



Published in final edited form as:

Proteomics. 2014 March ; 14(6): 763–773. doi:10.1002/pmic.201300398.

The Equine PeptideAtlas – a resource for developing proteomics-based veterinary research

Louise Bundgaard^{1,3}, Stine Jacobsen¹, Mette Aamand Sørensen¹, Zhi Sun², Eric W. Deutsch², Robert L. Moritz², and Emøke Bendixen^{3,*}

¹Department of Large Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

²Institute for Systems Biology, Seattle, WA, USA

³Department of Molecular Biology and Genetics, Faculty of Science and Technology, Aarhus University, Denmark

Abstract

Progress in mass spectrometry-based methods for veterinary research and diagnostics is lagging behind compared to the human research, and proteome data of domestic animals is still not well represented in open source data repositories. This is particularly true for the equine species. Here we present a first Equine PeptideAtlas encompassing high-resolution tandem mass spectrometry analyses of 51 samples representing a selection of equine tissues and body fluids from healthy and diseased animals. The raw data were processed through the Trans-Proteomic Pipeline to yield high quality identification of proteins and peptides. The current release comprises 24,131 distinct peptides representing 2636 canonical proteins observed at false discovery rates of 0.2 % at the peptide level and 1.4 % at the protein level. Data from the Equine PeptideAtlas are available for experimental planning, validation of new datasets, and as a proteomic data mining resource. The advantages of the Equine PeptideAtlas are demonstrated by examples of mining the contents for information on potential and well-known equine acute phase proteins, which have extensive general interest in the veterinary clinic. The extracted information will support further analyses, and emphasises the value of the Equine PeptideAtlas as a resource for the design of targeted quantitative proteomic studies.

Keywords

Acute phase proteins; Animal proteomics; Equine; PeptideAtlas; Proteotypic peptides

1 Introduction

The application of mass spectrometry (MS) based methods in veterinary research has only quite recently gained interest, and this is particularly true for equine studies [1-3]. Progress

*Correspondence: Emøke Bendixen, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10C, 8000 Aarhus C, Denmark, ebx@mb.au.dk.

The author's declare no conflict of interest.

in this field is greatly hampered by the limited size of the veterinary research domain, and by the fact that relatively few veterinary science laboratories have access to advanced MS instrumentation. It is nevertheless becoming clear that MS based approaches offers hitherto unmatched opportunities to identify, characterise, and quantify proteins, and circumvents the need for validated antibodies [4].

Successful planning of MS based experiments and progress in proteomics science is enabled by access to public repositories of MS data such as PeptideAtlas [5], PRIDE [6], and Global Proteome Machine Database [7], but the mammalian genus *Equus* is not yet represented well in any of the publicly available databases. The PeptideAtlas repository (<http://www.peptideatlas.org/>) provides a large-scale assembly of observed and validated MS derived data that is uniformly compiled through the use of the Trans-Proteomic Pipeline (TPP) [8] from a wide range of species, including both classic model organisms like the yeast *Saccharomyces cerevisiae* [5] and *Drosophila melanogaster* [9], but more recently also domesticated animal species like pig and cattle [10]. The raw MS data are processed through TPP to yield high quality identifications, generally with 1 % false discovery rate (FDR) at the protein level. Data in PeptideAtlas are freely available for the research community, for experimental planning, for validation of new datasets, and as a proteome data mining resource. The features are particularly useful in the design of targeted proteomic experiments, because the interface provides query tools that assist in evaluating candidate peptides suited for targeted quantitative proteotypic analyses, from the peptide repository of previously annotated MS experiments [5].

Here we present the first build of an Equine PeptideAtlas and demonstrate the use of this atlas by mining it for information about 3 commonly known equine acute phase proteins (APPs). APPs are a functionally heterogeneous group of proteins with the common feature that their concentrations change by at least 25 % during inflammation [11], which makes this group of proteins particularly well suited for monitoring infectious and inflammatory diseases. More than 30 APPs have been described in humans [12], but the APP response to inflammation is known to vary considerably between species [13] and so far only very few proteins have been observed to have APP properties in the horse. The Equine PeptideAtlas provides an important resource for establishing methods for future analyses of not only APPs, but also a wide range of less known proteins with relevance for scientific and clinical applications in the horse, and opens new avenues of research for advanced insight into the equine proteome. Our aim with this communication is to present contents, features, and workflow for extracting knowledge from the Equine PeptideAtlas.

2 Materials and Methods

2.1 Sample processing

Samples included a selection of equine tissues and body fluids from healthy and diseased animals (Table 1). The samples were either obtained with permission from the Danish Animal Experiments Inspectorate (license no. 2010/561–1882), or from horses dedicated to research after euthanasia. After collection, tissue samples were washed in PBS, snap frozen in liquid nitrogen, and stored at -80 °C until further processing. Tissue samples were homogenised in TES-buffer (10 mM tris, 1 mM EDTA, 0.25 M sucrose) for 3 × 20 sec and

30 Hz frequency (TissueLyser II; Qiagen, Hilden, Germany) followed by centrifugation (3,000 g at 4 °C for 30 min) to isolate the supernatant for further processing. Protein concentrations of supernatants and body fluids were determined using the Pierce BCA protein assay kit (Thermo Scientific, Waltham, Massachusetts) with BSA as standard, according to the manufacturer's protocol (<http://www.piercenet.com/instructions/2161296.pdf>). Protein (120 µg) was precipitated using ice-cold acetone, samples centrifuged (15,000 g, 10 min), and the pellet re-suspended in 20 µL of 0.5 M triethylammonium bicarbonate buffer, pH 8.5 (AB SCIEX, Framingham, Massachusetts). Proteins were denatured with 1 µL of 2 % SDS denaturant (AB SCIEX, Framingham, Massachusetts) and disulfide bonds reduced with 2 µL of 50 mM tris-(2-carboxyethyl)phosphine reagent (AB SCIEX, Framingham, Massachusetts) for 1 h at 60 °C. Cysteines were blocked with 1 µL 200 mM methyl methanethiosulfonate (AB SCIEX, Framingham, Massachusetts) for 10 min at room temperature, and samples were digested overnight with 10 µL of a 1 µg/µL trypsin stock solution (AB SCIEX, Framingham, Massachusetts). Each sample was passed through a 0.2-µm centrifuge filter (VWR, Radnor, Pennsylvania) at 10,000 g for 10 min and vacuum dried.

