning outbreaks. *Proteomics*. 8:4633-4636.

Durigova, M., H. Nagase, J.S. Mort, and P.J. Roughley. 2011. MMPs are less efficient than ADAMTS5 in cleaving aggrecan core protein. *Matrix Biol*. 30:145-153.

Durigova, M., P.J. Roughley, and J.S. Mort. 2008. Mechanism of proteoglycan aggregate degradation in cartilage stimulated with oncostatin M. *Osteoarthritis Cartilage*. 16:98-104.

Edman, P., and G. Begg. 1967. A protein sequenator. *Eur J Biochem*. 1:80-91.

El-Aneeda, A., B. Cohen, and J. Banoubc. 2009. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Applied Spectroscopy Reviews* 44.

Elfervig, M.K., R.D. Graff, G.M. Lee, S.S. Kelley, A. Sood, and A.J. Banes. 2001. ATP induces Ca(2+) signaling in human chondrons cultured in three-dimensional agarose films. *Osteoarthritis Cartilage*. 9:518-526.

English, J.L., Z. Kassiri, I. Koskivirta, S.J. Atkinson, M. Di Grappa, P.D. Soloway, H. Nagase, E. Vuorio, G. Murphy, and R. Khokha. 2006. Individual Timp deficiencies differentially impact pro-MMP-2 activation. *J Biol Chem*. 281:10337-10346.

Everts, V., W.S. Hou, X. Rialland, W. Tigchelaar, P. Saftig, D. Bromme, B.D. Gelb, and W. Beertsen. 2003. Cathepsin K deficiency in pycnodysostosis results in accumulation of non-digested phagocytosed collagen in fibroblasts. *Calcif Tissue Int*. 73:380-386.

Eyers, C.E., C. Lawless, D.C. Wedge, K.W. Lau, S.J. Gaskell, and S.J. Hubbard. 2011. CONSeQuence: prediction of reference peptides for absolute quantitative proteomics using consensus machine learning approaches. *Mol Cell Proteomics*. 10:M110 003384.

Eyre, D. 2002. Collagen of articular cartilage. *Arthritis Res*. 4:30-35.

Eyre, D.R. 1995. The specificity of collagen cross-links as markers of bone and connective tissue degradation. *Acta Orthop Scand Suppl*. 266:166-170.

Eyre, D.R., T. Pietka, M.A. Weis, and J.J. Wu. 2004. Covalent cross-linking of the NC1 domain of collagen type IX to collagen type II in cartilage. *J Biol Chem*. 279:2568-2574.

Eyre, D.R., and M.A. Weis. 2009. The Helix-II epitope: a cautionary tale from a cartilage biomarker based on an invalid collagen sequence. *Osteoarthritis Cartilage*. 17:423-426.

Fabunmi, R.P., G.K. Sukhova, S. Sugiyama, and P. Libby. 1998. Expression of tissue inhibitor of metalloproteinases-3 in human atheroma and regulation in lesion-associated cells: a potential protective mechanism in plaque stability. *Circ Res*. 83:270-278.

Felson, D.T. 2004. Risk factors for osteoarthritis: understanding joint vulnerability. *Clin Orthop Relat Res*:S16-21.

Fenn, J.B., M. Mann, C.K. Meng, S.F. Wong, and C.M. Whitehouse. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science*. 246:64-71.

Fernandes, J.C., J. Martel-Pelletier, and J.P. Pelletier. 2002. The role of cytokines in osteoarthritis pathophysiology. *Biorheology*. 39:237-246.

Finnson, K.W., Y. Chi, G. Bou-Gharios, A. Leask, and A. Philip. 2012. TGF-b signaling in cartilage homeostasis and osteoarthritis. *Front Biosci (Schol Ed)*. 4:251-268.

Folgueras, A.R., A.M. Pendas, L.M. Sanchez, and C. Lopez-Otin. 2004. Matrix metalloproteinases in cancer: from new functions to improved inhibition strategies. *Int J Dev Biol*. 48:411-424.

Font, B., D. Eichenberger, D. Goldschmidt, M.M. Boutillon, and D.J. Hulmes. 1998. Structural requirements for fibromodulin binding to collagen and the control of type I collagen fibrillogenesis--critical roles for disulphide bonding and the C-terminal region. *Eur J Biochem*. 254:580-587.

Fontan, P.A., V. Pancholi, M. Nociari, and V.A. Fischett. 2000. Antibodies to streptococcal surface enolase react with human alpha-enolase: implications in poststreptococcal sequelae. *J Infect Dis*. 182:1712-1721.

Forsyth, C.B., A. Cole, G. Murphy, J.L. Bienias, H.J. Im, and R.F. Loeser, Jr. 2005. Increased matrix metalloproteinase-13 production with aging by human articular chondrocytes in response to catabolic stimuli. *J Gerontol A Biol Sci Med Sci*. 60:1118-1124.

Fosang, A.J., K. Last, P. Gardiner, D.C. Jackson, and L. Brown. 1995. Development of a cleavage-site-specific monoclonal antibody for detecting metalloproteinase-

derived aggrecan fragments: detection of fragments in human synovial fluids. *Biochem J*. 310 ( Pt 1):337-343.

Fosang, A.J., K. Last, V. Knauper, G. Murphy, and P.J. Neame. 1996a. Degradation of cartilage aggrecan by collagenase-3 (MMP-13). *FEBS Lett*. 380:17-20.

Fosang, A.J., K. Last, and R.A. Maciewicz. 1996b. Aggrecan is degraded by matrix metalloproteinases in human arthritis. Evidence that matrix metalloproteinase and aggrecanase activities can be independent. *J Clin Invest*. 98:2292-2299.

Fosang, A.J., K. Last, H. Stanton, S.B. Golub, C.B. Little, L. Brown, and D.C. Jackson. 2010. Neoepitope antibodies against MMP-cleaved and aggrecanase-cleaved aggrecan. *Methods Mol Biol*. 622:312-347.

Fosang, A.J., P.J. Neame, K. Last, T.E. Hardingham, G. Murphy, and J.A. Hamilton. 1992. The interglobular domain of cartilage aggrecan is cleaved by PUMP, gelatinases, and cathepsin B. *J Biol Chem*. 267:19470-19474.

Fosang, A.J., F.M. Rogerson, C.J. East, and H. Stanton. 2008. ADAMTS-5: the story so far. *Eur Cell Mater*. 15:11-26.

Fosang, A.J., J.A. Tyler, and T.E. Hardingham. 1991. Effect of interleukin-1 and insulin like growth factor-1 on the release of proteoglycan components and hyaluronan from pig articular cartilage in explant culture. *Matrix*. 11:17-24.

Foss, E.J., D. Radulovic, S.A. Shaffer, D.R. Goodlett, L. Kruglyak, and A. Bedalov. 2011. Genetic variation shapes protein networks mainly through non-transcriptional mechanisms. *PLoS Biol*. 9:e1001144.

Fraga, M.F., and M. Esteller. 2007. Epigenetics and aging: the targets and the marks. *Trends Genet*. 23:413-418.

Fries, H.E., M. Aiello, and C.A. Evans. 2007. Use of a variably sensitive selected reaction monitoring method to extend the linear dynamic range for quantitative high-performance liquid chromatography/tandem mass spectrometry. *Rapid Commun Mass Spectrom*. 21:369-374.

Frisbie, D.D., M.W. Cross, and C.W. McIlwraith. 2006. A comparative study of articular cartilage thickness in the stifle of animal species used in human pre-clinical studies compared to articular cartilage thickness in the human knee. *Vet Comp Orthop Traumatol*. 19:142-146.

Fushimi, K., L. Troeberg, H. Nakamura, N.H. Lim, and H. Nagase. 2008. Functional differences of the catalytic and non-catalytic domains in human ADAMTS-4 and ADAMTS-5 in aggrecanolytic activity. *J Biol Chem*. 283:6706-6716.

Garcia, B.A., M.D. Platt, T.L. Born, J. Shabanowitz, N.A. Marcus, and D.F. Hunt. 2006. Protein profile of osteoarthritic human articular cartilage using tandem mass spectrometry. *Rapid Commun Mass Spectrom*. 20:2999-3006.

Garvican, E.R., A. Vaughan-Thomas, C. Redmond, and P.D. Clegg. 2008. MT3-MMP (MMP-16) is downregulated by in vitro cytokine stimulation of cartilage, but unaltered in naturally occurring equine osteoarthritis and osteochondrosis. *Connect Tissue Res*. 49:62-67.

Garvican, E.R., A. Vaughan-Thomas, C. Redmond, N. Gabriel, and P.D. Clegg. 2010. MMP-mediated collagen breakdown induced by activated protein C in equine cartilage is reduced by corticosteroids. *J Orthop Res*. 28:370-378.

Gendron, C., M. Kashiwagi, N.H. Lim, J.J. Enghild, I.B. Thogersen, C. Hughes, B. Caterson, and H. Nagase. 2007. Proteolytic activities of human ADAMTS-5: comparative studies with ADAMTS-4. *J Biol Chem*. 282:18294-18306.

Geng, Y., D. McQuillan, and P.J. Roughley. 2006. SLRP interaction can protect collagen fibrils from cleavage by collagenases. *Matrix Biol*. 25:484-491.

Gerber, S.A., J. Rush, O. Stemman, M.W. Kirschner, and S.P. Gygi. 2003. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A*. 100:6940-6945.

Gevaert, K., M. Goethals, L. Martens, J. Van Damme, A. Staes, G.R. Thomas, and J. Vandekerckhove. 2003. Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nat Biotechnol*. 21:566-569.

Geyer, M., S. Grassel, R.H. Straub, G. Schett, R. Dinser, J. Grifka, S. Gay, E. Neumann, and U. Muller-Ladner. 2009. Differential transcriptome analysis of intraarticular lesional vs intact cartilage reveals new candidate genes in osteoarthritis pathophysiology. *Osteoarthritis Cartilage*. 17:328-335.

Glant, T.T., E.I. Buzas, A. Finnegan, G. Negroiu, G. Cs-Szabo, and K. Mikecz. 1998. Critical roles of glycosaminoglycan side chains of cartilage proteoglycan (aggrecan) in antigen recognition and presentation. *J Immunol*. 160:3812-3819.

Goekoop, R.J., M. Kloppenburg, H.M. Kroon, M. Frolich, T.W. Huizinga, R.G. Westendorp, and J. Gussekloo. 2010. Low innate production of interleukin-1beta and interleukin-6 is associated with the absence of osteoarthritis in old age. *Osteoarthritis Cartilage*. 18:942-947.

Goessler, U.R., K. Bieback, P. Bugert, R. Naim, C. Schafer, H. Sadick, K. Hormann, and F. Riedel. 2005. Human chondrocytes differentially express matrix modulators during in vitro expansion for tissue engineering. *Int J Mol Med*. 16:509-515.

Goldring, M.B. 2000. The role of the chondrocyte in osteoarthritis. *Arthritis Rheum*. 43:1916-1926.

Goldring, M.B., and S.R. Goldring. 2007. Osteoarthritis. *J Cell Physiol*. 213:626-634.

Goldring, M.B., and K.B. Marcu. 2009. Cartilage homeostasis in health and rheumatic diseases. *Arthritis Res Ther*. 11:224.

Goldring, M.B., M. Otero, D.A. Plumb, C. Dragomir, M. Favero, K. El Hachem, K. Hashimoto, H.I. Roach, E. Olivotto, R.M. Borzi, and K.B. Marcu. 2011. Roles of inflammatory and anabolic cytokines in cartilage metabolism: signals and multiple effectors converge upon MMP-13 regulation in osteoarthritis. *Eur Cell Mater*. 21:202-220.

Goldring, S.R., and M.B. Goldring. 2004. The role of cytokines in cartilage matrix degeneration in osteoarthritis. *Clin Orthop Relat Res*:S27-36.

Golightly, Y.M., S.W. Marshall, L.F. Callahan, and K. Guskiewicz. 2009. Early-onset arthritis in retired National Football League players. *J Phys Act Health*. 6:638-643.

Gomez, D.E., D.F. Alonso, H. Yoshiji, and U.P. Thorgeirsson. 1997. Tissue inhibitors of metalloproteinases: structure, regulation and biological functions. *Eur J Cell Biol*. 74:111-122.

Gorman, M.O., A. Ivanov, P. Lavery, and M. Bennett. 2009. Validating a unique approach for label-free LCMS data analysis. *In* Cancer Proteomics, Dublin.

Grad, S., C.R. Lee, and M. Alini. 2006. Biology: mechanisms of cartilage breakdown and repair, .Basic Science, Clinical Repair and Reconstruction of Articular Cartilage Defects. B.M. Zanasi S, Marcacci M,, editor.

Grahame, T.J., and R.B. Schlesinger. 2012. Oxidative stress-induced telomeric erosion as a mechanism underlying airborne particulate matter-related cardiovascular disease. *Part Fibre Toxicol*. 9:21.

Grant, W.T., M.D. Sussman, and G. Balian. 1985. A disulfide-bonded short chain collagen synthesized by degenerative and calcifying zones of bovine growth plate cartilage. *J Biol Chem*. 260:3798-3803.

Greenbaum, D., C. Colangelo, K. Williams, and M. Gerstein. 2003. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol*. 4:117.

Groeneveld, T.W., M. Oroszlan, R.T. Owens, M.C. Faber-Krol, A.C. Bakker, G.J. Arlaud, D.J. McQuillan, U. Kishore, M.R. Daha, and A. Roos. 2005. Interactions of the extracellular matrix proteoglycans decorin and biglycan with C1q and collectins. *J Immunol*. 175:4715-4723.

