Published in final edited form as:

Proteomics. 2016 February ; 16(4): 634–644. doi:10.1002/pmic.201500195.

The Pig PeptideAtlas: a resource for systems biology in animal production and biomedicine

Marianne O. Hesselager¹, Marius C. Codrea², Zhi Sun³, Eric W. Deutsch³, Tue B. Bennike⁴, Allan Stensballe⁴, Louise Bundgaard⁵, Robert L. Moritz³, and Emøke Bendixen¹

¹Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark

²Quantitative Biology Center (QBiC), Eberhard Karls Universität, Tübingen, Germany

³Institute for Systems Biology, Seattle, WA, USA

⁴Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

⁵Department of Large Animal Sciences, University of Copenhagen, Copenhagen, Denmark

Abstract

Biological research of *Sus scrofa*, the domestic pig, is of immediate relevance for food production sciences, and for developing pig as a model organism for human biomedical research. Publicly available data repositories play a fundamental role for all biological sciences, and protein data repositories are in particular essential for the successful development of new proteomic methods. Cumulative proteome data repositories, including the PeptideAtlas, provide the means for targeted proteomics, system wide observations, and cross species observational studies, but pigs have so far been underrepresented in existing repositories.

We here present a significantly improved build of the Pig PeptideAtlas, which includes pig proteome data from 25 tissues and three body fluid types mapped to 7139 canonical proteins. The content of the Pig PeptideAtlas reflects actively ongoing research within the veterinary proteomics domain, and this manuscript demonstrates how the expression of isoform-unique peptides can be observed across distinct tissues and body fluids.

The Pig PeptideAtlas is a unique resource for use in animal proteome research, particularly biomarker discovery and for preliminary design of SRM assays, which are equally important for progress in research that supports farm animal production and veterinary health, as for developing pig models with relevance to human health research.

Keywords

Animal model; biomarker; Pig PeptideAtlas; proteomics; repositories

Corresponding author: Emøke Bendixen, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, 8000 Aarhus C, Denmark, ebx@mbg.au.dk, phone: +45 87155441.

The authors have declared no conflict of interest.

1. Introduction

Systems biology research of *Sus scrofa*, the domestic pig, is of immediate relevance for food production, animal husbandry and welfare. Utilizing a systems approach, a fundamental understanding of genetic and environmental factors influencing health, growth, nutrition, behaviour and fertility provides an essential platform for improving industrial farm animal production methods. Moreover, the domestic pig is an important model organism for human biomedical research, due to the close similarity on the gene and proteome level between pig and humans [1–4]. As a model organism, the pig is still a newcomer, partly because the pig genome was just recently published [5], but the diversity of available pig models is rapidly expanding. In particular for research in nutrition, inflammation and host-microbial cross-talk, pig models present unmatched modeling opportunities [3, 6], because the pig, like human, is an omnivore, with very similar nutritional requirements, digestive- and immune system and gut microbial components. Current nutrigenomic research aims to characterize complex relational networks between animal genomics, nutrition and gut microbes (metagenomics), which impact health, growth and metabolism of the individual animal. Such system-wide understanding is equally important for managing animal health in industrial pig production, as it is for facing the increasing global challenges to human health caused by e.g. malnutrition and food pathogens.

A key component to complex network analysis is readily available resources of multi-scalar data, of which proteomics plays a key role. Publicly available data repositories like the PeptideAtlas [7], PRIDE [8], ProteomeXchange [9] and the Global Proteome Machine Database [10], play a fundamental role in successful development of new proteomic methods and progress in biological sciences [11]. While individual proteome studies provide valuable new insights into specific molecular mechanisms (e.g. of particular diseases), cumulative proteome data repositories, including the PeptideAtlas, provide the means for targeted proteomics, system wide observations and cross species observational studies. However, unfortunately pig data has so far been underrepresented in these repositories, like many other relatively new model organism-species.

The PeptideAtlas project provides a large-scale assembly and highly stringent analysis of available LC-MS/MS-based shotgun proteomics data. Mammalian proteomes, and those of common model organisms like yeast (*Saccharomyces cerevisiae*) [12] and fly (*Drosophila melanogaster*) [13], are already well represented with over 70% proteome coverage for each in the PeptideAtlas. More recently, organisms of important economic value like the cow [14], horse [15] and honeybee [16] have been included in the PeptideAtlas project.

The PeptideAtlas provides a tool of choice for selecting proteotypic peptides for targeted quantitative analysis such as selected reaction monitoring (SRM) and data-independent approaches (DIA and SWATH) studies [7]. Targeted proteomics, using e.g. SRM has overcome some of the challenges of developing multiplexed assays for both high and low abundant proteins within complex biological samples [17]. The SRM approach is now widely used for the study of complex tissue samples and body fluids. For pig, a PeptideAtlas repository covering 2484 canonical proteins has been available since 2009. With the recent expansion of pig proteome research, we have for the first time the opportunity to collect data

across parallel projects and laboratories. This resulted in a significantly improved coverage of pig protein information, now collected in the PeptideAtlas. Here, we present the contents of the current Pig PeptideAtlas, including data from 25 tissues and three body fluid types mapped to 7139 canonical proteins. The content of the Pig PeptideAtlas reflects actively ongoing research within the veterinary proteomics domain, and covers extensively muscle, liver, gut and neural proteomes, as well as several body fluids relevant for diagnostic purposes. This article will focus on presenting proteins with immediate relevance to research in the closely connected pathways of inflammatory, metabolic and immune response biology. The representation of these central protein pathways in the Pig PeptideAtlas provides a resource for detecting an array of proteins playing significant roles in animal health under industrial farming systems, but also supporting a further development of pig models with relevance to human health research. The aim of the current paper is to illustrate how the Pig PeptideAtlas can be used to mine sequence specific information about proteins of particular interest. To present an example, we have chosen to thoroughly describe isoforms and tissue-specific expression of pig serum amyloid A (SAA) across the different tissues and body fluids represented in the Pig PeptideAtlas.

2. Materials and methods

2.1. Samples and sample processing

Samples in the Pig PeptideAtlas are from many different cohorts of animals, none of which had signs of infection at the time of sample collection. Samples were mainly collected from Duroc and Danish Landrace breeds. Synovial fluid originated from Yucatan minipigs. Pig retina, nerve, artery and plasma were obtained from a study performed with permission from the Danish Animal Experiments Inspectorate (permission no. 2013-15-2934-00775). Synovial fluid samples were obtained from a previous study [4], and the residual samples were obtained from former studies [18–23] or sampled from animals immediately after slaughter.

