Journal of proteome research

Article

pubs.acs.org/jpr

Terms of Use

A Normative Study of the Synovial Fluid Proteome from Healthy Porcine Knee Joints

Tue Bennike,[†,⊥] Ugur Ayturk,[‡,#] Carla M. Haslauer,[‡] John W. Froehlich,[§] Benedikt L. Proffen,[‡] Omar Barnaby,[†] Svend Birkelund,[⊥] Martha M. Murray,[‡] Matthew L. Warman,[‡,#,∥] Allan Stensballe,[⊥] and Hanno Steen*[,†]

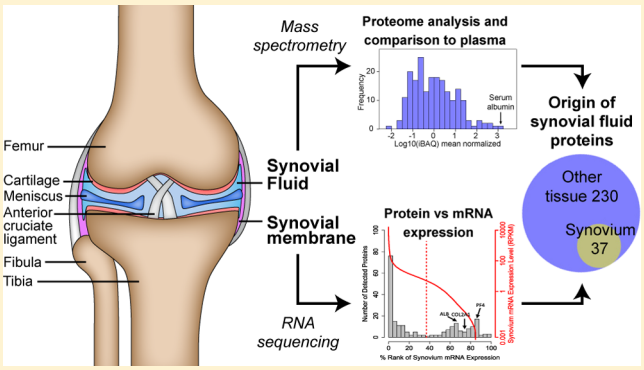
[†]Department of Pathology and Proteomics Center, [‡]Department of Orthopaedic Surgery, [§]Department of Urology, and [∥]Howard Hughes Medical Institute, Boston Children's Hospital, Boston, Massachusetts 02115, United States

[⊥]Department of Health Science and Technology, Aalborg University, Aalborg DK-9220, Denmark

[#]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, United States

Supporting Information

ABSTRACT: Synovial fluid in an articulating joint contains proteins derived from the blood plasma and proteins that are produced by cells within the joint tissues, such as synovium, cartilage, ligament, and meniscus. The proteome composition of healthy synovial fluid and the cellular origins of many synovial fluid components are not fully understood. Here, we present a normative proteomics study using porcine synovial fluid. Using our optimized method, we identified 267 proteins with high confidence in healthy synovial fluid. We also evaluated mRNA expression data from tissues that can contribute to the synovial fluid proteome, including synovium, cartilage, blood, and liver, to better estimate the relative contributions from these sources to specific synovial fluid components. We identified 113 proteins in healthy synovial fluid that appear to be primarily derived from plasma transudates, 37 proteins primarily derived from synovium, and 11 proteins primarily derived from cartilage. Finally, we compared the identified synovial fluid proteome to the proteome of human plasma, and we found that the two body fluids share many similarities, underlining the detected plasma derived nature of many synovial fluid components. Knowing the synovial fluid proteome of a healthy joint will help to identify mechanisms that cause joint disease and pathways involved in disease progression.



KEYWORDS: *Synovial fluid, synovium, plasma, porcine, human, proteomics, transcriptomics, origin, PTM, method optimization*

INTRODUCTION

Synovial fluid is present in all joint cavities, where it protects the articular cartilage surfaces, in part by reducing friction. Synovial fluid, furthermore, facilitates the transport of nutrients and waste products including proteins and metabolites between the vascularized synovium and the avascular cartilage.[1−3] Many components of synovial fluid are derived from blood plasma, and these two body fluids share many similarities in terms of protein composition.[4,5] However, synovial fluid also contains proteins secreted from the surrounding tissue, including the articular cartilage and synovium.[6] The protein concentration in synovial fluid from healthy knee joints is approximately 25 mg/mL, i.e., ∼$^1/_3$ of the concentration found in blood plasma, and albumin constitutes approximately 12 mg/mL.[1,3,7−10]

Joint diseases, in particular osteoarthritis (OA) and rheumatoid arthritis (RA), are the leading cause of disability in people over 55 years.[6] It has been estimated that, as of 2005, 27 million adults in the United States have clinical OA, and in 2009, OA was the fourth most common cause of hospitalization.[11,12] Furthermore, joint injuries, such as anterior cruciate tear, that predispose to precocious joint failure have become epidemic in young athletes.[13,14] Although changes in the composition of synovial fluid have been described in patients with joint disease, there do not exist reliable biomarkers for early disease diagnosis or biomarkers that accurately depict response to therapy.[1,15−17] As a consequence, OA is often not diagnosed before irreversible damage has occurred.[15,18,19] Since synovial fluid is in direct contact with the joint tissues, it provides an attractive source of biomarkers candidates for monitoring joint health and for furthering the understanding of the disease mechanisms.[3,18] Previous studies of synovial fluid have mainly been focused on various diseased states. However, the protein concentration, content, and volume of synovial fluid is known to change dramatically during active joint diseases, and few studies have been focused on synovial fluid in healthy

Received: June 12, 2014
Published: August 26, 2014

   dx.doi.org/10.1021/pr500587x | J. Proteome Res. 2014, 13, 4377−4387

state or the likely origins of the synovial fluid proteins.[3,17] Balakrishnan et al.[20] compared the human synovial fluid proteome from OA and RA patients. Synovial fluid was immune-depleted of the most abundant proteins, followed by liquid chromatography−tandem mass spectrometry (LC−MS/MS) analysis, which led to the identification of 575 proteins, of which 135 demonstrated a greater than 3-fold abundance change between the two groups. In a similar study, Mateos et al.[1] identified a total of 136 different proteins using two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) protein separation, followed by in-gel trypsin digestion and protein identification by LC−matrix-assisted laser desorption/ionization time-of-flight (MALDI TOF/TOF) MS. When assessing the total number of different proteins to be expected in synovial fluid, analyzing 2D-PAGE images seems to be a valid approach. Chen et al.[10] analyzed synovial fluid from patients with joint swelling using several precipitation strategies and detected 456 protein spots on 2D-PAGE images. Smith et al.[21] detected 1000 protein spots on 2D-PAGE images of synovial fluid from RA patients, which was estimated to represent 300 individual proteins. However, neither of the two latter studies identified the proteins within the protein spots. 2D-PAGE in-gel digestion strategies, while highly informative, are not suitable for high-throughput proteome analyses.[19] Therefore, we investigated the use of in-solution proteomics strategies with the aim to increase high-throughput protein identification and quantitation. In this normative proteomics study of synovial fluid from healthy porcine knee joints, we report results from applying our optimized method, using porcine synovial fluid. Domesticated or minipigs are a highly relevant anatomically large animal model organism to study several human diseases, including acute joint injury, OA, and other inflammatory disease.[22−26] The homology between the porcine genome and the human genome is conserved to a much greater extent than that between human and mouse or other rodents.[27] Furthermore, the porcine model has previously been used to study anterior cruciate ligament injury and surgical and tissue engineering approaches for healing.[28−30] We also report mRNA sequence data from healthy porcine synovium that we used to estimate the contribution to the synovial fluid proteome from this tissue, and we compare this to the relative contribution from plasma (e.g., proteins secreted into plasma by the liver or by blood cells) and articular cartilage.

EXPERIMENTAL PROCEDURES

Collection of Synovium and Synovial Fluid Samples

Six adolescent Yucatan minipigs (Coyote CCI, Douglas, MA), aged 12−15 months, were obtained for use in this study. All minipigs were housed and monitored by the Animal Resources at Boston Children's Hospital (ARCH) and handled according to approved Institutional Animal Care and Use Committee (IACUC) protocols. Minipigs were acclimated to the ARCH environment for a minimum of 3 days prior to experimental handling.

Nonbloody synovial fluid was obtained from the joints by aspiration using a 21 gauge needle. Samples were centrifuged at 3000*g* at room temperature for 10 min to pellet and remove cells and cellular debris. In some cases, 3 mL of sterile saline was injected into the knee joint to facilitate fluid extraction; after saline injection the knee was bent 10 times to ensure homogeneous fluid distribution and mixing. The saline/synovial fluid mix was then processed as above. Following centrifugation, the supernatants were stored at −80 °C. Furthermore, a human synovial fluid sample was obtained from a RA patient according to an approved IRB protocol (IRB-P00006443) to evaluate the integrity of the UniProt *Sus scrofa* protein database.

Euthanasia of the animals was induced by intramuscular injection of atropine (0.04 mg/kg), Telazol (4.4 mg/kg), and xylazine (2.2 mg/kg) and finalized by intravenous injection of Fatal Plus (86 mg/kg). At the time of euthanasia, synovia from the knee joints of the hind limbs were harvested. Care was taken to sample only the synovial membrane without any subintimal structures, such as fat or blood vessels. Each tissue specimen was snap frozen in liquid nitrogen and stored at −80 °C.

Protein Concentration

Total protein concentration for each sample (diluted 1:30 in water) was determined for normalization of sample material using a colorimetric (Bradford) protein assay kit (Bio-Rad, Hercules, CA) according to the manufacturer's instructions, with bovine serum albumin used as the standard.

SDS-PAGE

Thirty micrograms of total synovial fluid protein was prepared for sodium dodecyl sulfate (SDS)-PAGE in Laemmli sample buffer (Bio-Rad, Hercules, CA) according to the manufacturer's instructions. SeeBlue Plus2 pre-stained standard (Invitrogen, Carlsbad, CA) was used as the protein molecular weight standard. The sample was fractionated using NuPAGE 4−12% Bis-Tris minigels (Invitrogen) at 150 V for 65 min in MOPS SDS-running buffer (Invitrogen). The gel was stained using Coomassie blue, SimplyBlue SafeStain (Invitrogen), according to manufacturer's instructions.