2.2 Mass spectrometry analyses

Samples were analysed using one of three different approaches (Table 1). All samples except sample ID 7933, 7940, 7944, 7952, 7960 and 7969 were dissolved in a 0.03 % formic acid, 5 % acetonitrile solution and a volume corresponding to 50 µg of protein was fractionated by strong cation exchange (SCX) chromatography using an Agilent 1100 Series capillary HPLC equipped with a Zorbax Bio-SCX Series II, 0.8_50 mm column (Agilent Technologies, Santa Clara, California). Peptides were eluted with a gradient of sodium chloride increasing from 0-100 % in 65 min. Fractions were collected every 1 min for 60 min and were subsequently pooled into ten subsamples.

Samples analysed with the AB SCIEX Qstar Elite mass spectrometer (Table 1) were loaded onto a ReproSil-Pur C18-A1 precolumn (ID 100 µm × 2 cm, 5 µm, 120 Å; Thermo Scientific, Waltham, Massachusetts) and separated on an analytical ReproSil-Pur C18-A2 column (ID 75 µm × 10 cm, 3 µm, 120 Å; Thermo Scientific, Waltham, Massachusetts) using an Proxeon EASY- nLC (Thermo Scientific, Waltham, Massachusetts). The eluted peptides were sprayed through a nanospray emitter (PicoTip Emitter, SilicaTip, no coating, OD 360 µm, ID 20 µm; New Objective, Woburn, Massachusetts) directly into the Qstar Elite mass spectrometer (AB SCIEX, Framingham, Massachusetts). Separation was achieved by running a 73 min gradient from 0 % to 38 % solvent B (0.1 % formic acid, 90 % acetonitrile) followed by a 10 min wash in 100 % solvent B at a flow rate of 300 nL/min. An information dependent acquisition method was employed to automatically run experiments, operated under Analyst QS 2.0.

Samples analysed with the TripleTOF 5600 were, following micro-purification (StageTips C18, 20 µL tips; Thermo Scientific, Waltham, Massachusetts), dissolved in 0.1 % formic acid and processed on an EASY-nLC II (Thermo Scientific, Waltham, Massachusetts) with a Biosphere C18 precolumn (ID 100 µm × 2 cm, 5 µm, 120 Å; NanoSeparation, Nieuwkoop, Netherlands) and separated on an in-house packed C-18 analytical column (3 µm ReproSil-

Pur C18-AQ material packed in a PicoTip emitter, ID 75 $\mu\text{m} \times 10\text{ cm}$; New Objective, Woburn, Massachusetts) coupled in-line to a TripleTOF 5600 mass spectrometer (AB SCIEX, Framingham, Massachusetts). The samples were separated at a flow rate of 250 nL/min using a 50 min gradient from 5 % to 35 % solvent B followed by 10 min wash in 100 % solvent B with an information dependent acquisition method, operated under Analyst TF 1.5.1.

2.3 Data processing

The processing of the equine MS data followed the general workflow described in Deutsch *et al.* 2008 [16] using the TPP [8] as the main component. First, the vendor .wiff files from both the TripleTOF and Qstar Elite were converted to the mzML format [17] using the AB SCIEX converter version 1.3 beta. The sequence database used for searching was compiled as a non-redundant union of 45,741 equine sequences from UniProt, Ensembl, and UniGene. Unnamed proteins were annotated using Blast2GO (<http://www.blast2go.com>). The reference sequence database is available at ftp://ftp.peptideatlas.org/pub/PeptideAtlas/Repository/SearchDatabase/HORSE_May2011_targetdecoy.fasta.

Next the mzML files were searched against the non-redundant union equine database with both SEQUEST [18] and X!Tandem with the k score [19, 20]. For TripleTOF files, the parent mass error was set to ± 20 ppm. For Qstar Elite data, the parent mass error was set to -2.1 to +4.1 Da for X!Tandem and ± 3.0 Da for SEQUEST. For both instruments, a static modification of +45.987721 on cysteine was used, and a variable modification of +15.994915 on methionine was used. X!Tandem also searches some other artifactual modifications by default. 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 [20] and then the output of each search engine was modelled separately with the PeptideProphet [21] TPP tool. The search results from the two engines were combined with the iProphet [22] TPP tool for each experiment. Using the PeptideAtlas processing pipeline, all of the iProphet results were filtered at a variable probability threshold to maintain a constant FDR threshold for each experiment. All filtered data were assessed with the MAYU software [23] to calculate decoy-based FDRs at the peptide-spectrum match (PSM), as well as at distinct peptide and protein levels. All the results were loaded into the Equine PeptideAtlas (current build 2013-04) and made available at <http://www.peptideatlas.org/>.

2.4 Searching the Equine PeptideAtlas

In the Equine PeptideAtlas the **Browse protein** function under the **Queries** menu was used to search for specific proteins or groups of proteins. A distinctive part of the protein name was entered in the **Description Constraint** window, and all other constraints were left unspecified. The links related to each retrieved accession were used to mine the PeptideAtlas for further information about the protein of interest and for its related peptides. A detailed instruction in the use of the PeptideAtlas is given in Farrah *et al.* 2011 [24].