Samples were processed using one of two approaches:

- i. 10–20 mg of tissue samples were homogenized in 0.5 mL 5% sodium deoxycholate (SDC), pH 8.5, using bead beating [22]. The samples were cooled on ice between the runs. Retinal samples were transferred to YM-10 Spin filters (Millipore, Billerica, MA, USA) and buffer exchange performed to digestion buffer (1% SDC in 0.1M triethyleammonium bicarbonate (TEAB), pH 7.8). Cysteine residues were reduced at 37°C using 12 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 minutes and alkylated using 40 mM iodoacetamide for 30 minutes followed by two buffer exchanges to 0.5% SDC in 0.1M TEAB, pH 7.8 with centrifugation at $14,400 \times g$, in order to remove alkylation reagents and reduce the concentration of SDC to a concentration where trypsin retain optimal activity. A volume of 2 µg trypsin (Promega, Madison, WI, USA) were added to the samples (100 µg as determined by A280) and incubated overnight. The filtrate was recovered by centrifugation. SDC was precipitated with 5% formic acid (FA) and soluble peptides were recovered.

- ii. 200 mg of tissue samples were homogenized in 1 mL TES buffer (10 mM tris, 1 mM EDTA, 0.25 M sucrose) using a tissuelyzer for 3×20 sec and 30 Hz frequency (TissueLyser II, Qiagen, Hilden, Germany). The homogenates were centrifuged at $500 \times g$ for 30 min to isolate supernatant. Protein concentrations of the supernatants were determined using the Pierce BCA Protein Assay Kit (Thermo Scientific, Waltham, Massachusetts) with BSA as a standard according to manufacturer's protocol (<http://www.piercenet.com/instructions/2161296.pdf>). An aliquot of 120 μ g of protein was precipitated in ice-cold acetone overnight.

The precipitated protein pellets were re-suspended in 20 μ L of 0.5 M TEAB, pH 8.5 (AB SCIEX, Framingham, MA, USA). Proteins were denatured in 0.1% sodium dodecyl sulphate (SDS) (AB SCIEX), cysteine residues were reduced with 2.5 mM TCEP hydrochloride (AB SCIEX), incubated at 60 °C for 1 h, and blocked with 10 mM methylmethanethiosulfate (MMTS) (AB SCIEX) at 22°C for 10 min. The proteins were digested overnight at 37 °C with 1:10 w/w trypsin (AB SCIEX). Tryptic peptides were passed through a 0.2 μ m centrifuge filter (VWR, Radnor, Pennsylvania) at $10,000 \times g$ for 10 min, dried in a vacuum centrifuge, and stored at -80°C .

2.2. MS analysis

Up to three technical replicates of each sample were performed. Samples were analyzed using one of three approaches:

- i. Digested samples were dissolved in buffer A (0.03% FA and 5% acetonitrile (ACN)) in water. A volume corresponding to 50 μ g protein digest was loaded into an Agilent 1100 Series capillary HPLC system equipped with a Zorbax Bio strong cation exchange (SCX) Series II, 0.8×50 mm column (Agilent Technologies, Santa Clara, CA). Peptides were eluted with a gradient of 0–100% buffer B (0.03% FA, 5% ACN, and 1 M NaCl in water) in 65 min, at a flow rate of 15 μ L/min. 1-min fractions were collected and following pre-fractionation, these were pooled into a total of 10 samples in order to equalize peptide loads for the individual LC-MS/MS analyses.

Pooled peptide samples were further separated on a Proxeon EASY- nLC (Thermo Scientific, Waltham, Massachusetts). The HPLC system was equipped with an isocratic pump working at 2 μ L/min in buffer A (0.1% FA and 3% ACN in water) and sample was loaded into a ReproSil-Pur C18-A1 precolumn (ID 100 μ m $\times$ 2 cm, 5 μ m, 120 Å; from Thermo Scientific). The pre-column was switched into a nanoflow path (300 nL/min), and the peptides were further separated on a ReproSil-Pur C18-A2 column (ID 75 μ m $\times$ 10 cm, 3 μ m, 120 Å; (Thermo Scientific) by running a 73 min gradient from 0–38% buffer B (0.1% FA and 90% ACN in water) followed by a 10 min wash in 100% buffer B. The eluted peptides were sprayed through a nanospray emitter (PicoTip Emitter, SilicaTip, no coating, OD 360 μ m, ID 20 μ m; from New Objective, Woburn, MA, USA) directly into a Qstar Elite mass spectrometer (AB SCIEX). A data-dependent acquisition method was employed to automatically run experiments, operated under Analyst QS 2.0 (AB SCIEX).

- ii. Digested samples were micropurified (StageTips C18, 20 μ L tips; from Thermo Scientific) dissolved in buffer A (0.1% FA) and processed on an EASY-nLC II (Thermo Scientific) with a Biosphere C18 precolumn (id 100 μ m $\times$ 2 cm, 5 μ m, 120 $\text{\AA}$; NanoSeparation, Nieuwkoop, Netherlands) and separated on an in-house packed C18 analytical column (3 μ m ReproSil-Pur C18-AQ material packed in a PicoTip emitter, id 75 μ m $\times$ 10 cm; New Objective, Woburn, MA, USA) coupled in-line to a TripleTOF 5600 mass spectrometer (AB SCIEX). Samples were separated at a flow rate of 250 nL/min using a 50 min gradient from 5–35% buffer B (0.1% FA, 90% ACN) followed by 10 min wash in 100% buffer B with an data-dependent acquisition method, operated under Analyst TF 1.5.1 (AB SCIEX).
- iii. Peptide samples were supplemented with 20 fmol/ μ L BSA massprep digest (Waters, Milford, MA, USA) as internal standard, and five μ g sample was injected into a Dionex RSLC nanoUPLC system (Thermo Scientific) that was connected to Quadrupole Orbitrap mass spectrometers (Velos, XL and Q Exactive Plus, from Thermo Scientific) equipped with a NanoSpray Flex ion source (Thermo Scientific). The flow settings were 8 μ L/min for the sample loading onto a trapping column (Acclaim PepMap100 C18 5 μ m, Thermo Scientific). The nanoflow was set to 300 nL/min for the peptide separation on the analytical column, which was a 50 cm Acclaim Pepmap RSLC, 75 μ m ID connected with nanoviper fittings (Thermo Scientific). The nano-electrospray was done using a Picotip 'Silicatip' emitter from New Objectives (New objective). The LC buffers were buffer A (0.1% FA) and buffer B (0.1% FA, 99.9% ACN). The applied gradient was from 10–30% buffer B over 30 or 120 min.