Synovial Fluid Protein Digestion

Three trypsin digest protocols were evaluated:

(1). Filter-Aided Sample Preparation (FASP) Digestion. Performed using the FASP protein digestion kit (Protein Discovery, San Diego, CA) according to manufacturer's instructions using 30 kDa cutoff spin filters. Ninety micrograms of total synovial fluid protein was digested overnight at 37 °C with 2 μg of sequencing grade modified trypsin (Promega, Fitchburg, MA). To assess the need of glycan removal when working with synovial fluid, 500 U peptide-*N*4-(*N*-acetyl-beta-glucosaminyl)-asparagine amidase (PNGase F) (New England BioLabs, Ipswich, MA) was added to these samples prior to the trypsin digestion step, and samples were incubated overnight at 37 °C, after which the normal FASP protocol was continued. After trypsin digestion, the samples were desalted with TARGA C18 columns (Nest Group, Southborough, MA) and resuspended in 5% acetonitrile (ACN) and 5% formic acid (FA) prior to analysis.

(2). Urea In-Solution Digestion. Performed according to Gallien et al.[31] Ninety micrograms of total synovial fluid protein was diluted with 8 M urea in 100 mM ammonium bicarbonate to a final volume of 25 μL. The sample was reduced with dithiothreitol at a final concentration of 12 mM for 30 min at 37 °C and alkylated with iodoacetamide at a final concentration of 40 mM for 1 h at room temperature in the dark. The samples were diluted with 100 mM ammonium bicarbonate to a total volume of 100 μL, 2 μg of trypsin was added, and the sample was digested overnight at 37 °C. The samples were desalted

with TARGA C18 columns (Nest Group) and resuspended in 5% ACN, 5% FA prior to analysis.

(3). In-Gel Digestion. Three gel-lanes, each loaded with 150 μg total synovial fluid protein, were divided into 10 sections each and subjected to standard in-gel tryptic digestion as previously described,[32−34] followed by analysis.

Human Plasma Protein Digestion

Human plasma was acquired as part of an ongoing method optimization study using a deidentified, discarded cord plasma sample and thus are not considered research of human subjects. One hundred micrograms of plasma protein was digested using the FASP protein digestion kit (Protein Discovery, San Diego, CA) with a modification to the recommended protocol. This modification involved the use of a 10 kDa MWCO filter instead of the stock 30 MWCO filters. A trypsin/LysC mix (Promega, Madison, WI) was added to the FASP filter at 1:25 ratio for digestion. The samples were then incubated at 37 °C overnight, and the resulting peptides were recovered as recommended by the manufacturer's protocol. The peptides were desalted using Oasis HLB columns (Waters, Milford, MA) and resuspended in 2% ACN in 0.1% FA prior to analysis.

LC−MS/MS Measurement and Proteomics Data Analysis

Two different high-resolution/high-accuracy mass spectrometer systems were used for the shotgun proteomic analysis: (1) For post-translational modification (PTM) analysis and method optimization, synovial fluid samples were analyzed on a TripleTOF 5600 (AB Sciex, Framingham, MA) connected online with a nanoflow UPLC and a NanoFlex system (Eksigent/AB Sciex). The samples were loaded onto a 15 cm reversed-phase C18 200 μm chip with 2 μL/min in 100% solvent A (0.1% FA). The samples were then separated using a 15 cm reversed-phase C18 75 μm chip and eluted with a linear gradient of 2% solvent B (0.1% FA in ACN), which was raised to 35% solvent B over 120 min (60 min for in-gel digested samples) at a constant flow rate of 500 nL/min. (2) The six FASP digested synovial fluid samples used to determine the synovial fluid protein list and abundances as well as the trypsin-digested human plasma samples were analyzed on a Q Exactive (Thermo Scientific, Waltham, MA) connected online to an EASY-nUPLC 1000 (Thermo Scientific). The samples were loaded onto a 10.5 cm reversed-phase C18 PicoChip with a flow rate of approximately 1 μL/min in 98% solvent A and 2% solvent B and were eluted with eluent B using a linear gradient that was raised to 35% over 120 min at a constant flow rate of 300 nL/min.

The AB Sciex.wiff data files were analyzed using ProteinPilot 4.5 (rev. 1656, Paragon Algorithm 4.5.0.0). To identify the most commonly single observed PTMs, data files were searched in thorough mode with a focus on biological modifications in ProteinPilot to include more than 300 different PTMs.

The .raw data files from synovial fluid and human plasma analyzed on the Q Exactive were searched using MaxQuant 1.4.1.2.[35] All standard settings were employed with carbamidomethyl (C) as a static modification and deamidation (NQR), oxidation (M), and protein N-terminal acetylation included as variable modifications. Label-free quantitation of all proteins was performed in MaxQuant based on integrated precursor intensities. Protein abundances are represented as protein intensity-based absolute quantitation values (iBAQ) and are reported for all proteins having at least two quantifiable unique peptides in at least three LC−MS runs.[36]

The human plasma sample was searched against the UniProt *Homo sapiens* reference proteome database with isoforms (downloaded 7/18/2014, containing 89 032 entries). The porcine synovial fluid data was searched against the UniProt *S. scrofa* reference proteome database (downloaded 11/09/2013, containing 26 070 entries). The human RA synovial fluid, used to evaluate the UniProt *S. scrofa* database, was searched against all reviewed *H. sapiens* UniProt proteins (downloaded 08/10/2013, containing 20 277 entries). All proteins and peptides are reported below a 1% false discovery rate (FDR) cutoff, and protein posterior error probability (PEP, equivalent to expectancy) was investigated to ensure only confident protein identifications.[35] For the PTM analysis, the search results were analyzed using ProteinPilot Descriptive Statistics Template, version 3.001, and for the protein abundance analysis, the iBAQ values were analyzed using Perseus, version 1.4.1.3, and IBM SPSS Statistics (version 21). Venn diagrams were created with BioVenn[37] and Venny.[38]

Assignment of Formerly Glycosylated Asparagine Residues

Four criteria were required to assign N-glycosylation sites: (I) a 1% FDR cutoff to all peptide spectral matches (PSMs); (II) all site assignments required the presence of a consensus site (CS) for N-glycosylation, i.e., NX(S/T), where X may be any amino acid except proline; (III) once CS status was established for all peptide assignments, an asparagine deamidation at the asparagine within the CS was required; and (IV), finally, all true site assignments were required to come from sample preparations that were treated with PNGase F. The FDR of site assignment was estimated by evaluation of the random rate of site assignment among control samples that were not treated with PNGase F. In this way, the rate of PSMs leading to the identification of a deglycosylated peptide may be compared.

RNA Extraction

Total RNA was extracted from frozen synovium tissue using the PureLink RNA mini kit (Ambion, Austin, TX), treated with PureLink DNase I (Life Technologies, Grand Island, NY) according to the manufacturer's protocol, and quantified. Briefly, frozen tissue samples were placed in tubes containing 5 metal lysing beads (Metal Bead Lysing Matrix, MP Biomedicals, Solon, OH) and 1 mL of TRIzol (Life Technologies). Tissue was homogenized using a FastPrep24 (MP Biomedicals, Solon, OH) at 6 m/s for 40 s. This was repeated two additional times, with samples briefly placed in liquid nitrogen between homogenization runs. Following total RNA extraction and quantitation via spectrophotometry using a Nanodrop 1000 (Thermo Scientific, Wilmington, DE), samples were submitted for quality analysis on a Bioanalyzer model 2100 (Agilent Technologies, Santa Clara, CA). Each RNA sample had an RIN > 7.20, indicating that they were of sufficient quality for sequencing.[39]

Library Preparation

A mRNA library for each tissue sample was prepared as previously described using the TruSeq RNA sample preparation kit, v2 (Illumina, San Diego, CA).[40] Briefly, mRNA was enriched from total RNA, chemically fragmented, reverse-transcribed using random hexamers, and ligated to barcoded adapters per the manufacturer's instructions. The cDNA fragments were amplified via PCR, and the libraries were washed with AMPure XP beads (Beckman Coulter Inc., Danvers, MA) to remove primer dimers. One microliter aliquots from each library were run on a 4−20% TBE gel (Life

dx.doi.org/10.1021/pr500587x | J. Proteome Res. 2014, 13, 4377−4387

Technologies, Grand Island, NY) for verification. Equal amounts of DNA from separately barcoded cDNA libraries were pooled (n = 12 per lane) and sent for 50 base pair paired-end sequencing on an HiSeq 2000 (Illumina).

RNA-Seq Data Analysis

Reads were mapped to the pig genome (Susscr3) using RUM.[41] Known genes were annotated with an R script based on data available from ENSEMBL. Expression level of each gene was quantified by the reads per kilobase of exon per million mapped reads (RPKM). The repeatability of the RNA-seq data was measured by comparing libraries generated from the right and left knees of the same animal and calculating the Pearson correlation coefficient (R^2) with respect to the genome-wide RPKM values. The calculated RPKM values were averaged across 12 samples for each gene, and these mean values were used in ranking the genes with respect to their expression levels.