3 Results and discussion

3.1 Overall summary

The Equine PeptideAtlas is currently the largest public peptide repository for the mammalian genus *Equus*. It represents an extensive collection of observed peptides and corresponding proteins in 16 tissues and 7 body fluids from the horse, and it provides solid support for further in-depth research of a broad range of proteins.

The overall summary of the Equine PeptideAtlas protein and peptide coverage of individual tissue and body fluids is shown in Table 1. Starting from 2.2 million MS/MS spectra from 51 experiments, we have incorporated 543,093 PSMs with an FDR of 0.025 % at the PSM level into the PeptideAtlas build, identifying 24,131 distinct peptide sequences with a peptide-level FDR of 0.20 %. These peptides map to 2636 canonical proteins with a protein-level FDR of 1.4 %. An additional 412 proteins are labelled as “possibly distinguished”, which means that there is at least one peptide of high confidence that can distinguish a canonical protein sequence from its closely related family of proteins. The current subset of 2636 canonical proteins cover approximately 13 % of the 20,449 predicted equine proteins, based on the latest prediction of coding sequences from the completed equine genome assembly (http://www.ensembl.org/Equus_caballus/Info/Annotation/).

The PeptideAtlas is a dynamic repository and the coverage of the equine proteome is aimed to expand with the addition of other types of equine tissue and body fluids, as well as with improved MS methods. This expectation is illustrated also in the current build by the relatively higher number of contributing spectra from samples analysed by the AB SCIEX TripleTOF 5600 instrument compared to samples analysed by the older AB SCIEX Qstar Elite instrument.

The reference database is a union of 45,741 equine sequences from three databases; hence, the number of sequences is considerably higher than the predicted 20,449 equine proteins. Although redundant sequences were deleted before searches were made, still many related, although not identical, sequences remain and those sequences mapping to the same protein form a protein group (Table 2). These may either reflect different protein species, potential point mutations or lack of consensus about the sequence. It should be noted that despite the fact that the current reference genome reports a higher number of sequences than what has been predicted to be expressed as equine proteins, this reference database does not yet cover the complete equine genome. Moreover, a substantial part of the sequences in the reference database are derived from transcripts, which further adds to redundancy in protein assignments, and to fragment variants of the same sequence. Future progress in equine genome and proteome coverage will greatly improve the coverage of proteins and peptides, and a reprocessing of the existing data will likely improve coverage of as yet unclassified proteins.

3.2 The Equine PeptideAtlas protein and peptide view

For each protein, a dynamic protein view page provides an overview of basic information about the protein, including information on sequence coverage and of total observed peptides. The table of observed peptides includes ranking of the peptides by the empirical

observability score (EOS) and suitability score (ESS). The EOS reflects the likelihood for observing the given peptide/protein in a complex sample, while the ESS represents a measure for how suited the given peptide is as a significant proteotypic detection of the given protein from where it originated. The protein view page also provides a graphic depiction of the abundance of each peptide in each sample, and a link to Cytoscape [25], which can provide further information of the observed peptides presence in other related proteins. Moreover, information is given as to which specific sample and tissue the protein observation was made, and a section with links to the samples in which the protein was observed.

For each observed peptide, a peptide view page summarises the available information including basic information, genome mapping, and the alignment to the proteins from where the peptide may originate. Moreover, information is provided about the spectrum that supports the identification of the peptide and a list of samples, in which the peptide has been observed.

3.3 Mining information on acute phase proteins

This release of the Equine PeptideAtlas provides a comprehensive resource for mining e.g. information about APPs of horses. It is well documented that protein regulation as a response to acute inflammatory stimuli is considerably different across species [26], and knowledge of the equine APPs and their patterns of regulation is still limited. Recent research has demonstrated how MS can be used to study equine APPs in equine disorders such as endometritis [27], poisoning [28], and gastric ulcerations [29]. Table 2 presents a summary of known as well as suggested APPs for horses, which have been repeatedly observed in the PeptideAtlas, but further studies involving healthy and inflamed tissues are needed to confirm whether all these proteins indeed possess APP properties. A quantitative selected reaction monitoring (SRM) study of non-inflamed and inflamed tissues and body fluids could provide information needed to confirm the acute phase properties of these proteins, and the Equine PeptideAtlas hereby provides a unique resource for mining the information needed to design such SRM based studies. The EOS and ESS given in PeptideAtlas support selection of the best suited proteotypic peptides for proteins of interest and the experimentally observed spectra helps to select transitions with consistent presence and high-intensity [24]. The equine APPs SAA, fibrinogen and haptoglobin are among the observed proteins and their representation in the equine atlas will be discussed in details below, with the aim to illustrate the use of the PeptideAtlas.

3.4 Serum amyloid A

SAA has gained much recent interest in the equine clinic. SAA concentrations increase quickly in response to inflammation with a large amplitude, and was shown to be a very sensitive marker of equine inflammation and tissue damage, e.g. post surgery [15, 30], in horses with abdominal conditions [14], in equine joint diseases [31, 32], and in horses with respiratory disorders [33, 34]. Hepatocytes are the main source of SAA synthesis, but extrahepatic production of SAA was demonstrated in horses [35-37] and other species [36, 38].

From this PeptideAtlas it can be observed that equine SAA refers to a protein group with fourteen non-redundant sequences. Assessment of the sequences, using the protein group sequence alignment function in PeptideAtlas, suggests a further division into two subgroups here represented by PeptideAtlas protein names Q9N0Y1 (SAA subgroup 1) and ENSECAP0000009324 (SAA subgroup 2) (Fig. 1). Indeed, the N-terminal amino acids spanning 1-37 differ between the two subgroups with 54 % identity between the two representing proteins, compared to 88 % identity for the rest of the sequences.

