bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Metabolomic and Proteomic Stratification of Equine Osteoarthritis

James R Anderson[1], Marie M Phelan[2], Eva Caamaño-Gutiérrez[3], Peter D Clegg[1], Luis M Rubio-Martinez[1,4] and Mandy J Peffers[1]

[1]Musculoskeletal and Ageing Science, Institute of Life Course and Medical Sciences, University of Liverpool, Liverpool, UK

[2]NMR Metabolomics Facility, Technology Directorate & Department of Biochemistry & Systems Biology, Institute of Systems, Molecular and Integrative Biology, Liverpool, UK

[3]Computational Biology Facility, Technology Directorate & Department of Biochemistry and Systems Biology, Institute of Systems, Molecular and Integrative Biology, Liverpool, UK

[4]Equine Clinical Science, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Neston, UK

Corresponding author email address

James R Anderson janders@liverpool.ac.uk Tel: 01517949287

Running Title

Stratification of Equine Osteoarthritis

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abbreviations

ADAMTS, A disintegrin and metalloproteinase with thrombospondin motifs; ApoA1, Apolipoprotein A1; ApoA2, Apolipoprotein A-II; BLAST, Basic local alignment search tool; CPMG, Carr-Purcell-Meiboom-Gill; COMP, Cartilage oligomeric matrix protein; CD14, Cluster of differentiation 14; DAPC, Discriminant analysis of principal components; ECM, Extracellular matrix; FDR, False discovery rate; H & E, Haematoxylin and eosin; HDL, High density lipoprotein; HKJC, Hong Kong Jockey Club; Lasso, Least absolute shrinkage and selection operator; LBP, Lipopolysaccharide binding protein; MMP, Matrix metalloproteinase; MSI, Metabolomics Standards Initiative; MCP, Metacarpophalangeal; MTP, Metatarsophalangeal; OA, Osteoarthritis; POD, Palmar/plantar osteochondral disease; PFA, Paraformaldehyde; PCA, Principal component analysis; PQN, Probabilistic quotient normalisation; Saf O, Safranin O; SF, Synovial fluid; TIC, Total ion current.

Abstract

Osteoarthritis (OA) is characterised by loss of articular cartilage, synovial membrane dysfunction and subchondral sclerosis. Few studies have used a global approach to stratify equine synovial fluid (SF) molecular profiles according to OA severity. SF was collected from 58 metacarpophalangeal (MCP) and metatarsophalangeal joints of racing Thoroughbred horses (Hong Kong Jockey Club; HKJC) and 83 MCP joints of mixed breed horses from an abattoir and equine hospital (biobank). Joints were histologically and macroscopically assessed for OA severity. For proteomic analysis, native SF and SF loaded onto ProteoMiner™ equalisation columns, to deplete high

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

abundant proteins, were analysed using liquid chromatography-tandem mass spectrometry (LC-MS/MS) and label-free quantification. Validation of selected differentially expressed proteins was undertaken using clinical SF collected during diagnostic investigations. Native SF metabolites were analysed using 1D ^{1}H Nuclear Magnetic Resonance (NMR). 1,834 proteins and 40 metabolites were identified in equine SF. Afamin levels decreased with synovitis severity and four uncharacterised proteins decreased with OA severity. Gelsolin and lipoprotein binding protein decreased with OA severity and apolipoprotein A1 levels increased for mild and moderate OA. Within the biobank, glutamate levels decreased with OA severity and for the HKJC cohort, 2-aminobutyrate, alanine and creatine increased with severity. Proteomic and metabolomic integration was undertaken using linear regression via Lasso penalisation modelling, incorporating 29 variables (R^2=0.82) with principal component 2 able to discriminate advanced OA from earlier stages, predominantly driven by H9GZQ9, F6ZR63 and alanine. Combining biobank and HKJC datasets, discriminant analysis of principal components modelling prediction was good for mild OA (90%). This study has stratified equine OA using both metabolomic and proteomic SF profiles and identified a panel of markers of interest which may be applicable to grading OA severity. This is also the first study to undertake computational integration of NMR metabolomic and LC-MS/MS proteomic datasets of any biological system.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

The age-related degenerative musculoskeletal condition osteoarthritis (OA) is mainly characterised by articular cartilage degradation, synovitis, subchondral bone sclerosis and abnormal bone proliferation (1, 2). OA is an important welfare issue for equids, with up to 60% of lameness cases attributed to OA, leading to substantial morbidity and mortality (3–5). Although it is known that degradation of the extracellular matrix (ECM) is driven by increased enzymatic activity of multiple matrix metalloproteinases (MMPs) and a disintegrin and metalloproteinases with thrombospondin motifs (ADAMTSs), the underlying pathogenesis of OA is yet to be fully understood (6–8). Currently, equine OA is predominantly diagnosed through radiography, however due to the slow onset of the condition this often leads to substantial pathology of the joint and articular cartilage degradation prior to diagnosis (9). There is therefore a need to develop accurate biomarkers of early OA which can be applied to a clinical setting and allow for timely intervention, as well as improving our understanding of OA pathogenesis and identifying potentially novel therapeutic targets.

Currently, no equine OA-specific biomarkers have been identified to aid an early clinical diagnosis (10). For human OA diagnosis, increased synovial fluid (SF) abundances of both matrix metalloproteinases (MMPs) and cartilage oligomeric matrix protein (COMP) have been identified as markers of interest (11). For horses, MMP activity has also shown OA diagnostic potential, however studies investigating COMP levels have shown conflicting results with SF COMP abundance unable to stage equine OA (12–19). However, these markers are generated following significant joint pathology, including substantial articular cartilage degradation, and

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

thus there is a need to identify markers at an earlier disease stage, when intervention would be most beneficial.

Palmar/plantar osteochondral disease (POD), located at the distal condyles of metacarpal III and metatarsal III, is a highly prevalent pathology of racehorses, resultant of repetitive joint overload during cyclic locomotion at high-speed (20–22). POD lesions range in severity, from mild to end-stage disease, and as the disease originates within the subchondral bone, provides an effective model to investigate subchondral bone mediated OA (23).

The inner layer of the joint capsule consists of a one-cell thick lining of synoviocytes within a hyaluronic acid and collagen matrix, called the synovial membrane, which produces SF (24, 25). SF primarily acts as a lubricant within the joint, protecting hyaline articular cartilage surfaces (26). However, SF also provides a pool of nutrients for surrounding tissues and a medium of cellular communication with the semi-permeable synovial membrane allowing passive protein transfer and synoviocytes secreting regulatory cytokines and growth factors (24–28). Due to the close relationship, both in terms of location and biological communication that SF holds with surrounding tissues which are primarily altered during OA, and its accessibility, SF can provide a unique source of chemical information and holds great promise for biomarker discovery (10, 29).

Various studies have utilised ^{1}H nuclear magnetic resonance (NMR) to investigate metabolite markers associated with OA in various species, including pigs, dogs, humans and horses (10, 30–35). Lacitignola *et al.* identified a panel of ten metabolites which were elevated in equine OA SF compared to a normal control group. This panel included alanine, acetate, N-acetyl glucosamine, pyruvate, citrate,

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

creatine/creatinine, choline, glycerol, lactate and α-glucose. However, no studies to date have used ^{1}H NMR to stratify OA to identify changes to the metabolite profile at different stages of OA severity.

Several studies have used mass spectrometry (MS) based methodologies to investigate equine OA SF (36–40). The relatively recent development of liquid chromatography tandem mass spectrometry (LC-MS/MS) has enabled a method which can quickly, with high sensitivity, quantify proteins present within biological fluids of high complexity (25). Protein biomarker discovery is however hindered by the large protein concentration dynamic range exhibited by SF, meaning that highly abundant proteins can mask those of less abundance (41, 42). ProteoMiner™ protein enrichment columns (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK) have however been developed which utilise combinational ligand library technology, depleting highly abundant proteins and subsequently enriching low abundance proteins to reduce this dynamic range (43–45). This methodology has successfully been used to investigate OA in equine SF (38). However, this technique has not yet been utilised with SF to enable OA stratification, analysing SF at different OA severities.

Increased activity of MMPs, ADAMTSs, cathepsins and serine proteases during OA leads to cartilage breakdown and the generation of OA-specific peptide degradation products (neopeptides) (46–48). Numerous studies investigating equine tissue have identified potential OA neopeptides of interest using *ex-vivo* cartilage and SF (37, 38, 46, 49). As OA neopeptides have also been shown to be a driver of OA pain, identification and quantification of neopeptides has the potential to stratify OA, providing markers of early OA pathology and providing potential novel OA therapeutic and analgesic targets (50).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Within this study we have interrogated equine SF from horses with naturally occurring OA and those exhibiting POD pathology, as a subchondral bone mediated OA model using both ^{1}H NMR metabolomics and LC-MS/MS proteomics to stratify equine OA. This is the first study to use both techniques on the same samples and the first to statistically integrate NMR-led metabolomics with MS-based proteomics.

Experimental Procedures

Experimental Design and Statistical Rationale

A total of 141 separate equine SF samples were analysed during this study. Depending on the analysis undertaken, biological replicates ranged from n=8-34 per group for OA grading and n=17-38 per group for synovitis grading (Table 1). Joints with a microscopic or macroscopic OA grade of 0 were considered the control group, ranging from n=8-17. Metabolite and protein abundances were analysed by t-test or ANOVA, depending on the number of groups, with p values corrected for multiple testing using the Benjamini-Hochberg false discovery rate method. For NMR metabolomics and LC-MS/MS proteomics dataset integration, the Mahalanobis distance on principal components was calculated and Chi-squared testing undertaken to identify potential outliers. A linear regression, using Lasso (least absolute shrinkage and selection operator) penalisation to select the most important variables, was applied to combined datasets for both the biobank and Hong Kong Jockey Club (HKJC) sample groups. Additionally, correlations of all variables (proteins and metabolites) with macroscopic OA score were calculated, with significant variables ($p < 0.05$) included for further analysis. Significant variables across all analyses were then selected in a dataset, combining both biobank and

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

HKJC results, and discriminant analysis of principal components (DAPC) was used to predict different levels of disease severity.

Sample Collection and Processing

Equine post mortem samples were collected from a commercial abattoir, The Philip Leverhulme Equine Hospital, University of Liverpool and The Hong Kong Jockey Club (HKJC) Equine Hospital. Abattoir and Philip Leverhulme Equine Hospital samples were collected from horses of mixed breed and sex and represented naturally occurring OA disease. HKJC samples were exclusively from Thoroughbred racehorses (aged 3-10 years old) with a high prevalence of POD, providing a model for subchondral bone mediated OA.

University of Liverpool Biobank

Following euthanasia at the abattoir (F Drury & Sons, Swindon, UK) distal equine forelimbs were transported to the University of Liverpool. Horses euthanised at the Philip Leverhulme Equine Hospital, University of Liverpool, were processed on site. Within 8 hr of euthanasia the metacarpophalangeal (MCP) joint was opened aseptically and the distal metacarpal III photographed. SF was removed using a 10 ml syringe, transferred to a plain eppendorf, centrifuged at 2,540*g* and 4˚C for 5 min, supernatant removed, snap frozen in liquid nitrogen and stored at -80˚C (51) (Figure S1). SF was separated into separate 1 ml eppendorfs prior to snap freezing to optimise study design prior to multi ‘omics’ analysis and integration (52). Wedge sections of articular cartilage/subchondral bone (measuring 4.0 cm x 1.5 cm x 0.5 cm) were also sampled from the distal condylar region of metacarpal III, lateral to the sagittal ridge, and placed into 4% paraformaldehyde (PFA, Sigma-Aldrich,

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gillingham, UK) in phosphate buffered saline (PBS, Sigma-Aldrich) (Figure S2). Following decalcification in ethylenediaminetetraacetic acid (EDTA, Sigma-Aldrich), these were then sectioned and stained with Haematoxylin and eosin (H & E) (TCS Biosciences Ltd, Buckingham, UK) and Safranin O (BDH Chemicals, Poole, UK).

As biobank samples were obtained from a commercial abattoir and equine hospital, horses sampled were of mixed breed and sex and represented naturally occurring OA disease.

Hong Kong Jockey Club

HKJC samples were collected using the same collection protocols as used for the University of Liverpool equine biobank. In addition, synovial membranes were also dissected, fixed in 4% PFA and processed to prepare H & E stained histology slides. For HKJC samples, both MCP and metatarsophalangeal (MTP) joints were sampled. Following snap freezing with liquid nitrogen, frozen SF samples were shipped to Liverpool on dry ice.

Clinical Synovial Fluid Collection for ELISA

During clinical diagnostic investigations of horses presenting to The Philip Leverhulme Equine Hospital, University of Liverpool between 2014 and 2017, excess aspirated SF from live horses was collected and subsequently processed using the same processing steps as used for biobank and HKJC SF samples. Clinical OA was diagnosed via clinical examination, radiography and/or arthroscopy. Healthy/control SF samples were collected during clinical examinations, including diagnostic analgesia, whereby the joint involved was not identified as the cause of the current lameness. However, underlying joint pathology, not leading to a clinical

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

manifestation, could not be ruled out. These clinical SF samples were used for enzyme-linked immunosorbent assay (ELISA) validations of altered protein abundances.

Osteoarthritis Pathology: Macroscopic and Microscopic Scoring

Biobank distal metacarpal III articular surfaces were assessed for macroscopic OA pathology using the equine OARSI scoring scale (53). Histological sections were also assessed for microscopic OA related pathology using the microscopic aspect of the equine OARSI scoring scale. Within the HKJC group, distal metacarpal III samples were assessed macroscopically and microscopically for OA related and POD pathology using separate published scoring scales (21, 54). Marginal remodelling and dorsal impact injury categories were excluded from the macroscopic scoring scale as these could not be scored using the joint photographs which were provided. Synovial membrane histological sections were also scored according to synovitis severity (55).

Metabolomics

NMR Sample Preparation

SF was thawed over ice and centrifuged for 15 min at 13,000*g* and 4°C. 150 µl of supernatant was then diluted to produce a final volume containing 50% (v/v) SF, 40% (v/v) dd $^{1}H_2O$, 100 mM PO_4^{3-} pH 7.4 buffer (Na_2HPO_4, VWR International Ltd., Radnor, Pennsylvania, USA and NaH_2PO_4, Sigma-Aldrich) in deuterium oxide ($^{2}H_2O$, Sigma-Aldrich) and 0.0025% (v/v) sodium azide (NaN_3, Sigma-Aldrich). Prepared samples were then vortexed for 1 min, centrifuged for 2 min at 13,000*g* and 4˚C and, using a glass pipette, 195 µl transferred into 3 mm outer diameter NMR tubes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

NMR Acquisition

All SF samples were individually analysed. A 700 MHz NMR Bruker Avance III HD spectrometer with associated TCI cryoprobe and chilled Sample-Jet autosampler was used to acquire all spectra. 1D ^{1}H NMR spectra, using a Carr-Purcell-Meiboom-Gill (CPMG) filter to attenuate macromolecule (e.g. protein) signals, were acquired using a standard cpmgpr1d vendor pulse sequence. Spectral acquisition was carried out at 37˚C with a 4 s interscan delay, 32 transients and a 15 ppm spectral width. Software programmes Topsin 3.1 and IconNMR 4.6.7 were used for acquisition and processing, carrying out automated phasing, baseline correction and a standard vendor processing routine (exponential window function with 0.3 Hz line broadening).