In order to estimate the likely source of proteins detected in the synovial fluid, we additionally reviewed mouse liver,[42] mouse blood,[40] and human articular cartilage (authors' unpublished data) RNA-seq data. On the basis of homology information retrieved from the ENSEMBL database, we matched average RPKM values for each transcript with the proteins detected in the synovial fluid, and we identified the proteins for which all four mRNA expression values were available. We then determined the subset of proteins with a signal peptide. Among these proteins, we identified those with at least 2-fold higher mRNA levels in the synovium compared to all of the other individual tissues, and we then performed the same calculations for liver, cartilage, and blood.

RESULTS AND DISCUSSION

Optimizing Preparation and Digestion Methods for Synovial Fluid

Initially, we determined whether precipitating the synovial fluid proteins increased the diversity of the detectable protein bands by Coomassie-stained SDS-PAGE. However, similar to Chen et al.,[10] we found that precipitation appeared to reduce polypeptide diversity rather than increase it (data not shown); thus, we chose not to use precipitated samples for our subsequent analyses. Likewise, because albumin can function as a protein carrier for a wide range of proteins, we chose not to deplete albumin or other abundant proteins from our samples to avoid changing protein abundance profiles and removing important proteins that might bind the depletion targets.[1,10,43−45] Zhou et al.[46] reported that removing albumin led to the depletion of many proteins in human serum, including clinically useful biomarkers.

We compared urea in-solution digestion to FASP to find the most efficient digestion method without prior prefractionation for synovial fluid and used the robust and proven in-gel digestion as a reference. Urea in-solution digestion yielded the lowest number of 127 identified proteins. In contrast, FASP and in-gel-based digestion each yielded 227 protein identifications, indicating similar performance of the two digestion techniques. However, 170 of the 227 identified proteins were found in at least two of the three synovial fluid samples when using the FASP protocol in contrast to only 145 of the 227 proteins in the case of the in-gel digestion protocol (Figure 1). Furthermore, despite substantial overlap among the data sets, roughly 26% of proteins found in at least two of the three synovial fluid samples with FASP were missed by in-gel digestion, and 12% of the in-gel digestion proteins were missed

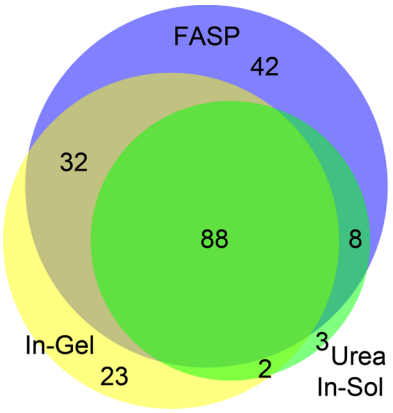


Figure 1. Digestion method evaluation. Number of proteins uniquely identified in at least two of the three synovial fluid samples, using FASP digestion, urea in-solution digestion, and in-gel digestion.

with FASP. Additionally, 17% of the proteins unique to the in-gel digestion were keratins, supporting the notion that in-gel digestion protocols can be prone to keratin contamination. Besides this difference, no other bias was found with regard to which groups of proteins were identified uniquely with the different digestion methods. This indicates that the FASP protocol gives more consistent data than the in-gel digestion protocol. It is possible that depletion strategies can increase the number of identified proteins further; however, this was not investigated in this study.

The UniProt *S. scrofa* reference proteome is not as well annotated as the human proteome. Therefore, in order to determine whether our proteomic method is as sensitive for detecting pig synovial fluid proteins as it would be for human synovial fluid proteins, we performed the same FASP digestion protocol on a synovial fluid sample from a RA patient. In the four technical replicates, we identified, on average, 173 different proteins in the human sample, which seems to be comparable to the 179 different proteins found, on average, in the individual porcine synovial fluid FASP LC−MS runs. The protein composition of synovial fluid is known to be altered during active joint diseases, so the protein overlap was not investigated.[3,17] Nonetheless, the similar number of identified proteins with the two databases indicates that the UniProt *S. scrofa* reference proteome is adequate for the analysis of the porcine synovial fluid.

Importance of Accounting for Post-Translational Modification in Protein Identification

Many PTMs occur in vivo or are introduced during sample preparation prior to MS analysis in vitro.[47−50] Therefore, to identify which PTMs should be considered for the database search of MS data from FASP-digested synovial fluid samples, we analyzed the identified peptides from this method in ProteinPilot Descriptive Statistics Template (Table 1). Several PTMs were identified, and all were likely artifacts from the sample preparation. Chloroacetamide can be used as an alternative alkylating agent to iodoacetamide to reduce the number of alkylation related artifacts.[47] However, due to the low number of detected peptides with these artificial modifications, this was not investigated further.

Synovial fluid is known to be rich in glycoproteins, which prompted us to evaluate the effect of removing N-linked glycans on the number of identifiable peptides and proteins.[2] We compared two synovial fluid samples using the FASP

dx.doi.org/10.1021/pr500587x | J. Proteome Res. 2014, 13, 4377−4387

Table 1. Most Commonly Observed PTMs of the FASP-Digested Synovial Fluid Samples Based on the Most Frequent Single Features

rank	modification and position (N-terminal; C-terminal; acceptor amino acid)	modification mass (Da)	no. MS/MS events	modified sites of possible (%)
1	Carbamidomethyl (C)	57.0215	13 201	100.0
2	Deamidated (N)	0.9840	1024	8.7
3	Acetyl at N-term	42.0106	563	2.9
4	Carbamidomethyl at N-term	57.0215	535	2.7
5	Glu → pyro-Glu at N-term	−18.0106	522	24.1
6	Deamidated (Q)	0.9840	440	2.9
7	Carbamyl (K)	43.0058	434	2.0
8	Carbamidomethyl (K)	57.0215	420	1.9
9	Carbamyl at N-term	43.0058	416	2.1

protocol, with and without deglycosylation by PNGase F (±). As expected, PSMs resulting in identification of deglycosylated peptides considered to be previously N-glycosylated was increased ∼40-fold in the PNGase F-treated samples in comparison to that of the untreated samples (Table 2).

Table 2. Summary of Overall Numbers of PSMs in the Deglycosylation Experiments

sample	total no. of PSMs	PSM identifying deglycosylated peptides	deglycosylated peptide PSMs/all PSMs (%)
+ PNGase F 1	4827	91	1.9
+ PNGase F 2	4878	86	1.8
− PNGase F 1	4068	3	0.07
− PNGase F 2	4372	2	0.05

Twenty-eight and 34 unique deglycosylated peptides in the biological repeats were identified using the four criteria described in the Experimental Procedures section, compared to only two peptides in each of the untreated samples (Supporting Information Table 1). However, all deglycosylated peptides belonged to proteins that had already been identified from other peptides and thus resulted in no new protein identifications. Twenty-eight and 34 unique deglycosylated peptides are relatively low numbers for an analysis of a biological fluid known to be rich in glycoproteins, which may be due to a number of factors: First, the proteome of synovial fluid may not be particularly complex. This hypothesis is consistent with the proteomics data we observed in this study and with previously reported 2D-PAGE analyses.[1,10,17] Second, N-glycosylation, which is the only glycosylation amenable to PNGase F treatment protocols, may not be highly prevalent in synovial fluid. Synovial fluid contains significant amounts of glycosaminoglycans, possibly lessening the biologic necessity for N-glycosylation.[51,52] Third, the finite dynamic range of the MS analysis might cause abundant peptides to hinder the detection of less abundant ones. A study that exclusively focuses on O-glycosylation in synovial fluid and combines data from depleted and nondepleted samples will likely identify more extensive glycosylation than that reported here.

Identifying the Transcriptome of Healthy Pig Synovium

In order to identify synovial fluid proteins that are secreted by the synovium cells, i.e., transcribed and translated, we generated sequencing data of mRNA as active transcripts for pig synovium. We prepared and sequenced knee synovium from the right and left legs of 6 pigs. On average, we obtained 12.4 million sequencing reads per library, of which 90% mapped to the pig genome and 81% aligned uniquely. We then calculated RPKM values for individual genes and compared their expression level within each individual animal's left and right knee (Figure 2a). Intra-animal Pearson correlation coefficients exceeded 0.85, indicative of reproducible data (Figure 2d). The top 20 abundant protein-coding mRNAs that did not originate from the ribosome or mitochondria exhibited minimal variation in abundance, as they ranked in the top 1% of all sample-specific RNA-seq data sets that we analyzed (Figure 2c).

Focusing on transcripts encoding proteins containing signal peptides, which are most likely secreted or targeted to cell surfaces, we observed high expression of known connective tissue proteins such as decorin (DCN), fibronectin (FN1), collagen type III (COL3A1), and clusterin (CLU) (Figure 2b). Importantly, the principal lubricating protein in synovial fluid, lubricin (PRG4) made by type B synoviocytes, was also highly expressed. From these data, we produced a pig synovium transcriptome database (Supporting Information Table 2), which is composed of transcripts encoding known and predicted proteins.