Sequence comparison to bovine SAA sequences (Fig. 1) indicates that equine SAA subgroup 1 and subgroup 2, respectively, has close similarity with the N-terminal sequences of bovine SAA produced hepatically (P35541) and extrahepatically (Q8SQ28). These findings are supported by results of McDonald *et al.* 2001 [36], and indicate that both hepatic and extrahepatic SAA protein isoforms are represented in the current Equine PeptideAtlas. However, since the two observed peptides for SAA subgroup 2 are also contained in the SAA subgroup 1 sequence (Fig. 1, supplementary file 1)) and at the same time, unique peptides have not yet been observed for SAA subgroup 2, we currently lack specific evidence for distinguishing the SAA subgroup 2 in the Equine PeptideAtlas build. In the horse, extrahepatic SAA synthesis has previously been detected in various equine tissues, such as the mammary gland [36] in joints with septic arthritis [35], and the uterus [37, 39]. Because none of the tissues where extrahepatic SAA have previously been observed are represented in the current Equine PeptideAtlas, we can not yet confirm the presence of the SAA subgroup 2 in the analysed samples. This may also be due to the relatively low sensitivity of the data dependent acquisition method by which the protein analyses were made. Differentiation between the two SAA subgroups can be achieved through SRM based methods targeting the proteotypic peptides located in the unique N-terminal segments of these two subgroups. This approach to specifically differentiate between SAA protein isoforms has so far only been described for the human SAA proteins [40]. The observed peptides for SAA were found in granulation tissue, interstitial fluid collected by microdialysis from a wound ('wound fluid'), necrotic caecum, obstructed jejunum, and in saliva from horses with systemic inflammation. In contrast, SAA was not detected in any of the 17 different tissues and fluids from healthy individuals. High SAA concentrations have previously been reported in horses with tissue trauma and inflammatory abdominal conditions [14, 15, 30], and the findings of the present study thus corroborate SAA as a marker of inflammation in the horse [11, 41].

3.5 Fibrinogen

Fibrinogen is widely used as a diagnostic marker of inflammation in the equine clinic [15, 42, 43]. It is a plasma glycoprotein composed of $\text{A}\alpha$, $\text{B}\beta$, and γ polypeptide chains, and it is converted to fibrin by the action of thrombin [44]. In the Equine PeptideAtlas fibrinogen chains were detected in most tissues and body fluids. Since fibrinogen is primarily synthesised in the liver, the localisation of fibrinogen in samples from outside the liver most likely is due to distribution through blood-circulation [44]. Fibrinogen is present in plasma of healthy horses, and in response to an inflammatory stimulus the plasma concentration increases 2-4 folds with a peak concentration at day 3-6 after the stimulus [15, 31, 45].

The Equine PeptideAtlas includes all three fibrinogen polypeptide chains, α , β , and γ , with sequence coverage at 49.8 %, 55.7 %, and 54.4 % respectively. Despite the fact that fibrinogen is widely regarded as a valuable diagnostic marker in the equine clinic, little is known about its sequence annotation. Fibrinogen is highly conserved throughout evolution, demonstrating close similarity between the equine and human sequences (Fig. 2) [46], and should be expected to also possess cross species conservation of functional domains. In humans only two short peptides distinguish fibrinogen and fibrin, namely, the fibrinopeptide A (FpA) and B (FpB) [47]. In the equine sequence gi|194208383 the position of the observed peptide EGEFLHEGGGVR corresponds to a segment of the human FpA, hence this peptide appears to be the only peptide that is suited for uniquely and unequivocally measuring fibrinogen levels in horse samples by MS. FpA is a widely used marker in human medicine and is considered to be an important and significant marker of prethrombotic and thrombotic states such as disseminated intravascular coagulation [48] and lung embolism [49]. It is also measured to monitor coagulation during pregnancy [50] and cancer [51], and it has been proposed to be a potential marker of inflammation in human peritonitis [52] and end-stage renal disease [53]. To our knowledge FpA has never been investigated in horses, but SRM based targeting of the potential fragment of equine FpA may be useful for future studies, and the Equine PeptideAtlas contains the information needed for designing and developing these assays.

Fibrinogen β chain did not contribute with any observed proteotypic peptides for fibrinogen, but the observed peptides cover a sequence corresponding to an interesting biomarker used in the human clinic, namely the β 15-42 (Fig. 2). This peptide is one of the earliest detectable peptides being released from the fibrin blood clot by the action of plasmin, and it has been suggested to play a role in angiogenesis [54] and inflammation [55], and it may protect the myocardium against ischemia-reperfusion injury [56]. The Equine PeptideAtlas contains information, which demonstrates that several peptides map to the equine sequence corresponding to the human β 15-42 (Fig. 2). These peptides were repeatedly observed in liver, granulation tissue, wound fluid, peritoneal fluid from horses with inflammatory abdominal conditions, and in follicular fluid from the ovary.

3.6 Haptoglobin

Haptoglobin is an APP, with serum concentrations increasing 5–10 fold as a result of inflammation and infection in horses [57] Lavoie-Lamoureux *et al.* 2012 [34] have found that haptoglobin is a marker of both acute and chronic systemic inflammation in equine heave, Pihl *et al.* 2013 [14] designated it as a potential marker in peritoneal fluid for inflammatory abdominal conditions, and Wolf *et al.* 2012 found that its concentration increased in uterine samples from mares upon endometritis [27]. Haptoglobin is a glycoprotein composed of two α and two β chains [58]. In the Equine PeptideAtlas the sequence coverage is 67.9 % and includes a total of 19 observed peptides with EOS = 0.3 only mapping to this protein group, and representing both the α and β chain. In humans, the α chain is represented by two different isoforms α -1 and α -2, and recent observations from studies of diabetes, coronary artery disease and HIV have suggested that patients with the α -2 type haptoglobin have a poor prognosis [59]. In ruminants the predicted amino acid sequence suggests a domain structure similar to that of human α -2 [60]. For other mammals

including the horse, the electrophoretic mobility of haptoglobin has been found to be similar to that of human α -1 [61].