The mass spectrometer was operated in data-dependent acquisition mode. A full MS scan in the mass range of 350 to 1400 m/z was acquired at a resolution of 70,000 with an automatic gain control (AGC) target of 1×10^6 and maximum fill time set to 100 ms. Instrument lock mass correction was applied using the contaminant ion at 391.28429 m/z. In each cycle, the mass spectrometer would trigger up to 12 MS/MS acquisition on abundant peptide precursors ions. The MS/MS scans were acquired with a dynamic mass range at a resolution of 17,500, and with an AGC target of 5×10^5 and max fill time of 50 ms. The precursor ions were isolated using a quadrupole isolation window of 2.0 m/z and the fragmented in the HCD trap with a normalized collision energy set to 30. The under-fill ratio was set to 1.0% with the intensity threshold at 1.0×10^5 . Apex triggering was 3 to 10 s, with charge and isotopes exclusion on. Dynamic exclusion was set to 30 s.

2.3. Data processing

All raw data can be accessed from the PeptideAtlas Repository (<http://www.peptideatlas.org/repository/>) by searching for pig. The processing of the pig MS data followed the general workflow described in Farrah et al. [24] using the Trans-Proteomic Pipeline [25] as the main component. First, the vendor .wiff files from both the TripleTOF and Qstar Elite were converted to the mzML format [26] using the AB SCIEX converter version 1.3 beta. The porcine search database was compiled as a nonredundant union of 26139 porcine sequences from UniProtKB, 24556 sequences from NCBI RefSeq databases (versions as of February 2014) and cRAP database (common contaminants). Since more

than 80% of the UniProtKB proteins were lacking informative annotations (mainly annotated as “uncharacterized protein”) we ran BLAST against the Homo Sapiens UniProtKB database to annotate pig proteins based on human orthologue protein names.

The search database is available at http://www.peptideatlas.org/builds/pig/201504/Pig_TargetDecoy.fasta. Next the mzML files were searched against the nonredundant union pig database with Comet [27]. For Q Exactive files, the parent mass error was set to ± 3 ppm. For Qstar Elite, Qstar XL, and TripleTOF data, the parent mass error was set to ± 1.1 Da. For QstarXL instrument runs, a static modification of +45.987721 for MMTS modification of cysteine was used, and for all other runs, static modification of + 57.021464 for iodoacetamide modification of cysteine was used. Variable protein N-terminal acetylation (+42.010565 and methionine oxidation (+15.9949) were allowed. The maximum number of missed cleavage sites was set to 2.

For each experiment, the output of the search engines was converted to the pepXML format [28] and then the output of each search engine was modeled separately with the PeptideProphet [29], iProphet [30] and ProteinProphet TPP [25, 31].

3. Results and discussion

3.1. Assembling the Pig PeptideAtlas

All individual iProphet files from 317 porcine samples (Supplementary table S1) were remapped using RefreshParser from the TPP toolkit to the PeptideAtlas reference database, which was compiled from the search database described above and *Sus scrofa* protein sequences from Ensembl release 74. All the iProphet results were filtered at a variable probability threshold to maintain a constant peptide-spectrum-match (PSM) FDR threshold of 0.0012 for each experiment. All filtered results were loaded into the Pig PeptideAtlas (build 2014-11) and made available at <http://www.peptideatlas.org>. This build contains 87908 distinct peptides at peptide-level FDR of 0.0011 and 7139 canonical proteins at protein-level FDR of 0.010 as reported by Mayu analysis [32]. Table 1 provides an overview of the different sample types presented in the Pig PeptideAtlas as well as the number of representations of each type.

3.2. Acute phase proteins in the Pig PeptideAtlas

The acute phase response is referred to as an early, unspecific reaction of an organism against infection and tissue damage. Thus, specific knowledge (time and place) of the abundances of individual acute phase proteins (APPs) has broad clinical and scientific relevance. The abundance of APPs changes during an inflammatory response, and APPs are commonly used for monitoring infectious and inflammatory disease in both veterinary [33] and human [34] clinics. The response magnitude of APPs differs according to the cause of reaction, e.g. different pathogens result in different APP responses.

Haptoglobin (Hp), pig major acute phase protein (pig-MAP), serum amyloid A (SAA), C-reactive protein (CRP) and alpha-2-HS-glycoprotein are all classified as positive porcine APPs (relative abundance increasing upon challenge) [35–38]. The negative porcine APPs (relative abundance decreasing upon challenge) include apolipoprotein A-I (Apo A-I) [39],

serum albumin [35, 36] and transthyretin [40]. Table 2 presents an overview of the eight most commonly studied porcine APPs, summarizing their observations across the porcine tissues represented in the PeptideAtlas.

Hp, Apo A-I and albumin are represented in all examined tissues, whereas SAA and CRP are rather specific for body fluids, and the minor presence of e.g. SAA in muscle may originate from blood circulating the muscle.

3.3. Mining the tissue specificity of serum amyloid A

Serum amyloid A and clinical diagnostics—SAA is a 9–14 kDa apolipoprotein, recognized as a major acute phase responder (concentration can change 100–1000 fold during an inflammatory response) to inflammation in most vertebrate species [41], including the pig [42, 43]. In pigs, as in most mammals, a four-loci gene cluster gives rise to distinct SAA isoforms [44, 45]. Distinguishing SAA isoforms and elucidating their expression patterns is of general interest for research as well as diagnostic applications, and already applied for human [46] diagnostics. However, distinguishing SAA isoforms is currently a challenging task, because isoform specific antibodies to these porcine proteins are not available.

The SAA protein will be mined in details to demonstrate the PeptideAtlas features that support differentiating between isoforms of closely related sequences. This knowledge is essential for clinical and research applications, where it is essential to distinguish between local and systemic inflammatory disorders, as well as monitoring stages of disease.

SAA genes are highly conserved, but different isoforms within the same species are more alike than orthologous isoforms between different species [47], thus, there is a need for species-specific assays to measure different SAA isoforms.

Serum amyloid A data in the pig PeptideAtlas—In all species studied so far, SAA is produced both hepatically and in other tissues, although in relatively low amounts in healthy states [48–50]. Searching the pig PeptideAtlas for “serum amyloid A” yields three canonical SAA protein isoforms; UniProt accessions Q2HXZ9 (SAA3), F1S9B8 (SAA4) and F1S9B9 (SAA2). In the literature, the porcine SAA nomenclature is inconsistent. Although Q2HXZ9 was originally published as SAA2 (GenBank accession number DQ367410) by Chang et al. [51], Q2HXZ9 is referred to as SAA3 in the present paper due to shared characteristics of SAA3 from other species as described by Olsen et al. [44]. UniProt accession F1S9B9 corresponds to SAA2 based on UniGene (Ssc. 17030) and F1S9B8 is SAA4 in Uniprot, Olsen et al. [44] sequenced SAA isoforms from cDNA from multiple porcine tissues and found transcripts corresponding to the same three isoforms as observed in the PeptideAtlas. The three sequences are aligned in Figure 1 and observed isoform-unique peptides are highlighted. In PeptideAtlas, peptides that map to only a single genome location, can be identified from the “N Gen Loc” column in the “Distinct Observed Peptides” view or in the “Sample peptide map”, denoted by an asterisk.