Groutars, R.G., J.F. Verzijlbergen, A.J. Muller, C.A. Ascoop, M.M. Tiel-van Buul, A.H. Zwinderman, N.M. van Hemel, and E.E. van der Wall. 2000. Prognostic value and quality of life in patients with normal rest thallium-201/stress technetium 99m-tetrofosmin dual-isotope myocardial SPECT. *J Nucl Cardiol*. 7:333-341.

Gruber, J., T.L. Vincent, M. Hermansson, M. Bolton, R. Wait, and J. Saklatvala. 2004. Induction of interleukin-1 in articular cartilage by explantation and cutting. *Arthritis Rheum*. 50:2539-2546.

Guo, D., W. Tan, F. Wang, Z. Lv, J. Hu, T. Lv, Q. Chen, X. Gu, B. Wan, and Z. Zhang. 2008. Proteomic analysis of human articular cartilage: identification of differentially expressed proteins in knee osteoarthritis. *Joint Bone Spine*. 75:439-444.

Gustafsson, J.O., M.K. Oehler, A. Ruszkiewicz, S.R. McColl, and P. Hoffmann. 2011. MALDI Imaging Mass Spectrometry (MALDI-IMS)-Application of Spatial Proteomics for Ovarian Cancer Classification and Diagnosis. *Int J Mol Sci*. 12:773-794.

Gygi, S.P., B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, and R. Aebersold. 1999a. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol*. 17:994-999.

Gygi, S.P., Y. Rochon, B.R. Franza, and R. Aebersold. 1999b. Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol*. 19:1720-1730.

Haglund, L., S.M. Bernier, P. Onnerfjord, and A.D. Recklies. 2008. Proteomic analysis of the LPS-induced stress response in rat chondrocytes reveals induction of innate immune response components in articular cartilage. *Matrix Biol*. 27:107-118.

Halasz, K., A. Kassner, M. Morgelin, and D. Heinegard. 2007. COMP acts as a catalyst in collagen fibrillogenesis. *J Biol Chem*. 282:31166-31173.

Hall, N. 2007. Advanced sequencing technologies and their wider impact in microbiology. *J Exp Biol*. 210:1518-1525.

Han, M., S.W. Lin, S.O. Smith, and T.P. Sakmar. 1996. The effects of amino acid replacements of glycine 121 on transmembrane helix 3 of rhodopsin. *J Biol Chem*. 271:32330-32336.

Han, X., L. He, B. Xin, and B. Shan. 2011. PEAKS PTM: Mass Spectrometry Based Identification of Peptides with Unspecified Modifications. *Journal of Proteomics Research*. 10:2930-2936.

Handley, C.J., M.T. Mok, M.Z. Ilic, C. Adcocks, D.J. Buttle, and H.C. Robinson. 2001. Cathepsin D cleaves aggrecan at unique sites within the interglobular domain and chondroitin sulfate attachment regions that are also cleaved when cartilage is maintained at acid pH. *Matrix Biol*. 20:543-553.

Hankin, J.A., S.E. Farias, R.M. Barkley, K. Heidenreich, L.C. Frey, K. Hamazaki, H.Y. Kim, and R.C. Murphy. 2011. MALDI mass spectrometric imaging of lipids in rat brain injury models. *J Am Soc Mass Spectrom*. 22:1014-1021.

Hardingham, T. 1981. Proteoglycans: their structure, interactions and molecular organization in cartilage. *Biochem Soc Trans*. 9:489-497.

Hardingham, T., and M. Bayliss. 1990. Proteoglycans of articular cartilage: changes in aging and in joint disease. *Semin Arthritis Rheum*. 20:12-33.

Hardingham, T.E., M. Beardmore-Gray, D.G. Dunham, and A. Ratcliffe. 1986. Cartilage proteoglycans. *Ciba Found Symp*. 124:30-46.

Hardingham, T.E., and H. Muir. 1972. The specific interaction of hyaluronic acid with cartillage proteoglycans. *Biochim Biophys Acta*. 279:401-405.

Hargrave, M., K. James, K. Nield, C. Toomes, K. Georgas, T. Sullivan, H.T. Verzijl, C.A. Oley, M. Little, P. De Jonghe, J.M. Kwon, H. Kremer, M.J. Dixon, V. Timmerman, T. Yamada, and P. Koopman. 2000. Fine mapping of the neurally expressed gene SOX14 to human 3q23, relative to three congenital diseases. *Hum Genet*. 106:432-439.

Harman, V. 2012. QcxonCAT expression. J. Peffers M, editor, Liverpool.

Harrow, J., A. Frankish, J.M. Gonzalez, E. Tapanari, M. Diekhans, F. Kokocinski, B.L. Aken, D. Barrell, A. Zadissa, S. Searle, I. Barnes, A. Bignell, V. Boychenko, T. Hunt, M. Kay, G. Mukherjee, J. Rajan, G. Despacio-Reyes, G. Saunders, C. Steward, R. Harte, M. Lin, C. Howald, A. Tanzer, T. Derrien, J. Chrast, N. Walters, S. Balasubramanian, B. Pei, M. Tress, J.M. Rodriguez, I. Ezkurdia, J. van Baren, M. Brent, D. Haussler, M. Kellis, A. Valencia, A. Reymond, M. Gerstein, R. Guigo, and T.J. Hubbard. 2012. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*. 22:1760-1774.

Hathout, Y. 2007. Approaches to the study of the cell secretome. *Expert Rev Proteomics*. 4:239-248.

Haudenschild, D.R., J. Chen, N. Pang, M.K. Lotz, and D.D. D'Lima. 2010. Rho kinase-dependent activation of SOX9 in chondrocytes. *Arthritis Rheum*. 62:191-200.

Heathfield, T.F., P. Onnerfjord, L. Dahlberg, and D. Heinegard. 2004. Cleavage of fibromodulin in cartilage explants involves removal of the N-terminal tyrosine sulfate-rich region by proteolysis at a site that is sensitive to matrix metalloproteinase-13. *J Biol Chem*. 279:6286-6295.

Hedbom, E., P. Antonsson, A. Hjerpe, D. Aeschlimann, M. Paulsson, E. Rosa-Pimentel, Y. Sommarin, M. Wendel, A. Oldberg, and D. Heinegard. 1992. Cartilage matrix proteins. An acidic oligomeric protein (COMP) detected only in cartilage. *J Biol Chem*. 267:6132-6136.

Hedbom, E., and D. Heinegard. 1989. Interaction of a 59-kDa connective tissue matrix protein with collagen I and collagen II. *J Biol Chem*. 264:6898-6905.

Hedbom, E., and D. Heinegard. 1993. Binding of fibromodulin and decorin to separate sites on fibrillar collagens. *J Biol Chem*. 268:27307-27312.

Heeren, R.M. 2012. Innovation in molecular imaging with mass spectrometry: running towards high resolution. *Osteoarthritis and cartilage*. 20:S1-S2.

Heeren, R.M.A., L.A. McDonnell, E.R. Amstalden van Hove, S.L. Luxembourg, A. A.F.M., and P. S.R. 2006. Why don't biologists use SIMS? : A critical evaluation of imaging MS. *Appl Surf. Sci.* 252:6827--6835.

Heidebrecht, F., A. Heidebrecht, I. Schulz, S.E. Behrens, and A. Bader. 2009. Improved semiquantitative Western blot technique with increased quantification range. *J Immunol Methods*. 345:40-48.

Heinegard, D. 1977. Polydispersity of cartilage proteoglycans. Structural variations with size and buoyant density of the molecules. *J Biol Chem*. 252:1980-1989.

Hellemans, J., G. Mortier, A. De Paepe, F. Speleman, and J. Vandesompele. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol*. 8:R19.

Helmick, C.G., D.T. Felson, R.C. Lawrence, S. Gabriel, R. Hirsch, C.K. Kwoh, M.H. Liang, H.M. Kremers, M.D. Mayes, P.A. Merkel, S.R. Pillemer, J.D. Reveille, and J.H. Stone. 2008. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. *Arthritis Rheum*. 58:15-25.

Henriques, C.M., J. Rino, R.J. Nibbs, G.J. Graham, and J.T. Barata. 2010. IL-7 induces rapid clathrin-mediated internalization and JAK3-dependent degradation of IL-7Ralpha in T cells. *Blood*. 115:3269-3277.

Henzel, W.J., T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley, and C. Watanabe. 1993. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci U S A*. 90:5011-5015.

Hermansson, M., J. Saklatvala, and R. Wait. 2007. Two-dimensional electrophoresis of proteins secreted from articular cartilage. *Methods Mol Med*. 136:349-359.

Hermansson, M., Y. Sawaji, M. Bolton, S. Alexander, A. Wallace, S. Begum, R. Wait, and J. Saklatvala. 2004. Proteomic analysis of articular cartilage shows increased type II collagen synthesis in osteoarthritis and expression of inhibin betaA (activin A), a regulatory molecule for chondrocytes. *J Biol Chem*. 279:43514-43521.

Hoch, J.M., C.G. Mattacola, J.M. Medina McKeon, J.S. Howard, and C. Lattermann. 2011. Serum cartilage oligomeric matrix protein (sCOMP) is elevated in patients with knee osteoarthritis: a systematic review and meta-analysis. *Osteoarthritis Cartilage*. 19:1396-1404.

Hoffman, K.D., R.R. Pool, and J.R. Pascoe. 1984. Degenerative joint disease of the proximal interphalangeal joints of the forelimbs of two young horses. *Equine Vet J*. 16:138-140.

Holden, P. 2012. Cartlage oligomeric matrix protein degradation. J. Peffers M, editor.

Holmskov, U., R. Malhotra, R.B. Sim, and J.C. Jensenius. 1994. Collectins: collagenous C-type lectins of the innate immune defense system. *Immunol Today*. 15:67-74.

Holtzer, H., J. Abbott, J. Lash, and S. Holtzer. 1960. The Loss of Phenotypic Traits by Differentiated Cells in Vitro, I. Dedifferentiation of Cartilage Cells. *Proc Natl Acad Sci U S A*. 46:1533-1542.

Homandberg, G.A. 1999. Potential regulation of cartilage metabolism in osteoarthritis by fibronectin fragments. *Front Biosci*. 4:D713-730.

Homandberg, G.A., G. Davis, C. Maniglia, and A. Shrikhande. 1997. Cartilage chondrolysis by fibronectin fragments causes cleavage of aggrecan at the same site as found in osteoarthritic cartilage. *Osteoarthritis Cartilage*. 5:450-453.

Homandberg, G.A., R. Meyers, M. Aydelotte, D. Tripier, and K.E. Kuettner. 1992. Isolation and characterization of an abundant elastase inhibitor from NaCl extracts of bovine nasal septa and articular cartilage. *Connect Tissue Res*. 28:289-305.

Hopewell, B., and J.P. Urban. 2003. Adaptation of articular chondrocytes to changes in osmolality. *Biorheology*. 40:73-77.

Horton, W.E., Jr., L. Feng, and C. Adams. 1998. Chondrocyte apoptosis in development, aging and disease. *Matrix Biol*. 17:107-115.

Hossain, M., D.T. Kaleta, E.W. Robinson, T. Liu, R. Zhao, J.S. Page, R.T. Kelly, R.J. Moore, K. Tang, D.G. Camp, 2nd, W.J. Qian, and R.D. Smith. 2011. Enhanced sensitivity for selected reaction monitoring mass spectrometry-based targeted proteomics using a dual stage electrodynamic ion funnel interface. *Mol Cell Proteomics*. 10:M000062-MCP000201.

Hou, W.S., Z. Li, F.H. Buttner, E. Bartnik, and D. Bromme. 2003. Cleavage site specificity of cathepsin K toward cartilage proteoglycans and protease complex formation. *Biol Chem*. 384:891-897.

Hoyland, J.A., J.T. Thomas, R. Donn, A. Marriott, S. Ayad, R.P. Boot-Handford, M.E. Grant, and A.J. Freemont. 1991. Distribution of type X collagen mRNA in normal and osteoarthritic human cartilage. *Bone Miner*. 15:151-163.

Huang da, W., B.T. Sherman, and R.A. Lempicki. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 4:44-57.

Huang, D.C., J.M. Adams, and S. Cory. 1998. The conserved N-terminal BH4 domain of Bcl-2 homologues is essential for inhibition of apoptosis and interaction with CED-4. *EMBO J*. 17:1029-1039.

Huang, D.W., B.T. Sherman, and R.A. Lempicki. 2008. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*. 4.

Huang, K., and L.D. Wu. 2008. Aggrecanase and aggrecan degradation in osteoarthritis: a review. *J Int Med Res*. 36:1149-1160.

Hubbard, R., M. Cooper, M. Antoniak, A. Venn, S. Khan, I. Johnston, S. Lewis, and J. Britton. 2000. Risk of cryptogenic fibrosing alveolitis in metal workers. *Lancet*. 355:466-467.

Hughes, C.E., B. Caterson, A.J. Fosang, P.J. Roughley, and J.S. Mort. 1995. Monoclonal antibodies that specifically recognize neoepitope sequences generated by 'aggrecanase' and matrix metalloproteinase cleavage of aggrecan: application to catabolism in situ and in vitro. *Biochem J*. 305 ( Pt 3):799-804.

Hughes, C.E., B. Caterson, R.J. White, P.J. Roughley, and J.S. Mort. 1992. Monoclonal antibodies recognizing protease-generated neoepitopes from cartilage proteoglycan degradation. Application to studies of human link protein cleavage by stromelysin. *J Biol Chem*. 267:16011-16014.

Hugle, T., J. Geurts, C. Nuesch, M. Muller-Gerbl, and V. Valderrabano. 2012. Aging and osteoarthritis: an inevitable encounter? *J Aging Res*. 2012:950192.