Identifying the Proteome of Healthy Pig Synovial Fluid

We opted for the FASP digestion strategy for the detailed triplicate proteomic analysis of synovial fluid from 6 healthy pigs. This choice was based on the number of identified proteins, sample amount requirement, processing time, and instrument time. We identified 374 different proteins in total using the UniProt *S. scrofa* database and the synovium transcriptome database, and 42 proteins were identified solely in the transcriptome database. This seemed to be a large fraction assuming an adequate performance of the UniProt pig database, so all identified proteins were sorted by PEP (equivalent to expectancy[35]), with a high PEP value indicating a lower confidence in the identification (Figure 3a). Proteins unique to the transcriptome database displayed high PEP values, indicating that the number of identified proteins is too low for the global FDR calculation. To address this issue, all proteins with PEP > 1×10^{-3} were removed from further analysis, which came to 92 identified proteins, of which 36 were unique to the transcriptome database. Furthermore, we manually assessed the fragment spectra for proteins identified based on a single peptide and removed poor spectra with many unassigned high-intensity signals, which resulted in one additional protein being removed. After these two filtering steps, 267 proteins remained, of which 6 were unique to the transcriptome database (Figure 3b). Five of the 6 unique proteins were immunoglobulin lambda like proteins (UPID: IGLV-7 to IGLV-11), and the identified peptides originate from a variable region of the lambda chain, which is not annotated in the UniProt database. The remaining protein, ribonuclease 4 (UPID: I3LDZ2_PIG), was present in the UniProt database, but it had not been identified. We included several keratins (10, 14, 3, and 75) in the synovial fluid proteome because keratin-coding mRNA was detected in the articular cartilage. However, we cannot preclude some of the keratins being contaminants from sample preparation.

Of the identified 267 different proteins in synovial fluid (Supporting Information Table 2), 194 proteins (73%) were identified in all pigs, indicating similar protein expression patterns between the pigs and a high method robustness. The

dx.doi.org/10.1021/pr500587x | J. Proteome Res. 2014, 13, 4377−4387

(a)

Library #	Number of Reads	Mapped (%)	Uniquely Mapped (%)
P1	11,040,002	92.3	83.4
P2	13,478,370	87.6	79.1
P3	11,536,658	89.9	81.1
P4	15,433,739	92.9	83.6
P5	11,443,699	87.6	79.0
P6	14,047,352	92.3	83.5
P7	13,243,709	84.2	76.0
P8	11,825,018	92.6	83.6
P9	11,046,003	89.1	79.9
P10	11,341,586	90.5	81.1
P11	13,169,629	92.2	83.8
P12	11,402,283	90.5	81.7
Average	12,417,337	90.1	81.3

(b)

Gene Symbol	Ensembl Symbol	Mean RPKM
EF1ALPHA	ENSSSCG00000004489	7138
DCN	ENSSSCG00000000915	6609
UBC	ENSSSCG00000029226	4898
CYTB	ENSSSCG00000018094	4684
FN1	ENSSSCG00000016174	4519
COL3A1	ENSSSCG00000016034	4141
CLU	ENSSSCG00000009668	4050
APOR	ENSSSCG00000015663	3910
uncharacterized gene	ENSSSCG00000000207	3501
MGP	ENSSSCG00000000606	3431
CST3	ENSSSCG00000029889	3268
PRG4	ENSSSCG00000029449	2940
FTH1	ENSSSCG00000014540	2796
TMSB10	ENSSSCG00000008245	2683
VIM	ENSSSCG00000011033	2547
FTL	ENSSSCG00000003153	2532
uncharacterized gene	ENSSSCG00000004579	2334
C1QC	ENSSSCG00000003525	2309
COL1A1	ENSSSCG00000017581	2297
S100A6	ENSSSCG00000021084	1933

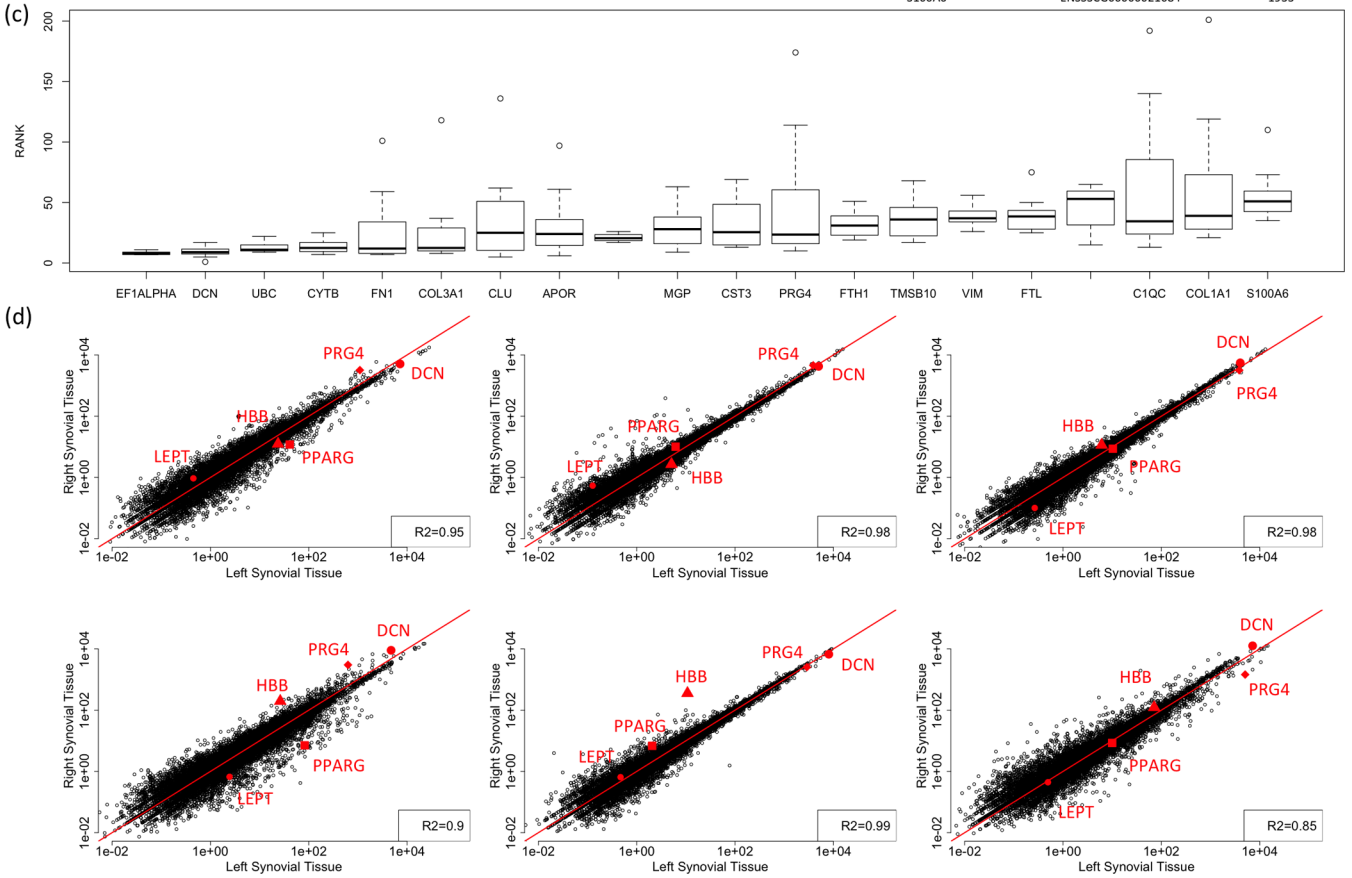


Figure 2. RNA-seq of pig synovium is repeatable and indicates high expression of several transcripts encoding secreted proteins. (a) Table indicating the total number of reads and percentages of successfully mapped reads. More than 90% of reads on average are mapped to the pig genome (Susscr3), and more than 80% of reads are mapped uniquely. (b) Top 20 protein coding genes (based on expression level measured with RPKM) that do not originate from the mitochondria or the ribosome. (c) Box plot indicating the variation in the expression-based rank order of genes listed in (b) in all 12 libraries analyzed. The highest variation is in the rank of PRG4; however, it remains within the top 1% of genes in all libraries. (d) Scatterplots indicating high similarity ($R^2 > 0.85$) between the synovial transcriptomes of paired left and right legs of all pigs. Data indicate RPKM, and each dot represents a single gene. Transcripts for the secreted proteins decorin (DCN) and lubricin (PRG4), red blood cell-derived beta hemoglobin (HBB), an adipocyte-derived transcription factor (PPARG), and secreted protein leptin (LEPT) are indicated with red colored symbols and closely follow the $y = x$ line (solid red line) in each plot. Importantly, transcripts for matrix-degrading enzymes and inflammatory cytokines are not abundant, clustering in the bottom left corner of the panels.

difference in identified proteins between the pigs likely originates from biological diversity, which is to be expected, and the finite dynamic range of the MS analysis.[53] To investigate the relative protein abundances in synovial fluid, we investigated individual synovial fluid protein abundance by calculating iBAQ values.[36] The protein abundance in synovial fluid, as estimated by the 203 iBAQ quantifiable proteins, spans 5 orders of magnitude (Figure 4). As expected, albumin was the most abundant protein, and no apparent bias could be found regarding the abundances of serum-derived proteins (RPKM < 5) compared to proteins expressed by the synovium in the joint (RPKM > 5) (see below).