In addition to confirming the presence of the three distinct isoforms, the PeptideAtlas also provides information on their tissue distribution. Table 3 gives the observations of the SAA

isoform-unique peptides (including miss-cleavages) across all porcine samples in the Pig PeptideAtlas. Also, the number of different biological samples in which these peptides were observed is reported. Only isoform-unique peptides are reported according to their specific representation in only four types of the biological tissues/body fluids represented in the Pig PeptideAtlas. SAA2 is found in synovial fluid and plasma samples, SAA3 is exclusively present in colostrum, even though it has previously been described as the main circulating form of SAA in pigs [52]. SAA4 is detected in skeletal muscle, synovial fluid and plasma. SAA4 is a constitutively expressed isoform and known to be present in synovial fluid in healthy porcine knee joints [4].

The liver is the major site for SAA synthesis upon inflammation, yet no SAA is found in any of the >60 liver samples present in the PeptideAtlas repository, thus this supports the assumption that all samples were collected from healthy animals.

Tissue enriched expression of serum amyloid A isoforms—Being able to measure APPs locally at sites of injury increase their clinical and diagnostic value [53]. In horse, SAA and Hp in peritoneal fluid is recognized as a valuable diagnostic marker for abdominal conditions [54], and the SAA concentration in synovial fluid was found to be a marker for infectious arthritis and tendovaginitis [55]. In cattle, SAA and Hp in combination have proven useful for distinguishing between acute and chronic inflammation [56].

Our observation of SAA3 being exclusive to porcine colostrum is well in line with reports of bovine, equine and ovine mammary-associated isoforms (M-SAA3) found in colostrum from healthy animals [49]. Also, previous studies of SAA in pigs have suggested the expression of a mammary-specific SAA isoform [57, 58]. Our observations across the pig PeptideAtlas clearly suggest a local secretion of an SAA3 isoform to the colostrum and not found in other tissues. However, in the PeptideAtlas, only a single unique peptide identifies the SAA3 isoform, thus, our observations may also reflect that specific SAA3 fragments, or modifications of SAA3 may be present in the colostrum.

This observation is interesting in the light that the pig mammary gland related SAA isoform has been observed to increase during pregnancy, peaking at day 15 of lactation [58]. Also, a specific peptide from SAA3 (DMREANYKNSDKYFHARGNYDAA) was found to act as a chemo-attractant for immune cells [58], indicating that the mammary-associated isoform of SAA3 may have an important biological role in host response against pathogens, both in the mammary gland as well as in the gut of the young piglet upon uptake of milk and colostrum proteins. Also in human milk, the role of SAA3 has been related to host response to pathogens [59–62]. Moreover, we have also previously observed bovine SAA3, expressed by mammary epithelial cells, and increased in abundance both in milk and mammary gland tissues after bacterial lipopolysaccharide [19] or *Streptococcus uberis* [63] challenge. Thus, monitoring this isoform could present a sensitive indicator for udder infections in dairy cows [63–65].

The Pig PeptideAtlas supports accumulating cross-species evidence that local expression of SAA isoforms have important biological relevance. Specifically, underpinning the role of SAA secretion to colostrum in protecting the neonate against pathogens.

3.4. Gut Inflammatory markers mapped in the Pig PeptideAtlas

Understanding inflammatory bowel disorders affect both pig and human and is actively pursued in the veterinary sciences as well as human clinical diagnostics, thus cross-species comparative systems biology will greatly improve progress to both areas [66]. Inflammatory states of the gut are closely correlated to the individual animal's susceptibility to gut pathogens (e.g. *E. coli*), which in modern pig production cause severe challenges to economic gain as well as for animal welfare and need for extensive use of antibiotics in the farm industry. Current pig models are available for monitoring pig gut inflammatory conditions [3], and these models have recently gained increasing relevance for modeling human conditions, because pig models offer an unmatched close similarity to human digestion tract, and offer access to samples that are very hard to obtain from human patients [3]. Thus, mining the Pig PeptideAtlas for inflammatory markers is of common interest for supporting ongoing research.

A selection of protein markers that are currently used for diagnosing and monitoring inflammatory bowel disease activity in humans are readily observed in the Pig PeptideAtlas (Supplementary table S2 shows descriptions and PeptideAtlas mappings of proteins known to be affected in human gut inflammation). Whether these proteins have relevance for monitoring pig gut health is not yet clear. However, one of the purposes of the Pig PeptideAtlas is to make it possible for a wide range of users to study these proteins, to support progress of clinically relevant markers which can be used to monitoring health and disease in pigs. For diagnosing gut inflammation in the human clinic, stool samples are preferred over serum and tissue samples, partly due to its non-invasive and continuous sampling opportunity, and also for its direct contact with the inflamed area. This is in contrast to serum-markers which to a larger extend can be increased on account of a variety of conditions [66].

Calprotectin, lactotransferrin and CRP are the most commonly referred human gut inflammatory markers. Calprotectin, a heterodimer formed by the S100A8 and S100A9 proteins, is a calcium- and zinc-binding protein occurring in large amounts in neutrophil granulocytes. Calprotectin can be searched in the Pig PeptideAtlas by S100A8 and S100A9. Searching the Pig PeptideAtlas for "S100A9" yields one entry (C3S7K6), while searching for "S100A8" yields two entries: C3S7K5 and K7GQW2. C3S7K5 is the canonical protein, while K7GQW2 is "NTT subsumed", meaning that all peptides mapped to this protein can also be explained by mappings to the canonical protein. K7GQW2 is identical to C3S7K5 except for an extended N-terminal, and therefore a variant originating from the same gene. Although no peptide mappings to this N-terminal are reported in the current build of the Pig PeptideAtlas, both entries are kept (as they are in the UniProtKB).

Because protein names are redundant and inconsistent, the PeptideAtlas search engine is designed to operate primarily by unique identifiers like UniProt accession numbers, gene names or PeptideAtlas accession numbers.