Hulejova, H., V. Baresova, Z. Klezl, M. Polanska, M. Adam, and L. Senolt. 2007. Increased level of cytokines and matrix metalloproteinases in osteoarthritic subchondral bone. *Cytokine*. 38:151-156.

Hyman, T., M. Shmuel, and Y. Altschuler. 2006. Actin is required for endocytosis at the apical surface of Madin-Darby canine kidney cells where ARF6 and clathrin regulate the actin cytoskeleton. *Mol Biol Cell*. 17:427-437.

icerydernet/agerelationship. 2011. Horse Human Age Relationship Chart

Ilic, M.Z., M.T. Mok, O.D. Williamson, M.A. Campbell, C.E. Hughes, and C.J. Handley. 1995. Catabolism of aggrecan by explant cultures of human articular cartilage in the presence of retinoic acid. *Arch Biochem Biophys*. 322:22-30.

Iliopoulos, D., V. Gkretsi, and A. Tsezou. 2010. Proteomics of osteoarthritic chondrocytes and cartilage. *Expert Rev Proteomics*. 7:749-760.

Imai, K., A. Hiramatsu, D. Fukushima, M.D. Pierschbacher, and Y. Okada. 1997. Degradation of decorin by matrix metalloproteinases: identification of the cleavage sites, kinetic analyses and transforming growth factor-beta1 release. *Biochem J*. 322 ( Pt 3):809-814.

Inerot, S., D. Heinegard, S.E. Olsson, H. Telhag, and L. Audell. 1991. Proteoglycan alterations during developing experimental osteoarthritis in a novel hip joint model. *J Orthop Res*. 9:658-673.

Ingvarsson, T., G. Hagglund, H. Lindberg, and L.S. Lohmander. 2000. Assessment of primary hip osteoarthritis: comparison of radiographic methods using colon radiographs. *Ann Rheum Dis*. 59:650-653.

Iqbal, J., J. Dudhia, J.L. Bird, and M.T. Bayliss. 2000. Age-related effects of TGF-beta on proteoglycan synthesis in equine articular cartilage. *Biochem Biophys Res Commun*. 274:467-471.

Irving, S.G., P.F. Zipfel, J. Balke, O.W. McBride, C.C. Morton, P.R. Burd, U. Siebenlist, and K. Kelly. 1990. Two inflammatory mediator cytokine genes are closely linked and variably amplified on chromosome 17q. *Nucleic Acids Res*. 18:3261-3270.

Ishihama, Y., Y. Oda, T. Tabata, T. Sato, T. Nagasu, J. Rappsilber, and M. Mann. 2005. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics*. 4:1265-1272.

Ismaiel, S., R.M. Atkins, M.F. Pearse, P.A. Dieppe, and C.J. Elson. 1992. Susceptibility of normal and arthritic human articular cartilage to degradative stimuli. *Br J Rheumatol*. 31:369-373.

Jakoby, T., B.H. van den Berg, and A. Tholey. 2012. Quantitative protease cleavage site profiling using tandem-mass-tag labeling and LC-MALDI-TOF/TOF MS/MS analysis. *J Proteome Res*. 11:1812-1820.

Jallali, N., H. Ridha, C. Thrasivoulou, C. Underwood, P.E. Butler, and T. Cowen. 2005. Vulnerability to ROS-induced cell death in ageing articular cartilage: the role of antioxidant enzyme activity. *Osteoarthritis Cartilage*. 13:614-622.

Janusz, M.J., M. Hare, S.L. Durham, J. Potempa, W. McGraw, R. Pike, J. Travis, and S.D. Shapiro. 1999. Cartilage proteoglycan degradation by a mouse transformed macrophage cell line is mediated by macrophage metalloelastase. *Inflamm Res*. 48:280-288.

Jariyachawalid, K., P. Laowanapiban, V. Meevootisom, and S. Wiyakrutta. 2012. Effective enhancement of Pseudomonas stutzeri D-phenylglycine aminotransferase functional expression in Pichia pastoris by co-expressing Escherichia coli GroEL-GroES. *Microb Cell Fact*. 11:47.

Jensen, L.J., M. Kuhn, M. Stark, S. Chaffron, C. Creevey, J. Muller, T. Doerks, P. Julien, A. Roth, M. Simonovic, P. Bork, and C. von Mering. 2009. STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res*. 37:D412-416.

Johnson, A., R. Smith, T. Saxne, M. Hickery, and D. Heinegard. 2004. Fibronectin fragments cause release and degradation of collagen-binding molecules from equine explant cultures. *Osteoarthritis Cartilage*. 12:149-159.

Jones, P.L., and R. Cote. 2008. The PRIDE proteomics identification database:data submission, query, and dataset comparison. *Methods Mol Biol*. 484:287-303.

Jormsjo, S., S. Ye, J. Moritz, D.H. Walter, S. Dimmeler, A.M. Zeiher, A. Henney, A. Hamsten, and P. Eriksson. 2000. Allele-specific regulation of matrix metalloproteinase-12 gene activity is associated with coronary artery luminal dimensions in diabetic patients with manifest coronary artery disease. *Circ Res*. 86:998-1003.

Ju, X.D., M. Deng, Y.F. Ao, C.L. Yu, J.Q. Wang, J.K. Yu, G.Q. Cui, and Y.L. Hu. 2010. The protective effect of tetramethylpyrazine on cartilage explants and chondrocytes. *J Ethnopharmacol*. 132:414-420.

Kafienah, W., D. Bromme, D.J. Buttle, L.J. Croucher, and A.P. Hollander. 1998. Human cathepsin K cleaves native type I and II collagens at the N-terminal end of the triple helix. *Biochem J*. 331 ( Pt 3):727-732.

Kajita, M., Y. Itoh, T. Chiba, H. Mori, A. Okada, H. Kinoh, and M. Seiki. 2001. Membrane-type 1 matrix metalloproteinase cleaves CD44 and promotes cell migration. *J Cell Biol*. 153:893-904.

Kamburov, A., C. Wierling, H. Lehrach, and R. Herwig. 2009. ConsensusPathDB--a database for integrating human functional interaction networks. *Nucleic Acids Res*. 37:D623-628.

Kamm, J.L., A.J. Nixon, and T.H. Witte. 2010. Cytokine and catabolic enzyme expression in synovium, synovial fluid and articular cartilage of naturally osteoarthritic equine carpi. *Equine Vet J*. 42:693-699.

Kane, D., L.E. Jensen, S. Grehan, A.S. Whitehead, B. Bresnihan, and O. Fitzgerald. 2004. Quantitation of metalloproteinase gene expression in rheumatoid and psoriatic arthritis synovial tissue distal and proximal to the cartilage-pannus junction. *J Rheumatol*. 31:1274-1280.

Kanu, A.B., P. Dwivedi, M. Tam, L. Matz, and H.H. Hill, Jr. 2008. Ion mobility-mass spectrometry. *J Mass Spectrom*. 43:1-22.

Kapranov, P., J. Cheng, S. Dike, D.A. Nix, R. Duttagupta, A.T. Willingham, P.F. Stadler, J. Hertel, J. Hackermuller, I.L. Hofacker, I. Bell, E. Cheung, J. Drenkow, E. Dumais, S. Patel, G. Helt, M. Ganesh, S. Ghosh, A. Piccolboni, V. Sementchenko, H. Tammana, and T.R. Gingeras. 2007. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science*. 316:1484-1488.

Karas, M., and F. Hillenkamp. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem*. 60:2299-2301.

Karsenty, G., and E.F. Wagner. 2002. Reaching a genetic and molecular understanding of skeletal development. *Dev Cell*. 2:389-406.

Kashiwagi, M., J.J. Enghild, C. Gendron, C. Hughes, B. Caterson, Y. Itoh, and H. Nagase. 2004. Altered proteolytic activities of ADAMTS-4 expressed by C-terminal processing. *J Biol Chem*. 279:10109-10119.

Kawcak, C.E., D.D. Frisbie, N.M. Werpy, R.D. Park, and C.W. McIlwraith. 2008. Effects of exercise vs experimental osteoarthritis on imaging outcomes. *Osteoarthritis Cartilage*. 16:1519-1525.

Keshishian, H., T. Addona, M. Burgess, E. Kuhn, and S.A. Carr. 2007. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics*. 6:2212-2229.

Kevorkian, L., D.A. Young, C. Darrah, S.T. Donell, L. Shepstone, S. Porter, S.M. Brockbank, D.R. Edwards, A.E. Parker, and I.M. Clark. 2004. Expression profiling of metalloproteinases and their inhibitors in cartilage. *Arthritis Rheum*. 50:131-141.

Kiani, C., L. Chen, Y.J. Wu, A.J. Yee, and B.B. Yang. 2002. Structure and function of aggrecan. *Cell Res*. 12:19-32.

Kidd, J.A., C. Fuller, and A.R.S. Barr. 2001. Osteoarthritis in the horse. *Equine veterinary Education*. 13:160-168.

Kito, K., K. Ota, T. Fujita, and T. Ito. 2007. A synthetic protein approach toward accurate mass spectrometric quantification of component stoichiometry of multiprotein complexes. *J Proteome Res*. 6:792-800.

Knauper, V., C. Lopez-Otin, B. Smith, G. Knight, and G. Murphy. 1996. Biochemical characterization of human collagenase-3. *J Biol Chem*. 271:1544-1550.

Kongcharoensombat, W., T. Nakasa, M. Ishikawa, A. Nakamae, M. Deie, N. Adachi, A. Mohamed, and M. Ochi. 2010. The effect of microRNA-21 on proliferation and matrix synthesis of chondrocytes embedded in atelocollagen gel. *Knee Surg Sports Traumatol Arthrosc*. 18:1679-1684.

Kriegsmann, M., E.H. Seeley, A. Schwarting, J. Kriegsmann, M. Otto, H. Thabe, B. Dierkes, C. Biehl, U. Sack, A. Wellmann, G.J. Kahaly, K. Schwamborn, and R.M. Caprioli. 2012. MALDI MS imaging as a powerful tool for investigating synovial tissue. *Scand J Rheumatol*. 41:305-309.

Kuettner, K.E., V.A. Memoli, B.U. Pauli, N.C. Wrobel, E.J. Thonar, and J.C. Daniel. 1982. Synthesis of cartilage matrix by mammalian chondrocytes in vitro. II. Maintenance of collagen and proteoglycan phenotype. *J Cell Biol*. 93:751-757.

Kurz, B., B. Jost, and M. Schunke. 2002. Dietary vitamins and selenium diminish the development of mechanically induced osteoarthritis and increase the expression of antioxidative enzymes in the knee joint of STR/1N mice. *Osteoarthritis Cartilage*. 10:119-126.

Kwaan, H.C. 1992. The plasminogen-plasmin system in malignancy. *Cancer Metastasis Rev*. 11:291-311.

Lai, W.M., J.S. Hou, and V.C. Mow. 1991. A triphasic theory for the swelling and deformation behaviors of articular cartilage. *J Biomech Eng*. 113:245-258.

Lamas, J.R., L. Rodriguez-Rodriguez, A.G. Vigo, R. Alvarez-Lafuente, P. Lopez-Romero, F. Marco, E. Camafeita, A. Dopazo, S. Callejas, E. Villafuertes, J.A. Hoyas, M.P.

Tornero-Esteban, E. Urcelay, and B. Fernandez-Gutierrez. 2010. Large-scale gene expression in bone marrow mesenchymal stem cells: a putative role for COL10A1 in osteoarthritis. *Ann Rheum Dis*. 69:1880-1885.

Lambrecht, S., G. Verbruggen, P. Verdonk, D. Elewaut, and D. Deforce. 2008. Differential proteome analysis of normal and osteoarthritic chondrocytes reveals distortion of vimentin network in osteoarthritis. *Osteoarthritis Cartilage*. 16.

Lange, V., P. Picotti, B. Domon, and R. Aebersold. 2008. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol*. 4:222.

Lark, M.W., E.K. Bayne, J. Flanagan, C.F. Harper, L.A. Hoerrner, N.I. Hutchinson, Singer, II, S.A. Donatelli, J.R. Weidner, H.R. Williams, R.A. Mumford, and L.S. Lohmander. 1997. Aggrecan degradation in human cartilage. Evidence for both matrix metalloproteinase and aggrecanase activity in normal, osteoarthritic, and rheumatoid joints. *J Clin Invest*. 100:93-106.

Lark, M.W., J.T. Gordy, J.R. Weidner, J. Ayala, J.H. Kimura, H.R. Williams, R.A. Mumford, C.R. Flannery, S.S. Carlson, M. Iwata, and et al. 1995. Cell-mediated catabolism of aggrecan. Evidence that cleavage at the "aggrecanase" site (Glu373-Ala374) is a primary event in proteolysis of the interglobular domain. *J Biol Chem*. 270:2550-2556.

Lawrence, R.C., D.T. Felson, C.G. Helmick, L.M. Arnold, H. Choi, R.A. Deyo, S. Gabriel, R. Hirsch, M.C. Hochberg, G.G. Hunder, J.M. Jordan, J.N. Katz, H.M. Kremers, and F. Wolfe. 2008. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum*. 58:26-35.

Lee, Y.L., B.C. Ahn, Y. Lee, S.W. Lee, J.Y. Cho, and J. Lee. 2011. Targeting of hepatocellular carcinoma with glypican-3-targeting peptide ligand. *J Pept Sci*. 17:763-769.

Lefebvre, V., C. Peeters-Joris, and G. Vaes. 1990. Modulation by interleukin 1 and tumor necrosis factor alpha of production of collagenase, tissue inhibitor of metalloproteinases and collagen types in differentiated and dedifferentiated articular chondrocytes. *Biochim Biophys Acta*. 1052:366-378.