Estimating the Relative Contributions of Plasma, Synovium, and Cartilage to the Synovial Fluid Proteome

Synovial fluid is a transudate of blood that has additional components added by the surrounding joint tissue including synovium, cartilage, and, in the knee, ligaments and menisci. To determine whether a likely tissue source for a synovial fluid protein can be predicted, we compared protein expression with

dx.doi.org/10.1021/pr500587x | J. Proteome Res. 2014, 13, 4377−4387

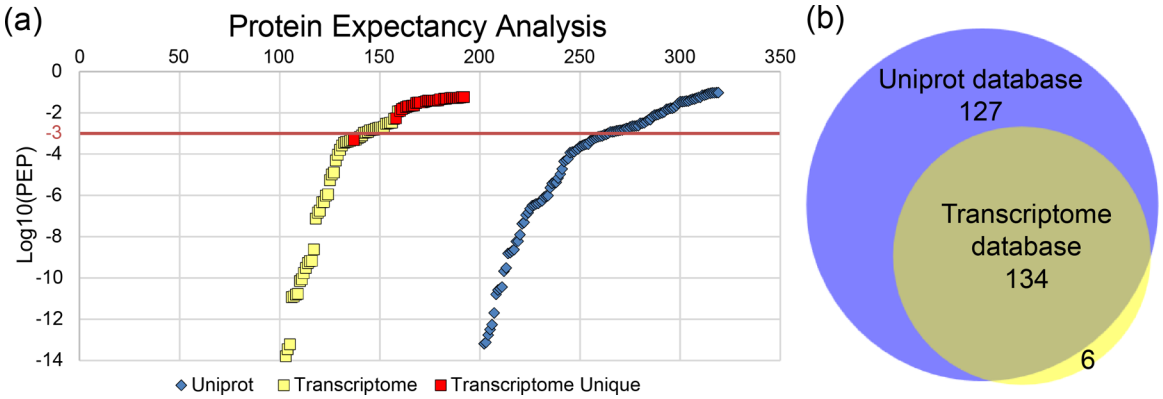


Figure 3. Analysis of identified synovial fluid proteins. (a) Identified proteins at 1% FDR sorted by PEP. Proteins unique to the transcriptome database (red squares) were, for the most part, identified with low confidence; hence, all proteins with PEP > 1×10^{-3} (red line) were removed from further analysis. (b) Synovial fluid proteins identified using the UniProt database and the transcriptome database fulfilling the filtering criteria. Of the 267 synovial fluid proteins, 140 had RPKM > 5 in the synovial mRNA libraries, indicating they are expressed by synovium.

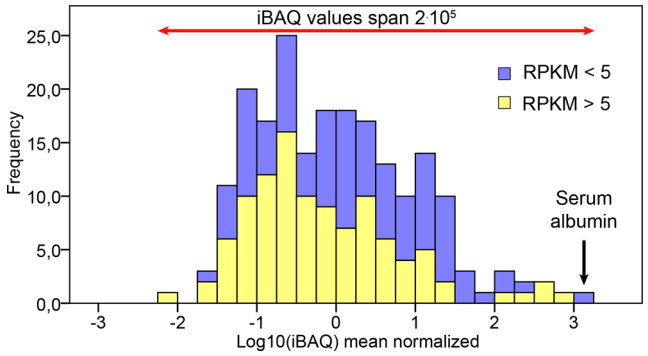


Figure 4. Protein abundance distribution in synovial fluid. Frequency histogram of the calculated iBAQ values for all proteomics quantifiable proteins. The relative proportion of proteins whose mRNA transcript abundances were less than or greater than 5 RPKM in the synovial membrane transcriptome are colored blue and yellow, respectively.

the pig synovium transcriptome, a human articular cartilage transcriptome we previously generated, and published mouse liver[42] and blood[40] transcriptomes (Supporting Information Table 2). As expected, albumin in synovial fluid solely derives from blood, as its transcript is highly expressed only in liver (Figure 5).

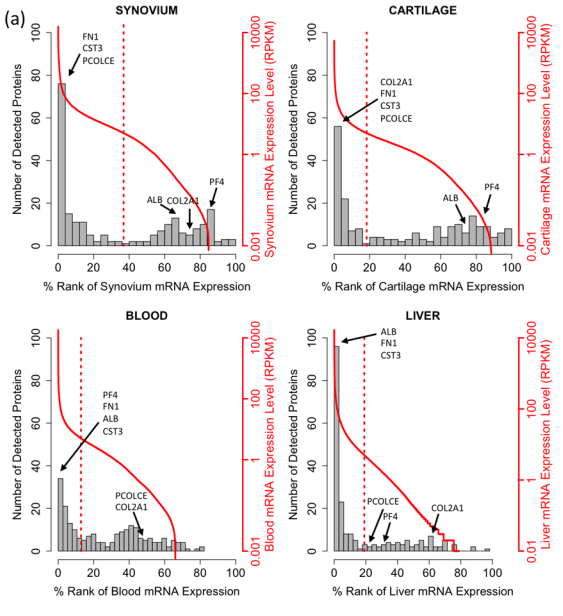
When we compared the proteins detected in the synovial fluid with the mRNA expression levels in four possible sources, we found that the majority of proteins corresponded to highly expressed transcripts regardless of the source and that liver and synovium likely made the greatest contribution (Figure 5a). Further analysis of our data revealed that 171 proteins detected in the synovial fluid had a signal peptide and that liver, blood, and articular cartilage mRNA expression data were available for 149 proteins (Supporting Information Table 2). Of these 149 common proteins in the synovial fluid proteome, 113 most likely derive from plasma transudate since their transcript abundance is 2-fold or higher in liver or blood than that in synovium or cartilage. In contrast, 37 synovial fluid proteins have transcripts whose expression in synovium is at least 2-fold higher than that in other tissues, while 11 proteins have at least 2-fold or higher transcript expression in articular cartilage. Several proteins serve as positive controls in validation of our estimations. For example, aggrecan, type II collagen, cartilage oligomeric matrix protein, and cartilage intermediate layer protein 2 are detected in the synovial fluid, and the corresponding transcripts are expressed at greater levels in articular cartilage than in other tissues. Similarly, proteins involved in post-translational modification of type I collagen (such as procollagen C-endopeptidase enhancer) and connective tissue markers (such as fibronectin 1 and clusterin) were expressed at higher levels in synovium than in other tissues. Thus, the putative origins of a substantial number of synovial proteome constituents can be predicted from the RNA sequencing data. For those 17 proteins whose mRNAs are expressed by multiple tissues, tissue-specific changes in mRNA abundance will need to be correlated with changes in synovial protein abundance to reliably determine their most likely tissue of origin.

We have additionally detected several proteins without a signal peptide, which are likely residual fragments that remained in the synovial fluid following apoptosis. Some examples of these are hemoglobin (HBA and HBB), which are expressed at extremely higher levels in blood than the remaining tissues, and beta-actin, a ubiquitously expressed protein whose mRNA was detected at high levels (RPKM > 180) in all four tissues. Also of interest are proteins not found in healthy pig synovial fluid and expressed at low levels in healthy synovium. Increased amounts of matrix metalloproteinases, including MMP13 and ADAMTS4, have been reported in synovial fluid from humans with OA, while proteins involved in the inflammatory/immune response, including complement activation, were found in patients with RA.[1] These proteins were neither detected in healthy pig synovial fluid nor were their transcripts abundant in healthy synovium. Therefore, in addition to monitoring changes in protein abundance for the healthy synovial fluid proteome, it will be of interest to compare proteomes and transcriptomes of healthy and diseased joints to identify pathways that may be integral to disease processes or mechanisms. Previous comparisons of synovial fluid obtained from patients with RA and OA identified 135 proteins at least 3-fold differentially abundant between the two groups.[20]

Comparison of Synovial Fluid and Plasma

Finally, to identify similarities between synovial fluid and blood plasma, from which many of the synovial fluid components were found to originate, we characterized the human plasma proteome (Supporting Information Table 3) and identified and quantified a total of 168 human plasma proteins. The detected iBAQ values of the human plasma proteins span 5 orders of

dx.doi.org/10.1021/pr500587x | J. Proteome Res. 2014, 13, 4377−4387



(b)

Ensembl Symbol	Gene Symbol	Description	Protein Expression (iBAQ)	Synovium mRNA Expression (RPKM)	Blood mRNA Expression (Ayturk et al., 2013)	Liver mRNA Expression (Mortazavi et al., 2008)	Articular Cartilage mRNA Expression (unpublished)	Secreted
ENSSSCG00000016174	FN1	Fibronectin 1	1.15E+08	4518.64	141.00	150.81	740.07	YES
ENSSSCG00000009668	CLU	Clusterin	4.20E+08	4050.09	20.26	803.91	306.74	YES
ENSSSCG00000015663	APOR	Apolipoprotein R precursor	5.87E+07	3910.00		not conserved		YES
ENSSSCG00000029889	CST3	Cystatin-C	8.83E+06	3268.22	59.87	135.88	93.31	YES
ENSSSCG00000029449	PRG4	Lubricin	1.73E+07	2940.02	34.45	not annotated	243.53	YES
ENSSSCG00000011033	VIM	Vimentin	2.17E+06	2547.27	59.74	8.81	459.73	NO
ENSSSCG00000003525	C1QC	Complement C1q subcomponent subunit C precursor	1.14E+07	2309.38	12.67	35.66	16.78	YES
ENSSSCG00000013907	UBAS2	Ubiquitin-60S ribosomal protein L40	2.12E+07	1884.50	13.89	87.10	43.51	NO
ENSSSCG00000003524	C1QA	Complement C1q subcomponent subunit A precursor	2.61E+06	1775.73	23.29	37.56	27.57	YES
ENSSSCG00000022236	FOLR1	Folate receptor 1	5.71E+06	1773.75	0.00	0.09	0.33	YES
ENSSSCG00000012119	TMSB4	Thymosin beta-4 Hematopoietic system regulatory peptide	7.02E+07	1519.66		not conserved	398.11	NO
ENSSSCG00000000916	LUM	Lumican	4.14E+07	1267.27	0.31	3.12	1314.07	YES
ENSSSCG00000013181	SERPING1	Plasma protease C1 inhibitor	1.59E+07	1080.94	1.20	392.03	35.22	YES
ENSSSCG00000027466	PCOLCE	Procollagen C-endopeptidase enhancer	3.41E+06	1022.94	0.11	3.09	80.54	YES
ENSSSCG00000007585	ACTB	Actin, beta	1.29E+08	945.24	197.39	180.25	290.65	NO
ENSSSCG00000016737	PPIA	Peptidyl-prolyl cis-trans isomerase A	8.20E+06	938.89	2.73	31.87	39.62	NO
ENSSSCG00000000694	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	6.64E+06	867.44	0	0.39	not conserved	NO
ENSSSCG00000026721	RNASE4	Ribonuclease 4 precursor	2.91E+06	845.00	5.66	232.03	19.22	YES
ENSSSCG00000028674	CR2	Complement receptor 2	7.77E+05	804.01	6.70	0.05	0.00	YES
ENSSSCG00000021208	SEPP1	Selenoprotein P, plasma 1	5.95E+06	801.43	22.05	2816.57	32.64	NO