The abundance of calprotectin and lactotransferrin increases in gut-tissue and stools during intestinal inflammation. As such, fecal calprotectin and fecal lactotransferrin are widely used for human diagnostics, to rule out inflammatory bowel disease [67]. The proteins withstand

degradation during storage at room temperature, as well as repeated freeze-thaw cycles, making shipping of stool samples feasible [66]. However, elevated levels of calprotectin in stool is not a unique marker for inflamed gut tissues, as it has also been associated with other clinical conditions, including the presence of polyps [68], use of non-steroidal anti-inflammatory drugs and normal ageing [69]. Lactotransferrin, an iron-binding glycoprotein expressed by activated neutrophils, is released upon tissue damage, where it contributes to modulate the inflammatory response, and as a part of the innate immune system delivers a first line defense against pathogens [70]. Finally, CRP is mainly hepatically expressed as a response to factors released by immune cells at the site of inflammation, and takes part in the innate immune system by activating the complement system [71, 72].

3.5. Immunological and metabolic pathways mapped in the Pig PeptideAtlas

The Pig PeptideAtlas has a rich representation of proteins and pathways, which are key to future proteome research of pig health and disease. These include pathways related to nutrition and metabolism and have immediate relevance to industrial animal production, e.g. improving food conversion rates, and they are directly related to economic gain in meat production. Also, protein pathways connected to immunology and host response have wide interest for improving health in pig production, and thus are targeted through much ongoing pig proteome research. The Pig PeptideAtlas includes observations of the major pathways targeted through pig proteome research, as summarized in the Supplementary table S3. Proteins mapped in the Pig PeptideAtlas and involved in KEGG pathways can be visualized by accessing the pig build and the [Query] [Pathways] menu (https://db.systemsbiology.net/sbeams/cgi/PeptideAtlas/showPathways?atlas_build_id=439) and [View Data] for PeptideAtlas, which then shows PeptideAtlas data overlayed on KEGG pathways.

4. Concluding remarks

The recent update of the Pig PeptideAtlas presents a collection of 87908 high-confidence peptide and 7139 protein identifications from data collected across projects and laboratories. Here, we have presented the content of the latest build and specifically demonstrated how the PeptideAtlas provides access to observing tissue specific expression of acute phase protein isoforms. The Pig PeptideAtlas provides access to information about a wide range of proteins relevant to ongoing research in the inflammatory, metabolic, and immune response biology. PeptideAtlas information of these pathways offers a basis for detecting proteins playing significant roles in animal health.

Single amino acid variants have been mined in detail in the Pig PeptideAtlas to demonstrate how isoform-unique peptides can be observed at the sample level in the PeptideAtlas, and thus assist in studying isoform-specific expression in distinct tissues. The Pig PeptideAtlas is a unique resource for use in animal research at the proteome level with particular relevance to biomarker discovery and for preliminary design of SRM assays.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank Dave Campbell for providing important solutions to pathway mappings and unique peptide mappings, and to Dorte Thomassen for invaluable help with tissue sampling and sample preparation. The authors also wish to acknowledge the support from the Danish Meat Association and the Danish Ministry of Agriculture and Fisheries, and to thank Dr. Benedict Kjaergaard, Dr. Henrik Vorum, Dr. Lasse Cehofski and Dr. Hanno Steen for help establishing the pig sample material. This study was funded by the Danish Council for Independent Research (grant number 11-106956) and by a Danish Government PhD grant. Additionally, the Obelske family foundation, the SparNord foundation and the Svend Andersen foundation are acknowledged for grants supporting the analytical platform (AS grants). Z.S., E.W.D., and R.L.M. were funded in part by the National Institutes of Health NHGRI grant RC2HG005805, NIGMS grants R01GM087221, S10RR027584 and 2P50GM076547 to the Center for Systems Biology and NIBIB grant U54EB020406, and the National Science Foundation MRI grant 0923536.

Abbreviations

ACN	acetonitrile
AGC	automatic gain control
Apo A-I	apolipoprotein A-I
APPs	acute phase proteins
CRP	C-reactive protein
FA	formic acid
Hp	haptoglobin
M-SAA3	mammary-associated isoforms
MMTS	methylmethanethiosulfate
pig-MAP	pig major acute phase protein
PSM	peptide-spectrum-match
SAA	serum amyloid A
SCX	strong cation exchange
SDC	sodium deoxycholate
SDS	sodium dodecyl sulphate
SRM	selected reaction monitoring
TCEP	tris(2-carboxyethyl)phosphine
TEAB	triethyleammonium bicarbonate

References

1. Bassols A, Costa C, Eckersall PD, Osada J, et al. The Pig as an Animal Model for Human Pathologies: A Proteomics Perspective. *Proteomics Clin Appl.* 2014; 8:715–731. [PubMed: 25092613]
2. Meurens F, Summerfield A, Nauwynck H, Saif L, Gerdt V. The Pig: A Model for Human Infectious Diseases. *Trends Microbiol.* 2012; 20:50–57. [PubMed: 22153753]