Leijten, J.C., J. Emons, C. Sticht, S. van Gool, E. Decker, A. Uitterlinden, G. Rappold, A. Hofman, F. Rivadeneira, S. Scherjon, J.M. Wit, J. van Meurs, C.A. van Blitterswijk, and M. Karperien. 2012. Gremlin 1, frizzled-related protein, and Dkk-1 are key regulators of human articular cartilage homeostasis. *Arthritis Rheum*. 64:3302-3312.

Leinweber, B.D., G. Tsaprailis, T.J. Monks, and S.S. Lau. 2009. Improved MALDI-TOF imaging yields increased protein signals at high molecular mass. *J Am Soc Mass Spectrom*. 20:89-95.

Li, Y.F., R.J. Arnold, H. Tang, and P. Radivojac. 2010. The importance of peptide detectability for protein identification, quantification, and experiment design in MS/MS proteomics. *J Proteome Res*. 9:6288-6297.

Libicher, M., M. Ivancic, M. Hoffmann, and W. Wenz. 2005. Early changes in experimental osteoarthritis using the Pond-Nuki dog model: technical procedure and initial results of in vivo MR imaging. *Eur Radiol*. 15:390-394.

Little, C.B., C.R. Flannery, C.E. Hughes, A. Goodship, and B. Caterson. 2005. Cytokine induced metalloproteinase expression and activity does not correlate with focal susceptibility of articular cartilage to degeneration. *Osteoarthritis Cartilage*. 13:162-170.

Liu, C., V.M. Rangnekar, E. Adamson, and D. Mercola. 1998. Suppression of growth and transformation and induction of apoptosis by EGR-1. *Cancer Gene Ther*. 5:3-28.

Livak, K.J., and T.D. Schmittgen. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 25:402-408.

Loeser, R.F. 2002. Integrins and cell signaling in chondrocytes. *Biorheology*. 39:119-124.

Loeser, R.F. 2009. Aging and osteoarthritis: the role of chondrocyte senescence and aging changes in the cartilage matrix. *Osteoarthritis Cartilage*. 17:971-979.

Loeser, R.F. 2010. Age-related changes in the musculoskeletal system and the development of osteoarthritis. *Clin Geriatr Med*. 26:371-386.

Loeser, R.F., C.S. Carlson, M. Del Carlo, and A. Cole. 2002. Detection of nitrotyrosine in aging and osteoarthritic cartilage: Correlation of oxidative damage with the presence of interleukin-1beta and with chondrocyte resistance to insulin-like growth factor 1. *Arthritis Rheum*. 46:2349-2357.

Loeser, R.F., A.L. Olex, M.A. McNulty, C.S. Carlson, M.F. Callahan, C.M. Ferguson, J. Chou, X. Leng, and J.S. Fetrow. 2012. Microarray analysis reveals age-related differences in gene expression during the development of osteoarthritis in mice. *Arthritis Rheum*. 64:705-717.

Loeser, R.F., and G. Shanker. 2000. Autocrine stimulation by insulin-like growth factor 1 and insulin-like growth factor 2 mediates chondrocyte survival in vitro. *Arthritis Rheum*. 43:1552-1559.

Lohmander, L.S. 2000. What can we do about osteoarthritis? *Arthritis Res*. 2:95-100.

Lohmander, L.S., K.D. Brandt, S.A. Mazzuca, B.P. Katz, S. Larsson, A. Struglics, and K.A. Lane. 2005. Use of the plasma stromelysin (matrix metalloproteinase 3) concentration to predict joint space narrowing in knee osteoarthritis. *Arthritis Rheum*. 52:3160-3167.

Lohmander, L.S., L.A. Hoerrner, and M.W. Lark. 1993a. Metalloproteinases, tissue inhibitor, and proteoglycan fragments in knee synovial fluid in human osteoarthritis. *Arthritis Rheum*. 36:181-189.

Lohmander, L.S., P.J. Neame, and J.D. Sandy. 1993b. The structure of aggrecan fragments in human synovial fluid. Evidence that aggrecanase mediates cartilage degradation in inflammatory joint disease, joint injury, and osteoarthritis. *Arthritis Rheum*. 36:1214-1222.

Lombard, C., J. Saulnier, and J. Wallach. 2005. Assays of matrix metalloproteinases (MMPs) activities: a review. *Biochimie*. 87:265-272.

Lombardi, G., D. Burzyn, J. Mundinano, P. Berguer, P. Bekinschtein, H. Costa, L.F. Castillo, A. Goldman, R. Meiss, I. Piazzon, and I. Nepomnaschy. 2005. Cathepsin-L influences the expression of extracellular matrix in lymphoid organs and plays a role in the regulation of thymic output and of peripheral T cell number. *J Immunol*. 174:7022-7032.

Long, D., S. Blake, X.Y. Song, M. Lark, and R.F. Loeser. 2008. Human articular chondrocytes produce IL-7 and respond to IL-7 with increased production of matrix metalloproteinase-13. *Arthritis Res Ther*. 10:R23.

Lougheed, J.C., J.M. Holton, T. Alber, J.F. Bazan, and T.M. Handel. 2001. Structure of melanoma inhibitory activity protein, a member of a recently identified family of secreted proteins. *Proc Natl Acad Sci U S A*. 98:5515-5520.

Loulakis, P., A. Shrikhande, G. Davis, and C.A. Maniglia. 1992. N-terminal sequence of proteoglycan fragments isolated from medium of interleukin-1-treated articular-cartilage cultures. Putative site(s) of enzymic cleavage. *Biochem J*. 284 ( Pt 2):589-593.

Lucchinetti, E., C.S. Adams, W.E. Horton, Jr., and P.A. Torzilli. 2002. Cartilage viability after repetitive loading: a preliminary report. *Osteoarthritis Cartilage*. 10:71-81.

Ma, B., E.B. Landman, R.L. Miclea, J.M. Wit, E.C. Robanus-Maandag, J.N. Post, and M. Karperien. 2012a. WNT Signaling and Cartilage: Of Mice and Men. *Calcif Tissue Int*.

Ma, B., C.A. van Blitterswijk, and M. Karperien. 2012b. A Wnt/beta-catenin negative feedback loop inhibits interleukin-1-induced matrix metalloproteinase expression in human articular chondrocytes. *Arthritis Rheum*. 64:2589-2600.

Ma, B., K. Zhang, C. Hendrie, C. Liang, M. Li, A. Doherty-Kirby, and G. Lajoie. 2003. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectrom*. 17:2337-2342.

Ma, W.J., X. Guo, J.T. Liu, R.Y. Liu, J.W. Hu, A.G. Sun, Y.X. Yu, and M.J. Lammi. 2011. Proteomic changes in articular cartilage of human endemic osteoarthritis in China. *Proteomics*. 11:2881-2890.

MacDonald, M.H., S.M. Stover, N.H. Willits, and H.P. Benton. 1992. Regulation of matrix metabolism in equine cartilage explant cultures by interleukin 1. *Am J Vet Res*. 53:2278-2285.

Madsen, S.H., E.U. Sumer, A.C. Bay-Jensen, B.C. Sondergaard, P. Qvist, and M.A. Karsdal. 2010. Aggrecanase- and matrix metalloproteinase-mediated aggrecan degradation is associated with different molecular characteristics of aggrecan and separated in time ex vivo. *Biomarkers*. 15:266-276.

Mahrus, S., J.C. Trinidad, D.T. Barkan, A. Sali, A.L. Burlingame, and J.A. Wells. 2008. Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. *Cell*. 134:866-876.

Majumdar, M.K., R. Askew, S. Schelling, N. Stedman, T. Blanchet, B. Hopkins, E.A. Morris, and S.S. Glasson. 2007. Double-knockout of ADAMTS-4 and ADAMTS-5 in mice results in physiologically normal animals and prevents the progression of osteoarthritis. *Arthritis Rheum*. 56:3670-3674.

Malemud, C.J., N. Islam, and T.M. Haqqi. 2003. Pathophysiological mechanisms in osteoarthritis lead to novel therapeutic strategies. *Cells Tissues Organs*. 174:34-48.

Maly, M.R. 2008. Abnormal and cumulative loading in knee osteoarthritis. *Curr Opin Rheumatol*. 20:547-552.

Mankin, H.J., and L. Lippiello. 1970. Biochemical and metabolic abnormalities in articular cartilage from osteo-arthritic human hips. *J Bone Joint Surg Am*. 52:424-434.

Mann, H.H., S. Ozbek, J. Engel, M. Paulsson, and R. Wagener. 2004. Interactions between the cartilage oligomeric matrix protein and matrilins. Implications for matrix assembly and the pathogenesis of chondrodysplasias. *J Biol Chem*. 279:25294-25298.

Mansson, B., D. Carey, M. Alini, M. Ionescu, L.C. Rosenberg, A.R. Poole, D. Heinegard, and T. Saxne. 1995. Cartilage and bone metabolism in rheumatoid arthritis. Differences between rapid and slow progression of disease identified by serum markers of cartilage metabolism. *J Clin Invest*. 95:1071-1077.

Marioni, J.C., C.E. Mason, S.M. Mane, M. Stephens, and Y. Gilad. 2008. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res*. 18:1509-1517.

Martel-Pelletier, J. 2004. Pathophysiology of osteoarthritis. *Osteoarthritis Cartilage*. 12 Suppl A:S31-33.

Martel-Pelletier, J., R. McCollum, N. Fujimoto, K. Obata, J.M. Cloutier, and J.P. Pelletier. 1994. Excess of metalloproteases over tissue inhibitor of metalloprotease may contribute to cartilage degradation in osteoarthritis and rheumatoid arthritis. *Lab Invest*. 70:807-815.

Martin, I., M. Jakob, D. Schafer, W. Dick, G. Spagnoli, and M. Heberer. 2001. Quantitative analysis of gene expression in human articular cartilage from normal and osteoarthritic joints. *Osteoarthritis Cartilage*. 9:112-118.

Martin, J.A., and J.A. Buckwalter. 2001. Roles of articular cartilage aging and chondrocyte senescence in the pathogenesis of osteoarthritis. *Iowa Orthop J*. 21:1-7.

Martin, J.A., S.M. Ellerbroek, and J.A. Buckwalter. 1997. Age-related decline in chondrocyte response to insulin-like growth factor-I: the role of growth factor binding proteins. *J Orthop Res*. 15:491-498.

Mateos, J., L. Lourido, P. Fernandez-Puente, V. Calamia, C. Fernandez-Lopez, N. Oreiro, C. Ruiz-Romero, and F.J. Blanco. 2012. Differential protein profiling of synovial fluid from rheumatoid arthritis and osteoarthritis patients using LC-MALDI TOF/TOF. *J Proteomics*. 75:2869-2878.

Matkovich, S.J., Y. Zhang, D.J. Van Booven, and G.W. Dorn, 2nd. 2010. Deep mRNA sequencing for in vivo functional analysis of cardiac transcriptional regulators: application to Galphaq. *Circ Res*. 106:1459-1467.

Matsuka, Y.V., M.M. Migliorini, and K.C. Ingham. 1997. Cross-linking of fibronectin to C-terminal fragments of the fibrinogen alpha-chain by factor XIIIa. *J Protein Chem*. 16:739-745.

McAlinden, A., J. Dudhia, M.C. Bolton, P. Lorenzo, D. Heinegard, and M.T. Bayliss. 2001. Age-related changes in the synthesis and mRNA expression of decorin and aggrecan in human meniscus and articular cartilage. *Osteoarthritis Cartilage*. 9:33-41.

McAnulty, R.J.a.L.G.J. 1990. In vivo measurement of collagen metabolism in cartilage and bone. *Methods in cartilage research*:140-142.

McCachren, S.S. 1991. Expression of metalloproteinases and metalloproteinase inhibitor in human arthritic synovium. *Arthritis Rheum*. 34:1085-1093.

McCarthy, H.E., J.J. Bara, K. Brakspear, S.K. Singhrao, and C.W. Archer. 2012. The comparison of equine articular cartilage progenitor cells and bone marrow-derived stromal cells as potential cell sources for cartilage repair in the horse. *Vet J*. 192:345-351.

McClean, L., J.L. Hurst, C. Gaskell, J. Lewis, and R.J. Beynon. 2007. Characterisation of cauxin in the urine of domestic and big cats. *J. Chem. Ecol.* 33:1997-2009.

McDonald, L., and R.J. Beynon. 2006. Positional proteomics: preparation of amino-terminal peptides as a strategy for proteome simplification and characterization. *Nat Protoc*. 1:1790-1798.

McDonald, L., D.H. Robertson, J.L. Hurst, and R.J. Beynon. 2005. Positional proteomics: selective recovery and analysis of N-terminal proteolytic peptides. *Nat Methods*. 2:955-957.

McIlwraith, C.W. 2012. The horse as a model of naturally occurring osteoarthritis *Bone and Joint Research*. 1:297-309.

McIlwraith, C.W., D.D. Frisbie, C.E. Kawcak, C.J. Fuller, M. Hurtig, and A. Cruz. 2010. The OARSI histopathology initiative - recommendations for histological assessments of osteoarthritis in the horse. *Osteoarthritis Cartilage*. 18 Suppl 3:S93-105.

Meistermann, H., J.L. Norris, H.R. Aerni, D.S. Cornett, A. Friedlein, A.R. Erskine, A. Augustin, M.C. De Vera Mudry, S. Ruepp, L. Suter, H. Langen, R.M. Caprioli, and A. Ducret. 2006. Biomarker discovery by imaging mass spectrometry: transthyretin is a biomarker for gentamicin-induced nephrotoxicity in rat. *Mol Cell Proteomics*. 5:1876-1886.