(c)

Ensembl Symbol	Gene Symbol	Description	Protein Expression (iBAQ)	Synovium mRNA Expression (RPKM)	Blood mRNA Expression (Ayturk et al., 2013)	Liver mRNA Expression (Mortazavi et al., 2008)	Articular Cartilage mRNA Expression (unpublished)	Secreted
ENSSSCG00000008948	ALB	Serum albumin precursor	3.68E+10	0.14	75.46	33249.73	0.05	YES
ENSSSCG00000027332	IGKV-11	Immunoglobulin kappa variable 6-21 orthologue	2.50E+09	1.97	not conserved		0	YES
ENSSSCG00000014626	HPX	Hemopexin	2.21E+09	0.73	10.6	2087.17	0.43	YES
ENSSSCG00000011799	AHSG	Alpha-2-HS-glycoprotein	1.87E+09	0	14.50	6224.46	0	YES
ENSSSCG00000002486	SERPINA3-1	Serpin peptidase inhibitor, clade A orthologue	7.48E+08	0	0.79	154.34	9.21	YES
ENSSSCG00000002749	HP	Haptoglobin	6.89E+08	0.64	37.16	3174.88	0.02	YES
ENSSSCG00000000660	A2M	Alpha-2-macroglobulin	6.84E+08	1.65	0.01	0.24	37.1	YES
ENSSSCG00000000660	A1BG	Alpha-1B-glycoprotein	6.75E+08	0.03	0	0.06	0.09	YES
ENSSSCG00000023686	TTR	Transthyretin	5.63E+08	0	12.05	5867.50	0.02	YES
ENSSSCG00000030371	Novel Gene	Serpin peptidase inhibitor, clade A orthologue	5.59E+08	0.01	0.79	154.34	9.21	YES
ENSSSCG00000027609	GC	group-specific component (vitamin D binding protein) orthologue	5.09E+08	0	7.95	3698.15	0	NO
ENSSSCG00000006355	APOA2	Apolipoprotein A-II	5.04E+08	0.19	10.23	12696.14	0.01	YES
ENSSSCG00000011801	HRG	Histidine-rich glycoprotein	4.92E+08	0	0.94	739.65	0.01	YES
ENSSSCG00000011700	CP	Ceruloplasmin (ferroxidase)	3.86E+08	2.12	0.34	200.59	2.63	YES
ENSSSCG00000005488	Novel Gene	Orosomucoid orthologue	3.67E+08	0	2.09	413.1	0.06	YES
ENSSSCG00000002476	SERPINA1	Alpha-1-antitrypsin precursor	3.05E+08	0.12	0	0.83	8.46	YES
ENSSSCG00000015493	SERPINC1	Antithrombin-III	2.68E+08	0.92	2.67	991.18	0.07	YES
ENSSSCG00000027290	Novel Gene	Apolipoprtoein H orthologue	2.52E+08	0	2.98	2038.85	0	YES
ENSSSCG00000002487	SERPINA3-2	Alpha-1-antichymotrypsin 2	2.47E+08	0.01	0.79	154.34	9.21	YES
ENSSSCG00000015069	APOC3	Apolipoprotein C-III	2.37E+08	0	3.18	2572.71	0	YES

Figure 5. mRNA expression levels associated with the proteins detected with MS. (a) Tissue-specific histograms indicating the number of detected proteins corresponding to each transcript. Genes are ranked based on mRNA expression along the *x* axis, and the secondary *y* axis (in red) indicates the expression levels in RPKM (in log scale). (b) Table indicating the top 20 mRNAs in synovium with detectable protein products in synovial fluid. Gene expression data derived from published and unpublished studies on tissues with potential influence on protein expression repertoire of serum and synovial fluid are also presented in the three columns on the right. Blood and liver mRNA data originate from mice, and articular cartilage mRNA data originate from humans. (c) Table indicating the top 20 proteins with no detectable mRNA expression in the synovium: note that the majority of mRNA species corresponding to these proteins are abundantly expressed by liver cells.

Table 3. Top Ten Most Abundant Proteins in Human Plasma and Porcine Synovial Fluid Ranked from the Most Abundant[a]

rank	human plasma	human UPID	porcine synovial fluid	human UPID
1	Serum albumin	P02768	Serum albumin	P02768
2	Ig gamma-1 chain C region	P01857	Hemoglobin subunit alpha	P69905
3	Ig kappa chain C region	P01834	Ig lambda chain C region	P0CG04
4	Apolipoprotein A-I	P02647	Hemoglobin subunit beta	P68871
5	Ig lambda-2 chain C regions	P0CG05	Serotransferrin	P02787
6	Ig gamma-2 chain C region	P01859	Immunoglobulin lambda variable 8-61	Q5NV62
7	Serotransferrin	P02787	Immunoglobulin kappa variable 6−21	P01834
8	Alpha-1-antitrypsin	P01009	Apolipoprotein A-I	P02647
9	Apolipoprotein A-II	P02652	Hemopexin	P02790
10	Alpha-2-macroglobulin	P01023	Alpha-2-HS-glycoprotein	P02765

[a]As determined by iBAQ values.

magnitude, similar to what was found for the synovial fluid, demonstrating the large degree of protein abundance variation encountered in the two body fluids. However, the actual range of protein abundances in plasma and synovial fluid is expected to be several orders of magnitude larger, due to the limited dynamic range of the MS analysis.[54] Probing the top 10 most abundant proteins in both fluids, the most abundant protein is serum albumin, as expected (Table 3). Other shared high-abundance proteins include serotransferrin, apolipoprotein A, and several immunoglobulins. Serotransferrin and apoliopoprotein A are binding and transporting proteins of iron and lipids, respectively.[55,56] The high abundance of the transporter proteins and immunoglobulins points to the shared functions of the two body fluids as a transport media of cellular products. The main difference between the high-abundance proteins in the plasma and synovial fluid proteomes is the high abundance of hemoglobin in the synovial fluid, which is detected with a lower abundance in plasma (rank 25). In both fluids, the protein is a likely contaminant from red blood cells prior to fluid centrifugation. Fibrinogens, involved in blood clotting, were identified with high abundance in plasma in contrast to that in synovial fluid, as expected.

Our findings demonstrate the high degree of similarity shared by the synovial fluid proteome and the plasma proteome as well as the serum-derived nature of many synovial fluid components. The expected high degree of similarity found between human plasma and porcine synovial fluid, furthermore, indicates that porcine model systems of synovial joints are suitable for studies focusing on human joint diseases.[57]

■ CONCLUSIONS

We evaluated different proteomic strategies for analyzing healthy synovial fluid. We utilized porcine synovial fluid because this animal is used to model human joint disease. In this study, we developed a FASP-based analysis pipeline that consumes only 0.1% of the starting sample material and 10% of the MS instrument time used by, e.g., Balakrishnan;[20] thus, our method lends itself to clinical proteomics studies on larger

dx.doi.org/10.1021/pr500587x | J. Proteome Res. 2014, 13, 4377−4387

cohorts. Applying this fast and efficient workflow, we completed a normative proteomics study of synovial fluid from healthy knee joints and identified more than 250 proteins using very stringent identification criteria. Preprocessing with PNGase F to remove N-glycosylation was not found to be essential for the proteomic analysis of synovial fluid. Through an analysis of high-occurrence peptide PTMs, we found that proteomics projects using the FASP protocol for synovial fluid analysis, and possibly for other biological samples as well, should include peptide N-terminal acetylation, deamidation of asparginine and glutamine, and peptide N-terminal pyro-glutamic acid in the data analysis for comprehensive peptide identification.

Going beyond the conventional LC−MS-based proteome mapping, we also classified and predicted the likely tissue origins for the majority of the detected proteins in healthy synovial fluid by cross-referencing the proteome with the RNA transcriptomes from synovium, cartilage, blood, and liver. Although many proteins derive from plasma transudate, as expected, an important subfraction appears to be solely expressed by synovium or cartilage. Changes in the abundance of these latter proteins in synovial fluid or in blood may be useful biomarkers of disease onset or progression. For synovial fluid proteins whose transcripts are expressed in multiple tissues, it will be necessary to correlate tissue-specific changes in mRNA expression with changing synovial fluid protein abundance in order to determine these proteins' principal tissues of origin. Finally, as many proteins were found to likely originate from plasma, we compared the identified porcine synovial fluid proteome to the proteome of human plasma. We found that the two body fluids share many similarities in terms of protein functions and localizations, underlining the detected plasma derived nature of many synovial fluid components.