Haptoglobin was observed in all tissue and body fluids analysed (Fig. 3), but with more consistent representation of haptoglobin peptides in samples from horses with local and systemic inflammation, such as wound fluid, granulation tissue, necrotic caecum, obstructed jejunum, as well as saliva and peritoneal fluid. These observations are consistent with that haptoglobin is classified as an APP in the horse [31, 57]. The graphic depiction also indicates a high content of haptoglobin in follicular fluid from the ovary. In humans, haptoglobin, together with other APPs, has been observed in higher concentrations in follicular fluid from mature oocytes than in serum [62], which is in concordance with the description of the mammalian ovulation as an inflammation-like event [63].

The proteotypic peptide of fibrinogen A α was also observed in follicular fluid from the equine ovary. Fibrinogen has been described as a component of the follicular fluid in horses [64] and is expressed in the granulosa cells of the bovine ovary [65]. Both haptoglobin and fibrinogen has been mentioned as potential markers for oocyte quality for human in-vitro fertilisation [66]. Assisted reproductive technology is of major importance for managing equine reproduction, and the Equine PeptideAtlas contains information to design new methods for further investigating markers relevant for this field.

4 Concluding remarks

The Equine PeptideAtlas described here provides a comprehensive collection of high-confidence peptide and protein identifications from MS analyses of a broad range of tissue and body fluid samples from the horse. This is a unique resource for use in biomarker discovery and for design of SRM assays. In this paper we have illustrated this relevance by mining for information on three examples selected from well-known equine APPs, namely, SAA, fibrinogen, and haptoglobin, and for candidate proteins with acute phase properties in other mammalian species. Our observations demonstrate hitherto unmatched possibilities for in-depth studies of equine APPs, which represent a resource for discovery of potential diagnostic markers, of which some are already in use for human clinical applications.

The Equine PeptideAtlas is a dynamic repository, which will continually grow in size and complexity, as more genome and proteome data becomes available, and as mapping of the equine reference proteome is completed and well annotated. This collection of information about identified protein and peptides, and all raw MS data are publicly accessible from www.peptideatlas.org.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The study was funded by a Danish Government PhD grant. Additional funding was generously provided by the following Danish funds: The Hartmanns Foundation, The Horse Levy Foundation, The Carlsberg Foundation for KVL, and The Raben-Levetzkauske Foundation. Zhi Sun, Eric W. Deutsch, and Robert L. Moritz are supported in

part with federal funds from the National Science Foundation MRI (grant No. 0923536), by the American Recovery and Reinvestment Act (ARRA) funds through National Institutes of Health grant number RC2 HG005805 from the National Human Genome Research Institute, the National Institute of General Medical Sciences under grant No. R01 GM087221 and 2P50 GM076547/Center for Systems Biology the EU FP7 grant 'ProteomeXchange' [grant number 260558], and by the Luxembourg Centre for Systems Biomedicine and the University of Luxembourg. The authors further wish to thank laboratory technician Dorte Thomassen for invaluable help with sample preparation and analysis, and Marius Codrea for help with design of the reference database.

Abbreviations

APPs	Acute phase proteins
EOS	Empirical observability score
ESS	Empirical suitability score
FDR	False discovery rate
FpA	Fibrinopeptide A
FpB	Fibrinopeptide B
PSM	Peptide spectrum match
SAA	Serum amyloid A
SCX	Strong cation exchange
SRM	Selected reaction monitoring
TPP	Trans-proteomic pipeline

References

1. Barton C, Beck P, Kay R, Teale P, Roberts J. Multiplexed LC-MS/MS analysis of horse plasma proteins to study doping in sport. *Proteomics*. 2009; 9:3058–3065. [PubMed: 19526555]
2. Bouwman FG, van Ginneken MM, Noben JP, Royackers E, et al. Differential expression of equine muscle biopsy proteins during normal training and intensified training in young standardbred horses using proteomics technology. *Comp. Biochem. Physiol. Part D. Genomics Proteomics*. 2010; 5:55–64. [PubMed: 20374942]
3. Clutterbuck AL, Smith JR, Allaway D, Harris P, et al. High throughput proteomic analysis of the secretome in an explant model of articular cartilage inflammation. *J. Proteomics*. 2011; 74:704–715. [PubMed: 21354348]
4. Saalmuller A, Lunney JK, Daubenberger C, Davis W, et al. Summary of the animal homologue section of HLDA8. *Cell. Immunol*. 2005; 236:51–58. [PubMed: 16198325]
5. Deutsch EW. The PeptideAtlas Project. *Methods Mol. Biol*. 2010; 604:285–296. [PubMed: 20013378]
6. 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]
7. Craig R, Cortens JP, Beavis RC. Open source system for analyzing, validating, and storing protein identification data. *J. Proteome Res*. 2004; 3:1234–1242. [PubMed: 15595733]
8. Deutsch EW, Mendoza L, Shteynberg D, Farrah T, et al. A guided tour of the Trans-Proteomic Pipeline. *Proteomics*. 2010; 10:1150–1159. [PubMed: 20101611]
9. 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]
10. 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]