Metabolite Annotation and Identification

All 1D ^{1}H NMR spectra were scrutinised to make sure that the minimum reporting standards were met, as outlined by the Metabolomics Society (56). These quality control criteria included flat baseline correction, water suppression, and consistent line widths. Spectra which did not meet these minimum requirements were removed from all subsequent analyses. Spectra were aligned to a single formate peak at 8.46 ppm. Chenomx NMR Suite 8.2 (330-mammalian metabolite library) software was used to carry out metabolite annotations and relative abundances. Metabolite identifications were confirmed where possible using in-house 1D ^{1}H NMR metabolite spectral library standards.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Proteomics

Synovial Fluid Processing and Protein Assay

SF was thawed on ice and centrifuged at 4°C at 14,000*g* for 10 min. The supernatant was treated with 1µg/ml of hyaluronidase (bovine origin, Sigma-Aldrich) at 37˚C for 1 hr, centrifuged at 1,000*g* for 5 min, supernatant removed and 1 ml centrifuged through a polypropylene microcentrifuge tube filter with 0.22 µm pore cellulose acetate membrane (Costar Spin-X, Corning, Tokyo, Japan) at 5,000*g* for 15 min to remove remaining insoluble particulates (29). A Pierce® 660 nm protein assay (Thermo Scientific, Waltham, Massachusetts, USA) was used to determine SF protein concentrations.

ProteoMiner™ Column Processing

2 mg of protein was loaded onto ProteoMiner™ Small Capacity bead columns (Bio-Rad Laboratories Ltd., Hemel Hempstead, UK) to achieve peptide-based depletion of the most abundant proteins. SF samples were rotated for 2 hr at room temperature and centrifuged at 1,000*g* for 60 s. ProteoMiner™ beads were then washed in PBS, rotated for 5 min and centrifuged at 1,000*g* for 60 s. The wash step was repeated two further times.

Protein Digestion

To assess both high and low abundance synovial proteins, both native and ProteoMiner™ processed SF were analysed. For native SF, 100 µg of protein was used for each protein trypsin digestion. 25 mM ammonium bicarbonate (Fluka Chemicals Ltd., Gillingham, UK) containing 0.05% (w/v) RapiGest (Waters, Elstree,

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Hertfordshire, UK) was added to both native SF and peptide bound ProteoMiner™ beads to produce a final volume of 160 µl and heated at 80°C for 10 min. DL-Dithiothreitol (Sigma-Aldrich) was added (3 mM final concentration) and incubated at 60°C for 10 min. Iodoacetamide (Sigma-Aldrich) was added (9 mM final concentration) and incubated at room temperature for 30 min in the dark. ProteoMiner™ processed samples then underwent an additional step which entailed the addition of 2 µg of Lys-C endopeptidase (FUJIFILM Wako Pure Chemical, Osaka, Japan) and incubation at 37˚C for 4 hr (51). 2 µg of proteomics grade trypsin (Sigma-Aldrich) was added to all samples and rotated for 16 hr at 37°C, followed by a second trypsin supplementation for 2 hr incubation. Digests were centrifuged at 1,000*g* for 1 min, supernatant removed, trifluoroacetic acid (TFA, Sigma-Aldrich) added (0.5% (v/v) final concentration) and rotated for 30 min at 37°C. Digests were then centrifuged at 13,000*g* for 15 min at 4°C and the supernatant removed and stored at -80°C. To ensure complete protein digestion, 5 µl of each digest was analysed by 1D SDS PAGE and stained with either Coomassie Blue (Bio-Rad) or silver stain (Thermo Scientific) following manufacturer instructions (data not shown).

Sample Processing for Neopeptide Analysis

After 4 hr of the 16 hr tryptic digestion of ProteoMiner™ processed samples, 10 µl was removed and supplemented with TFA and stored at -80°C using the same protocol as mentioned above.

Label Free LC-MS/MS

All SF digest samples underwent randomisation and were individually analysed using LC-MS/MS on an UltiMate 3000 Nano LC System (Dionex/Thermo Scientific)

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

coupled to a Q Exactive™ Quadrupole-Orbitrap instrument (Thermo Scientific). Full LC-MS/MS instrument methods are described in the supplemental data. Tryptic peptides, which were equivalent to 200 ng of protein, were loaded onto the column and run over 60 min, 90 min and 120 min LC gradients for 4 hr Lys-C + 4 hr tryptic digest ProteoMiner™ samples, native SF and 4 hr Lys-C + 16 hr + 2 hr tryptic digest ProteoMiner™ processed samples respectively. Representative ion chromatograms are shown in Figure S3.

LC-MS/MS Spectra Processing and Protein Identification

Progenesis™ QI 2.0 (Nonlinear Dynamics, Waters) software was used to process raw spectral files and undertake spectral alignment, peak picking, total protein abundance normalisation and peptide/protein quantification. The top ten spectra for each feature were then exported with peptide and protein identifications carried out with PEAKS® Studio 8.0 (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada) using the reviewed *Equus caballus* database (downloaded 29th July 2016, 22,694 sequences). Search parameters included: precursor mass error tolerance, 10.0 ppm; fragment mass error tolerance, 0.01 Da; precursor mass search type, monoisotopic; enzyme, trypsin; maximum missed cleavages, 1; non-specific cleavage, none; fixed modifications, carbamidomethylation; variable modifications, oxidation or hydroxylation and oxidation (methionine). The false discovery rate (FDR) was set to 1% with only proteins identified with > 2 unique peptides used for quantitation.

Semi-Tryptic Peptide Identification

A ‘semi-tryptic’ search was undertaken to identify potential synovial neopeptides. PEAKS® search parameters were kept the same as those used for protein

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

identifications, apart from 'Non-specific Cleavage' which was changed from 'none' to 'one'. The exported 'peptide ion measurements' file from Progenesis™ was then analysed using an online neopeptide analyser software tool to identify only semi-tryptic peptides and perform normalisation (57).

Batch Corrections

ProteoMiner™ processed samples for both the biobank and HKJC cohorts were run in three separate batches on the Q Exactive™ for protein analysis. To eliminate batch effects on the final analysis, a COMBAT batch correction was applied (Figure S4) (58). Metabolomic spectra were also acquired over 3 batches for both biobank and HKJC cohorts and data also underwent COMBAT batch correction.

ELISA Protein Validations

Differentially expressed proteins apolipoprotein A1 (ApoA1), gelsolin and lipopolysaccharide binding protein (LBP) were selected for ELISA as commercially available kits were compatible with equine samples. Equine ApoA1 (MBS034194, MyBioSource Inc., San Diego, California, USA) and equine LBP (MBS062216, MyBioSource) utilised sandwich ELISA technology using undiluted SF whilst equine gelsolin (CSB-EL009965HO, Cusabio Technology LLC, Houston, Texas, USA) was a competitive inhibition assay with SF diluted 1:1250. 3-6 dependent (and independent where stated) SF samples were analysed per group. Aliquots of 100 µl were analysed in duplicate for each sample with absorbance measured at 450 nm and protein concentrations calculated from assay standard curves. Protein concentrations were normalised to total protein.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Separate Dataset Statistical Analysis

SF metabolite abundances underwent median normalisation and protein abundances were normalised to the total ion current (TIC). Before multivariate analysis, both metabolite and protein datasets underwent Pareto scaling (59). All principal component analysis (PCA), ANOVA and t-tests of metabolite, protein and neopeptide abundances were conducted using MetaboAnalyst 4.0 (http://www.metaboanalyst.ca). ELISA t-tests and ANOVAs were performed using Minitab version 17. For protein analysis an additional filter of > 2 fold abundance was implemented. A p value of < 0.05 was considered statistically significant following correction for multiple testing using the Benjamini-Hochberg false discovery rate method (60). All box plots were produced using SPSS 24.

NMR metabolomics and LC-MS/MS Proteomics Integration

Metabolomics and proteomics datasets were initially integrated and analysed separately for the biobank and HKJC, then finally all datasets integrated to produce an overall model.

Only datasets for horses which had both metabolite and protein abundance values were used for integration. Proteomics datasets were again normalised to the TIC whilst NMR datasets were normalised via probabilistic quotient normalisation (PQN) (61). When combining ProteoMiner™ processed SF and native SF protein abundances for the same SF sample, in which the same protein had been identified in both datasets, the higher abundance was included for analysis. The Mahalanobis distance on principal components was calculated and Chi-squared testing undertaken to identify potential outliers. When combining metabolite and protein

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

variables, categorisation was carried out in accordance to macroscopic OA scoring. A linear regression, using Lasso penalisation to select the most important variables, was applied to combined datasets for both the biobank and HKJC sample groups. Additionally, correlations of all variables (proteins and metabolites) with macroscopic OA score were calculated using the Spearman coefficient, with significant variables after correcting for False Discovery Rate (Adjusted p value < 0.05) included for further analysis. Significant variables across all analyses were then applied to a collective dataset, combing both biobank and HKJC results. This dataset was stratified into healthy, mild OA and severe OA and a DAPC model was created to predict the disease severity. The number of parameters of the model were chosen after a cross-validation process calculated using the function xvalDapc and splitting the data into train/test subsets at a 80%/20% proportion respectively. All integration analyses were undertaken using standard analytical routines within the software R and packages adegenet, ggplot2 and reshape (62–65).

Uncharacterised Proteins

Within this study, proteins which were considered uncharacterised were also analysed using BLAST (Basic local alignment search tool) to assess similarity of their amino acid sequence to characterised proteins within the *Equus caballus* database as well as other species (66). Search parameters included: Matrix, blosum62; threshold, 0.1 E; filtering, none; gapped, true.

Pathway Analysis

Pathway analysis was conducted on proteins and metabolites which were considered significant variables for DAPC when carrying out NMR metabolomic and

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

LC-MS/MS proteomic dataset integration. Metabolite pathway analysis was conducted using the online tool KEGG for *Equus caballus* with protein pathways also analysed using KEGG, via the STRING database (67, 68). A filter of a minimum of two metabolites or proteins was set for inclusion of the relevant biological pathway.

Results

Stratification Groups

Histological microscopic OA scores vs macroscopic scores for biobank and HKJC samples both showed a weak positive correlation with R^2 coefficient values of 0.12 and 0.17 respectively (Figure S5). Joints were assigned into severity groups according to microscopic OA pathology, macroscopic OA pathology and synovitis scores (Table 1). Although scored using different scoring systems, overall the HKJC samples exhibited a more severe OA phenotype. Macroscopically, scores were on average 30% of the highest possible score, compared to 14% within the biobank and microscopic scores 31% compared to 15% within the biobank. All scoring is shown in Tables S6-S10. Overall, the HKJC dataset was a younger cohort than the biobank cohort, with average ages of 6.6 ± 1.9 years and 14.3 ± 7.6 years respectively (Table S1).

NMR Metabolomics

Overall, both the biobank and HKJC SF 1D ^{1}H NMR spectra produced similar profiles, although quantile plots revealed the biobank group exhibited more variation between samples (Figure S6). In total, 40 metabolites were identified within equine SF (Table 2). For the biobank cohort, unsupervised multivariate PCA did not identify separation between OA grades based on microscopic or macroscopic scoring

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Figure 1). However, when stratified according to macroscopic grade, glutamate levels were found to be differentially abundant, with lower levels identified at grades 2 and 3 compared to 1. PCA also did not identify clear separation between severity grades of OA or synovitis for HKJC SF samples (Figure 2). When stratified according to macroscopic OA pathology, three metabolites were found to be differentially abundant. 2-aminobutyrate levels were increased at grades 1 and 2 compared to 0 and alanine and creatine levels were increased with OA severity.

LC-MS/MS Proteomics

Within the biobank cohort, 74 native SF samples were analysed with 68 of these additionally processed using ProteoMiner™ columns. For the HKJC group, 56 native SF samples were analysed with 55 of these additionally having also undergone ProteoMiner™ processing. In total, across all samples, 1,834 proteins were identified (Figure 3). A combination of an increase in LC gradient length and ProteoMiner™ processing resulted in a 168% increase in the overall number of identified proteins compared to native SF analysis.

Following PCA analyses, HKJC native SF samples 95 and 122 were identified as outliers and removed from further analyses. For biobank and HKJC groups, when categorised according to macroscopic OA severity, PCA identified that increased OA severity resulted in less variation between samples (Figure 4). However, this was not evident when categorised according to microscopic OA grading. For native SF biobank samples categorised according to macroscopic grading, the abundance of three uncharacterised proteins and immunoglobulin kappa constant decreased with increasing OA severity. An uncharacterised protein (a member of the superfamily containing a leucine-rich repeat) and ApoA1 abundances both increased initially with

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

OA pathology but then returned towards baseline with greater OA severity. Microscopic OA categorisation for native SF biobank as well as macroscopic and microscopic OA categorisation of native SF HKJC samples did not identify any differentially abundant proteins. The proteomes of native SF HKJC samples, categorised according to synovitis grade, were not separated via PCA (Figure 5). However, the abundance of afamin was identified as differentially abundant between low and high grade synovitis, decreasing with synovitis severity. ProteoMiner™ processing of both biobank and HKJC SF did not identify any differentially abundant proteins or distinct proteome clusters (PCA) for any OA or synovitis categorisations (Figures S7 and S8). However, LC-MS/MS analysis of ProteoMiner™ processed biobank SF identified trends in gelsolin and LBP abundance, decreasing with increasing OA severity (Figure 6). ELISA analysis of gelsolin using the same SF samples that were analysed by LC-MS/MS corroborated this finding and was additionally supported by an independent clinical cohort, with gelsolin abundance lower in clinical OA cases compared to controls. LBP ELISA analysis of both dependent and independent SF samples also supported the trend identified via LC-MS/MS, although statistical significance was not reached. For biobank native and ProteoMiner™ processed SF we identified a trend that ApoA1 increased initially with OA severity and subsequently decreased. This trend was supported by ApoA1 ELISA analysis using dependent SF samples; however statistical significance was not reached.

Semi-Tryptic Peptide Profiles

No semi-tryptic peptides arising from ECM proteins were identified which were more highly abundant at more severe OA levels and thus no potential neopeptide

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

biomarkers were identified. Neither for the biobank or HKJC groups were distinct profiles identified between severities of OA (Figure S9). For the HKJC group there was less intra-group variation identified with increased levels of OA and synovitis pathology. However, conversely for the biobank group, samples with no OA pathology showed the least variation.

NMR metabolomics and LC-MS/MS Proteomics Integration

For the HKJC dataset, linear regression using Lasso penalisation modelling was unable to select a suitable number of parameters without overfitting the model. However, correlation analysis of all variables (proteins and metabolites) identified 58 significant variables, with a range in correlation from -0.48 to 0.42 (Table S2). Using these selected variables, PCA identified less variation between samples with more severe macroscopic OA scores, although their grouping could not clearly be distinguished from lower scoring samples (Figure 7).