Melching, L.I., W.D. Fisher, E.R. Lee, J.S. Mort, and P.J. Roughley. 2006. The cleavage of biglycan by aggrecanases. *Osteoarthritis Cartilage*. 14:1147-1154.

Melching, L.I., and P.J. Roughley. 1989. The synthesis of dermatan sulphate proteoglycans by fetal and adult human articular cartilage. *Biochem J*. 261:501-508.

Mengshol, J.A., K.S. Mix, and C.E. Brinckerhoff. 2002. Matrix metalloproteinases as therapeutic targets in arthritic diseases: bull's-eye or missing the mark? *Arthritis Rheum*. 46:13-20.

Miller, E.J., J.E. Finch, Jr., E. Chung, W.T. Butler, and P.B. Robertson. 1976. Specific cleavage of the native type III collagen molecule with trypsin. Similarity of the cleavage products to collagenase-produced fragments and primary structure at the cleavage site. *Arch Biochem Biophys*. 173:631-637.

Miller, M.C., J.B. Resnick, B.T. Smith, and C.M. Lovett, Jr. 1996. The bacillus subtilis dinR gene codes for the analogue of Escherichia coli LexA. Purification and characterization of the DinR protein. *J Biol Chem*. 271:33502-33508.

Milner, J.M., S.F. Elliott, and T.E. Cawston. 2001. Activation of procollagenases is a key control point in cartilage collagen degradation: interaction of serine and metalloproteinase pathways. *Arthritis Rheum*. 44:2084-2096.

Milner, J.M., A. Patel, R.K. Davidson, T.E. Swingler, A. Desilets, D.A. Young, E.B. Kelso, S.T. Donell, T.E. Cawston, I.M. Clark, W.R. Ferrell, R. Plevin, J.C. Lockhart, R. Leduc, and A.D. Rowan. 2010. Matriptase is a novel initiator of cartilage matrix degradation in osteoarthritis. *Arthritis Rheum*.

Mirzaei, H., J.K. McBee, J. Watts, and R. Aebersold. 2008. Comparative evaluation of current peptide production platforms used in absolute quantification in proteomics. *Mol Cell Proteomics*. 7:813-823.

Misumi, K., V. Vilim, P.D. Clegg, C.C. Thompson, and S.D. Carter. 2001. Measurement of cartilage oligomeric matrix protein (COMP) in normal and diseased equine synovial fluids. *Osteoarthritis Cartilage*. 9:119-127.

Mitchell, P.G., H.A. Magna, L.M. Reeves, L.L. Lopresti-Morrow, S.A. Yocum, P.J. Rosner, K.F. Geoghegan, and J.E. Hambor. 1996. Cloning, expression, and type II collagenolytic activity of matrix metalloproteinase-13 from human osteoarthritic cartilage. *J Clin Invest*. 97:761-768.

Monfort, J., G. Tardif, P. Reboul, F. Mineau, P. Roughley, J.P. Pelletier, and J. Martel-Pelletier. 2006. Degradation of small leucine-rich repeat proteoglycans by matrix metalloprotease-13: identification of a new biglycan cleavage site. *Arthritis Res Ther*. 8:R26.

Montanaro, L., D. Trere, and M. Derenzini. 2008. Nucleolus, ribosomes, and cancer. *Am J Pathol*. 173:301-310.

Morgenroth, D.C., A.C. Gellhorn, and P. Suri. 2012. Osteoarthritis in the disabled population: a mechanical perspective. *PM R*. 4:S20-27.

Morgunova, E., A. Tuuttila, U. Bergmann, M. Isupov, Y. Lindqvist, G. Schneider, and K. Tryggvason. 1999. Structure of human pro-matrix metalloproteinase-2: activation mechanism revealed. *Science*. 284:1667-1670.

Morris, H.R., T. Paxton, M. Panico, R. McDowell, and A. Dell. 1997. A novel geometry mass spectrometer, the Q-TOF, for low-femtomole/attomole-range biopolymer sequencing. *J Protein Chem*. 16:469-479.

Mort, J.S., and D.J. Buttle. 1997. Cathepsin B. *Int J Biochem Cell Biol*. 29:715-720.

Mort, J.S., G.R. Dodge, P.J. Roughley, J. Liu, S.J. Finch, G. DiPasquale, and A.R. Poole. 1993. Direct evidence for active metalloproteinases mediating matrix degradation in interleukin 1-stimulated human articular cartilage. *Matrix*. 13:95-102.

Mort, J.S., A.D. Recklies, and A.R. Poole. 1984. Extracellular presence of the lysosomal proteinase cathepsin B in rheumatoid synovium and its activity at neutral pH. *Arthritis Rheum*. 27:509-515.

Mortazavi, A., B.A. Williams, K. McCue, L. Schaeffer, and B. Wold. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 5:621-628.

Moskalewski, S., I. Adamiec, and A. Golaszewska. 1979. Maturation of rabbit auricular chondrocytes grown in vitro in monolayer culture. *Am J Anat*. 155:339-348.

Mow, V.C., A. Ratcliffe, and A.R. Poole. 1992. Cartilage and diarthrodial joints as paradigms for hierarchical materials and structures. *Biomaterials*. 13:67-97.

Mow, V.C., C.C. Wang, and C.T. Hung. 1999. The extracellular matrix, interstitial fluid and ions as a mechanical signal transducer in articular cartilage. *Osteoarthritis Cartilage*. 7:41-58.

Mueller, M.B., and R.S. Tuan. 2011. Anabolic/Catabolic balance in pathogenesis of osteoarthritis: identifying molecular targets. *PM R*. 3:S3-11.

Muir, H. 1995. The chondrocyte, architect of cartilage. Biomechanics, structure, function and molecular biology of cartilage matrix macromolecules. *Bioessays*. 17:1039-1048.

Mundermann, A., C.O. Dyrby, T.P. Andriacchi, and K.B. King. 2005. Serum concentration of cartilage oligomeric matrix protein (COMP) is sensitive to physiological cyclic loading in healthy adults. *Osteoarthritis Cartilage*. 13:34-38.

Muro, E.M., N. Mah, and M.A. Andrade-Navarro. 2011. Functional evidence of post-transcriptional regulation by pseudogenes. *Biochimie*. 93:1916-1921.

Murphy, G., H. Stanton, S. Cowell, G. Butler, V. Knauper, S. Atkinson, and J. Gavrilovic. 1999. Mechanisms for pro matrix metalloproteinase activation. *APMIS*. 107:38-44.

Nagase, H., and K. Brew. 2003. Designing TIMP (tissue inhibitor of metalloproteinases) variants that are selective metalloproteinase inhibitors. *Biochem Soc Symp*:201-212.

Nagase, H., and M. Kashiwagi. 2003. Aggrecanases and cartilage matrix degradation. *Arthritis Res Ther*. 5:94-103.

Nagase, H., and J.F. Woessner, Jr. 1999. Matrix metalloproteinases. *J Biol Chem*. 274:21491-21494.

Narla, A., and B.L. Ebert. 2010. Ribosomopathies: human disorders of ribosome dysfunction. *Blood*. 115:3196-3205.

Nawrat, P., A. Surazynski, ., E. Karna, and J.A. Palka. 2005. The effect of hyaluronic acid on interleukin- 1-induced deregulation of collagen metabolism in cultured human skin fibroblasts. *Pharmacol Res*. 51:473-477.

Neidhart, M., N. Hauser, M. Paulsson, P.E. DiCesare, B.A. Michel, and H.J. Hauselmann. 1997. Small fragments of cartilage oligomeric matrix protein in synovial fluid and serum as markers for cartilage degradation. *Br J Rheumatol*. 36:1151-1160.

Nelson, F., L. Dahlberg, S. Laverty, A. Reiner, I. Pidoux, M. Ionescu, G.L. Fraser, E. Brooks, M. Tanzer, L.C. Rosenberg, P. Dieppe, and A. Robin Poole. 1998. Evidence for altered synthesis of type II collagen in patients with osteoarthritis. *J Clin Invest*. 102:2115-2125.

Neundorf, R.H., M.B. Lowerison, A.M. Cruz, J.J. Thomason, B.J. McEwen, and M.B. Hurtig. 2010. Determination of the prevalence and severity of metacarpophalangeal joint osteoarthritis in Thoroughbred racehorses via quantitative macroscopic evaluation. *Am J Vet Res*. 71:1284-1293.

O'Connor, L., A. Strasser, L.A. O'Reilly, G. Hausmann, J.M. Adams, S. Cory, and D.C. Huang. 1998. Bim: a novel member of the Bcl-2 family that promotes apoptosis. *EMBO J*. 17:384-395.

O'Farrell, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*. 250:4007-4021.

O'Loughlin, A., D.J. Lynn, M. McGee, S. Doyle, M. McCabe, and B. Earley. 2012. Transcriptomic analysis of the stress response to weaning at housing in bovine leukocytes using RNA-seq technology. *BMC Genomics*. 13:250.

Ohta, S., K. Imai, K. Yamashita, T. Matsumoto, I. Azumano, and Y. Okada. 1998. Expression of matrix metalloproteinase 7 (matrilysin) in human osteoarthritic cartilage. *Lab Invest*. 78:79-87.

Okada, Y., and G. Hashimoto. 2001. Degradation of extracellular matrix by matrix metalloproteinases and joint destruction. *Seikagaku*. 73:1309-1321.

Oleksyszyn, J., and A.J. Augustine. 1996. Plasminogen modulation of IL-1-stimulated degradation in bovine and human articular cartilage explants. The role of the

endogenous inhibitors: PAI-1, alpha 2-antiplasmin, alpha 1-PI, alpha 2-macroglobulin and TIMP. *Inflamm Res*. 45:464-472.

Ong, S.E., L.J. Foster, and M. Mann. 2003. Mass spectrometric-based approaches in quantitative proteomics. *Methods*. 29:124-130.

Otterness, I.G., J.T. Downs, C. Lane, M.L. Bliven, H. Stukenbrok, D.N. Scampoli, A.J. Milici, and P.S. Mezes. 1999. Detection of collagenase-induced damage of collagen by 9A4, a monoclonal C-terminal neoepitope antibody. *Matrix Biol*. 18:331-341.

Otto, A., J. Bernhardt, M. Hecker, U. Völker, and D. Becher. 2012. Systems Biology of Bacteria. *In* Proteomics - From Relative to Absolute Quantification for Systems Biology Approaches. H. Wipat, editor.

Page-McCaw, A., A.J. Ewald, and Z. Werb. 2007. Matrix metalloproteinases and the regulation of tissue remodelling. *Nat Rev Mol Cell Biol*. 8:221-233.

Pan, S., R. Aebersold, R. Chen, J. Rush, D.R. Goodlett, M.W. McIntosh, J. Zhang, and T.A. Brentnall. 2009. Mass spectrometry based targeted protein quantification: methods and applications. *J Proteome Res*. 8:787-797.

Pan, W., S. Zhu, M. Yuan, H. Cui, L. Wang, X. Luo, J. Li, H. Zhou, Y. Tang, and N. Shen. 2010. MicroRNA-21 and microRNA-148a contribute to DNA hypomethylation in lupus CD4+ T cells by directly and indirectly targeting DNA methyltransferase 1. *J Immunol*. 184:6773-6781.

Paulsson, M., M. Morgelin, H. Wiedemann, M. Beardmore-Gray, D. Dunham, T. Hardingham, D. Heinegard, R. Timpl, and J. Engel. 1987. Extended and globular protein domains in cartilage proteoglycans. *Biochem J*. 245:763-772.

Peffers, M.J., R.J. Beynon, J.S. Selley, and P.D. Clegg. 2012. Proteomic characterisation and quantification of an in-vitro early equine osteoarthritis model. *Osteoarthritis and Cartilage*. 20.

Peffers, M.J., P.I. Milner, S.R. Tew, and P.D. Clegg. 2010. Regulation of SOX9 in normal and osteoarthritic equine articular chondrocytes by hyperosmotic loading. *Osteoarthritis Cartilage*. 18:1502-1508.

Pei, B., C. Sisu, A. Frankish, C. Howald, L. Habegger, X.J. Mu, R. Harte, S. Balasubramanian, A. Tanzer, M. Diekhans, A. Reymond, T.J. Hubbard, J. Harrow, and M.B. Gerstein. 2012. The GENCODE pseudogene resource. *Genome Biol*. 13:R51.

Peng, M.Y., Z.H. Wang, C.Y. Yao, L.N. Jiang, ., J. Q.L., J. Wang, and B.Q. Li. 2008. Interleukin 17-producing gamma delta T cells increased in patients with active pulmonary tuberculosis. *Cell Mol Immunol*. 5:203-208.

Perkins, D.N., D.J. Pappin, D.M. Creasy, and J.S. Cottrell. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 20:3551-3567.

Peters, H.C., T.J. Otto, J.T. Enders, W. Jin, B.R. Moed, and Z. Zhang. 2011. The protective role of the pericellular matrix in chondrocyte apoptosis. *Tissue Eng Part A*. 17:2017-2024.

Picotti, P., H. Lam, D. Campbell, E.W. Deutsch, H. Mirzaei, J. Ranish, B. Domon, and R. Aebersold. 2008. A database of mass spectrometric assays for the yeast proteome. *Nat Methods*. 5:913-914.

Polacek, M., J.A. Bruun, O. Johansen, and I. Martinez. 2010a. Differences in the secretome of cartilage explants and cultured chondrocytes unveiled by SILAC technology. *J Orthop Res*.

Polacek, M., J.A. Bruun, O. Johansen, and I. Martinez. 2010b. Differences in the secretome of cartilage explants and cultured chondrocytes unveiled by SILAC technology. *J Orthop Res*. 28:1040-1049.

Poliseno, L., L. Salmena, J. Zhang, B. Carver, W.J. Haveman, and P.P. Pandolfi. 2010. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*. 465:1033-1038.