Our methods and the resulting proteome and transcriptome data sets will be useful when comparing porcine synovial fluid in healthy and diseased states. Importantly, the high homology between pigs and humans should make these methods and data sets valuable for human studies.

■ ASSOCIATED CONTENT

Supporting Information

Table 1 contains all identified glycosylated peptides. Table 2 contains all identified proteins, iBAQ and RPKM values (when available), sequence coverage, and imported transcriptome information from associated mouse tissues. Table 3 contains all identified human plasma proteins. Table 4 contains additional information for all proteins identified from a single peptide. This material is available free of charge via the Internet at http://pubs.acs.org. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository[58−61] with the data set identifier PXD000935.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +1-617-919-2629. Fax: +1-617-730-0148. E-mail: hanno.steen@childrens.harvard.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This investigation was supported by National Institutes of Health under NIAMS AR054099 (M.M.M.) and AR050180 (M.L.W.) and by the Ruth L. Kirschstein National Research Service Award (F32 AR061186) (C.M.H.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The overseas collaboration was supported by the Knud Hojgaards Foundation Denmark, the Lundbeck Foundation Denmark, the Oticon Foundation Denmark, and the Otto Monsteds Foundation Denmark. The authors would like to thank the ARCH staff, Dr. Arthur Nedder, Kathryn Mullen, Dana Bolgen, and Courtney White for their assistance and care in handling the minipigs, Saima Ahmed, Nino Esile, and Kevin Broadbelt for assistance in the proteomics laboratory, and Lau Sennels for help with the proteomics data analysis.

■ ABBREVIATIONS

2D-PAGE, two-dimensional polyacrylamide gel electrophoresis; ACN, acetonitrile; CLU, clusterin; COL3A1, collagen type III; CS, consensus site; DCN, decorin; FA, formic acid; FASP, filter-aided sample preparation; FDR, false discovery rate; FN1, fibronectin; HBA, alpha hemoglobin; HBB, beta hemoglobin; iBAQ, intensity-based absolute quantitation; LC, liquid chromatography; LEPT, secreted protein leptin; MALDI, matrix-assisted laser desorption/ionization; MS, mass spectrometry; OA, osteoarthritis; PEP, posterior error probability; PNGase F, peptide-*N*4-(*N*-acetyl-beta-glucosaminyl) asparagine amidase; PPARG, adipocyte-derived transcription factor; PRG4, lubricin; PSM, peptide spectral matches; PTM, post-translational modification; RA, rheumatoid arthritis; RPKM, reads per kilobase of exon per million mapped reads; SDS, sodium dodecyl sulfate; TOF, time-of-flight

■ REFERENCES

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

(2) Swann, D. A.; Hendren, R. B.; Radin, E. L.; Sotman, S. L.; Duda, E. A. The lubricating activity of synovial fluid glycoproteins. *Arthritis Rheum.* **1981**, *24*, 22−30.

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

(4) Schmid, K.; Macnair, M. B. Characterization of the proteins of certain postmortem human synovial fluids. *J. Clin. Invest.* **1958**, *37*, 708−718.

(5) Schmid, K.; Macnair, M. B. Characterization of the proteins of human synovial fluid in certain disease states. *J. Clin. Invest.* **1956**, *35*, 814−824.

(6) Ruiz-Romero, C.; Calamia, V.; Carreira, V.; Mateos, J.; Fernández, P.; Blanco, F. J. Strategies to optimize two-dimensional gel electrophoresis analysis of the human joint proteome. *Talanta* **2010**, *80*, 1552−1560.

(7) Ropes, M. W.; Rossmeisl, E. C.; Bauer, W. The origin and nature of normal human synovial fluid. *J. Clin. Invest.* **1940**, *19*, 795−799.

(8) Platt, D.; Pigman, W.; Holley, H. L.; Patton, F. M. An electrophoretic study of normal and post-mortem human and bovine synovial fluids. *Arch. Biochem. Biophys.* **1956**, *64*, 152−163.

(9) Levick, J. R. Permeability of rheumatoid and normal human synovium to specific plasma proteins. *Arthritis Rheum.* **1981**, *24*, 1550−1560.