For the biobank dataset, Lasso penalisation produced a model incorporating 29 variables (Table S3). This produced a good model (R^2 = 0.82) with PC2 able to discriminate more advanced stages of OA from the earlier stages (Figure 8). This model was predominantly driven by two uncharacterised proteins (H9GZQ9 and F6ZR63) and alanine. Correlation analysis of all variables (proteins and metabolites) identified 32 significant variables, with a range in correlation from -0.44 to 0.49 (Table S4).

Combining both the biobank and HKJC datasets, including significant protein and metabolite variables, DAPC analysis produced a model whereby OA severity prediction was good for mild OA (90%) although this was reduced significantly for healthy (57%) and severe OA (35%) classifications (Figure 9). Linear discriminant 1

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

was driven predominantly by neural EGFL like 2, serum albumin and alanine whilst linear discriminant 2 was driven by periostin, gelsolin and myocilin.

Uncharacterised Proteins

BLAST analysis of amino acid sequences of 12 uncharacterised proteins included within this study identified various related characterised proteins. The characterised protein with the highest percentage amino acid sequence similarity for each uncharacterised protein is shown in Table S5.

Pathway Analysis

Pathway analysis conducted on proteins and metabolites which were considered significant variables during DAPC modelling, when carrying out NMR metabolomic and LC-MS/MS proteomic dataset integration, identified the complement and coagulation cascades pathway and ABC transporters pathway to be the most represented (Tables 3 and 4). These pathways consisted of 7/96 proteins and 4/9 metabolites included within the pathway analysis respectively.

Discussion

Equine OA is currently predominantly diagnosed through radiography and clinical examination, however, due to the slow onset of the condition this often leads to substantial pathology of the joint and articular cartilage degradation prior to diagnosis (9). Presently, no equine OA-specific biomarkers have been identified which are able to accurately diagnose and stratify OA with none used to aid an earlier clinical diagnosis (10). Within this study we have investigated equine SF of varying

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

severities of associated OA pathology using both NMR-led metabolomics and MS-based proteomics approaches.

Although both the biobank and HKJC datasets produced a degree of OA stratification, markers of interest were generally distinct. This may be reflective of the more severe OA phenotype demonstrated by the HKJC horses, whilst markers identified within the biobank represent a less severe form of OA. Alternatively, the differences in markers identified may be due to the varying OA aetiologies, with the biobank representing a naturally occurring OA whilst the HKJC horses demonstrate a POD model for subchondral bone mediated OA, which is associated with trauma and overload (23). Additionally, the biobank cohort is on average an older donor group than the HKJC cohort which may also account for differences between the datasets. ^{1}H NMR SF spectra were more variable between biobank donors compared to the HKJC donors. This is likely because the biobank is a more heterogeneous population, including age, breed, work loads and diet. The HKJC group is a much more controlled sample set, with all horses housed together, all Thoroughbred racehorses, generally fed and trained similarly and all of a similar age at euthanasia. These differentials may also reflect the different markers of interest identified between the biobank and HKJC groups.

Proteomic analysis of biobank samples identified several proteins which were able to discriminate a healthy phenotype from early OA changes. Although these proteins were largely uncharacterised, BLAST analysis of the amino acid sequences of three uncharacterised proteins, decreasing in abundance with OA severity, identified high levels of similarity to immunoglobulin gamma 1 heavy chain constant region and immunoglobulin kappa chain V-III region MOPC 63 proteins. However, abundances of these uncharacterised proteins could not be validated using antibody

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

methodologies. Recently, a study investigating the SF proteome of a surgery-induced OA model in rabbits identified reduced levels of immunoglobulin heavy chain protein compared to sham controls (69). Immunoglobulins have previously been identified within superficial articular cartilage layers of a proportion of OA patients as well as elevated levels found within the synovial membrane of dogs diagnosed with cranial cruciate ligament rupture (70, 71). Thus, reduction of these immunoglobulins within the SF may be reflective of their translocation to surrounding articular tissues.

Glutamate is an excitatory amino acid neurotransmitter within the central nervous system, although evidence also suggests glutamate operates through intercellular signally cascades, as an autocrine/paracrine factor, in non-neuronal tissues (72). A self-sufficient glutamate signalling machinery has been identified within chondrocytes with a peripheral NMDA receptor proposed to have a role within inflammation and cartilage degradation (73). Within our study, the biobank group showed a reduction in synovial glutamate levels at higher OA severities. Previously, elevated levels of glutamate have been identified within OA SF of humans and within an OA rat models (74–76). However, no increase in glutamate was identified with equine OA SF using ^{1}H NMR, although it is not clear whether glutamate was identified during this study (32). The disparity between the lower glutamate levels identified within the biobank group and elevated levels within the literature may be reflective of the subtler OA phenotype exhibited by the biobank samples, and they may be indicative of glutamate levels at an earlier OA severity. However, the HKJC dataset, reflective of a higher grade OA, did not identify differential abundance. Therefore, at this stage the relationship between synovial glutamate abundance and equine OA remains inconclusive.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Creatine is a nonessential amino acid involved in cellular energy metabolism, maintaining cellular adenosine triphosphate (ATP) levels, in particular within the muscle and brain (77). Creatine SF levels increased with OA severity, which is supported by previous human and equine studies which also identified elevated SF creatine in OA SF (32, 78). Approximately 95% of stored creatine is located within skeletal muscle (79). Thus, given the association of muscle atrophy with OA, this elevation in synovial creatine may be reflective of an associated muscle mass loss (80).

Within this study, alanine levels were identified to be increasing in SF in accordance with OA severity, which was also observed by Lacitignolia *et al.* (32). Previously, depleted alanine abundance has been identified within human OA cartilage using high resolution magic angle spinning NMR spectroscopy (81). As alanine is one of the main amino acid residues which constitutes collagen, it may be that the reduction in alanine abundance identified with OA cartilage is resultant of degradation of the cartilage collagen framework, which are subsequently released into the SF resulting in the elevated synovial abundance within this study (81, 82).

Gelsolin (82-84 kDa) is a multifunctional, calcium ion-regulated actin filament severing, capping, and nucleating protein which is involved in the determination of cell shape, secretion and chemotaxis (83, 84). Previous studies of gelsolin on different biological systems have identified gelsolin as a potential predictor of both inflammation and tissue injury (85–87). Within rheumatoid arthritis patients, reduced circulating levels of plasma gelsolin have been identified (88). The authors propose this may be due to gelsolin redistribution to the affected joint space, binding to a plasma factor, reduced production or increased degradation. Within this study, SF gelsolin abundance was identified as decreasing with OA severity. Gelsolin was also

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

identified as having a significant correlation with OA severity and was found to be a highly influential factor when a model combining both biobank and HKJC datasets was developed. These results are supported by a mouse model whereby gelsolin knockout mice resulted in arthritis exacerbation (89). Additionally, within a mouse model of pain and acute inflammation, exogenous delivery of gelsolin was identified to have effective analgesic and anti-inflammatory properties (90). Exogenous gelsolin administration has also been shown to have chondroprotective properties, nullifying the effect of interleukin-1β and OA SF on anabolic gene expression and increased glycosaminoglycan deposition in chondrocytes and protection of the integrity of murine cartilage following intra-articular injection (84). Therefore, gelsolin has potential as a biomarker of equine OA as well as an OA therapeutic target, and potential analgesic.

LBP is an endogenous protein which binds to lipopolysaccharides and catalytically delivers monomeric liposaccharides to cluster of differentiation 14 (CD14) protein (91, 92). Previously, serum and synovial levels of LBP were not identified to be differentially expressed between human degenerative arthropathy patients and control samples (93). However, increased plasma LBP levels have recently been described to predict knee OA progression (94). Within this current study, the trends identified suggest a decreasing SF abundance of LBP with increasing OA severity. It has previously been identified that mononuclear cell activation, induced by lipopolysaccharides, is enhanced by low LBP concentrations (95). Activation of monocytes and macrophages by lipopolysaccharides leads to the secretion of tumour necrosis factor alpha and interleukin-1 beta, two pro-inflammatory cytokines which are central to OA pathogenesis (96, 97). Thus, decreasing synovial LBP levels may have a role in OA development. With CD14 also identified as a key variable

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

within the biobank and HKJC combined model, this provides further support that this pathway is involved in equine OA pathogenesis.

Elevations in synovial ApoA1 were identified in mild and moderate OA within the biobank group although abundance decreased for severe OA. Elevations in OA SF have previously been identified in horses and dogs (39, 98). ApoA1 is the primary protein component of high density lipoproteins (HDLs) and involved in HDL binding to ATP-binding cassette (ABC) transporters as well as being a lecithin cholesterol acyl transferase cofactor (99–102). ApoA1 has previously been found to induce the expression of interleukin-6, MMP-1 and MMP-3 in chondrocytes and synoviocytes through toll-like receptor 4 with the same study identifying a dissociation between the relationship of ApoA1 and HDLs in OA SF (103). When combining datasets within this study, apolipoprotein A-II (ApoA2) was also found to be an important variable. ApoA2 has previously been identified to be involved in the acute phase response which is associated with reactive amyloid A amyloidosis, via lipoprotein conformational changes, an associated complication of rheumatoid arthritis (104). Thus, this study provides further evidence that OA is a metabolic syndrome with disruption of lipid homeostasis due to alterations to apolipoprotein activity (103).

Synovitis has previously been identified as an important aspect of OA pathogenesis (105–107). This study identified reduced synovial abundance of afamin in high grade synovitis compared to low grade. Similarly, reduced afamin levels have been recorded in equine OA SF compared to healthy joints (39). In a previous study however, elevated levels of afamin were identified within human knee OA SF (108). Afamin is a vitamin E binding glycoprotein and a member of the albumin gene family (109). Afamin forms a 1:1 complex with various hydrophobic Wnt proteins, solubilising the proteins and producing a biologically active complex (110). A growing

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

body of evidence has identified that the Wnt/β-catenin signalling cascade is likely to have a central role within OA pathogenesis (111). Thus, a reduction in synovial abundance of afamin may be reflective of a translocation following Wnt solubilisation to surrounding articular tissues, i.e. the synovium. However, it should be noted that a differential abundance of afamin was not identified when categorised according to OA severity.

For the HKJC the semi-tryptic profile was most consistent between samples in the groups with the most severe OA and synovitis pathology. This suggests, as expected, the OA phenotype is driving the semi-tryptic peptide profile, most likely due to an increase in enzymatic activity leading to ECM degradation fragments (112). However, this was not evident within the biobank SF sample set. This may be because the OA pathology identified within this group was far subtler than that identified within the HKJC samples and thus the level of pathology present was not severe enough to drive a global change within the semi-tryptic peptide profile.

Computational integration of separate ‘omics’ datasets can work synergistically, greatly enhancing the information that can be obtained from separate analyses (113). However, this poses the challenge of developing multi ‘omic’ integration techniques. Previously, a study carried out MS proteomic analysis comparing early and late stage human OA SF and combined this with transcriptomics of articular tissues to identify the source of differentially abundant synovial proteins (114). However, no previous studies have combined NMR metabolomics and MS proteomic datasets.

Integrating the metabolomic and proteomic datasets separately for the biobank and HKJC both identified groupings associated with OA severity. Although correlation

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

values of variables to OA severity were generally low for the HKJC samples, stratification was still evident. Modelling was most robust for the biobank dataset, with the application of a Lasso model, with the model's strength being the separation of more severe OA grades. However, within this sample set there were relatively few samples which were assigned to the higher OA severity grades compared to lower and is therefore a limitation of the model. Combining all results from both the biobank and HKJC datasets produced a model which was found to be highly accurate in correctly assigning samples within the mild OA group based on their integrated metabolomic and proteomic profiles. This is promising as mild OA recognition is the current unmet clinical need, aiding in early diagnosis and allowing for timely OA interventional management. However, the model is dominated by predominantly mild OA graded samples, and thus this will inevitably introduce bias into the model and make a correct reclassification of mild OA more likely.

Within the combined dataset DAPC model, periostin was the principal driver of linear discriminant 2. Periostin is a 90 kDa matricellular protein which has regulatory functions in cell differentiation, cell adhesion and ECM organisation (115, 116). In human medicine, a positive correlation has been identified between both plasma and SF periostin abundance and knee OA severity (117). *In-vivo* results have identified that interleukin-13 may induce the production of periostin during OA, with periostin subsequently stimulating the production of MMPs within synoviocytes (118). In women, serum periostin levels were found to be associated with both prevalence and the risk of development/progression of knee OA (119). Thus for equids, periostin may also provide a potential therapeutic target and prognostic indicator.

Protein pathway analysis identified the 'complement and coagulation cascades' pathway to be altered during OA pathology. Additionally, the uncharacterised protein

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

F6ZR63, the second most important variable of influence driving the biobank Lasso model, was found to have a high level of amino acid sequence similarity to complement factor H-like isoform X1. The complement system is an important component of the innate immune response, encompassing various roles including opsonisation initiation, pathogen phagocytosis, the inflammatory response and terminating within cell lysis (120, 121). However, a growing body of evidence has identified a role of complement activation within OA pathogenesis, which is further supported by this study (120). Complement has been identified to have a role in the degradation of cartilage ECM, synovitis and osteophyte formation, with complement split fragment C3a found to upregulate gene expression of tumour necrosis factor-α and Interleukin-1β, pro-inflammatory cytokines central to OA pathogenesis (97, 120). Therefore, targeting the complement cascade may provide a novel therapeutic target for OA treatment.

Within experimental arthritis models, coagulation and fibrinolysis pathways are known to play a role within disease pathogenesis, with these cascades also found to be activated within both the joint and circulation of degenerative and inflammatory arthropathies (122). Within the synovium of patients diagnosed with OA, the coagulation factor fibrinogen is present throughout the tissue with elevations in tissue factor, a coagulation initiator, identified near endothelial cells (123). Previously, fibrinogen has been identified as a potential target for arthritis therapy, with removal of fibrinogen-leukocyte integrin receptor αMβ2 interactions limiting inflammatory processes whilst maintaining normal fibrinogen coagulation function (124). Thus these pathways may also hold potential as a novel target for OA therapy.

Metabolite pathway analysis identified the ‘ABC transporters’ pathway to be altered during OA pathology. ABC transporters utilise ATP hydrolysis to import or export

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

molecules across cell membranes (125). ABC exporters have an important role in the export of cholesterol, fatty acid and lipids from cells, with dysregulation of this pathway underlying numerous diseases. The Wnt/β-catenin signalling cascade, with a proposed central role within OA pathogenesis, has previously been identified as a regulator of ABC transporters (126). Additionally, the ABC transporter MRP5 has been found to be the principal exporter of hyaluronan from its site of synthesis within the cell to the ECM (127). However, as only four key metabolites were identified within this pathway, caution should be applied when analysing these results to avoid over interpretation.