Poole, A.R., T. Kojima, T. Yasuda, F. Mwale, M. Kobayashi, and S. Laverty. 2001. Composition and structure of articular cartilage: a template for tissue repair. *Clin Orthop Relat Res*:S26-33.

Poole, A.R., F. Nelson, L. Dahlberg, E. Tchetina, M. Kobayashi, T. Yasuda, S. Laverty, G. Squires, T. Kojima, W. Wu, and R.C. Billinghurst. 2003. Proteolysis of the collagen fibril in osteoarthritis. *Biochem Soc Symp*:115-123.

Poole, C.A. 1997. Articular cartilage chondrons: form, function and failure. *J Anat*. 191:1-13.

Poole, C.A., S. Ayad, and J.R. Schofield. 1988. Chondrons from articular cartilage: I. Immunolocalization of type VI collagen in the pericellular capsule of isolated canine tibial chondrons. *J Cell Sci*. 90 ( Pt 4):635-643.

Pratt, J.M., D.M. Simpson, M.K. Doherty, J. Rivers, S.J. Gaskell, and R.J. Beynon. 2006. Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nat Protoc*. 1:1029-1043.

Price, J.S., S. Wang-Weigand, R. Bohne, L.D. Kozaci, and A.P. Hollander. 1999. Retinoic acid-induced type II collagen degradation does not correlate with matrix metalloproteinase activity in cartilage explant cultures. *Arthritis Rheum*. 42:137-147.

Pritzker, K.P., S. Gay, S.A. Jimenez, K. Ostergaard, J.P. Pelletier, P.A. Revell, D. Salter, and W.B. van den Berg. 2006. Osteoarthritis cartilage histopathology: grading and staging. *Osteoarthritis Cartilage*. 14:13-29.

Pullig, O., G. Weseloh, A.R. Klatt, R. Wagener, and B. Swoboda. 2002a. Matrilin-3 in human articular cartilage: increased expression in osteoarthritis. *10*:253-263.

Pullig, O., G. Weseloh, A.R. Klatt, R. Wagener, and B. Swoboda. 2002b. Matrilin-3 in human articular cartilage: increased expression in osteoarthritis. *Osteoarthritis Cartilage*. 10:253-263.

Rabilloud, T. 1996. Solubilization of proteins for electrophoretic analyses. *Electrophoresis*. 17:813-829.

Radin, E.L., and R.M. Rose. 1986. Role of subchondral bone in the initiation and progression of cartilage damage. *Clin Orthop Relat Res*:34-40.

Radinsky, R., K.S. Flickinger, M.A. Kosir, L. Zardi, and L.A. Culp. 1990. Adhesion of Kirsten-ras+ tumor-progressing and Kirsten-ras- revertant 3T3 cells on fibronectin proteolytic fragments. *Cancer Res*. 50:4388-4400.

Ratcliffe, A., J.A. Tyler, and T.E. Hardingham. 1986. Articular cartilage cultured with interleukin 1. Increased release of link protein, hyaluronate-binding region and other proteoglycan fragments. *Biochem J*. 238:571-580.

Rauser, S., C. Marquardt, B. Balluff, S.O. Deininger, C. Albers, E. Belau, R. Hartmer, D. Suckau, K. Specht, M.P. Ebert, M. Schmitt, M. Aubele, H. Hofler, and A. Walch. 2010. Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry. *J Proteome Res*. 9:1854-1863.

Reboul, P., J.P. Pelletier, G. Tardif, J.M. Cloutier, and J. Martel-Pelletier. 1996. The new collagenase, collagenase-3, is expressed and synthesized by human chondrocytes but not by synoviocytes. A role in osteoarthritis. *J Clin Invest*. 97:2011-2019.

Reed, C.C., and R.V. Iozzo. 2002. The role of decorin in collagen fibrillogenesis and skin homeostasis. *Glycoconj J*. 19:249-255.

Ren, D., S. Hou, H. Wang, D. Luo, and L. Zhang. 2006. Evaluation of RGD modification on collagen matrix. *Artif Cells Blood Substit Immobil Biotechnol*. 34:293-303.

Richter, H., and H. Hormann. 1982. Early and late cathepsin D-derived fragments of fibronectin containing the C-terminal interchain disulfide cross-link. *Hoppe Seylers Z Physiol Chem*. 363:351-364.

Rivers, J., C. Hughes, T. McKenna, Y. Woolerton, J.P. Vissers, J.I. Langridge, and R.J. Beynon. 2011. Asymmetric proteome equalization of the skeletal muscle proteome using a combinatorial hexapeptide library. *PLoS One*. 6:e28902.

Rivers, J., D.M. Simpson, D.H. Robertson, S.J. Gaskell, and R.J. Beynon. 2007. Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. *Mol Cell Proteomics*. 6:1416-1427.

Roberts, C.R., J.S. Mort, and P.J. Roughley. 1987. Treatment of cartilage proteoglycan aggregate with hydrogen peroxide. Relationship between observed degradation products and those that occur naturally during aging. *Biochem J*. 247:349-357.

Robinson, M.D., D.J. McCarthy, and G.K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 26:139-140.

Robinson, M.D., and A. Oshlack. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 11:R25.

Rodriguez-Manzaneque, J.C., J. Westling, S.N. Thai, A. Luque, V. Knauper, G. Murphy, J.D. Sandy, and M.L. Iruela-Arispe. 2002. ADAMTS1 cleaves aggrecan at multiple sites and is differentially inhibited by metalloproteinase inhibitors. *Biochem Biophys Res Commun*. 293:501-508.

Rossdale, P.D., R. Hopes, N.J. Digby, and K. offord. 1985. Epidemiological study of wastage among racehorses 1982 and 1983. *Vet Rec*. 116:66-69.

Roughley, P.J., Q. Nguyen, J.S. Mort, C.E. Hughes, and B. Caterson. 1993. Proteolytic degradation in human articular cartilage: its relationship to stromelysin. *Agents Actions Suppl*. 39:149-159.

Roughley, P.J., R.J. White, G. Cs-Szabo, and J.S. Mort. 1996. Changes with age in the structure of fibromodulin in human articular cartilage. *Osteoarthritis Cartilage*. 4:153-161.

Rousseau, J.C., and P.D. Delmas. 2007. Biological markers in osteoarthritis. *Nat Clin Pract Rheumatol*. 3:346-356.

Ruiz-Romero, C., and F.J. Blanco. 2009. The role of proteomics in osteoarthritis pathogenesis research. *Curr Drug Targets*. 10:543-556.

Ruiz-Romero, C., and F.J. Blanco. 2010. Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis Cartilage*. 18:500-509.

Ruiz-Romero, C., V. Calamia, J. Mateos, V. Carreira, M. Martinez-Gomariz, M. Fernandez, and F.J. Blanco. 2009. Mitochondrial dysregulation of osteoarthritic human articular chondrocytes analyzed by proteomics: a decrease in mitochondrial superoxide dismutase points to a redox imbalance. *Mol Cell Proteomics*. 8:172-189.

Sadowski, T., and J. Steinmeyer. 2001. Effects of non-steroidal antiinflammatory drugs and dexamethasone on the activity and expression of matrix metalloproteinase-1, matrix metalloproteinase-3 and tissue inhibitor of metalloproteinases-1 by bovine articular chondrocytes. *Osteoarthritis Cartilage*. 9:407-415.

Sahota, I.S., H.R. Ravensbergen, M.S. McGrath, and V.E. Claydon. 2012. Cerebrovascular responses to orthostatic stress after spinal cord injury. *J Neurotrauma*. 29:2446-2456.

Sajdera, S.W., and V.C. Hascall. 1969. Proteinpolysaccharide complex from bovine nasal cartilage. A comparison of low and high shear extraction procedures. *J Biol Chem*. 244:77-87.

Salmena, L., L. Poliseno, Y. Tay, L. Kats, and P.P. Pandolfi. 2011. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. 146:353-358.

Salminen, H., E. Vuorio, and A.M. Saamanen. 2001. Expression of Sox9 and type IIA procollagen during attempted repair of articular cartilage damage in a transgenic mouse model of osteoarthritis. *Arthritis Rheum*. 44:947-955.

Samilson, R.L., and V. Prieto. 1983. Dislocation arthropathy of the shoulder. *J Bone Joint Surg Am*. 65:456-460.

Sandell, L.J. 1994. In situ expression of collagen and proteoglycan genes in notochord and during skeletal development and growth. *Microsc Res Tech*. 28:470-482.

Sandell, L.J., and T. Aigner. 2001. Articular cartilage and changes in arthritis. An introduction: cell biology of osteoarthritis. *Arthritis Res*. 3:107-113.

Sandy, J.D. 2006. A contentious issue finds some clarity: on the independent and complementary roles of aggrecanase activity and MMP activity in human joint aggrecanolysis. *Osteoarthritis Cartilage*. 14:95-100.

Sandy, J.D., R.E. Boynton, and C.R. Flannery. 1991. Analysis of the catabolism of aggrecan in cartilage explants by quantitation of peptides from the three globular domains. *J Biol Chem*. 266:8198-8205.

Sandy, J.D., C.R. Flannery, P.J. Neame, and L.S. Lohmander. 1992. The structure of aggrecan fragments in human synovial fluid. Evidence for the involvement in osteoarthritis of a novel proteinase which cleaves the Glu 373-Ala 374 bond of the interglobular domain. *J Clin Invest*. 89:1512-1516.

Sandy, J.D., A.H.K. Plaas, and T.J. Koob. 1995. *Acta Orthopaedics Scandananavia 266* 26-32.

Sandy, J.D., and C. Verscharen. 2001. Analysis of aggrecan in human knee cartilage and synovial fluid indicates that aggrecanase (ADAMTS) activity is responsible for the catabolic turnover and loss of whole aggrecan whereas other protease activity is required for C-terminal processing in vivo. *Biochem J*. 358:615-626.

Sapolsky, A.I., R.D. Altman, J.F. Woessner, and D.S. Howell. 1973. The action of cathepsin D in human articular cartilage on proteoglycans. *J Clin Invest*. 52:624-633.

Saxne, T., and D. Heinegard. 1992. Cartilage oligomeric matrix protein: a novel marker of cartilage turnover detectable in synovial fluid and blood. *Br J Rheumatol*. 31:583-591.

Schilling, O., and C.M. Overall. 2008. Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. *Nat Biotechnol*. 26:685-694.

Schonherr, E., P. Witsch-Prehm, B. Harrach, H. Robenek, J. Rauterberg, and H. Kresse. 1995. Interaction of biglycan with type I collagen. *J Biol Chem*. 270:2776-2783.

Schulz, R., S. Hohle, G. Zernia, M. Zscharnack, J. Schiller, A. Bader, K. Arnold, and D. Huster. 2006. Analysis of extracellular matrix production in artificial cartilage constructs by histology, immunocytochemistry, mass spectrometry, and NMR spectroscopy. *J Nanosci Nanotechnol*. 6:2368-2381.

Scuderi, G.J., S.R. Golish, F.F. Cook, J.M. Cuellar, R.P. Bowser, and L.S. Hanna. 2011. Identification of a novel fibronectin-aggrecan complex in the synovial fluid of knees with painful meniscal injury. *J Bone Joint Surg Am*. 93:336-340.

Seeley, E.H., and R.M. Caprioli. 2008. Molecular imaging of proteins in tissues by mass spectrometry. *Proc Natl Acad Sci U S A*. 105:18126-18131.

Selsted, M.E., Y.Q. Tang, W.L. Morris, P.A. McGuire, M.J. Novotny, W. Smith, A.H. Henschen, and J.S. Cullor. 1996. Purification, primary structures, and antibacterial activities of beta-defensins, a new family of antimicrobial peptides from bovine neutrophils. *J Biol Chem*. 271:16430.

Shakibaei, M., T. John, C. Seifarth, and A. Mobasheri. 2007. Resveratrol inhibits IL-1 beta-induced stimulation of caspase-3 and cleavage of PARP in human articular chondrocytes in vitro. *Ann N Y Acad Sci* 1095:554-563.

Shepherd, D.E., and B.B. Seedhom. 1999. Thickness of human articular cartilage in joints of the lower limb. *Ann Rheum Dis*. 58:27-34.

Shiio, Y., and R. Aebersold. 2006. Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry. *Nat Protoc*. 1:139-145.

Shimazaki, A., M.O. Wright, K. Elliot, D.M. Salter, and S.J. Millward-Sadler. 2006. Calcium/calmodulin-dependent protein kinase II in human articular chondrocytes. *Biorheology*. 43:223-233.

Shinmei, M., K. Masuda, T. Kikuchi, Y. Shimomura, and Y. Okada. 1991. Production of cytokines by chondrocytes and its role in proteoglycan degradation. *J Rheumatol Suppl*. 27:89-91.

Shohani, B., M. Orazizadeh, M. Hashemitabar, and D. Heinegard. 2010. Degradation of extracellular matrix molecules in interleukin-1alpha treated bovine nasal cartilage. *Iran Biomed J*. 14:158-163.

Siepen, J.A., E.J. Keevil, D. Knight, and S.J. Hubbard. 2007. Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. *J Proteome Res*. 6:399-408.

Silva, J.C., R. Denny, C.A. Dorschel, M. Gorenstein, I.J. Kass, G.Z. Li, T. McKenna, M.J. Nold, K. Richardson, P. Young, and S. Geromanos. 2005. Quantitative proteomic analysis by accurate mass retention time pairs. *Anal Chem*. 77:2187-2200.

Simpson, D.M., and R.J. Beynon. 2012. QconCATs: design and expression of concatenated protein standards for multiplexed protein quantification. *Anal Bioanal Chem*. 404:977-989.