3. Jiang P, Sangild PT. Intestinal Proteomics in Pig Models of Necrotising Enterocolitis, Short Bowel Syndrome and Intrauterine Growth Restriction. *Proteomics Clin Appl*. 2014; 8:700–714. [PubMed: 24634357]
4. Bennike T, Ayturk U, Haslauer CM, Froehlich JW, et al. A Normative Study of the Synovial Fluid Proteome from Healthy Porcine Knee Joints. *J Proteome Res*. 2014; 13:4377–4387. [PubMed: 25160569]
5. Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, et al. Analyses of Pig Genomes Provide Insight into Porcine Demography and Evolution. *Nature*. 2012; 491:393–398. [PubMed: 23151582]
6. Cecilian F, Restelli L, Lecchi C. Proteomics in Farm Animals Models of Human Diseases. *Proteomics Clin Appl*. 2014; 8:677–688. [PubMed: 24595991]
7. Deutsch EW, Lam H, Aebersold R. PeptideAtlas: A Resource for Target Selection for Emerging Targeted Proteomics Workflows. *EMBO Rep*. 2008; 9:429–434. [PubMed: 18451766]
8. Vizcaino JA, Cote R, Reisinger F, Foster JM, et al. A Guide to the Proteomics Identifications Database Proteomics Data Repository. *Proteomics*. 2009; 9:4276–4283. [PubMed: 19662629]
9. Vizcaino JA, Deutsch EW, Wang R, Csordas A, et al. ProteomeXchange Provides Globally Co-Ordinated Proteomics Data Submission and Dissemination. *Nat Biotechnol*. 2014; 32:223–226. [PubMed: 24727771]
10. Craig R, Cortens J, Beavis R. Open Source System for Analyzing, Validating, and Storing Protein Identification Data. *Journal of Proteome Research*. 2004; 3:1234–1242. [PubMed: 15595733]
11. Vizcaino JA, Foster JM, Martens L. Proteomics Data Repositories: Providing a Safe Haven for Your Data and Acting as a Springboard for further Research. *Journal of Proteomics*. 2010; 73:2136–2146. [PubMed: 20615486]
12. King NL, Deutsch EW, Ranish JA, Nesvizhskii AI, et al. Analysis of the *Saccharomyces Cerevisiae* Proteome with PeptideAtlas. *Genome Biol*. 2006; 7:R106. [PubMed: 17101051]
13. Loevenich SN, Brunner E, King NL, Deutsch EW, et al. The *Drosophila Melanogaster* PeptideAtlas Facilitates the use of Peptide Data for Improved Fly Proteomics and Genome Annotation. *BMC Bioinformatics*. 2009; 10:59. [PubMed: 19210778]
14. Bislev SL, Deutsch EW, Sun Z, Farrah T, et al. A Bovine PeptideAtlas of Milk and Mammary Gland Proteomes. *Proteomics*. 2012; 12:2895–2899. [PubMed: 22837157]
15. Bundgaard L, Jacobsen S, Sorensen MA, Sun Z, et al. The Equine PeptideAtlas: A Resource for Developing Proteomics-Based Veterinary Research. *Proteomics*. 2014; 14:763–773. [PubMed: 24436130]
16. Chan QW, Parker R, Sun Z, Deutsch EW, Foster LJ. A Honey Bee (*Apis Mellifera* L) PeptideAtlas Crossing Castes and Tissues. *BMC Genomics*. 2011; 12 290-2164-12-290.
17. Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R. Full Dynamic Range Proteome Analysis of *S. Cerevisiae* by Targeted Proteomics. *Cell*. 2009; 138:795–806. [PubMed: 19664813]
18. Danielsen M, Hornshøj H, Siggers RH, Jensen BB, et al. Effects of Bacterial Colonization on the Porcine Intestinal Proteome. *J Proteome Res*. 2007; 6:2596–2604. [PubMed: 17542629]
19. Danielsen M, Codrea MC, Ingvarsen KL, Friggens NC, et al. Quantitative Milk Proteomics--Host Responses to Lipopolysaccharide-Mediated Inflammation of Bovine Mammary Gland. *Proteomics*. 2010; 10:2240–2249. [PubMed: 20352626]
20. Danielsen M, Pedersen LJ, Bendixen E. An in Vivo Characterization of Colostrum Protein Uptake in Porcine Gut during Early Lactation. *J Proteomics*. 2011; 74:101–109. [PubMed: 20831910]
21. Hornshøj H, Bendixen E, Conley LN, Andersen PK, et al. Transcriptomic and Proteomic Profiling of Two Porcine Tissues using High-Throughput Technologies. *BMC Genomics*. 2009; 10 30-2164-10-30.
22. Cehofski LJ, Kruse A, Kjærgaard B, Stensballe A, et al. Dye Free Porcine Model of Experimental Branch Retinal Vein Occlusion – a Suitable Approach for Retinal Proteomics. *Journal of Ophthalmology*. 2015 in press.
23. Cehofski LJ, Kruse A, Kjærgaard B, Stensballe A, et al. Proteins Involved in Focal Adhesion Signaling Pathways are Differentially Regulated in Experimental Branch Retinal Vein Occlusion. *Exp Eye Res*. 2015; 138:87–95. [PubMed: 26086079]

24. Farrah T, Deutsch EW, Omenn GS, Sun Z, et al. State of the Human Proteome in 2013 as Viewed through PeptideAtlas: Comparing the Kidney, Urine, and Plasma Proteomes for the Biology- and Disease-Driven Human Proteome Project. *J Proteome Res.* 2014; 13:60–75. [PubMed: 24261998]
25. Deutsch EW, Mendoza L, Shteynberg D, Slagel J, et al. Trans-Proteomic Pipeline, a Standardized Data Processing Pipeline for Large-Scale Reproducible Proteomics Informatics. *Proteomics Clin Appl.* 2015
26. Deutsch, E. Hubbard, SJ.; Jones, AR., editors. Humana Press; 2010. p. 319-331.
27. Eng JK, Jahan TA, Hoopmann MR. Comet: An Open-Source MS/MS Sequence Database Search Tool. *Proteomics.* 2013; 13:22–24. [PubMed: 23148064]
28. Keller A, Eng J, Zhang N, Li X, Aebersold R. A Uniform Proteomics MS/MS Analysis Platform Utilizing Open XML File Formats. *Molecular Systems Biology.* 2005; 1:2005.0017. [PubMed: 16729052]
29. Keller A, Nesvizhskii A, Kolker E, Aebersold R. Empirical Statistical Model to Estimate the Accuracy of Peptide Identifications made by MS/MS and Database Search. *Anal Chem.* 2002; 74:5383–5392. [PubMed: 12403597]
30. Shteynberg D, Deutsch EW, Lam H, Eng JK, et al. iProphet: Multi-Level Integrative Analysis of Shotgun Proteomic Data Improves Peptide and Protein Identification Rates and Error Estimates. *Molecular & Cellular Proteomics.* 2011; 10
31. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A Statistical Model for Identifying Proteins by Tandem Mass Spectrometry. *Anal Chem.* 2003; 75:4646–4658. [PubMed: 14632076]
32. Reiter L, Claassen M, Schrimpf SP, Jovanovic M, et al. Protein Identification False Discovery Rates for very Large Proteomics Data Sets Generated by Tandem Mass Spectrometry. *Molecular & Cellular Proteomics.* 2009; 8:2405–2417. [PubMed: 19608599]
33. Petersen HH, Nielsen JP, Heegaard PM. Application of Acute Phase Protein Measurements in Veterinary Clinical Chemistry. *Vet Res.* 2004; 35:163–187. [PubMed: 15099494]
34. Gabay C, Kushner I. Acute-Phase Proteins and Other Systemic Responses to Inflammation. *N Engl J Med.* 1999; 340:448–454. [PubMed: 9971870]
35. Parra MD, Fuentes P, Tecles F, Martinez-Subiela S, et al. Porcine Acute Phase Protein Concentrations in Different Diseases in Field Conditions. *J Vet Med B Infect Dis Vet Public Health.* 2006; 53:488–493. [PubMed: 17123428]
36. Heegaard PMH, Klausen J, Nielsen JP, González-Ramón N, et al. The Porcine Acute Phase Response to Infection with *Actinobacillus Pleuropneumoniae*. Haptoglobin, C-Reactive Protein, Major Acute Phase Protein and Serum Amyloid A Protein are Sensitive Indicators of Infection. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology.* 1998; 119:365–373.
37. Pomorska-Mol M, Markowska-Daniel I, Kwit K, Stepniewska K, Pejsak Z. C-Reactive Protein, Haptoglobin, Serum Amyloid A and Pig Major Acute Phase Protein Response in Pigs Simultaneously Infected with H1N1 Swine Influenza Virus and *Pasteurella Multocida*. *BMC Vet Res.* 2013; 9 14-6148-9-14.
38. Thongboonkerd V, Chiangjong W, Mares J, Moravec J, et al. Altered Plasma Proteome during an Early Phase of Peritonitis-Induced Sepsis. *Clin Sci(Lond).* 2009; 116:721–730. [PubMed: 19006484]
39. Carpintero R, Pineiro M, Andres M, Iturralde M, et al. The Concentration of Apolipoprotein A-I Decreases during Experimentally Induced Acute-Phase Processes in Pigs. *Infect Immun.* 2005; 73:3184–3187. [PubMed: 15845530]
40. Campbell FM, Waterston M, Andresen LO, Sorensen NS, et al. The Negative Acute Phase Response of Serum Transthyretin Following *Streptococcus Suis* Infection in the Pig. *Vet Res.* 2005; 36:657–664. [PubMed: 15955288]
41. Uhlar CM, Whitehead AS. Serum Amyloid A, the Major Vertebrate Acute-Phase Reactant. *European Journal of Biochemistry.* 1999; 265:501–523. [PubMed: 10504381]
42. Skovgaard K, Mortensen S, Boye M, Poulsen KT, et al. Rapid and Widely Disseminated Acute Phase Protein Response After Experimental Bacterial Infection of Pigs. *Vet Res.* 2009; 40:23. [PubMed: 19236838]