Further Work

Several proteins of interest identified within this study were uncharacterised. BLAST amino acid sequence analysis of similar proteins provided further information on their potential function, however this could be taken further by conducting 3D modelling of these proteins to help confirm function and potential interactions with other proteins/molecules. Additionally, developing monoclonal antibodies which are specific to these uncharacterised proteins would allow this methodology to provide an orthologous method to confirm protein abundances. Within the cohorts used in this study, there was an over-representation of mild OA diagnoses opposed to more severe phenotypes. Although this is advantageous in terms of identifying early OA markers, in order to fully stratify OA, the addition of a greater number of donors with higher grade OA would provide a more comprehensive synovial profile of the metabolite and protein profiles at differing OA grades and provide more robust models. As this study adopted a cross-sectional approach, further work of interest would be to conduct repeated measures within a longitudinal study to discover

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

whether the identified markers of interest validate this study's findings. Also, given the differential abundance of ApoA1 and LBP identified within this study, it would be of interest to further interrogate SF, using both NMR and MS based approaches, to investigate the lipid profile and identify changes during OA progression.

Conclusions

In conclusion, this study has stratified equine OA using both metabolomic and proteomic SF profiles and identified a panel of markers of interest which may be applicable to grading OA severity. This is also the first study to undertake computational integration of NMR metabolomic and LC-MS/MS proteomic datasets of any biological system.

Acknowledgements

Dr James Anderson was funded through a Horse Trust PhD studentship (G1015) and Dr Mandy Peffers funded through a Wellcome Trust Intermediate Clinical Fellowship (107471/Z/15/Z). Software licenses for data analysis used in the Shared Research Facility for NMR metabolomics were funded by the Medical Research Council (MRC) Clinical Research Capabilities and Technologies Initiative (MR/M009114/1). ELISA protein validation experiments were funded by the Institute of Veterinary Science, University of Liverpool. This work was also supported by the MRC and Versus Arthritis as part of the Medical Research Council Versus Arthritis Centre for Integrated Research into Musculoskeletal Ageing (CIMA) [MR/R502182/1] and a Technology Directorate Voucher (Faculty of Health and Life Sciences, University of Liverpool). The MRC Versus Arthritis Centre for Integrated Research into Musculoskeletal Ageing is a collaboration between the Universities of Liverpool,

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Sheffield and Newcastle. The authors would like to thank Professor Christopher Riggs, Hong Kong Jockey Club, for providing samples for analysis, staff at both the F Drury and Sons Abattoir, Swindon and Hong Kong Jockey Club for assistance in sample collection and processing, Ms Valerie Tilston for preparation of histology slides, Mr Phillip Dyer for assisting with ProteoMiner™ column processing, members of the Centre for Protein Research, University of Liverpool, including Professor Rob Beynon and Dr Philip Brownridge for guidance and access to their mass spectrometry facilities, Mr Jake Ellis, Cardiff University, for undertaking NMR file depositions and Dr Ben McDermott, Dr Elizabeth Barr, Ms Amy McDermott, Mr Aibek Smagul and Mr Colin Anderson for conducting pathology scoring.

Ethics

Hong Kong Jockey Club samples were collected under the regulations of the Hong Kong Jockey Club with owner consent. Abattoir samples were collected as a by-product of the agricultural industry, and hence ethical approval was not required. Ethical collection of excess clinical equine synovial fluid from live horses and collection of post mortem equine synovial fluid was approved by The University of Liverpool Veterinary Research Ethics Committee, approval references VREC175 and VREC561.

Data Availability

1D ^{1}H NMR spectra, together with annotated metabolite HMDB IDs and annotation level, is available within the MetaboLights repository (www.ebi.ac.uk/metabolights/MTBLS1645) (128).

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Due to the COVID-19 pandemic, we are unfortunately unable to upload mass spectrometry proteomics data at present.

References

1. Kramer, C. M., Tsang, A. S., Koenig, T., Jeffcott, L. B., Dart, C. M., and Dart, A. J. (2014) Survey of the therapeutic approach and efficacy of pentosan polysulfate for the prevention and treatment of equine osteoarthritis in veterinary practice in Australia. *Aust Vet J*. **92**, 482–487

2. Truong, L.-H., Kuliwaba, J. S., Tsangari, H., and Fazzalari, N. L. (2006) Differential gene expression of bone anabolic factors and trabecular bone architectural changes in the proximal femoral shaft of primary hip osteoarthritis patients. *Arthritis Res. Ther.* **8**, R188–R188

3. Ireland, J. L., Clegg, P. D., McGowan, C. M., McKane, S. A., Chandler, K. J., and Pinchbeck, G. L. (2012) Disease prevalence in geriatric horses in the United Kingdom: veterinary clinical assessment of 200 cases. *Equine Vet J*. **44**, 101–106

4. Ireland, J. L., Clegg, P. D., McGowan, C. M., Platt, L., and Pinchbeck, G. L. (2011) Factors associated with mortality of geriatric horses in the United Kingdom. *Prev Vet Med*. **101**, 204–218

5. Caron, J., and Genovese, R. (2003) Principles and practices of joint disease treatment. in *Diagnosis and management of lameness in the horse*, 2nd Ed. (Ross, M., and Dyson, S. eds), pp. 746–764, W.B. Saunders, Philadelphia

6. Li, Y., Xu, L., and Olsen, B. R. (2007) Lessons from genetic forms of osteoarthritis for the pathogenesis of the disease. *Osteoarthr. Cartil.* **15**, 1101–

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

5

7. Struglics, A., Larsson, S., Pratta, M. A., Kumar, S., Lark, M. W., and Lohmander, L. S. (2006) Human osteoarthritis synovial fluid and joint cartilage contain both aggrecanase- and matrix metalloproteinase-generated aggrecan fragments. *Osteoarthr. Cartil.* **14**, 101–113

8. Marhardt, K., and Muurahainen, N. (2015) Development of a Disease-Modifying OA Drug (DMOAD) in Knee Osteoarthritis: The Example of Sprifermin. *Drug Res. (Stuttg).* **65**, S13–S13

9. Brommer, H., van Weeren, P. R., and Brama, P. A. (2003) New approach for quantitative assessment of articular cartilage degeneration in horses with osteoarthritis. *Am J Vet Res*. **64**, 83–87

10. Anderson, J. R., Phelan, M. M., Clegg, P. D., Peffers, M. J., and Rubio-Martinez, L. M. (2018) Synovial Fluid Metabolites Differentiate between Septic and Nonseptic Joint Pathologies. *J. Proteome Res.* **17**, 2735–2743

11. Balakrishnan, L., Nirujogi, R. S., Ahmad, S., Bhattacharjee, M., Manda, S. S., Renuse, S., Kelkar, D. S., Subbannayya, Y., Raju, R., Goel, R., Thomas, J. K., Kaur, N., Dhillon, M., Tankala, S. G., Jois, R., Vasdev, V., Ramachandra, Y., Sahasrabuddhe, N. A., Prasad Ts, K., Mohan, S., Gowda, H., Shankar, S., and Pandey, A. (2014) Proteomic analysis of human osteoarthritis synovial fluid. *Clin Proteomics*. **11**, 6

12. Taylor, S. E., Weaver, M. P., Pitsillides, A. A., Wheeler, B. T., Wheeler-Jones, C. P. D., Shaw, D. J., and Smith, R. K. W. (2006) Cartilage oligomeric matrix protein and hyaluronan levels in synovial fluid from horses with osteoarthritis of

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the tarsometatarsal joint compared to a control population. *Equine Vet. J.* **38**, 502–7

13. Zrimsek, P., Kadunc Kos, V., Mrkun, J., and Kosec, M. (2007) Diagnostic Value of Matrix Metalloproteinases MMP-2 and MMP-9 in Synovial Fluid for Identifying Osteoarthritis in the Distal Interphalangeal Joint in Horses. *Acta Vet. Brno*. **76**, 87–95

14. Arai, K., Misumi, K., Carter, S. D., Shinbara, S., Fujiki, M., and Sakamoto, H. (2005) Analysis of cartilage oligomeric matrix protein (COMP) degradation and synthesis in equine joint disease. *Equine Vet. J.* **37**, 31–36

15. Bertuglia, A., Pagliara, E., Grego, E., Ricci, A., and Brkljaca-Bottegaro, N. (2016) Pro-inflammatory cytokines and structural biomarkers are effective to categorize osteoarthritis phenotype and progression in Standardbred racehorses over five years of racing career. *BMC Vet. Res.* **12**, 246

16. Clegg, P. D., Coughlan, A. R., Riggs, C. M., and Carter, S. D. (1997) Matrix metalloproteinases 2 and 9 in equine synovial fluids. *Equine Vet. J.* **29**, 343–348

17. Misumi, K., Vilim, V., Clegg, P. D., Thompson, C. C. M., and Carter, S. D. (2001) Measurement of cartilage oligomeric matrix protein (COMP) in normal and diseased equine synovial fluids. *Osteoarthr. Cartil.* **9**, 119–127

18. Skiöldebrand, E., Lorenzo, P., Zunino, L., Rucklidge, G. J., Sandgren, B., Carlsten, J., and Ekman, S. (2001) Concentration of collagen, aggrecan and cartilage oligomeric matrix protein (COMP) in synovial fluid from equine middle carpal joints. *Equine Vet. J.* **33**, 394–402

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

19. Yamanokuchi, K., Tagami, M., Nishimatsu, E., Shimizu, Y., Hirose, Y., Komatsu, K., and Misumi, K. (2009) Sandwich ELISA system for cartilage oligomeric matrix protein in equine synovial fluid and serum. *Equine Vet. J.* **41**, 41–46

20. Pinchbeck, G. L., Clegg, P. D., Boyde, A., and Riggs, C. M. (2013) Pathological and clinical features associated with palmar/plantar osteochondral disease of the metacarpo/metatarsophalangeal joint in Thoroughbred racehorses. *Equine Vet. J.* **45**, 587–592

21. Barr, E. D., Pinchbeck, G. L., Clegg, P. D., Boyde, A., and Riggs, C. M. (2009) Post mortem evaluation of palmar osteochondral disease (traumatic osteochondrosis) of the metacarpo/metatarsophalangeal joint in Thoroughbred racehorses. *Equine Vet. J.* **41**, 366–71

22. Kawcak, C. E., McIlwraith, C. W., Norrdin, R. W., Park, R. D., and Steyn, P. S. (2000) Clinical effects of exercise on subchondral bone of carpal and metacarpophalangeal joints in horses. *Am. J. Vet. Res.* **61**, 1252–8

23. Barr, E. D. (2010) *The Association of Bone and Cartilage in Matrix Proteolysis of Articular Cartilage, and its Role in Palmar/Plantar Osteochondral Disease in the Thoroughbred Racehorse*. Ph.D. thesis, University of Liverpool

24. Gibson, D. S., and Rooney, M. E. (2007) The human synovial fluid proteome: A key factor in the pathology of joint disease. *Proteomics - Clin. Appl.* **1**, 889–899

25. Mahendran, S. M., Oikonomopoulou, K., Diamandis, E. P., and Chandran, V. (2017) Synovial fluid proteomics in the pursuit of arthritis mediators: An

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

evolving field of novel biomarker discovery. *Crit. Rev. Clin. Lab. Sci.* **54**, 495–505

26. Hui, A. Y., McCarty, W. J., Masuda, K., Firestein, G. S., and Sah, R. L. (2012) A systems biology approach to synovial joint lubrication in health, injury, and disease. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **4**, 15–37

27. Blewis, M. E., Nugent-Derfus, G. E., Schmidt, T. A., Schumacher, B. L., and Sah, R. L. (2007) A model of synovial fluid lubricant composition in normal and injured joints. *Eur Cell Mater*. **13**, 26–39

28. Tamer, T. M. (2013) Hyaluronan and synovial joint: function, distribution and healing. *Interdiscip. Toxicol.* **6**, 111–25

29. Mateos, J., Lourido, L., Fernandez-Puente, P., Calamia, V., Fernandez-Lopez, C., Oreiro, N., Ruiz-Romero, C., and Blanco, F. J. (2012) Differential protein profiling of synovial fluid from rheumatoid arthritis and osteoarthritis patients using LC-MALDI TOF/TOF. *J Proteomics*. **75**, 2869–2878

30. Damyanovich, A. Z., Staples, J. R., Chan, A. D., and Marshall, K. W. (1999) Comparative study of normal and osteoarthritic canine synovial fluid using 500 MHz 1H magnetic resonance spectroscopy. *J Orthop Res*. **17**, 223–231

31. Hugle, T., Kovacs, H., Heijnen, I. A., Daikeler, T., Baisch, U., Hicks, J. M., and Valderrabano, V. (2012) Synovial fluid metabolomics in different forms of arthritis assessed by nuclear magnetic resonance spectroscopy. *Clin Exp Rheumatol*. **30**, 240–245

32. Lacitignola, L., Fanizzi, F. P., Francioso, E., and Crovace, A. (2008) 1H NMR investigation of normal and osteo-arthritic synovial fluid in the horse. *Vet.*

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Comp. Orthop. Traumatol. **21**, 85–88

33. Mickiewicz, B., Heard, B. J., Chau, J. K., Chung, M., Hart, D. A., Shrive, N. G., Frank, C. B., and Vogel, H. J. (2015) Metabolic profiling of synovial fluid in a unilateral ovine model of anterior cruciate ligament reconstruction of the knee suggests biomarkers for early osteoarthritis. *J. Orthop. Res.* **33**, 71–77

34. Mickiewicz, B., Kelly, J. J., Ludwig, T. E., Weljie, A. M., Wiley, J. P., Schmidt, T. A., and Vogel, H. J. (2015) Metabolic analysis of knee synovial fluid as a potential diagnostic approach for osteoarthritis. *J Orthop Res*. **33**, 1631–1638

35. Anderson, J. R., Chokesuwattanaskul, S., Phelan, M. M., Welting, T. J. M., Lian, L.-Y., Peffers, M. J., and Wright, H. L. (2018) 1 H NMR Metabolomics Identifies Underlying Inflammatory Pathology in Osteoarthritis and Rheumatoid Arthritis Synovial Joints. *J. Proteome Res.* **17**, 3780–3790

36. Svala, E., Jin, C., Rüetschi, U., Ekman, S., Lindahl, A., Karlsson, N. G., and Skiöldebrand, E. (2017) Characterisation of lubricin in synovial fluid from horses with osteoarthritis. *Equine Vet. J.* **49**, 116–123

37. Skiöldebrand, E., Ekman, S., Mattsson Hultén, L., Svala, E., Björkman, K., Lindahl, A., Lundqvist, A., Önnerfjord, P., Sihlbom, C., and Rüetschi, U. (2017) Cartilage oligomeric matrix protein neoepitope in the synovial fluid of horses with acute lameness: A new biomarker for the early stages of osteoarthritis. *Equine Vet. J.* **49**, 662–667

38. Peffers, M. J., McDermott, B., Clegg, P. D., and Riggs, C. M. (2015) Comprehensive protein profiling of synovial fluid in osteoarthritis following protein equalization. *Osteoarthr. Cartil.* **23**, 1204–1213

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

39. Chiaradia, E., Pepe, M., Tartaglia, M., Scoppetta, F., D'Ambrosio, C., Renzone, G., Avellini, L., Moriconi, F., Gaiti, A., Bertuglia, A., Beccati, F., and Scaloni, A. (2012) Gambling on putative biomarkers of osteoarthritis and osteochondrosis by equine synovial fluid proteomics. *J. Proteomics*. **75**, 4478–4493