11. Jacobsen S, Andersen PH. The acute phase protein serum amyloid A (SAA) as a marker of inflammation in horses. *Equine Vet Educ.* 2007; 19:36–46.
12. Gabay C, Kushner I. Acute-phase proteins and other systemic responses to inflammation. *N. Engl. J. Med.* 1999; 340:448–454. [PubMed: 9971870]
13. Petersen HH, Nielsen JP, Heegaard PM. Application of acute phase protein measurements in veterinary clinical chemistry. *Vet. Res.* 2004; 35:163–187. [PubMed: 15099494]
14. 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]
15. Jacobsen S, Nielsen JV, Kjelgaard-Hansen M, Toelboell T, et al. Acute phase response to surgery of varying intensity in horses: a preliminary study. *Vet. Surg.* 2009; 38:762–769. [PubMed: 19674420]
16. 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]
17. Deutsch EW. Mass spectrometer output file format mzML. *Methods Mol. Biol.* 2010; 604:319–331. [PubMed: 20013381]
18. Yates JR 3rd, Eng JK, McCormack AL, Schieltz D. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal. Chem.* 1995; 67:1426–1436. [PubMed: 7741214]
19. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics.* 2004; 20:1466–1467. [PubMed: 14976030]
20. Keller A, Eng J, Zhang N, Li XJ, Aebersold R. A uniform proteomics MS/MS analysis platform utilizing open XML file formats. *Mol. Syst. Biol.* 2005; 1:2005.0017. [PubMed: 16729052]
21. Keller A, Nesvizhskii AI, 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]
22. 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. *Mol. Cell. Proteomics.* 2011; 10:M111.007690. [PubMed: 21876204]
23. 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. *Mol. Cell. Proteomics.* 2009; 8:2405–2417. [PubMed: 19608599]
24. Farrah T, Deutsch EW, Aebersold R. Using the Human Plasma PeptideAtlas to study human plasma proteins. *Methods Mol. Biol.* 2011; 728:349–374. [PubMed: 21468960]
25. Shannon P, Markiel A, Ozier O, Baliga NS, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003; 13:2498–2504. [PubMed: 14597658]
26. Murata H, Shimada N, Yoshioka M. Current research on acute phase proteins in veterinary diagnosis: an overview. *Vet. J.* 2004; 168:28–40. [PubMed: 15158206]
27. Wolf CA, Maslchitzky E, Gregory RM, Jobim MI, Mattos RC. Effect of corticotherapy on proteomics of endometrial fluid from mares susceptible to persistent postbreeding endometritis. *Theriogenology.* 2012; 77:1351–1359. [PubMed: 22225686]
28. Moore RE, Knottenbelt D, Matthews JB, Beynon RJ, Whitfield PD. Biomarkers for ragwort poisoning in horses: identification of protein targets. *BMC Vet. Res.* 2008; 4:30–6148-4-30. [PubMed: 18691403]
29. Taharaguchi S, Nagano A, Okai K, Miyasho T, et al. Detection of an isoform of alpha(1)-antitrypsin in serum samples from foals with gastric ulcers. *Vet. Rec.* 2007; 161:338–342. [PubMed: 17827473]
30. Jacobsen S, Jensen JC, Frei S, Jensen AL, Thoenfer MB. Use of serum amyloid A and other acute phase reactants to monitor the inflammatory response after castration in horses: a field study. *Equine Vet. J.* 2005; 37:552–556. [PubMed: 16295934]
31. Hulten C, Gronlund U, Hirvonen J, Tulamo RM, et al. Dynamics in serum of the inflammatory markers serum amyloid A (SAA), haptoglobin, fibrinogen and alpha2-globulins during induced noninfectious arthritis in the horse. *Equine Vet. J.* 2002; 34:699–704. [PubMed: 12455841]

32. 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]
33. Hobo S, Niwa H, Anzai T. Evaluation of serum amyloid A and surfactant protein D in sera for identification of the clinical condition of horses with bacterial pneumonia. *J. Vet. Med. Sci.* 2007; 69:827–830. [PubMed: 17827889]
34. Lavoie-Lamoureux A, Leclere M, Lemos K, Wagner B, Lavoie JP. Markers of systemic inflammation in horses with heaves. *J. Vet. Intern. Med.* 2012; 26:1419–1426. [PubMed: 22925172]
35. 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]
36. 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]
37. 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]
38. Kjelgaard-Hansen M, Christensen MB, Lee MH, Jensen AL, Jacobsen S. Serum amyloid A isoforms in serum and synovial fluid from spontaneously diseased dogs with joint diseases or other conditions. *Vet. Immunol. Immunopathol.* 2007; 117:296–301. [PubMed: 17451811]
39. Christoffersen M, Baagoe CD, Jacobsen S, Bojesen AM, et al. Evaluation of the systemic acute phase response and endometrial gene expression of serum amyloid A and pro- and anti-inflammatory cytokines in mares with experimentally induced endometritis. *Vet. Immunol. Immunopathol.* 2010; 138:95–105. [PubMed: 20728224]
40. Sung HJ, Jeon SA, Ahn JM, Seul KJ, et al. Large-scale isotype-specific quantification of Serum amyloid A 1/2 by multiple reaction monitoring in crude sera. *J. Proteomics.* 2012; 75:2170–2180. [PubMed: 22300576]
41. 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]
42. Borges AS, Divers TJ, Stokol T, Mohammed OH. Serum iron and plasma fibrinogen concentrations as indicators of systemic inflammatory diseases in horses. *J. Vet. Intern. Med.* 2007; 21:489–494. [PubMed: 17552456]
43. Copas VE, Durham AE, Stratford CH, McGorum BC, et al. In equine grass sickness, serum amyloid A and fibrinogen are elevated, and can aid differential diagnosis from non-inflammatory causes of colic. *Vet. Rec.* 2013; 172:395. [PubMed: 23428423]
44. Weisel JW. Fibrinogen and fibrin. *Adv. Protein Chem.* 2005; 70:247–299. [PubMed: 15837518]
45. Andersen SA, Petersen HH, Ersboll AK, Falk-Ronne J, Jacobsen S. Vaccination elicits a prominent acute phase response in horses. *Vet. J.* 2012; 191:199–202. [PubMed: 21371917]
46. Sobel JH, Thibodeau CA, Kolks MA, Canfield RE. Isolation and partial structural characterization of an equine fibrinogen CNBr fragment that exhibits immunologic cross-reactivity with an A alpha-chain cross-linking region of human fibrinogen. *Biochemistry.* 1990; 29:8907–8916. [PubMed: 2271566]
47. Riedel T, Suttner J, Brynda E, Houska M, et al. Fibrinopeptides A and B release in the process of surface fibrin formation. *Blood.* 2011; 117:1700–1706. [PubMed: 21106983]
48. Boisclair MD, Ireland H, Lane DA. Assessment of hypercoagulable states by measurement of activation fragments and peptides. *Blood Rev.* 1990; 4:25–40. [PubMed: 2182146]
49. Gando S, Tedo I. Diagnostic and prognostic value of fibrinopeptides in patients with clinically suspected pulmonary embolism. *Thromb. Res.* 1994; 75:195–202. [PubMed: 7974393]
50. Hellgren M. Hemostasis during normal pregnancy and puerperium. *Semin. Thromb. Hemost.* 2003; 29:125–130. [PubMed: 12709915]
51. Gouin-Thibault I, Samama MM. Laboratory diagnosis of the thrombophilic state in cancer patients. *Semin. Thromb. Hemost.* 1999; 25:167–172. [PubMed: 10357084]