(10) Chen, C.; Hsu, C. C.; Yeh, W. L.; Lin, H. C.; Hsieh, S. Y.; Lin, S. C.; Chen, T. T.; Chen, M.; Tang, S. Optimizing human synovial fluid preparation for two-dimensional gel electrophoresis. *Proteome Sci.* **2011**, *9*, 65.
(11) Murphy, L.; Helmick, C. G. The impact of osteoarthritis in the United States: a population-health perspective. *Am. J. Nurs.* **2012**, *112*, S13−S19.
(12) Lawrence, R. C.; Felson, D. T.; Helmick, C. G.; Arnold, L. M.; Choi, H.; Deyo, R. A.; Gabriel, S.; Hirsch, R.; Hochberg, M. C.; Hunder, G. G.; et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: part II. *Arthritis Rheum.* **2008**, *58*, 26−35.
(13) Dragoo, J. L.; Braun, H. J.; Durham, J. L.; Chen, M. R.; Harris, A. H. S. Incidence and risk factors for injuries to the anterior cruciate ligament in national collegiate athletic association football data from the 2004−2005 through 2008−2009 National Collegiate Athletic Association Injury Surveillance System. *Am. J. Sports Med.* **2012**, *40*, 990−995.
(14) Brown, T. D.; Johnston, R. C.; Saltzman, C. L.; Marsh, J. L.; Buckwalter, J. A. Posttraumatic osteoarthritis: a first estimate of incidence, prevalence, and burden of disease. *J. Orthop. Trauma* **2006**, *20*, 739−744.
(15) Mobasheri, A. Osteoarthritis year 2012 in review: biomarkers. *Osteoarthritis Cartilage* **2012**, 1−14.
(16) Ritter, S. Y.; Subbaiah, R.; Bebek, G.; Crish, J.; Scanzello, C. R.; Krastins, B.; Sarracino, D.; Lopez, M. F.; Crow, M. K.; Aigner, T.; et al. Proteomic analysis of synovial fluid from the osteoarthritic knee: comparison with transcriptome analyses of joint tissues. *Arthritis Rheum.* **2013**, *65*, 981−992.
(17) Gobezie, R.; Kho, A.; Krastins, B.; Sarracino, D. A.; Thornhill, T. S.; Chase, M.; Millett, P. J.; Lee, D. M. High abundance synovial fluid proteome: distinct profiles in health and osteoarthritis. *Arthritis Res. Ther.* **2007**, *9*, 36.
(18) Chiaradia, E.; Pepe, M.; Tartaglia, M.; Scoppetta, F.; D'Ambrosio, C.; Renzone, G.; Avellini, L.; Moriconi, F.; Gaiti, A.; Bertuglia, A.; et al. Gambling on putative biomarkers of osteoarthritis and osteochondrosis by equine synovial fluid proteomics. *J. Proteomics* **2012**, *75*, 4478−4493.
(19) Ruiz-Romero, C.; Blanco, F. J. Proteomics role in the search for improved diagnosis, prognosis and treatment of osteoarthritis. *Osteoarthritis Cartilage* **2010**, *18*, 500−509.
(20) Balakrishnan, L.; Bhattacharjee, M.; Ahmad, S.; Nirujogi, R. S.; Renuse, S.; Subbannayya, Y.; Marimuthu, A.; Srikanth, S. M.; Raju, R.; Dhillon, M. Differential proteomic analysis of synovial fluid from rheumatoid arthritis and osteoarthritis patients. *Clin. Proteomics* **2014**, *11*, 1.
(21) Smith, M. A.; Bains, S. K.; Betts, J. C.; Choy, E. H. S.; Zanders, E. D. Use of two-dimensional gel electrophoresis to measure changes in synovial fluid proteins from patients with rheumatoid arthritis treated with antibody to CD4. *Clin. Diagn. Lab. Immunol.* **2001**, *8*, 105−111.
(22) Yang, C.; Li, L.; Xue, Y.; Zhao, Z.; Zhao, T.; Jia, Y.; Rong, R.; Xu, M.; Nicholson, M. L.; Zhu, T. Innate immunity activation involved in unprotected porcine auto-transplant kidneys preserved by naked caspase-3 siRNA. *J. Transl. Med.* **2013**, *11*, 210.
(23) Maeyama, A.; Hoshino, Y.; Debandi, A.; Kato, Y.; Saeki, K.; Asai, S.; Goto, B.; Smolinski, P.; Fu, F. H. Evaluation of rotational instability in the anterior cruciate ligament deficient knee using triaxial accelerometer: a biomechanical model in porcine knees. *Knee Surg., Sports Traumatol., Arthrosc.* **2011**, *19*, 1233−1238.
(24) Wei, L.; Fleming, B. C.; Sun, X.; Teeple, E.; Wu, W.; Jay, G. D.; Elsaid, K. A.; Luo, J.; Machan, J. T.; Chen, Q. A Comparison of Differential Biomarkers of Osteoarthritis with and without Post-traumatic Injury in the Hartley Guinea Pig Model. *J. Orthop. Res.* **2010**, *28*, 900−906.
(25) Meurens, F.; Summerfield, A.; Nauwynck, H.; Saif, L.; Gerdts, V. The pig: a model for human infectious diseases. *Trends Microbiol.* **2012**, *20*, 50−57.
(26) De Almeida, A. M.; Bendixen, E. Pig proteomics: a review of a species in the crossroad between biomedical and food sciences. *J. Proteomics* **2012**, *75*, 4296−4314.
(27) Humphray, S. J.; Scott, C. E.; Clark, R.; Marron, B.; Bender, C.; Camm, N.; Davis, J.; Jenks, A.; Noon, A.; Patel, M.; et al. A high utility integrated map of the pig genome. *Genome Biol.* **2007**, *8*, R139.
(28) Murray, M. M.; Fleming, B. C. Use of a bioactive scaffold to stimulate ACL healing also minimizes post-traumatic osteoarthritis after surgery. *Am. J. Sports Med.* **2013**, *41*, 1762−1770.
(29) Vavken, P.; Fleming, B. C.; Mastrangelo, A. N.; Machan, J. T.; Murray, M. M. Biomechanical outcomes after bio-enhanced anterior cruciate ligament repair and anterior cruciate ligament reconstruction are equal in a porcine model. *Arthroscopy* **2012**, *28*, 672−680.
(30) Murray, M. M.; Magarian, E. M.; Harrison, S. L.; Mastrangelo, A. N.; Zurakowski, D.; Fleming, B. C. The effect of skeletal maturity on functional healing of the anterior cruciate ligament. *J. Bone Jt. Surg. Am.* **2010**, *92*, 2039−2049.
(31) Gallien, S.; Duriez, E.; Crone, C.; Kellmann, M.; Moehring, T.; Domon, B. Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Mol. Cell. Proteomics* **2012**, *11*, 1709−1723.
(32) Steen, H.; Küster, B.; Fernandez, M.; Pandey, A.; Mann, M. Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. *Anal. Chem.* **2001**, *73*, 1440−1448.
(33) Neubauer, G.; Mann, M. Mapping of phosphorylation sites of gel-isolated proteins by nanoelectrospray tandem mass spectrometry: potentials and limitations. *Anal. Chem.* **1999**, *71*, 235−242.
(34) Shevchenko, An.; Wilm, M.; Vorm, O.; Jensen, O. N.; Podtelejnikov, A. V.; Neubauer, G.; Shevchenko, Al.; Mortensen, P.; Mann, M. A strategy for identifying gel-separated proteins in sequence databases by MS alone. *Biochem. Soc. Trans.* **1996**, *24*, 893−896.
(35) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367−1372.
(36) Schwanhäusser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. Global quantification of mammalian gene expression control. *Nature* **2011**, *473*, 337−342.
(37) Hulsen, T.; de Vlieg, J.; Alkema, W. BioVenn − a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **2008**, *9*, 488.
(38) J. C., Oliveros. VENNY. An interactive tool for comparing lists with Venn Diagrams http://bioinfogp.cnb.csic.es/tools/venny/.
(39) Schroeder, A.; Mueller, O.; Stocker, S.; Salowsky, R.; Leiber, M.; Gassmann, M.; Lightfoot, S.; Menzel, W.; Granzow, M.; Ragg, T. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol. Biol.* **2006**, *7*, 3.
(40) Ayturk, U. M.; Jacobsen, C. M.; Christodoulou, D. C.; Gorham, J.; Seidman, J. G.; Seidman, C. E.; Robling, A. G.; Warman, M. L. An RNA-seq protocol to identify mRNA expression changes in mouse diaphyseal bone: applications in mice with bone property altering Lrp5 mutations. *J. Bone Miner. Res.* **2013**, *28*, 2081−2093.
(41) Grant, G. R.; Farkas, M. H.; Pizarro, A. D.; Lahens, N. F.; Schug, J.; Brunk, B. P.; Stoeckert, C. J.; Hogenesch, J. B.; Pierce, E. A. Comparative analysis of RNA-seq alignment algorithms and the RNA-seq unified mapper (RUM). *Bioinformatics* **2011**, *27*, 2518−2528.
(42) Mortazavi, A.; Williams, B. A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat. Methods* **2008**, *5*, 621−628.
(43) Desrosiers, R. R.; Beaulieu, É.; Buchanan, M.; Béliveau, R. Proteomic analysis of human plasma proteins by two-dimensional gel electrophoresis and by antibody arrays following depletion of high-abundance proteins. *Cell Biochem. Biophys.* **2007**, *49*, 182−195.
(44) Schussler, G. C. The thyroxine-binding proteins. *Thyroid* **2000**, *10*, 141−149.
(45) Dea, M. K.; Hamilton-Wessler, M.; Ader, M.; Moore, D.; Schäffer, L.; Loftager, M.; Vølund, A.; Bergman, R. N. Albumin binding of acylated insulin (NN304) does not deter action to stimulate glucose uptake. *Diabetes* **2002**, *51*, 762−769.

(46) Zhou, M.; Lucas, D. A.; Chan, K. C.; Issaq, H. J.; Petricoin, E. F.; Liotta, L. A.; Veenstra, T. D.; Conrads, T. P. An investigation into the human serum "interactome". *Electrophoresis* **2004**, *25*, 1289−1298.
(47) Nielsen, M. L.; Vermeulen, M.; Bonaldi, T.; Cox, J.; Moroder, L.; Mann, M. Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nat. Methods* **2008**, *5*, 459−460.
(48) Kollipara, L.; Zahedi, R. P. Protein carbamylation: in vivo modification or in vitro artefact? *Proteomics* **2013**, *13*, 941−944.
(49) Bennike, T.; Lauridsen, K. B.; Olesen, M. K.; Andersen, V.; Birkelund, S.; Stensballe, A. Optimizing the identification of citrullinated peptides by mass spectrometry: utilizing the inability of trypsin to cleave after citrullinated amino acids. *J. Proteomics Bioinform.* **2013**, *06*, 288−295.
(50) Bennike, T.; Birkelund, S.; Stensballe, A.; Andersen, V. Biomarkers in inflammatory bowel diseases: current status and proteomics identification strategies. *World J. Gastroenterol.* **2014**, *20*, 3231−3244.
(51) Bensouyad, A.; Hollander, A. P.; Dularay, B.; Bedwell, A. E.; Cooper, R. A.; Hutton, C. W.; Dieppe, P. A.; Elson, C. J. Concentrations of glycosaminoglycans in synovial fluids and their relation with immunological and inflammatory mediators in rheumatoid arthritis. *Ann. Rheum. Dis.* **1990**, *49*, 301−307.
(52) Nakano, T.; Aherne, F. X.; Thompson, J. R. Relative amounts of chondroitin sulfate and hyaluronic acid in synovial fluid from normal and osteochondrotic swine joints. *Can. J. Comp. Med.* **1984**, *48*, 434−436.
(53) Geiger, T.; Velic, A.; Macek, B.; Lundberg, E.; Kampf, C.; Nagaraj, N.; Uhlen, M.; Cox, J.; Mann, M. Initial quantitative proteomic map of 28 mouse tissues using the SILAC mouse. *Mol. Cell. Proteomics* **2013**, *12*, 1709−1722.
(54) Hortin, G. L.; Sviridov, D. The dynamic range problem in the analysis of the plasma proteome. *J. Proteomics* **2010**, *73*, 629−636.
(55) Breslow, J. L.; Ross, D.; McPherson, J.; Williams, H.; Kurnit, D.; Nussbaum, A. L.; Karathanasis, S. K.; Zannis, V. I. Isolation and characterization of cDNA clones for human apolipoprotein A-I. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 6861−6865.
(56) Crichton, R. R.; Charloteaux-Wauters, M. Iron transport and storage. *Eur. J. Biochem.* **1987**, *164*, 485−506.
(57) Swindle, M. M.; Makin, A.; Herron, A. J.; Clubb, F. J.; Frazier, K. S. Swine as models in biomedical research and toxicology testing. *Vet. Pathol.* **2012**, *49*, 344−356.
(58) Vizcaíno, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Ríos, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **2014**, *32*, 223−226.
(59) Vizcaíno, J. A.; Côté, R. G.; Csordas, A.; Dianes, J. A.; Fabregat, A.; Foster, J. M.; Griss, J.; Alpi, E.; Birim, M.; Contell, J.; et al. The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* **2013**, *41*, D1063−1069.
(60) Wang, R.; Fabregat, A.; Ríos, D.; Ovelleiro, D.; Foster, J. M.; Côté, R. G.; Griss, J.; Csordas, A.; Perez-Riverol, Y.; Reisinger, F.; et al. PRIDE Inspector: a tool to visualize and validate MS proteomics data. *Nat. Biotechnol.* **2012**, *30*, 135−137.
(61) Côté, R. G.; Griss, J.; Dianes, J. A.; Wang, R.; Wright, J. C.; van den Toorn, H. W. P.; van Breukelen, B.; Heck, A. J. R.; Hulstaert, N.; Martens, L.; et al. The PRoteomics IDEntification (PRIDE) converter 2 framework: an improved suite of tools to facilitate data submission to the PRIDE database and the ProteomeXchange consortium. *Mol. Cell. Proteomics* **2012**, *11*, 1682−1689.