40. Anderson, J. R., Smagul, A., Simpson, D., Clegg, P. D., Rubio-Martinez, L. M., and Peffers, M. J. (2019) The synovial fluid proteome differentiates between septic and nonseptic articular pathologies. *J. Proteomics*. **202**, 103370

41. Roche, S., Tiers, L., Provansal, M., Piva, M. T., and Lehmann, S. (2006) Interest of major serum protein removal for Surface-Enhanced Laser Desorption/Ionization - Time Of Flight (SELDI-TOF) proteomic blood profiling. *Proteome Sci*. **4**, 20

42. Puangpila, C., Mayadunne, E., and El Rassi, Z. (2015) Liquid phase based separation systems for depletion, prefractionation, and enrichment of proteins in biological fluids and matrices for in-depth proteomics analysis-An update covering the period 2011-2014. *Electrophoresis*. **36**, 238–252

43. Fonslow, B. R., Carvalho, P. C., Academia, K., Freeby, S., Xu, T., Nakorchevsky, A., Paulus, A., and Yates, J. R. (2011) Improvements in Proteomic Metrics of Low Abundance Proteins through Proteome Equalization Using ProteoMiner Prior to MudPIT. *J. Proteome Res.* **10**, 3690–3700

44. Furka, A., Sebestyen, F., Asgedom, M., and Dibo, G. (1991) General method for rapid synthesis of multicomponent peptide mixtures. *Int J Pept Protein Res*. **37**, 487–493

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

45. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M., and Knapp, R. J. (1991) A new type of synthetic peptide library for identifying ligand-binding activity. *Nature*. **354**, 82–84

46. Peffers, M. J., Thornton, D. J., and Clegg, P. D. (2016) Characterization of neopeptides in equine articular cartilage degradation. *J. Orthop. Res.* **34**, 106–120

47. Polur, I., Lee, P. L., Servais, J. M., Xu, L., and Li, Y. (2010) Role of HTRA1, a serine protease, in the progression of articular cartilage degeneration. *Histol. Histopathol.* **25**, 599–608

48. Ben-Aderet, L., Merquiol, E., Fahham, D., Kumar, A., Reich, E., Ben-Nun, Y., Kandel, L., Haze, A., Liebergall, M., Kosińska, M. K., Steinmeyer, J., Turk, B., Blum, G., and Dvir-Ginzberg, M. (2015) Detecting cathepsin activity in human osteoarthritis via activity-based probes. *Arthritis Res. Ther.* **17**, 69

49. Peffers, M. J., Cillero-Pastor, B., Eijkel, G. B., Clegg, P. D., and Heeren, R. M. (2014) Matrix assisted laser desorption ionization mass spectrometry imaging identifies markers of ageing and osteoarthritic cartilage. *Arthritis Res. Ther.* **16**, R110

50. Miller, R. E., Ishihara, S., Tran, P. B., Golub, S. B., Last, K., Miller, R. J., Fosang, A. J., and Malfait, A.-M. (2018) An aggrecan fragment drives osteoarthritis pain through Toll-like receptor 2. *JCI Insight*. **3**, 1–9

51. Anderson, J. R., Phelan, M. M., Rubio-Martinez, L. M., Fitzgerald, M. M., Jones, S. W., Clegg, P. D., and Peffers, M. J. (2020) Optimization of Synovial Fluid Collection and Processing for NMR Metabolomics and LC-MS/MS

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Proteomics. *J. Proteome Res.* 10.1021/acs.jproteome.0c00035

52. Cavill, R., Jennen, D., Kleinjans, J., and Briedé, J. J. (2016) Transcriptomic and metabolomic data integration. *Brief. Bioinform.* **17**, 891–901

53. McIlwraith, C. W., Frisbie, D. D., Kawcak, C. E., Fuller, C. J., Hurtig, M., and Cruz, A. (2010) The OARSI histopathology initiative – recommendations for histological assessments of osteoarthritis in the horse. *Osteoarthr. Cartil.* **18**, S93–S105

54. Little, C. B., Smith, M. M., Cake, M. A., Read, R. A., Murphy, M. J., and Barry, F. P. (2010) The OARSI histopathology initiative – recommendations for histological assessments of osteoarthritis in sheep and goats. *Osteoarthr. Cartil.* **18**, S80–S92

55. Krenn, V., Morawietz, L., Burmester, G.-R., Kinne, R. W., Mueller-Ladner, U., Muller, B., and Haupl, T. (2006) Synovitis score: discrimination between chronic low-grade and high-grade synovitis. *Histopathology*. **49**, 358–364

56. Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., Reily, M. D., Thaden, J. J., and Viant, M. R. (2007) Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics*. **3**, 211–221

57. Peffers, M., Jones, A. R., McCabe, A., and Anderson, J. (2017) Neopeptide Analyser: A software tool for neopeptide discovery in proteomics data.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Wellcome Open Res. **2**, 24

58. Johnson, W. E., Li, C., and Rabinovic, A. (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. **8**, 118–127

59. Worley, B., and Powers, R. (2013) Multivariate Analysis in Metabolomics. *Curr. Metabolomics*. **1**, 92–107

60. Benjamini, Y., and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B*. **57**, 289–300

61. Kohl, S. M., Klein, M. S., Hochrein, J., Oefner, P. J., Spang, R., and Gronwald, W. (2012) State-of-the art data normalization methods improve NMR-based metabolomic analysis. *Metabolomics*. **8**, 146–160

62. Jombart, T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. **22**, 1403–1405

63. Wickham, H. (2007) Reshaping data with the reshape package. *J. Stat. Softw.* **21**, 1–20

64. Wickham, H. (2016) *ggplot2: elegant graphics for data analysis*, Use R!, Springer International Publishing, 10.1007/978-3-319-24277-4

65. Team_R_Core (2019) R: A language and environment for statistical computing. *R Found. Stat. Comput. Vienna, Austria*. [online] https://www.eea.europa.eu/data-and-maps/indicators/oxygen-consuming-substances-in-rivers/r-development-core-team-2006 (Accessed April 17, 2020)

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

66. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410

67. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., and von Mering, C. (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447-52

68. Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **27**, 29–34

69. Luo, Q., Qin, X., Qiu, Y., Hou, L., and Yang, N. (2018) The change of synovial fluid proteome in rabbit surgery-induced model of knee osteoarthritis. *Am. J. Transl. Res.* **10**, 2087–2101

70. Cooke, T. D., Hurd, E. R., Jasin, H. E., Bienenstock, J., and Ziff, M. (1975) Identification of immunoglobulins and complement in rheumatoid articular collagenous tissues. *Arthritis Rheum.* **18**, 541–551

71. Lawrence, D., Bao, S., J. Canfield, P., Allanson, M., and Husband, A. J. (1998) Elevation of immunoglobulin deposition in the synovial membrane of dogs with cranial cruciate ligament rupture. *Vet. Immunol. Immunopathol.* **65**, 89–96

72. Wen, Z.-H., Chang, Y.-C., and Jean, Y.-H. (2015) Excitatory amino acid glutamate: role in peripheral nociceptive transduction and inflammation in experimental and clinical osteoarthritis. *Osteoarthr. Cartil.* **23**, 2009–2016

73. Piepoli, T., Mennuni, L., Zerbi, S., Lanza, M., Rovati, L. C., and Caselli, G.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(2009) Glutamate signaling in chondrocytes and the potential involvement of NMDA receptors in cell proliferation and inflammatory gene expression. *Osteoarthr. Cartil.* **17**, 1076–1083

74. McNearney, T., Speegle, D., Lawand, N., Lisse, J., and Westlund, K. N. (2000) Excitatory amino acid profiles of synovial fluid from patients with arthritis. *J. Rheumatol.* **27**, 739–45

75. Lawand, N. B., McNearney, T., and Westlund, K. N. (2000) Amino acid release into the knee joint: key role in nociception and inflammation. *Pain*. **86**, 69–74

76. Jean, Y.-H., Wen, Z.-H., Chang, Y.-C., Huang, G.-S., Lee, H.-S., Hsieh, S.-P., and Wong, C.-S. (2005) Increased concentrations of neuro-excitatory amino acids in rat anterior cruciate ligament-transected knee joint dialysates: A microdialysis study. *J. Orthop. Res.* **23**, 569–575

77. da Silva, R. P., Clow, K., Brosnan, J. T., and Brosnan, M. E. (2014) Synthesis of guanidinoacetate and creatine from amino acids by rat pancreas. *Br. J. Nutr.* **111**, 571–577

78. Ciurtin, C., Cojocaru, V. M., Miron, I. M., Preda, F., Milicescu, M., Bojinca, M., Costan, O., Nicolescu, A., Deleanu, C., Kovacs, E., and Stoica, V. (2006) Correlation between different components of synovial fluid and pathogenesis of rheumatic diseases. *Rom J Intern Med*. **44**, 171–181

79. Snow, R. J., and Murphy, R. M. (2001) Creatine and the creatine transporter: A review. *Mol. Cell. Biochem.* **224**, 169–181

80. Porter, M. (2005) Equine rehabilitation therapy for joint disease. *Vet. Clin. North Am. Equine Pract.* **21**, 599–607, vi

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

81. Shet, K., Siddiqui, S. M., Yoshihara, H., Kurhanewicz, J., Ries, M., and Li, X. (2012) High-resolution magic angle spinning NMR spectroscopy of human osteoarthritic cartilage. *NMR Biomed.* **25**, 538–44

82. Huster, D., Schiller, J., Naji, L., Kaufmann, J., and Arnold, K. (2004) NMR Studies of Cartilage – Dynamics, Diffusion, Degradation, pp. 465–503, Springer, Berlin, Heidelberg, 10.1007/978-3-540-40024-0_13

83. Yin, H. L., Kwiatkowski, D. J., Mole, J. E., and Cole, F. S. (1984) Structure and biosynthesis of cytoplasmic and secreted variants of gelsolin. *J. Biol. Chem.* **259**, 5271–6

84. Kaneva, M. K., Greco, K. V, Headland, S. E., Montero-Melendez, T., Mori, P., Greenslade, K., Pitzalis, C., Moore, A., and Perretti, M. (2017) Identification of Novel Chondroprotective Mediators in Resolving Inflammatory Exudates. *J. Immunol.* **198**, 2876–2885

85. Piktel, E., Levental, I., Durnaś, B., Janmey, P., Bucki, R., Piktel, E., Levental, I., Durnaś, B., Janmey, P. A., and Bucki, R. (2018) Plasma Gelsolin: Indicator of Inflammation and Its Potential as a Diagnostic Tool and Therapeutic Target. *Int. J. Mol. Sci.* **19**, 2516

86. DiNubile, M. J., Stossel, T. P., Ljunghusen, O. C., Ferrara, J. L. M., and Antin, J. H. (2002) Prognostic implications of declining plasma gelsolin levels after allogeneic stem cell transplantation. *Blood*. **100**, 4367–71

87. Ito, H., Kambe, H., Kimura, Y., Nakamura, H., Hayashi, E., Kishimoto, T., Kishimoto, S., and Yamamoto, H. (1992) Depression of plasma gelsolin level during acute liver injury. *Gastroenterology*. **102**, 1686–92

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

88. Osborn, T. M., Verdrengh, M., Stossel, T. P., Tarkowski, A., and Bokarewa, M. (2008) Decreased levels of the gelsolin plasma isoform in patients with rheumatoid arthritis. *Arthritis Res. Ther.* **10**, R117

89. Aidinis, V., Carninci, P., Armaka, M., Witke, W., Harokopos, V., Pavelka, N., Koczan, D., Argyropoulos, C., Thwin, M.-M., Möller, S., Waki, K., Kazunori, W., Gopalakrishnakone, P., Ricciardi-Castagnoli, P., Thiesen, H.-J., Hayashizaki, Y., and Kollias, G. (2005) Cytoskeletal rearrangements in synovial fibroblasts as a novel pathophysiological determinant of modeled rheumatoid arthritis. *PLoS Genet.* **1**, 455–466

90. Gupta, A. K., Parasar, D., Sagar, A., Choudhary, V., Chopra, B. S., Garg, R., Ashish, and Khatri, N. (2015) Analgesic and Anti-Inflammatory Properties of Gelsolin in Acetic Acid Induced Writhing, Tail Immersion and Carrageenan Induced Paw Edema in Mice. *PLoS One*. **10**, e0135558

91. Ranoa, D. R. E., Kelley, S. L., and Tapping, R. I. (2013) Human lipopolysaccharide-binding protein (LBP) and CD14 independently deliver triacylated lipoproteins to Toll-like receptor 1 (TLR1) and TLR2 and enhance formation of the ternary signaling complex. *J. Biol. Chem.* **288**, 9729–41

92. Citronberg, J. S., Wilkens, L. R., Lim, U., Hullar, M. A. J., White, E., Newcomb, P. A., Le Marchand, L., and Lampe, J. W. (2016) Reliability of plasma lipopolysaccharide-binding protein (LBP) from repeated measures in healthy adults. *Cancer Causes Control*. **27**, 1163–6

93. Heumann, D., Bas, S., Gallay, P., Le Roy, D., Barras, C., Mensi, N., Glauser, M. P., and Vischer, T. (1995) Lipopolysaccharide binding protein as a marker

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of inflammation in synovial fluid of patients with arthritis: correlation with interleukin 6 and C-reactive protein. *J. Rheumatol.* **22**, 1224–9

94. Huang, Z. Y., Perry, E., Huebner, J. L., Katz, B., Li, Y.-J., and Kraus, V. B. (2018) Biomarkers of inflammation – LBP and TLR- predict progression of knee osteoarthritis in the DOXY clinical trial. *Osteoarthr. Cartil.* 10.1016/j.joca.2018.08.005

95. Lamping, N., Dettmer, R., Schröder, N. W., Pfeil, D., Hallatschek, W., Burger, R., and Schumann, R. R. (1998) LPS-binding protein protects mice from septic shock caused by LPS or gram-negative bacteria. *J. Clin. Invest.* **101**, 2065–2071

96. Schumann, R., Rietschel, E., and Loppnow, H. (1994) The role of CD14 and lipopolysaccharide-binding protein (LBP) in the activation of different cell types by endotoxin. *Med. Microbiol. Immunol.* **183**, 279–297

97. Wojdasiewicz, P., Poniatowski, Ł. A., and Szukiewicz, D. (2014) The role of inflammatory and anti-inflammatory cytokines in the pathogenesis of osteoarthritis. *Mediators Inflamm.* **2014**, 561459

98. Shahid, M. (2018) *Investigations on the quantitative and qualitative protein content in serum and synovial fluid of dogs with osteoarthritis*. Ph.D. thesis, Freien Universität Berlin

99. Luc, G., Majd, Z., Poulain, P., Elkhalil, L., and Fruchart, J. C. (1996) Interstitial fluid apolipoprotein A-II: an association with the occurrence of myocardial infarction. *Atherosclerosis*. **127**, 131–7

100. Ramella, N. A., Andújar, I., Ríos, J. L., Rosú, S. A., Tricerri, M. A., and

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Schinella, G. R. (2018) Human apolipoprotein A-I Gly26Arg stimulation of inflammatory responses via NF-kB activation: Potential roles in amyloidosis? *Pathophysiol. Off. J. Int. Soc. Pathophysiol.* 10.1016/j.pathophys.2018.08.002