52. Schoffel U, Zeller T, Lausen M, Ruf G, Farthmann EH. Monitoring of the inflammatory response in early peritonitis. *Am. J. Surg.* 1989; 157:567–572. [PubMed: 2786351]
53. Nelson K, Thethi I, Cunanan J, Hoppensteadt D, et al. Upregulation of surrogate markers of inflammation and thrombogenesis in patients with ESRD: pathophysiologic and therapeutic implications. *Clin. Appl. Thromb. Hemost.* 2011; 17:302–304. [PubMed: 21159713]
54. Erban JK, Wagner DD. A 130-kDa protein on endothelial cells binds to amino acids 15–42 of the B beta chain of fibrinogen. *J. Biol. Chem.* 1992; 267:2451–2458. [PubMed: 1370821]
55. Skogen WF, Senior RM, Griffin GL, Wilner GD. Fibrinogen-derived peptide B beta 1–42 is a multidomained neutrophil chemoattractant. *Blood.* 1988; 71:1475–1479. [PubMed: 3359049]
56. Petzelbauer P, Zacharowski PA, Miyazaki Y, Friedl P, et al. The fibrin-derived peptide Bbeta15–42 protects the myocardium against ischemia-reperfusion injury. *Nat. Med.* 2005; 11:298–304. [PubMed: 15723073]
57. Taira T, Fujinaga T, Okumura M, Yamashita K, et al. Equine haptoglobin: isolation, characterization, and the effects of ageing, delivery and inflammation on its serum concentration. *J. Vet. Med. Sci.* 1992; 54:435–442. [PubMed: 1643167]
58. Wicher KB, Fries E. Evolutionary aspects of hemoglobin scavengers. *Antioxid. Redox Signal.* 2010; 12:249–259. [PubMed: 19650711]
59. Quaye IK. Haptoglobin, inflammation and disease. *Trans. R. Soc. Trop. Med. Hyg.* 2008; 102:735–742. [PubMed: 18486167]
60. Morimatsu M, Syuto B, Shimada N, Fujinaga T, et al. Isolation and characterization of bovine haptoglobin from acute phase sera. *J. Biol. Chem.* 1991; 266:11833–11837. [PubMed: 1904872]
61. Schwantes AR, Tondo CV. Preliminary data on haptoglobin in vertebrates. *Braz J Med Biol Res.* 1969; 2:105–107.
62. Angelucci S, Ciavardelli D, Di Giuseppe F, Eleuterio E, et al. Proteome analysis of human follicular fluid. *Biochim. Biophys. Acta.* 2006; 1764:1775–1785. [PubMed: 17067859]
63. Richards JS, Russell DL, Ochsner S, Espey LL. Ovulation: new dimensions and new regulators of the inflammatory-like response. *Annu. Rev. Physiol.* 2002; 64:69–92. [PubMed: 11826264]
64. Yamada M, Gentry PA. The hemostatic profile of equine ovarian follicular fluid. *Thromb. Res.* 1995; 77:45–54. [PubMed: 7701476]
65. Parrott JA, Whaley PD, Skinner MK. Extrahepatic expression of fibrinogen by granulosa cells: potential role in ovulation. *Endocrinology.* 1993; 133:1645–1649. [PubMed: 8404605]
66. Estes SJ, Ye B, Qiu W, Cramer D, et al. A proteomic analysis of IVF follicular fluid in women <or=32 years old. *Fertil. Steril.* 2009; 92:1569–1578. [PubMed: 18980758]

**Figure 1.**

A) Alignment of sequences assumed to represent two subgroups of equine SAA. EQ_1 represents SAA subgroup 1 (Q9N0Y1), EQ_2 represents SAA subgroup 2 (ENSECAP0000009324). The similarity of amino acids to the left of the arrow is 54%, compared to 88% for the rest of the sequences. Observed peptides are marked in bold. B1) Alignment of the equine SAA subgroup 1 with hepatically produced bovine SAA (BO_1) (P35541). The similarity of amino acids to the left of the arrow is 61%. B2) Alignment of the equine SAA subgroup 2 with extrahepatically produced bovine SAA (BO_2)(Q8SQ28). The similarity of amino acids to the left of the arrow is 81%. Signal peptides are marked in italic. For the horse they are presumed by comparison with bovine SAA. CLUSTAL 2.1 multiple sequence alignment (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) were used for alignment.