43. Skovgaard K, Mortensen S, Boye M, Hedegaard J, Heegaard PM. Hepatic Gene Expression Changes in Pigs Experimentally Infected with the Lung Pathogen *Actinobacillus Pleuropneumoniae* as Analysed with an Innate Immunity Focused Microarray. *Innate Immun.* 2010; 16:343–353. [PubMed: 19710094]
44. Olsen HG, Skovgaard K, Nielsen OL, Leifsson PS, et al. Organization and Biology of the Porcine Serum Amyloid A (SAA) Gene Cluster: Isoform Specific Responses to Bacterial Infection. *PLoS One.* 2013; 8:e76695. [PubMed: 24146912]
45. Soler L, Molenaar A, Merola N, Eckersall PD, et al. Why Working with Porcine Circulating Serum Amyloid A is a Pig of a Job. *J Theor Biol.* 2013; 317:119–125. [PubMed: 23073471]
46. Sung HJ, Ahn JM, Yoon YH, Rhim TY, et al. Identification and Validation of SAA as a Potential Lung Cancer Biomarker and its Involvement in Metastatic Pathogenesis of Lung Cancer. *J Proteome Res.* 2011; 10:1383–1395. [PubMed: 21141971]
47. Uhlar CM, Burgess CJ, Sharp PM, Whitehead AS. Evolution of the Serum Amyloid A (SAA) Protein Superfamily. *Genomics.* 1994; 19:228–235. [PubMed: 8188253]
48. Jacobsen S, Niewold TA, Halling-Thomsen M, Nanni S, et al. Serum Amyloid A Isoforms in Serum and Synovial Fluid in Horses with Lipopolysaccharide-Induced Arthritis. *Vet Immunol Immunopathol.* 2006; 110:325–330. [PubMed: 16337010]
49. McDonald TL, Larson MA, Mack DR, Weber A. Elevated Extrahepatic Expression and Secretion of Mammary-Associated Serum Amyloid A 3 (M-SAA3) into Colostrum. *Vet Immunol Immunopathol.* 2001; 83:203–211. [PubMed: 11730930]
50. Berg LC, Thomsen PD, Andersen PH, Jensen HE, Jacobsen S. Serum Amyloid A is Expressed in Histologically Normal Tissues from Horses and Cattle. *Vet Immunol Immunopathol.* 2011; 144:155–159. [PubMed: 21783263]
51. Chang WC, Chen CH, Cheng WTK, Ding ST. The Effect of Dietary Docosahexaenoic Acid Enrichment on the Expression of Porcine Hepatic Genes. *Asian-Aust J Anim Sci.* 2007; 20:768–774.
52. Soler L, Luyten T, Stinckens A, Buys N, et al. Serum Amyloid A3 (SAA3), Not SAA1 Appears to be the Major Acute Phase SAA Isoform in the Pig. *Vet Immunol Immunopathol.* 2011; 141:109–115. [PubMed: 21439655]
53. Kjelgaard-Hansen M, Jacobsen S. Assay Validation and Diagnostic Applications of Major Acute-Phase Protein Testing in Companion Animals. *Clin Lab Med.* 2011; 31:51–70. [PubMed: 21295722]
54. Pihl TH, Andersen PH, Kjelgaard-Hansen M, Morck NB, Jacobsen S. Serum Amyloid A and Haptoglobin Concentrations in Serum and Peritoneal Fluid of Healthy Horses and Horses with Acute Abdominal Pain. *Vet Clin Pathol.* 2013; 42:177–183. [PubMed: 23577834]
55. Jacobsen S, Thomsen MH, Nanni S. Concentrations of Serum Amyloid A in Serum and Synovial Fluid from Healthy Horses and Horses with Joint Disease. *Am J Vet Res.* 2006; 67:1738–1742. [PubMed: 17014325]
56. Horadagoda NU, Knox KM, Gibbs HA, Reid SW, et al. Acute Phase Proteins in Cattle: Discrimination between Acute and Chronic Inflammation. *Vet Rec.* 1999; 144:437–441. [PubMed: 10343375]
57. Soler, L.; Cerón, JJ.; Gutiérrez, A.; Lecchi, C., et al. Rodrigues, P.; Eckersall, D.; de Almeida, A., editors. Wageningen Academic Publishers; 2012. p. 106-110.
58. de Jesus Rodriguez B, Chevalere C, Henry G, Molle D, et al. Identification in Milk of a Serum Amyloid A Peptide Chemoattractant for B Lymphoblasts. *BMC Immunol.* 2009; 10 4-2172-10-4.
59. Mack DR, Michail S, Wei S, McDougall L, Hollingsworth MA. Probiotics Inhibit Enteropathogenic *E. Coli* Adherence in Vitro by Inducing Intestinal Mucin Gene Expression. *Am J Physiol.* 1999; 276:G941–50. [PubMed: 10198338]
60. Dai D, Nanthkumar NN, Newburg DS, Walker WA. Role of Oligosaccharides and Glycoconjugates in Intestinal Host Defense. *J Pediatr Gastroenterol Nutr.* 2000; 30(Suppl 2):S23–33. [PubMed: 10749398]
61. Mack DR, McDonald TL, Larson MA, Wei S, Weber A. The Conserved TFLK Motif of Mammary-Associated Serum Amyloid A3 is Responsible for Up-Regulation of Intestinal MUC3 Mucin Expression in Vitro. *Pediatr Res.* 2003; 53:137–142. [PubMed: 12508093]