101. Silver, D. L., Wang, N., Xiao, X., and Tall, A. R. (2001) High density lipoprotein (HDL) particle uptake mediated by scavenger receptor class B type 1 results in selective sorting of HDL cholesterol from protein and polarized cholesterol secretion. *J. Biol. Chem.* **276**, 25287–93

102. Rader, D. J. (2003) Regulation of reverse cholesterol transport and clinical implications. *Am. J. Cardiol.* **92**, 42J-49J

103. de Seny, D., Cobraiville, G., Charlier, E., Neuville, S., Lutteri, L., Le Goff, C., Malaise, D., Malaise, O., Chapelle, J.-P., Relic, B., and Malaise, M. G. (2015) Apolipoprotein-A1 as a Damage-Associated Molecular Patterns Protein in Osteoarthritis: Ex Vivo and In Vitro Pro-Inflammatory Properties. *PLoS One*. **10**, e0122904

104. Yang, M., Liu, Y., Dai, J., Li, L., Ding, X., Xu, Z., Mori, M., Miyahara, H., Sawashita, J., and Higuchi, K. (2018) Apolipoprotein A-II induces acute-phase response associated AA amyloidosis in mice through conformational changes of plasma lipoprotein structure. *Sci. Rep.* **8**, 5620

105. Mathiessen, A., and Conaghan, P. G. (2017) Synovitis in osteoarthritis: current understanding with therapeutic implications. *Arthritis Res. Ther.* **19**, 18

106. Berenbaum, F. (2013) Osteoarthritis as an inflammatory disease (osteoarthritis is not osteoarthrosis!). *Osteoarthr. Cartil.* **21**, 16–21

107. Sellam, J., and Berenbaum, F. (2010) The role of synovitis in pathophysiology

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and clinical symptoms of osteoarthritis. *Nat. Rev. Rheumatol.* **6**, 625–635

108. Ritter, S. Y., Collins, J., Krastins, B., Sarracino, D., Lopez, M., Losina, E., and Aliprantis, A. O. (2014) Mass spectrometry assays of plasma biomarkers to predict radiographic progression of knee osteoarthritis. *Arthritis Res. Ther.* **16**, 456

109. Dieplinger, H., and Dieplinger, B. (2015) Afamin — A pleiotropic glycoprotein involved in various disease states. *Clin. Chim. Acta*. **446**, 105–110

110. Mihara, E., Hirai, H., Yamamoto, H., Tamura-Kawakami, K., Matano, M., Kikuchi, A., Sato, T., and Takagi, J. (2016) Active and water-soluble form of lipidated Wnt protein is maintained by a serum glycoprotein afamin/α-albumin. *Elife*. 10.7554/eLife.11621

111. Zhou, Y., Wang, T., Hamilton, J. L., and Chen, D. (2017) Wnt/β-catenin Signaling in Osteoarthritis and in Other Forms of Arthritis. *Curr. Rheumatol. Rep.* **19**, 53

112. Lotz, M., Martel-Pelletier, J., Christiansen, C., Brandi, M.-L., Bruyère, O., Chapurlat, R., Collette, J., Cooper, C., Giacovelli, G., Kanis, J. A., Karsdal, M. A., Kraus, V., Lems, W. F., Meulenbelt, I., Pelletier, J.-P., Raynauld, J.-P., Reiter-Niesert, S., Rizzoli, R., Sandell, L. J., Van Spil, W. E., and Reginster, J.-Y. (2013) Value of biomarkers in osteoarthritis: current status and perspectives. *Ann. Rheum. Dis.* **72**, 1756–63

113. Peffers, M. J., Smagul, A., and Anderson, J. R. (2019) Proteomic analysis of synovial fluid: current and potential uses to improve clinical outcomes. *Expert Rev. Proteomics*. **16**, 287–302

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

114. Ritter, S. Y., Subbaiah, R., Bebek, G., Crish, J., Scanzello, C. R., Krastins, B., Sarracino, D., Lopez, M. F., Crow, M. K., Aigner, T., Goldring, M. B., Goldring, S. R., Lee, D. M., Gobezie, R., and Aliprantis, A. O. (2013) Proteomic analysis of synovial fluid from the osteoarthritic knee: comparison with transcriptome analyses of joint tissues. *Arthritis Rheum.* **65**, 981–92

115. Li, W., Gao, P., Zhi, Y., Xu, W., Wu, Y., Yin, J., and Zhang, J. (2015) Periostin: its role in asthma and its potential as a diagnostic or therapeutic target. *Respir. Res.* **16**, 57

116. Kudo, A., and Kii, I. (2018) Periostin function in communication with extracellular matrices. *J. Cell Commun. Signal.* **12**, 301–308

117. Honsawek, S., Wilairatana, V., Udomsinprasert, W., Sinlapavilawan, P., and Jirathanathornnukul, N. (2015) Association of plasma and synovial fluid periostin with radiographic knee osteoarthritis: Cross-sectional study. *Jt. Bone Spine*. **82**, 352–355

118. Tajika, Y., Moue, T., Ishikawa, S., Asano, K., Okumo, T., Takagi, H., and Hisamitsu, T. (2017) Influence of Periostin on Synoviocytes in Knee Osteoarthritis. *In Vivo*. **31**, 69–77

119. Rousseau, J. C., Sornay-Rendu, E., Bertholon, C., Garnero, P., and Chapurlat, R. (2015) Serum periostin is associated with prevalent knee osteoarthritis and disease incidence/progression in women: the OFELY study. *Osteoarthr. Cartil.* **23**, 1736–1742

120. Silawal, S., Triebel, J., Bertsch, T., and Schulze-Tanzil, G. (2018) Osteoarthritis and the Complement Cascade. *Clin. Med. Insights. Arthritis*

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Musculoskelet. Disord. **11**, 1179544117751430

121. Horwitz, J. K., Chun, N. H., and Heeger, P. S. (2019) Complement and Transplantation. *Clin. Lab. Med.* **39**, 31–43

122. So, A. K., Varisco, P.-A., Kemkes-Matthes, B., Herkenne-Morard, C., Chobaz-Peclat, V., Gerster, J.-C., and Busso, N. (2003) Arthritis is linked to local and systemic activation of coagulation and fibrinolysis pathways. *J. Thromb. Haemost.* **1**, 2510–2515

123. Weinberg, J. B., Pippen, A. M. M., and Greenberg, C. S. (1991) Extravascular fibrin formation and dissolution in synovial tissue of patients with osteoarthritis and rheumatoid arthritis. *Arthritis Rheum.* **34**, 996–1005

124. Raghu, H., and Flick, M. (2011) Targeting the Coagulation Factor Fibrinogen for Arthritis Therapy. *Curr. Pharm. Biotechnol.* **12**, 1497–1506

125. Rees, D. C., Johnson, E., and Lewinson, O. (2009) ABC transporters: the power to change. *Nat. Rev. Mol. Cell Biol.* **10**, 218–27

126. Takamatsu, A., Ohkawara, B., Ito, M., Masuda, A., Sakai, T., Ishiguro, N., and Ohno, K. (2014) Verapamil Protects against Cartilage Degradation in Osteoarthritis by Inhibiting Wnt/β-Catenin Signaling. *PLoS One*. **9**, e92699

127. Schulz, T., Schumacher, U., and Prehm, P. (2007) Hyaluronan export by the ABC transporter MRP5 and its modulation by intracellular cGMP. *J. Biol. Chem.* **282**, 20999–1004

128. Haug, K., Cochrane, K., Nainala, V. C., Williams, M., Chang, J., Jayaseelan, K. V., and O'Donovan, C. (2019) MetaboLights: a resource evolving in response

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to the needs of its scientific community. *Nucleic Acids Res.* 10.1093/nar/gkz1019

129. Salek, R. M., Steinbeck, C., Viant, M. R., Goodacre, R., and Dunn, W. B. (2013) The role of reporting standards for metabolite annotation and identification in metabolomic studies. *Gigascience*. **2**, 13

Tables

Table 1. Stratification groups according to microscopic osteoarthritis, macroscopic osteoarthritis and synovitis pathology used for synovial fluid LC-MS/MS proteomic and NMR metabolomic analysis.

Biobank

Metabolomics	OARSI Microscopic OA Group			OARSI Macroscopic OA Group			
	0	1	2	0	1	2	3
Number of Donors	8	34	28	14	27	21	14
Joint	8 x MCP	34 x MCP	28 x MCP	14 x MCP	27 x MCP	21 x MCP	14 x MCP
Mean Age (years)	8	15	17	9	15	16	19
Sex (M/F)	1 F, 7 Unknown	6 M, 10 F, 18 Unknown	12 M, 10 F, 6 Unknown	2 M, 5 F, 7 Unknown	7 M, 12 F, 8 Unknown	9 M, 5 F, 7 Unknown	6 M, 3 F, 5 Unknown

Proteomics	OARSI Microscopic OA Group			OARSI Macroscopic OA Group			
	0	1	2	0	1	2	3
Number of Donors	8	34	28	11	25	18	12
Joint	8 x MCP	34 x MCP	28 x MCP	11 x MCP	25 x MCP	18 x MCP	12 x MCP
Mean Age (years)	8	15	17	9	15	16	20
Sex (M/F)	1 F, 7 Unknown	6 M, 10 F, 18 Unknown	12 M, 10 F, 6 Unknown	1 M, 3 F, 7 Unknown	6 M, 11 F, 8 Unknown	6 M, 5 F, 7 Unknown	5 M, 2 F, 5 Unknown
Mean Protein Concentration (mg/ml)	5	5	4	5	4	5	6

Hong Kong Jockey Club

Metabolomics	Microscopic OA Group			Macroscopic OA Group			Synovitis Group		
	0	1	2	0	1	2	0	1	2
Number of Donors	17	20	13	13	25	12	1	38	17
Joint	10 x MCP, 7 x MTP	8 x MCP, 12 x MTP	9 x MCP, 4 x MTP	5 x MCP, 8 x MTP	13 x MCP, 12 x MTP	11 x MCP, 11 x MTP	1 x MTP	22 x MCP, 16 x MTP	10 x MCP, 7 x MTP
Mean Age (years)	6	7	7	5	7	7	6	7	7
Sex (M/F)	UNKNOWN			UNKNOWN			UNKNOWN		

Proteomics	Microscopic OA Group			Macroscopic OA Group			Synovitis Group		
	0	1	2	0	1	2	0	1	2
Number of Donors	16	20	13	13	24	12	1	34	19
Joint	9 x MCP, 7 x MTP	8 x MCP, 12 x MTP	9 x MCP, 4 x MTP	5 x MCP, 8 x MTP	12 x MCP, 12 x MTP	11 x MCP, 11 x MTP	1 x MTP	18 x MCP, 16 x MTP	10 x MCP, 9 x MTP
Mean Age (years)	6	7	7	5	7	7	6	7	7
Sex (M/F)	UNKNOWN			UNKNOWN			UNKNOWN		
Mean Protein Concentration (mg/ml)	10	9	13	7	9	17	16	9	12

Abbreviations: OA = Osteoarthritis; Joints, MCP = Metacarpophalangeal joint, MTP = Metatarsophalangeal joint; Sex, M = male, F = Female.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. Synovial fluid metabolites identified using Chenomx software. Metabolites which had also undergone identification using a 1D ^{1}H NMR in-house spectral library were assigned to Metabolomics Standards Initiative (MSI) level 1 (129).

Database Identifier	Metabolite Identification	Reliability
HMDB00650	2-Aminobutyrate	MS Level 2
HMDB00357	3-Hydroxybutyrate	MS Level 2
HMDB00754	3-Hydroxyisovalerate	MS Level 2
HMDB01149	5-Aminolevulinate	MS Level 2
HMDB00042	Acetate	MS Level 1
HMDB00194	Anserine	MS Level 2
HMDB00043	Betaine	MS Level 1
HMDB00097	Choline	MS Level 1
HMDB00094	Citrate	MS Level 1
HMDB00064	Creatine	MS Level 1
HMDB01511	Creatine phosphate	MS Level 2
HMDB00562	Creatinine	MS Level 1
HMDB00122	D-Glucose	MS Level 1
HMDB04983	Dimethyl sulfone	MS Level 2
HMDB00142	Formate	MS Level 2
HMDB00123	Glycine	MS Level 1
HMDB00128	Guanidoacetate	MS Level 2
HMDB00172	Isoleucine	MS Level 1
HMDB00190	Lactate	MS Level 1
HMDB00161	L-Alanine	MS Level 1
HMDB00062	L-Carnitine	MS Level 2
HMDB00174	L-Fucose	MS Level 2
HMDB00148	L-Glutamate	MS Level 1
HMDB00641	L-Glutamine	MS Level 1
HMDB00177	L-Histidine	MS Level 1
HMDB00687	L-Leucine	MS Level 1
HMDB00159	L-Phenylalanine	MS Level 1
HMDB00158	L-Tyrosine	MS Level 1
HMDB00883	L-Valine	MS Level 1
HMDB00691	Malonate	MS Level 2
HMDB01844	Methylsuccinate	MS Level 2
HMDB31419	N-Nitrosodimethylamine	MS Level 2
HMDB00895	O-Acetylcholine	MS Level 2
HMDB00243	Pyruvate	MS Level 1
HMDB00254	Succinate	MS Level 1
HMDB00294	Urea	MS Level 1
HMDB00296	Uridine	MS Level 1

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

HMDB00292	Xanthine	MS Level 2
HMDB00001	π-Methylhistidine	MS Level 2
HMDB00001	τ-Methylhistidine	MS Level 2

Metabolomics Standards Initiative definitions: MS Level 1 = Identified metabolite using two or more orthogonal properties of an authentic chemical standard analysed in the researcher's laboratory. MS Level 2 = Putatively annotated metabolite which does not require matching to data for authentic chemical standards acquired within the same laboratory.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 3. Pathway analysis conducted on proteins which were considered significant variables during DAPC modelling, when carrying out NMR metabolomic and LC-MS/MS proteomic dataset integration.

Pathway ID	Pathway Description	Number of Proteins in Pathway	False Discovery Rate
4610	Complement and coagulation cascades	7	0.00000573
4145	Phagosome	6	0.00498
4390	Hippo signalling pathway	5	0.0188

Table 4. Pathway analysis conducted on metabolites which were considered significant variables during DAPC modelling, when carrying out NMR metabolomic and LC-MS/MS proteomic dataset integration.