Fibrinopeptide A α

gi 194208383	(EQ)	<i>MFSMRI</i> FCLVLSVVG T-----IG EGEF LHEGGG V RGPRVVERQQSACK ETDWPFC SDE D W
P02671	(HUMAN)	<i>MFSMRI</i> VCLVLSVVGTA WTAD SGEGDFLAEGGGV R GPRVVERHQ SACK SDSWPFC S DE D W

Fibrinopeptide B β

ENSECAP0000008406	(EQ)	- <i>QRMVSWNFQKFK</i> TMKLLFLLLLCVSVVRSQALDYDHEVTFDAR SHRPLDKKREEPL SLK
P02675	(HUMAN)	<i>MKRMVSWSFHKL</i> KTMKLLLLLLCVFLVKSQGVNDNEEGFFSARGHR PLDKKREEA PSLR

ENSECAP0000008406	(EQ)	PAPPPISGPGYRPRPAK VAANKKKVER KAPDAGGC VHAD PD LGVL CPTGCQLQD TLVKQE
P02675	(HUMAN)	PAPPPISGGYR ARPAKAAATQ KKVER KAPDAGGC LHAD PD LGVL CPTGCQLQE ALLQQE

Fibrinopeptide γ

gi 194208381	(EQ)	<i>MSWSLYPRNLILYFY</i> TLFLLSSTCLAYVATRDNCCILDERFGSYCPTTCGIADFLSNYQT
P02679	(HUMAN)	<i>MSWSLHPRNLILYFY</i> ALLFLSSTCVAYVATRDNCCILDERFGSYCPTTCGIADFLSTYQT

Figure 2.

Comparison of N-terminal sequences of equine (EQ) and human fibrinopeptide A α , B β , and γ . The known human sequence annotations are given below the human sequence. Observed equine peptides are shown in bold.

Signal peptides are marked in italic. For the horse they are presumed by comparison with the human sequences. CLUSTAL 2.1 multiple sequence alignment (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) were used for alignment.



Figure 3. Graphical depiction of the most highly observed peptides for haptoglobin in each sample. The colour intensity of the square reflects the abundance of observations in the given sample, with white colour as no observation and black as very high abundance.

Table 1

An overview of the tissue and body fluid protein and peptide coverage in the Equine PeptideAtlas.

Sample	Analysis approach and sample ID number		Number of peptides	Number of proteins	Number of unique proteins
Tissue	Qstar Elite ^d	TripleTOF 5600 ^d			
Diaphragm	7934, 7938		493	2644	66
Wound biopsy	7947, 7953, 7954	7948	601	3062	64
Heart	7970		392	2000	16
Spleen	7924, 7935	7925	1989	11964	921
Lymph node	7956, 7958		661	2159	53
Kidney	7932, 7966		1101	4755	158
Ovary follicle wall	7968		279	894	2
Tongue	7943		362	1473	6
Jejunum	7959, 7973	7940 ^a	977	4254	74
Jejunum - obstructed	7936, 7945		514	2653	32
Ileum	7942, 7950		774	2513	35
Caecum	7931, 7967	7952 ^a	705	3031	8
Caecum - necrotic	7930, 7964	7933 ^a , 7974	412	1900	20
Colon	7926		909	3472	21
Liver	7963, 7972	7939	1341	7897	104
Fat	7946, 7962	7944 ^a	611	2830	26
Fluid					
Synovial fluid	7971		136	841	5
Wound fluid ^e	7929, 7941		285	1362	13
Serum	7961, 7965	7960 ^a	175	1473	18
Ovary follicle fluid	7927, 7957		152	970	11
Saliva	7951, 7955		172	525	17
Saliva - systemic inflammation	7928, 7949		272	1114	60
Peritoneal fluid - peritonitis	7937	7969 ^a	153	1164	6
Total		51 samples	24,131^b	2636^c	

^aNot fractionated by strong cation exchange before analysis^bNumber of distinct peptides with false discovery rate of 0.2%^cNumber of canonical proteins with false discovery rate of 1.4%^dAB SCIEX, Framingham, Massachusetts, USA^eWound interstitial fluid obtained by microdialysis

Table 2

Well-known and potential acute phase proteins for the equine species found by mining the Equine PeptideAtlas. Each protein description corresponds to a protein group, designated by accession number. A protein group includes related sequences and may reflect different protein species, potential point mutations or lack of consensus about the sequence. The number of sequence-unique protein sequences in each group is given.

PeptideAtlas Protein Description	PeptideAtlas Protein Group	Sequence unique proteins in group
Alpha-1-acid-glycoprotein (Orosomuroid)	ENSECAP00000005812	1
Alpha-1-antichymotrypsin (serpin-A-3)	ENSECAP00000011686	2
Alpha-1-antitrypsin (serpin-A-1)	B5BV13	22
Alpha-2-macroglobulin	gi 194211675	17
Angiotensinogen	ENSECAP00000013428	2
C1 inhibitor	ENSECAP00000008891	1
C4b-binding protein	ENSECAP00000011409	2
Ceruloplasmin	Q9N1X0	2
Complement C3	Q9XS82	9
Complement C4	ENSECAP00000007506	4
Complement C9	ENSECAP00000006188	2
C-reactive protein	ENSECAP00000022415	1
Factor B	ENSECAP00000010339	1
Ferritin heavy chain	Q8MIP0	8
Ferritin light chain	P02791	18
Fibrinogen A α	O97641	2
Fibrinogen B β	ENSECAP00000008406	3
Fibrinogen γ	O02682	5
Fibronectin	Q28377	7
Haptoglobin	ENSECAP00000019071	2
Hemopexin	ENSECAP00000014242	1
Inter- α -trypsin inhibitor	P04365	4
Inter- α -trypsin inhibitor H1	ENSECAP00000022393	2
Inter- α -trypsin inhibitor H2	A6P3C5	5
Inter- α -trypsin inhibitor H3	ENSECAP00000004389	3
Inter- α -trypsin inhibitor H4	gi 194221223	2
Mannose-binding lectin	ENSECAP00000015873	2
Plasminogen	P80010	6
Plasminogen activator inhibitor 1	ENSECAP00000013972	2
Prothrombin	gi 194217874	2
Secreted phospholipase A2	ENSECAP00000008204	2
Serum amyloid A (SAA)	P19857	14
Vitronectin	ENSECAP00000013752	2