Simpson, K.L., A.D. Whetton, and C. Dive. 2009. Quantitative mass spectrometry-based techniques for clinical use: biomarker identification and quantification. *J Chromatogr B Analyt Technol Biomed Life Sci*. 877:1240-1249.

Smith, K.J., A.L. Bertone, S.E. Weisbrode, and M. Radmacher. 2006. Gross, histologic, and gene expression characteristics of osteoarthritic articular cartilage of the metacarpal condyle of horses. *Am J Vet Res*. 67.

So, A., A.M. Chamot, V. Peclat, and J.C. Gerster. 1999. Serum MMP-3 in rheumatoid arthritis: correlation with systemic inflammation but not with erosive status. *Rheumatology (Oxford)*. 38:407-410.

Sofat, N. 2009. Analysing the role of endogenous matrix molecules in the development of osteoarthritis. *Int J Exp Pathol*. 90:463-479.

Song, R.H., M.D. Tortorella, A.M. Malfait, J.T. Alston, Z. Yang, E.C. Arner, and D.W. Griggs. 2007. Aggrecan degradation in human articular cartilage explants is mediated by both ADAMTS-4 and ADAMTS-5. *Arthritis Rheum*. 56:575-585.

Sonkoly, E., and A. Pivarcsi. 2009. Advances in microRNAs: implications for immunity and inflammatory diseases. *J Cell Mol Med*. 13:24-38.

Spector, T.D., F. Cicuttini, J. Baker, J. Loughlin, and D. Hart. 1996. Genetic influences on osteoarthritis in women: a twin study. *BMJ*. 312:940-943.

Stanton, H., S.B. Golub, F.M. Rogerson, K. Last, C.B. Little, and A.J. Fosang. 2011. Investigating ADAMTS-mediated aggrecanolysis in mouse cartilage. *Nat Protoc*. 6:388-404.

Stauber, J., R. Lemaire, J. Franck, D. Bonnel, D. Croix, R. Day, M. Wisztorski, I. Fournier, and M. Salzet. 2008. MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. *J Proteome Res*. 7:969-978.

Steffensen, B., H.F. Bigg, and C.M. Overall. 1998. The involvement of the fibronectin type II-like modules of human gelatinase A in cell surface localization and activation. *J Biol Chem*. 273:20622-20628.

Stevens, A.L., J.S. Wishnok, D.H. Chai, A.J. Grodzinsky, and S.R. Tannenbaum. 2008. A sodium dodecyl sulfate-polyacrylamide gel electrophoresis-liquid chromatography tandem mass spectrometry analysis of bovine cartilage tissue response to mechanical compression injury and the inflammatory cytokines tumor necrosis factor alpha and interleukin-1beta. *Arthritis Rheum*. 58:489-500.

Stevens, A.L., J.S. Wishnok, F.M. White, G. A.J., and T. S.R. 2009. Mechanical injury and cytokines cause loss of cartilage integrity and upregulate proteins associated with catabolism, immunity, inflammation, and repair. *Mol Cell Proteomics*. 8:1475-1489.

Stockwell, R.A. 1971. The ultrastructure of cartilage canals and the surrounding cartilage in the sheep fetus. *J Anat*. 109:397-410.

Stove, J., K. Huch, K.P. Gunther, and H.P. Scharf. 2000. Interleukin-1beta induces different gene expression of stromelysin, aggrecan and tumor-necrosis-factor-stimulated gene 6 in human osteoarthritic chondrocytes in vitro. *Pathobiology*. 68:144-149.

Struglics, A., and M. Hansson. 2012. MMP proteolysis of the human extracellular matrix protein aggrecan is mainly a process of normal turnover. *Biochem J*. 446:213-223.

Struglics, A., S. Larsson, M.A. Pratta, S. Kumar, M.W. Lark, and L.S. Lohmander. 2006. Human osteoarthritis synovial fluid and joint cartilage contain both aggrecanase- and matrix metalloproteinase-generated aggrecan fragments. *Osteoarthritis Cartilage*. 14:101-113.

Suri, P., D.C. Morgenroth, and D.J. Hunter. 2012. Epidemiology of osteoarthritis and associated comorbidities. *PM R*. 4:S10-19.

Svensson, L., I. Narlid, and A. Oldberg. 2000. Fibromodulin and lumican bind to the same region on collagen type I fibrils. *FEBS Lett*. 470:178-182.

Swingler, T.E., J.G. Waters, R.K. Davidson, C.J. Pennington, X.S. Puente, C. Darrah, A. Cooper, S.T. Donell, G.R. Guile, W. Wang, and I.M. Clark. 2009. Degradome expression profiling in human articular cartilage. *Arthritis Res Ther*. 11:R96.

Taban, I.M., A.F. Altelaar, Y.E. van der Burgt, L.A. McDonnell, R.M. Heeren, J. Fuchser, and G. Baykut. 2007. Imaging of peptides in the rat brain using MALDI-FTICR mass spectrometry. *J Am Soc Mass Spectrom*. 18:145-151.

Takats, Z., J.M. Wiseman, B. Gologan, and R.G. Cooks. 2004. Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science*. 306:471-473.

Tam, E.M., Y.I. Wu, G.S. Butler, M.S. Stack, and C.M. Overall. 2002. Collagen binding properties of the membrane type-1 matrix metalloproteinase (MT1-MMP) hemopexin C domain. The ectodomain of the 44-kDa autocatalytic product of MT1-MMP inhibits cell invasion by disrupting native type I collagen cleavage. *J Biol Chem*. 277:39005-39014.

Taniguchi, N., B. Carames, L. Ronfani, U. Ulmer, S. Komiya, M.E. Bianchi, and M. Lotz. 2009. Aging-related loss of the chromatin protein HMGB2 in articular cartilage is linked to reduced cellularity and osteoarthritis. *Proc Natl Acad Sci U S A*. 106:1181-1186.

Taylor, S.E., A. Vaughan-Thomas, D.N. Clements, G. Pinchbeck, L.C. Macrory, R.K. Smith, and P.D. Clegg. 2009. Gene expression markers of tendon fibroblasts in normal and diseased tissue compared to monolayer and three dimensional culture systems. *BMC Musculoskelet Disord*. 10:27.

Tchetina, E.V., G. Squires, and A.R. Poole. 2005. Increased type II collagen degradation and very early focal cartilage degeneration is associated with upregulation of chondrocyte differentiation related genes in early human articular cartilage lesions. *J Rheumatol*. 32:876-886.

Tetlow, L.C., D.J. Adlam, and D.E. Woolley. 2001. Matrix metalloproteinase and proinflammatory cytokine production by chondrocytes of human osteoarthritic cartilage: associations with degenerative changes. *Arthritis Rheum*. 44:585-594.

Tew, S.R., and P.D. Clegg. 2010. Post-transcriptional gene regulation in chondrocytes. *Biochem Soc Trans*. 38:1627-1631.

Theodoropoulos, J.S., P.M. Wolin, and D.W. Taylor. 2009. Arthroscopic release of flexor hallucis longus tendon using modified posteromedial and posterolateral portals in the supine position. *Foot (Edinb)*. 19:218-221.

Timmer, J.C., M. Enoksson, E. Wildfang, W. Zhu, Y. Igarashi, J.B. Denault, Y. Ma, B. Dummitt, Y.H. Chang, A.E. Mast, A. Eroshkin, J.W. Smith, W.A. Tao, and G.S. Salvesen. 2007. Profiling constitutive proteolytic events in vivo. *Biochem J*. 407:41-48.

Todhunter, R.J., and G. Lust. 1992. Equine Surgery. *In* Synovial joint anatomy, biology and pathobiology. J.A. Auer, editor. W.B. Saunders Co, Philadelphia. 844-866.

Toncheva, A., M. Remichkova, K. Ikonomova, P. Dimitrova, and N. Ivanovska. 2009. Inflammatory response in patients with active and inactive osteoarthritis. *Rheumatol Int*. 29:1197-1203.

Tonner, P., V. Srinivasasainagendra, S. Zhang, and D. Zhi. 2012. Detecting transcription of ribosomal protein pseudogenes in diverse human tissues from RNA-seq data. *BMC Genomics*. 13:412.

Tortorella, M., M. Pratta, R.Q. Liu, I. Abbaszade, H. Ross, T. Burn, and E. Arner. 2000a. The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for aggrecan substrate recognition and cleavage. *J Biol Chem*. 275:25791-25797.

Tortorella, M.D., T.C. Burn, M.A. Pratta, I. Abbaszade, J.M. Hollis, R. Liu, S.A. Rosenfeld, R.A. Copeland, C.P. Decicco, R. Wynn, A. Rockwell, F. Yang, J.L. Duke, K. Solomon, H. George, R. Bruckner, H. Nagase, Y. Itoh, D.M. Ellis, H. Ross, B.H. Wiswall, K. Murphy, M.C. Hillman, Jr., G.F. Hollis, R.C. Newton, R.L. Magolda, J.M. Trzaskos, and E.C. Arner. 1999. Purification and cloning of aggrecanase-1: a member of the ADAMTS family of proteins. *Science*. 284:1664-1666.

Tortorella, M.D., R.Q. Liu, T. Burn, R.C. Newton, and E. Arner. 2002. Characterization of human aggrecanase 2 (ADAM-TS5): substrate specificity studies and comparison with aggrecanase 1 (ADAM-TS4). *Matrix Biol*. 21:499-511.

Tortorella, M.D., and A.M. Malfait. 2008. Will the real aggrecanase(s) step up: evaluating the criteria that define aggrecanase activity in osteoarthritis. *Curr Pharm Biotechnol*. 9:16-23.

Tortorella, M.D., M. Pratta, R.Q. Liu, J. Austin, O.H. Ross, I. Abbaszade, T. Burn, and E. Arner. 2000b. Sites of aggrecan cleavage by recombinant human aggrecanase-1 (ADAMTS-4). *J Biol Chem*. 275:18566-18573.

Tseng, S., A.H. Reddi, and P.E. Di Cesare. 2009. Cartilage Oligomeric Matrix Protein (COMP): A Biomarker of Arthritis. *Biomark Insights*. 4:33-44.

Tsuchiya, K., W.J. Maloney, T. Vu, A.R. Hoffman, P. Huie, R. Sibley, D.J. Schurman, and R.L. Smith. 1997. Osteoarthritis: differential expression of matrix metalloproteinase-9 mRNA in nonfibrillated and fibrillated cartilage. *J Orthop Res*. 15:94-100.

Tsuchiya, K., W.J. Maloney, T. Vu, A.R. Hoffman, D.J. Schurman, and R.L. Smith. 1996. RT-PCR analysis of MMP-9 expression in human articular cartilage chondrocytes and synovial fluid cells. *Biotech Histochem*. 71:208-213.

van den Bemt, P.M., A.C. Egberts, A.W. Lenderink, J.M. Verzijl, K.A. Simons, W.S. van der Pol, and H.G. Leufkens. 2000. Risk factors for the development of adverse drug events in hospitalized patients. *Pharm World Sci*. 22:62-66.

van den Berg, B.H., and A. Tholey. 2012. Mass spectrometry-based proteomics strategies for protease cleavage site identification. *Proteomics*. 12:516-529.

Van der Rest, M., L.C. Rosenberg, B.R. Olsen, and A.R. Poole. 1986. Chondrocalcin is identical with the C-propeptide of type II procollagen. *Biochem J*. 237:923-925.

van Meurs, J.B., P.L. van Lent, A.A. van de Loo, A.E. Holthuysen, E.K. Bayne, Singer, II, and W.B. van den Berg. 1999. Increased vulnerability of postarthritic cartilage to a

second arthritic insult: accelerated MMP activity in a flare up of arthritis. *Ann Rheum Dis*. 58:350-356.

van Saase, J.L., L.K. van Romunde, A. Cats, J.P. Vandenbroucke, and H.A. Valkenburg. 1989. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. *Ann Rheum Dis*. 48:271-280.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 3:RESEARCH0034.

Verzijl, N., J. DeGroot, Z.C. Ben, O. Brau-Benjamin, A. Maroudas, R.A. Bank, J. Mizrahi, C.G. Schalkwijk, S.R. Thorpe, J.W. Baynes, J.W. Bijlsma, F.P. Lafeber, and J.M. TeKoppele. 2002. Crosslinking by advanced glycation end products increases the stiffness of the collagen network in human articular cartilage: a possible mechanism through which age is a risk factor for osteoarthritis. *Arthritis Rheum*. 46:114-123.

Verzijl, N., J. DeGroot, E. Oldehinkel, R.A. Bank, S.R. Thorpe, J.W. Baynes, M.T. Bayliss, J.W. Bijlsma, F.P. Lafeber, and J.M. Tekoppele. 2000a. Age-related accumulation of Maillard reaction products in human articular cartilage collagen. *Biochem J*. 350 Pt 2:381-387.

Verzijl, N., J. DeGroot, S.R. Thorpe, R.A. Bank, J.N. Shaw, T.J. Lyons, J.W. Bijlsma, F.P. Lafeber, J.W. Baynes, and J.M. TeKoppele. 2000b. Effect of collagen turnover on the accumulation of advanced glycation end products. *J Biol Chem*. 275:39027-39031.

Vincourt, J.B., P. Gillet, A.C. Rat, F. Guillemin, P. Netter, D. Mainard, J. Magdalou, and 2012. Measurement of matrilin-3 levels in human serum and synovial fluid using a competitive enzymelinked immunosorbent assay. *Osteoarthritis Cartilage*.

Voorhees, T., J. Chang, Y. Yao, M.H. Kaplan, C.H. Chang, and J.B. Travers. 2011. Dendritic cells produce inflammatory cytokines in response to bacterial products from Staphylococcus aureus-infected atopic dermatitis lesions. *Cell Immunol*. 267:17-22.