Pathway ID	Pathway Description	Number of Metabolites in Pathway
2010	ABC transporters	4
4978	Mineral absorption	3
10	Glycolysis / Gluconeogenesis	2
4974	Protein digestion and absorption	2
340	Histidine metabolism	2
4066	HIF-1 signalling pathway	2
4976	Bile secretion	2
4922	Glucagon signalling pathway	2
970	Aminoacyl-tRNA biosynthesis	2
1230	Biosynthesis of amino acids	2

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figures and Figure Legends

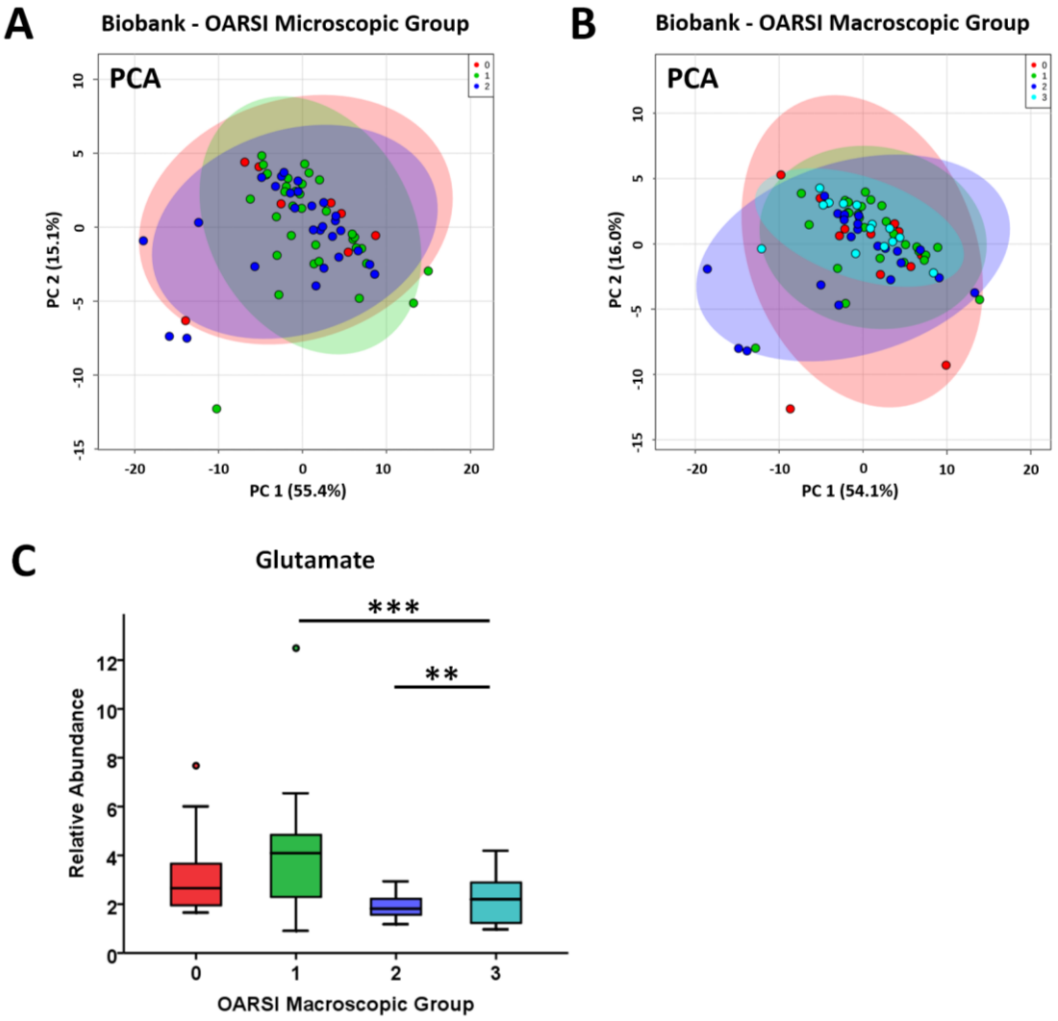

Figure 1. Principal component analysis (PCA) of metabolite profiles of equine biobank synovial fluid grouped according to (A) OARSI microscopic (n=70) and (B) OARSI macroscopic (n=76) grading. (C) Glutamate abundances according to macroscopic osteoarthritis grading. ANOVA: ** = $p < 0.01$ and *** = $p < 0.001$.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

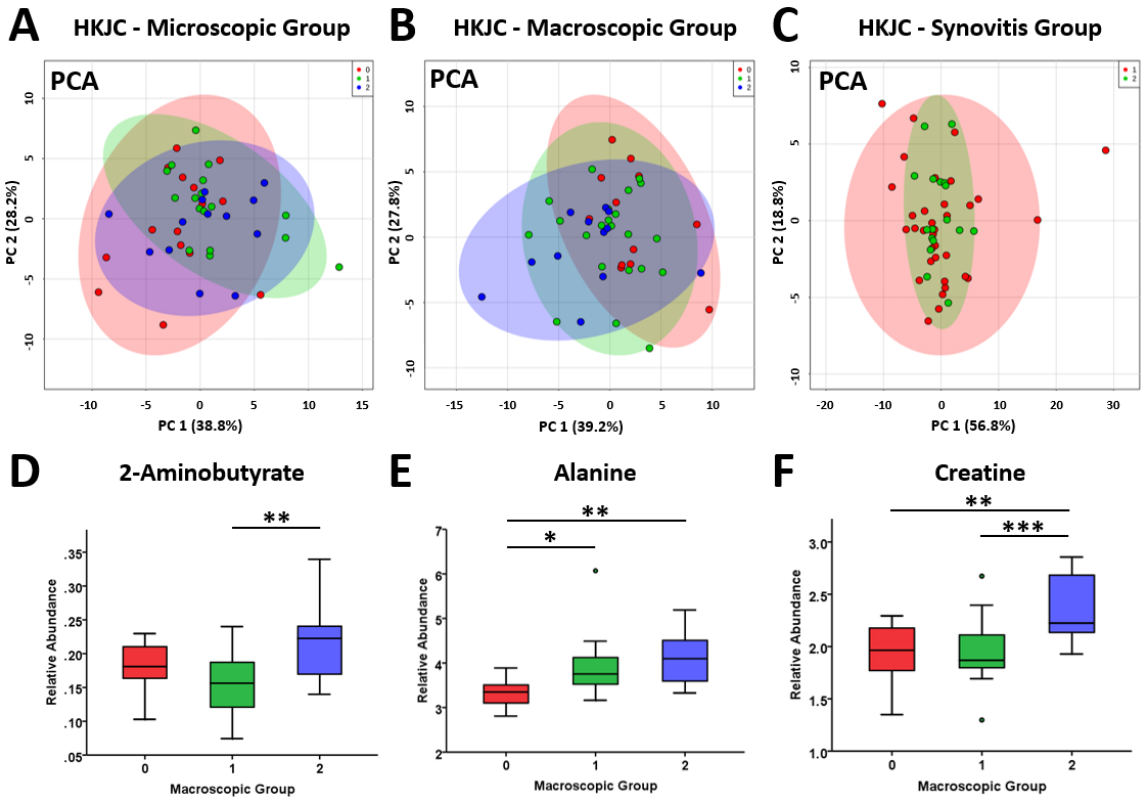

Figure 2. Principal component analysis (PCA) of metabolite profiles of equine Hong Kong Jockey Club (HKJC) synovial fluid (SF) grouped according to (A) OARSI microscopic (n=50), (B) OARSI macroscopic (n=50) and (C) synovitis (n=56) grading. For OARSI macroscopic grading, (D) 2-aminobutyrate, (E) alanine and (F) creatine were identified as being differentially abundant between osteoarthritis severity grades. ANOVA: * = $p < 0.05$, ** = $p < 0.01$ and *** = $p < 0.001$.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	NATIVE	PROTEOMINER™
No. Samples	130	124
LC gradient length (mins)	90	120
Total no. proteins identified	621	1,666
Mean no. proteins identified/sample	182	357
Median no. proteins identified/sample	165	294
Lowest no. proteins identified in a sample	115	124
Highest no. proteins identified in a sample	407	870

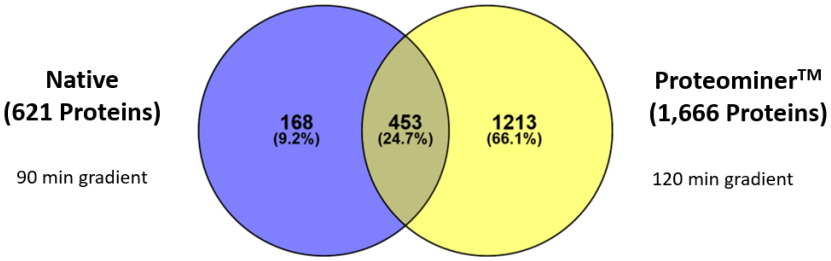

Figure 3. Number of proteins identified within native and ProteoMiner™ processed equine synovial fluid using liquid chromatography-tandem mass spectrometry. In total, 1,834 different proteins were identified. Search criteria included > 2 unique peptides with an FDR of 1%. LC = liquid chromatography.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

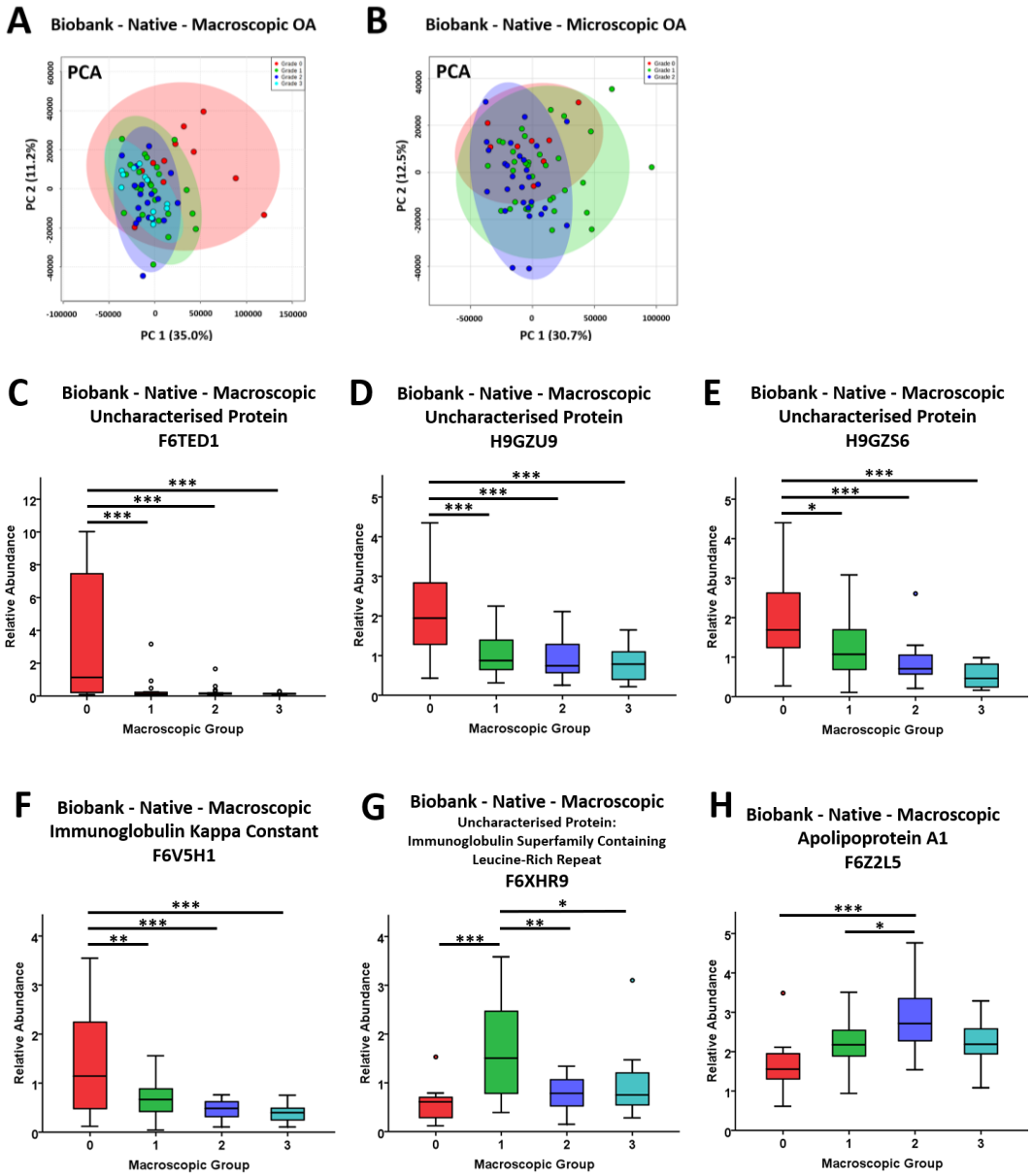

Figure 4. Principal component analysis (PCA) of the biobank native synovial fluid proteome categorised by (A) macroscopic osteoarthritis (OA) grade and (B) microscopic OA grade using LC-MS/MS. (C-H) Differentially expressed proteins when categorised according to macroscopic OA grade. ANOVA: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Macroscopic OA, n=66, Microscopic OA, n=70.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

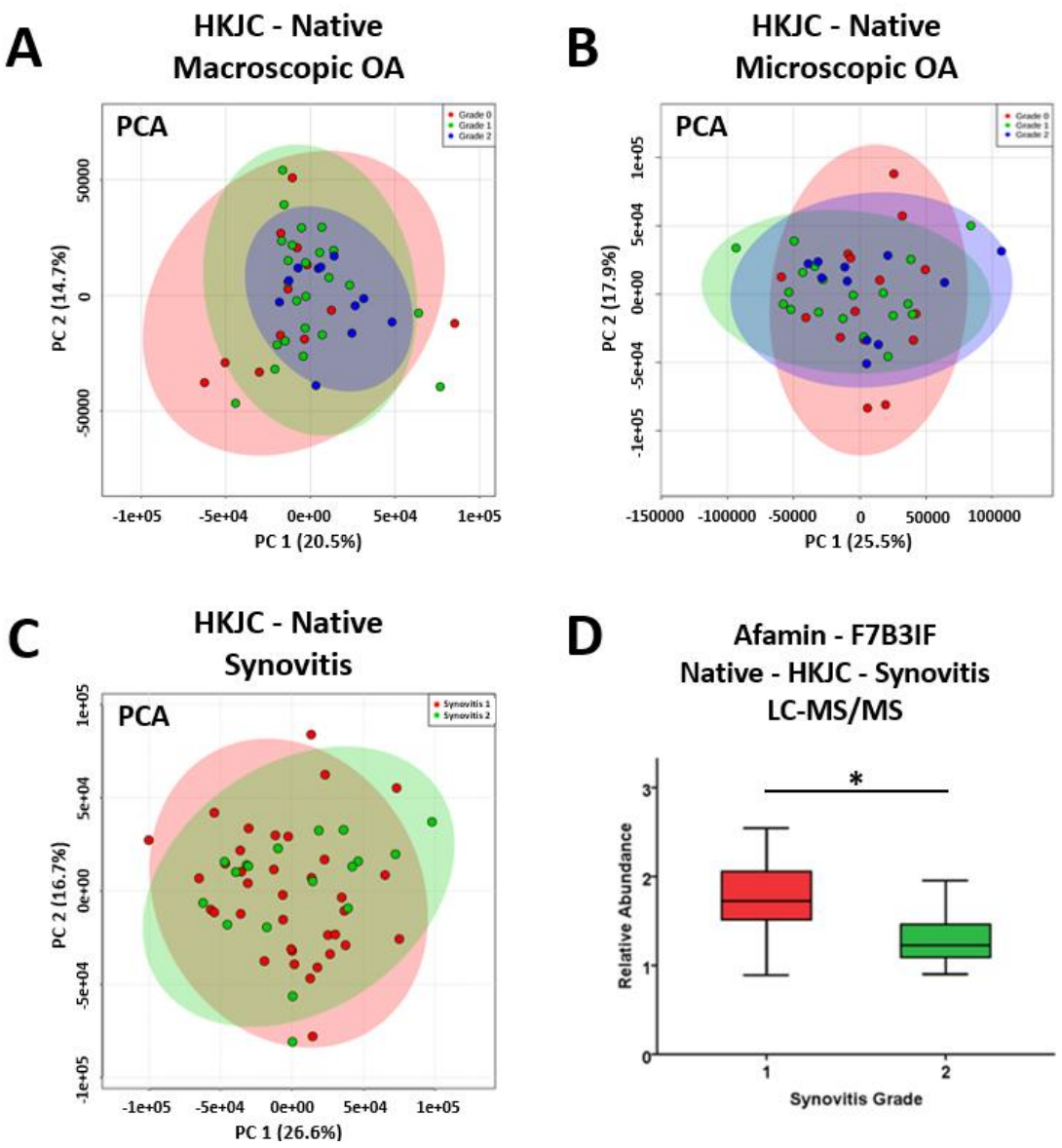

Figure 5. Principal component analysis (PCA) of the Hong Kong Jockey Club (HKJC) native synovial fluid proteome profile categorised by (A) macroscopic osteoarthritis (OA) grade (n=49), (B) microscopic OA grade (n=49) and (C) synovitis grade (n=53) using LC-MS/MS. (D) Differential expression of afamin identified between synovitis grade 1 and 2. t-test: * = $p < 0.05$.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