62. Larson MA, Wei SH, Weber A, Mack DR, McDonald TL. Human Serum Amyloid A3 Peptide Enhances Intestinal MUC3 Expression and Inhibits EPEC Adherence. *Biochem Biophys Res Commun.* 2003; 300:531–540. [PubMed: 12504116]
63. Bislev SL, Kusebauch U, Codrea MC, Beynon RJ, et al. Quantotypic Properties of QconCAT Peptides Targeting Bovine Host Response to *Streptococcus Uberis*. *J Proteome Res.* 2012; 11:1832–1843. [PubMed: 22256911]
64. Larson MA, Weber A, Weber AT, McDonald TL. Differential Expression and Secretion of Bovine Serum Amyloid A3 (SAA3) by Mammary Epithelial Cells Stimulated with Prolactin Or Lipopolysaccharide. *Vet Immunol Immunopathol.* 2005; 107:255–264. [PubMed: 15996754]
65. Weber A, Weber AT, McDonald TL, Larson MA. *Staphylococcus Aureus* Lipotechoic Acid Induces Differential Expression of Bovine Serum Amyloid A3 (SAA3) by Mammary Epithelial Cells: Implications for Early Diagnosis of Mastitis. *Vet Immunol Immunopathol.* 2006; 109:79–83. [PubMed: 16139367]
66. 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. [PubMed: 24696607]
67. Bennike TB, Carlsen TG, Ellingsen T, Bonderup OK, et al. Neutrophil Extracellular Traps in Ulcerative Colitis – A Proteome Analysis of Intestinal Biopsies. *Inflamm Bowel Dis.*
68. Mendoza JL, Abreu MT. Biological Markers in Inflammatory Bowel Disease: Practical Consideration for Clinicians. *Gastroenterol Clin Biol.* 2009; 33(Suppl 3):S158–73. [PubMed: 20117339]
69. Tibble JA, Sigthorsson G, Foster R, Scott D, et al. High Prevalence of NSAID Enteropathy as shown by a Simple Faecal Test. *Gut.* 1999; 45:362–366. [PubMed: 10446103]
70. Schoepfer AM, Trummler M, Seeholzer P, Seibold-Schmid B, Seibold F. Discriminating IBD from IBS: Comparison of the Test Performance of Fecal Markers, Blood Leukocytes, CRP, and IBD Antibodies. *Inflamm Bowel Dis.* 2008; 14:32–39. [PubMed: 17924558]
71. Zimmermann O, Li K, Zaczekiewicz M, Graf M, et al. C-Reactive Protein in Human Atherogenesis: Facts and Fiction. *Mediators of Inflammation.* 2014; 2014:Article ID 561428, 6.
72. Kaplan MH, Volanakis JE. Interaction of C-Reactive Protein Complexes with the Complement System. I. Consumption of Human Complement Associated with the Reaction of C-Reactive Protein with Pneumococcal C-Polysaccharide and with the Choline Phosphatides, Lecithin and Sphingomyelin. *J Immunol.* 1974; 112:2135–2147. [PubMed: 4151108]

Statement of significance of the study

The publicly available PeptideAtlas project has been essential for technological and scientific progress of human proteome research, but has so far hosted very few pig proteins and proteome datasets. We have recently collected data across several laboratories and many parallel projects, in order to increase the access to pig proteins and peptides. This atlas has now been made publicly available at www.peptideatlas.org. and includes data from 25 tissues and three body fluid types mapped to 7139 canonical proteins. This manuscript presents for the first time the contents of the Pig PeptideAtlas, and presents in particular the mining of acute phase proteins across a wide range of porcine tissues and body fluid.

We believe that this resource will be of great value to researchers from for both veterinary and human biomedical sciences

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Overview of sample type representations in the Pig PeptideAtlas

	Synovial fluid	Plasma	Milk (colostrum)	Skeletal muscles	Gut cross-section	Liver	Heart	Lung	Neural tissues	Retina	Testis	Gut epithelium (ileum)	Gut epithelium (jejunum)	Spleen	Kidney	Artery	Tongue
Number of sample representations	3	1	20	99	33	74	12	5	32	1	10	8	12	3	2	1	1

Table 2

PeptideAtlas representations of a panel of porcine acute phase proteins in the different sample types

	Synovial fluid	Plasma	Milk (colostrum)	Skeletal muscles ^a	Gut cross-section	Liver	Heart	Lung	Neural tissues ^d	Retina	Testis	Gut epithelium (ileum)	Gut epithelium (jejunum)	Spleen	Kidney	Artery	Tongue
Haptoglobin	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Alpha-2-HS-glycoprotein	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Apolipoprotein A-I	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Serum albumin	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
pIg MAP1	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Alpha-1-antitrypsin	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Transthyretin	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Serum amyloid A	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
C-reactive protein	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

^aSkeletal muscles and neural tissues represent a panel of samples of different origin within these types of tissues.

Table 3
Total PeptideAtlas mappings of SAA isoform-unique peptides and their observations

Isoform	UniProt accession	Total peptide observations	Isoform-unique peptide	Observations in PA ^a			
				Colostrum	Skeletal muscle	Synovial fluid (SF)	Plasma
SAA2	F1S9B9	95	QWLSFLGEAYEGAK	0	0	6	3
SAA3	Q2HXZ9	30	HGDSGHGVEDSRADQAANAWGR	7	0	0	0
SAA4	F1S9B8	56	IISNVGEYFQGLLQYLGSSSREEDQVSNR	0	0	42	0
SAA4	F1S9B8	56	EAVQGASDLWR	0	3	2	1
SAA4	F1S9B8	56	EANYQNSGR	0	0	3	0
SAA4	F1S9B8	56	GPGGIWAAK	0	3	2	0

^aNumbers refer to observations of the isoform-unique peptide (inclusive miss-cleavages). Only sample types in which the peptides are observed are shown.