Wachsmuth, L., B. Bau, Z. Fan, A. Pecht, N. Gerwin, and T. Aigner. 2004. ADAMTS-1, a gene product of articular chondrocytes in vivo and in vitro, is downregulated by interleukin 1beta. *J Rheumatol*. 31:315-320.

Wade, C.M., E. Giulotto, S. Sigurdsson, M. Zoli, S. Gnerre, F. Imsland, T.L. Lear, D.L. Adelson, E. Bailey, R.R. Bellone, H. Blocker, O. Distl, R.C. Edgar, M. Garber, T. Leeb, E. Mauceli, J.N. MacLeod, M.C. Penedo, J.M. Raison, T. Sharpe, J. Vogel, L. Andersson, D.F. Antczak, T. Biagi, M.M. Binns, B.P. Chowdhary, S.J. Coleman, G. Della Valle, S. Fryc, G. Guerin, T. Hasegawa, E.W. Hill, J. Jurka, A. Kiialainen, G. Lindgren, J. Liu, E. Magnani, J.R. Mickelson, J. Murray, S.G. Nergadze, R. Onofrio, S. Pedroni, M.F. Piras, T. Raudsepp, M. Rocchi, K.H. Roed, O.A. Ryder, S. Searle, L. Skow, J.E. Swinburne, A.C. Syvanen, T. Tozaki, S.J. Valberg, M. Vaudin, J.R. White, M.C. Zody, E.S. Lander, and K. Lindblad-Toh. 2009. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*. 326:865-867.

Wang, S.H., S.C. Kuo, and S.C. Chen. 2003. High-performance liquid chromatography determination of methionine adenosyltransferase activity using catechol-O-methyltransferase-coupled fluorometric detection. *Anal Biochem*. 319:13-20.

Wang, T., C.Y. Wen, C.H. Yan, W.W. Lu, and K.Y. Chiu. 2013. Spatial and temporal changes of subchondral bone proceed to microscopic articular cartilage degeneration in guinea pigs with spontaneous osteoarthritis. *Osteoarthritis Cartilage*. 21:574-581.

Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 10:57-63.

Webb-Robertson, B.J., W.R. Cannon, C.S. Oehmen, A.R. Shah, V. Gurumoorthi, M.S. Lipton, and K.M. Waters. 2008. A support vector machine model for the prediction of proteotypic peptides for accurate mass and time proteomics. *Bioinformatics*. 24:1503-1509.

Weiss, P.H., and L. Klein. 1969. The quantitative relationship of urinary peptide hydroxyproline excretion to collagen degradation. *J Clin Invest*. 48:1-10.

Wester, L., P. Olofsson, S.M. Ibrahim, and R. Holmdahl. 2003. Chronicity of pristane-induced arthritis in rats is controlled by genes on chromosome 14. *J Autoimmun*. 21:305-313.

Whiteaker, J.R., L. Zhao, S.E. Abbatiello, M. Burgess, E. Kuhn, C. Lin, M.E. Pope, M. Razavi, N.L. Anderson, T.W. Pearson, S.A. Carr, and A.G. Paulovich. 2011. Evaluation of large scale quantitative proteomic assay development using peptide affinity-based mass spectrometry. *Mol Cell Proteomics*. 10:M110 005645.

Whitfield, J.F. 2008. The solitary (primary) cilium--a mechanosensory toggle switch in bone and cartilage cells. *Cell Signal*. 20:1019-1024.

Wiberg, C., A.R. Klatt, R. Wagener, M. Paulsson, J.F. Bateman, D. Heinegard, and M. Morgelin. 2003. Complexes of matrilin-1 and biglycan or decorin connect collagen VI microfibrils to both collagen II and aggrecan. *J Biol Chem*. 278:37698-37704.

Wiese, S., K.A. Reidegeld, H.E. Meyer, and B. Warscheid. 2007. Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. *Proteomics*. 7:340-350.

Wilkins, M.R., C. Pasquali, R.D. Appel, K. Ou, O. Golaz, J.C. Sanchez, J.X. Yan, A.A. Gooley, G. Hughes, I. Humphery-Smith, K.L. Williams, and D.F. Hochstrasser. 1996. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y)*. 14:61-65.

Williams, R., I.M. Khan, K. Richardson, L. Nelson, H.E. McCarthy, T. Analbelsi, S.K. Singhrao, G.P. Dowthwaite, R.E. Jones, D.M. Baird, H. Lewis, S. Roberts, H.M. Shaw, J. Dudhia, J. Fairclough, T. Briggs, and C.W. Archer. 2010. Identification and clonal characterisation of a progenitor cell sub-population in normal human articular cartilage. *PLoS One*. 5:e13246.

Wilson, R., and J.F. Bateman. 2008. A robust method for proteomic characterization of mouse cartilage using solubility-based sequential fractionation and two-dimensional gel electrophoresis. *Matrix Biol*. 27:709-712.

Wilson, R., D. Belluoccio, C.B. Little, A.J. Fosang, and J.F. Bateman. 2008. Proteomic characterization of mouse cartilage degradation in vitro. *Arthritis Rheum*. 58:3120-3131.

Wilson, R., A.F. Diseberg, L. Gordon, S. Zivkovic, L. Tatarczuch, E.J. Mackie, J.J. Gorman, and J.F. Bateman. 2010a. Comprehensive profiling of cartilage extracellular matrix formation and maturation using sequential extraction and label-free quantitative proteomics. *Mol Cell Proteomics*.

Wilson, R., A.F. Diseberg, L. Gordon, S. Zivkovic, L. Tatarczuch, E.J. Mackie, J.J. Gorman, and J.F. Bateman. 2010b. Comprehensive profiling of cartilage extracellular matrix formation and maturation using sequential extraction and label-free quantitative proteomics. *Mol Cell Proteomics*. 9:1296-1313.

Wilson, R., J.M. Whitelock, and J.F. Bateman. 2009. Proteomics makes progress in cartilage and arthritis research. *Matrix Biol*. 28:121-128.

Wiseman, J.M., D.R. Ifa, Q. Song, and R.G. Cooks. 2006. Tissue imaging at atmospheric pressure using desorption electrospray ionization (DESI) mass spectrometry. *Angew Chem Int Ed Engl*. 45:7188-7192.

Wong, J.W., and G. Cagney. 2010. An overview of label-free quantitation methods in proteomics by mass spectrometry. *Methods Mol Biol*. 604:273-283.

Wu, J., W. Liu, A. Bemis, E. Wang, Y. Qiu, E.A. Morris, C.R. Flannery, and Z. Yang. 2007. Comparative proteomic characterization of articular cartilage tissue from normal donors and patients with osteoarthritis. *Arthritis Rheum*. 56:3675-3684.

Wu, J.J., and D.R. Eyre. 1989. Covalent interactions of type IX collagen in cartilage. *Connect Tissue Res*. 20:241-246.

Wu, W., R.C. Billinghurst, I. Pidoux, J. Antoniou, D. Zukor, M. Tanzer, and A.R. Poole. 2002. Sites of collagenase cleavage and denaturation of type II collagen in aging and osteoarthritic articular cartilage and their relationship to the distribution of matrix metalloproteinase 1 and matrix metalloproteinase 13. *Arthritis Rheum*. 46:2087-2094.

Yamanaka, H., K. Makino, M. Takizawa, H. Nakamura, N. Fujimoto, H. Moriya, R. Nemori, H. Sato, M. Seiki, and Y. Okada. 2000. Expression and tissue localization of membrane-types 1, 2, and 3 matrix metalloproteinases in rheumatoid synovium. *Lab Invest*. 80:677-687.

Yan, J., L. Liu, X. Li, F. Wang, T. Zhu, P. Yuan, and Q. Zhang. 2006. [A potential use of collagen-hyaluronan-chondroitin sulfate tri-copolymer scaffold for cartilage tissue engineering]. *Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi*. 20:130-133.

Yanagida, M., G. Jung, Y. Tanaka, S. Sone, M. Fujishiro, K. Ikeda, K. Nozawa, H. Kaneko, Y. Takasaki, H. Ogawa, K. Takamori, and I. Sekigawa. 2012. Serum proteome analysis in patients with rheumatoid arthritis receiving therapy with etanercept, a chimeric tumor necrosis factor-alpha receptor. *Int J Rheum Dis*. 15:486-495.

Yang, W., D. Wang, and A. Richmond. 1999. Role of clathrin-mediated endocytosis in CXCR2 sequestration, resensitization, and signal transduction. *J Biol Chem*. 274:11328-11333.

Yao, J.G., L.B. Gao, Y.G. Liu, J. Li, and G.F. Pang. 2008. Genetic variation in interleukin-10 gene and risk of oral cancer. *Clin Chim Acta*. 388:84-88.

Yates, K.E., S. Shortkroff, and R.G. Reish. 2005. Wnt influence on chondrocyte differentiation and cartilage function. *DNA Cell Biol*. 24:446-457.

Yin, W., J.I. Park, and R.F. Loeser. 2009. Oxidative stress inhibits insulin-like growth factor-I induction of chondrocyte proteoglycan synthesis through differential regulation of phosphatidylinositol 3-Kinase-Akt and MEK-ERK MAPK signaling pathways. *J Biol Chem*. 284:31972-31981.

Yoshihara, Y., H. Nakamura, K. Obata, H. Yamada, T. Hayakawa, K. Fujikawa, and Y. Okada. 2000. Matrix metalloproteinases and tissue inhibitors of metalloproteinases in synovial fluids from patients with rheumatoid arthritis or osteoarthritis. *Ann Rheum Dis*. 59:455-461.

Yuasa, T., T. Otani, T. Koike, M. Iwamoto, and M. Enomoto-Iwamoto. 2008. Wnt/beta-catenin signaling stimulates matrix catabolic genes and activity in articular chondrocytes: its possible role in joint degeneration. *Lab Invest*. 88:264-274.

Yudoh, K., K. Shishido, H. Murayama, M. Yano, K. Matsubayashi, H. Takada, H. Nakamura, K. Masuko, T. Kato, and K. Nishioka. 2007. Water-soluble C60 fullerene prevents degeneration of articular cartilage in osteoarthritis via down-regulation of chondrocyte catabolic activity and inhibition of cartilage degeneration during disease development. *Arthritis Rheum*. 56:3307-3318.

Zack, M.D., E.C. Arner, C.P. Anglin, J.T. Alston, A.M. Malfait, and M.D. Tortorella. 2006. Identification of fibronectin neoepitopes present in human osteoarthritic cartilage. *Arthritis Rheum*. 54:2912-2922.

Zardi, L., B. Carnemolla, E. Balza, L. Borsi, P. Castellani, M. Rocco, and A. Siri. 1985. Elution of fibronectin proteolytic fragments from a hydroxyapatite chromatography column. A simple procedure for the purification of fibronectin domains. *Eur J Biochem*. 146:571-579.

Zhang, G., M.M. Miyamoto, and M.J. Cohn. 2006a. Lamprey type II collagen and Sox9 reveal an ancient origin of the vertebrate collagenous skeleton. *Proc Natl Acad Sci U S A*. 103:3180-3185.

Zhang, G., A. Sun, W. Li, T. Liu, and Z. Su. 2006b. Mass spectrometric analysis of enzymatic digestion of denatured collagen for identification of collagen type. *J Chromatogr A*. 1114:274-277.

Zhang, J., L. Xin, B. Shan, W. Chen, M. Xie, D. Yuen, W. Zhang, Z. Zhang, G.A. Lajoie, and B. Ma. 2012a. PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics*. 11:M111 010587.

Zhang, L., M. Yang, P. Marks, L.M. White, M. Hurtig, Q.S. Mi, G. Divine, and G. Gibson. 2012b. Serum non-coding RNAs as biomarkers for osteoarthritis progression after ACL injury. *Osteoarthritis Cartilage*. 20:1631-1637.

Zhang, R., A. Barton, J. Bittenden, J. Huang, and D. Crowther. 2010. Evaluation for Computational Platforms of LC-MS Based Label-Free Quantitative Proteomics: A Global View.

. *Proteomics & Bioinformatics* 3:260-265.

Zhen, E.Y., I.J. Brittain, D.A. Laska, P.G. Mitchell, E.U. Sumer, M.A. Karsdal, and K.L. Duffin. 2008. Characterization of metalloprotease cleavage products of human articular cartilage. *Arthritis Rheum*. 58:2420-2431.

Zheng, D., A. Frankish, R. Baertsch, P. Kapranov, A. Reymond, S.W. Choo, Y. Lu, F. Denoeud, S.E. Antonarakis, M. Snyder, Y. Ruan, C.L. Wei, T.R. Gingeras, R. Guigo, J. Harrow, and M.B. Gerstein. 2007. Pseudogenes in the ENCODE regions: consensus annotation, analysis of transcription, and evolution. *Genome Res*. 17:839-851.

Zhu, M., M. Chen, M. Zuscik, Q. Wu, Y.J. Wang, R.N. Rosier, R.J. O'Keefe, and D. Chen. 2008. Inhibition of beta-catenin signaling in articular chondrocytes results in articular cartilage destruction. *Arthritis Rheum*. 58:2053-2064.

Zhu, W., J.W. Smith, and C.M. Huang. 2010. Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol*. 2010:840518.

Zhu, Y., L.J. Sandell, and J.R. Matyas. 2001. Differential expression of type IIA collagen in early osteoarthritis. *Transactions Orthopaedic Research Society*. 26:664–669.

Ziyan, W., Y. Shuhua, W. Xiufang, and L. Xiaoyun. 2011. MicroRNA-21 is involved in osteosarcoma cell invasion and migration. *Med Oncol*. 28:1469-1474.