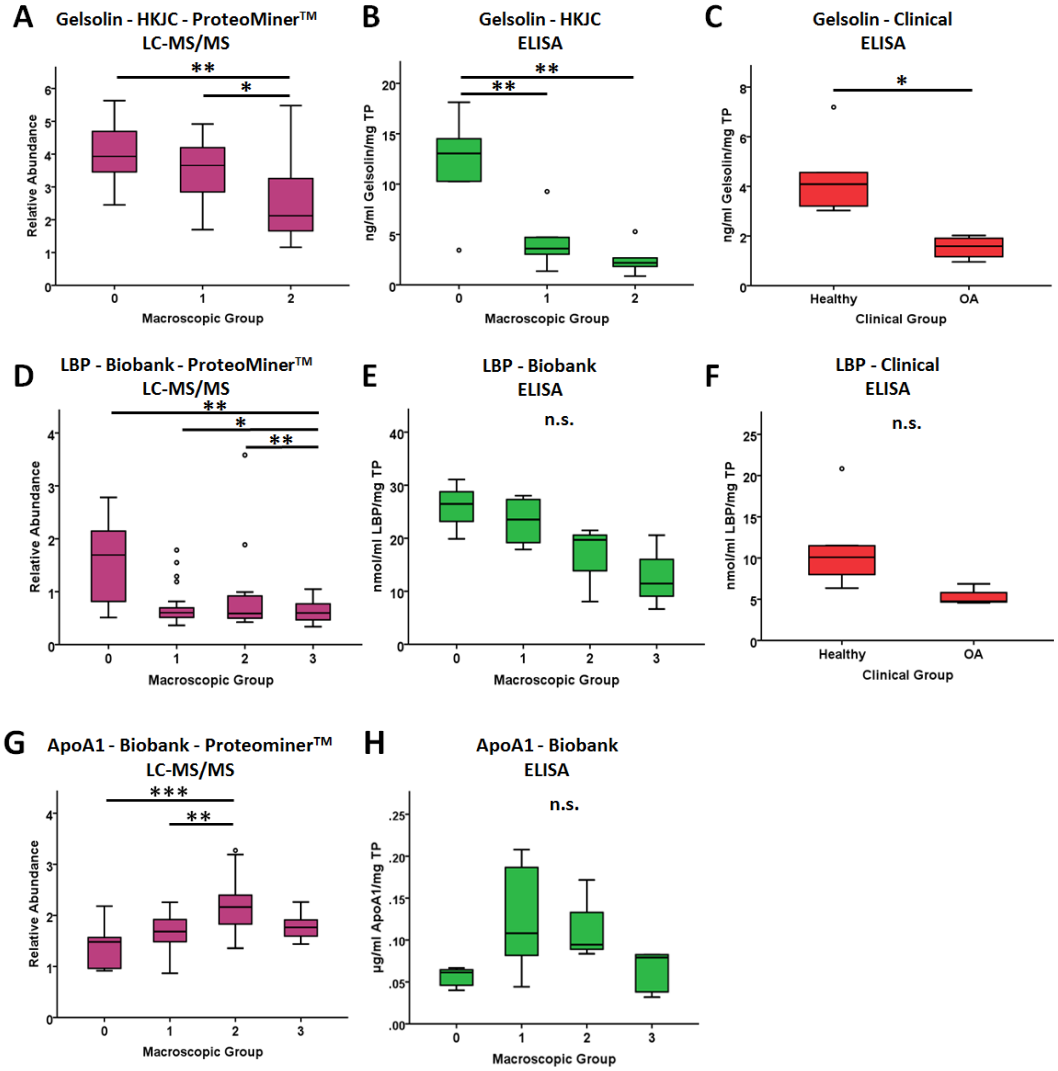

Figure 6. Abundances of gelsolin, lipopolysaccharide binding protein (LBP) and apolipoprotein A1 (ApoA1) within equine synovial fluid via LC-MS/MS (A, D & G) and ELISA validation using dependent (B, E & H) and independent (C & F) synovial fluid samples according to osteoarthritis (OA) grade. LC-MS/MS; HKJC, n=47; Biobank, n=60; ELISA, n=3-6/group. ANOVA and t-tests: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. LC-MS/MS p values corrected for multiple testing within sample, but not for entire dataset.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

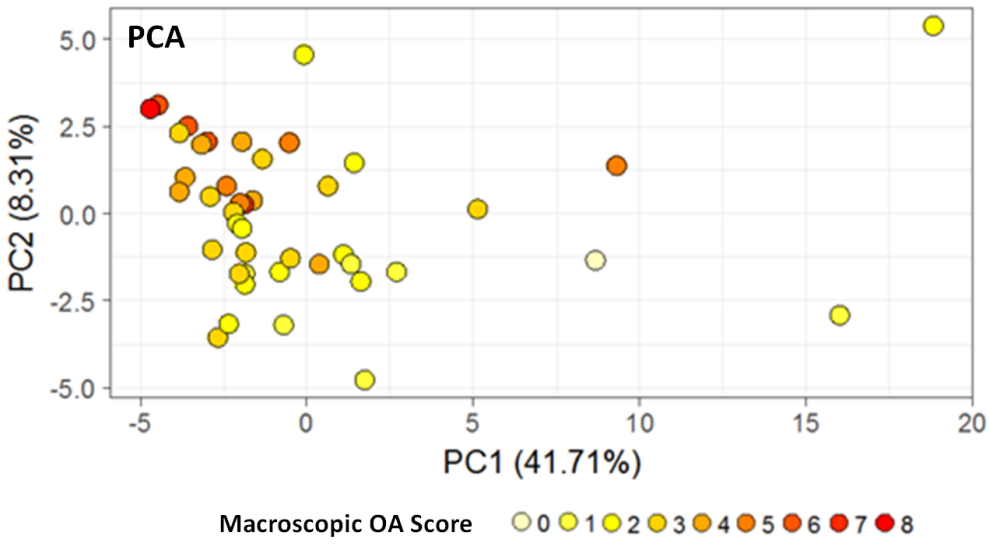

Figure 7. Principal component analysis (PCA) using 58 selected variables (proteins and metabolites) for the Hong Kong Jockey Club synovial fluid combined datasets (n=43) following batch correction, grouped according to macroscopic osteoarthritis score. OA = osteoarthritis.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

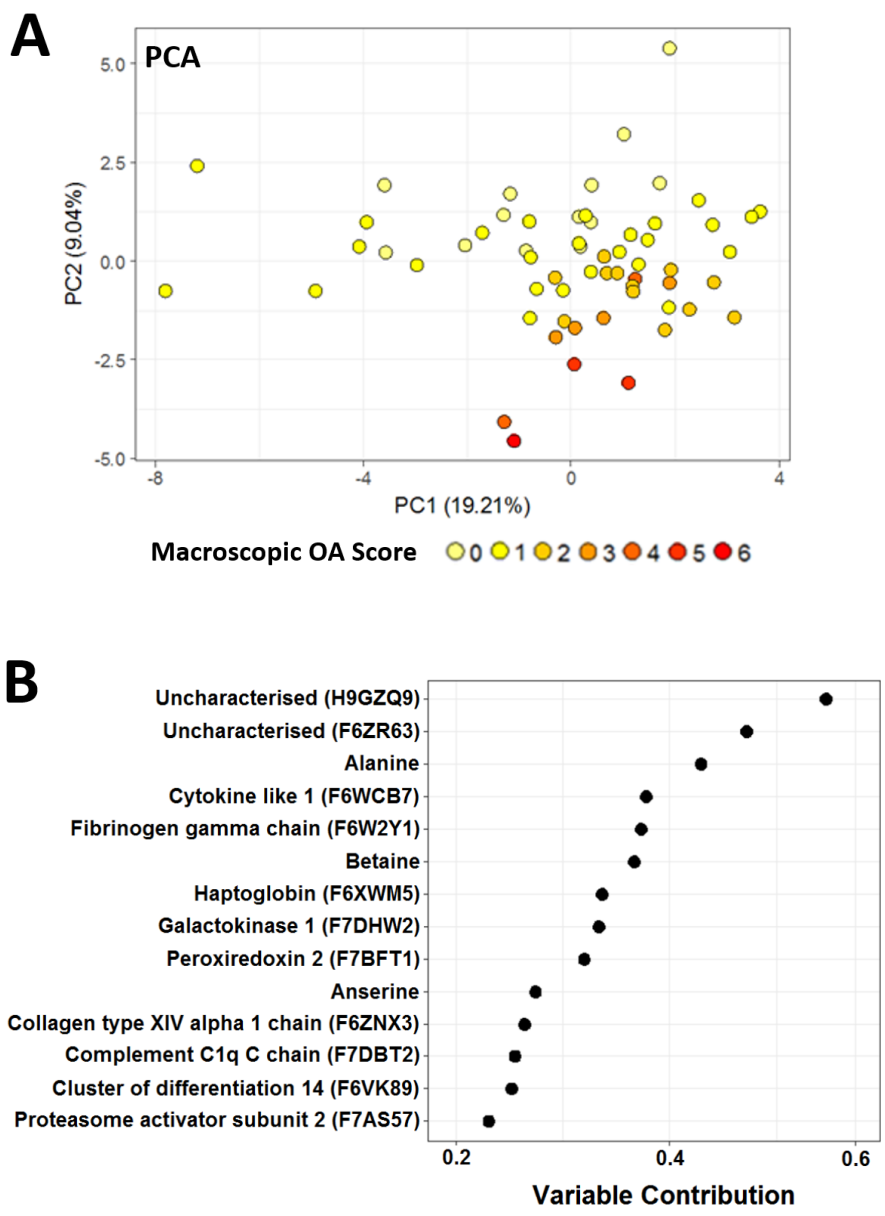

Figure 8. (A) Principal component analysis (PCA) following Lasso model selection using the top 29 variables of influence for biobank synovial fluid, integrating metabolite and protein abundances and grouped according to macroscopic osteoarthritis scoring. (B) Top 14 variables of influence contributing to the model. n=60.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.04.077305; this version posted May 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

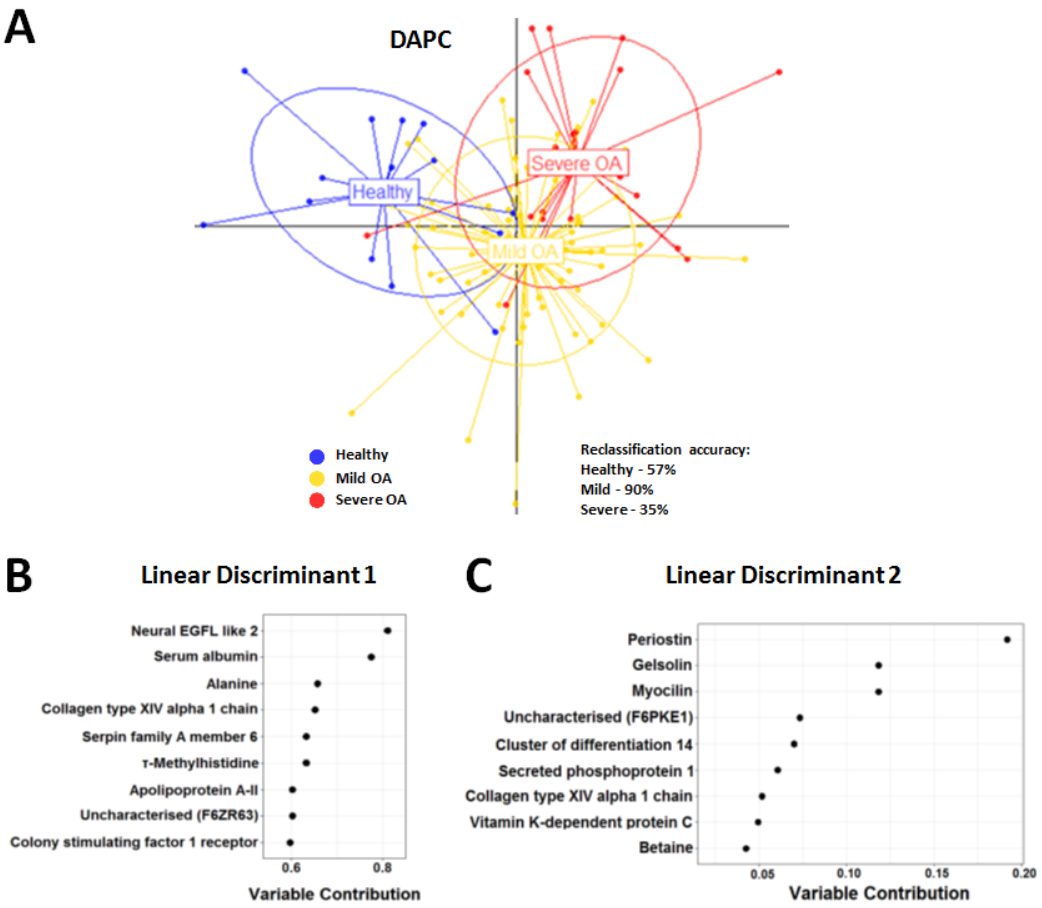

Figure 9. (A) Discriminant analysis of principal components (DAPC) of selected variables (proteins and metabolites) for a combined biobank and Hong Kong Jockey Club synovial fluid dataset (n=103), built using 20 principal components and two linear discriminants. (B) Top 25% of variables for the DAPC model contributing to linear discriminant 1 and (C) linear discriminant 2. Macroscopic osteoarthritis (OA) scoring; Healthy = 0, Mild OA = 1-3, Severe OA = 4-8.