Identification and Quantification of Proteins from Preparative Partition Chromatography and Peptides from Organic Extraction of Fetal Versus Adult Bovine Serum using Nano-Spray-LC-ESI-MS/MS

by

Zhuo Zhen Chen

B.Sc., Ryerson University, 2017

A thesis submitted to Ryerson University
in partial fulfillment of the requirements for the degree of
Master of Science
in the program of Molecular Science

Toronto, Ontario, Canada

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

Identification and Quantification of Proteins from Preparative Partition Chromatography and Peptides from Organic Extraction of Fetal Versus Adult Bovine Serum using Nano-Spray-LC-ESI-MS/MS

Zhuo Zhen Chen

Master of Science, Molecular Science, Ryerson University, 2019

ABSTRACT

Blood proteins communicate with many different cells, tissue and organs; perform key functions in the immune system and may be of particular biological complexity. One of the most widely used blood products in the laboratory is fetal bovine serum for cell culture. There are ethical and practical concerns regarding the use of fetal serum from animals and alternative serum-free replacements have been attempted using platelet lysates. Previous biochemical experiments have shown that FBS apparently contained factors such as alpha-feto protein (AFP) and insulin-like growth factors that may support the indefinite cell growth and division of certain cell lines. It is presumed that a set of as yet undefined growth factors transform cells growth resulting in rapid proliferation. Cultured Raw cells 264.7 in adult bovine serum multiplied slowly and differentiated into elongated cells with a dendritic shape, which died after the first few generations. On the contrary, in fetal bovine serum, cultured cells multiplied rapidly and formed many smaller cells with a rounded shape through many cell passages. Three independent batches of fetal bovine serum were tested on Raw cells 264.7 macrophages to confirm that they supported cell growth in culture compared to three independent batches of adult bovine serum. The intact proteins of each serum sample were separated by partition chromatography into 16 fractions with an increasing step gradient of salts over quaternary amine resin (proteomics). The endogenous peptides were precipitated with 90% of acetonitrile and extracted into 10 fractions

with a decreasing step gradient of acetonitrile in water (peptidomics). Trypsin digested intact proteins and endogenous peptides were then analyzed on a fresh C18 nano-HPLC column with random and independent sampling by LC-ESI-MS/MS. The fractionated mass spectra were identified with SEQUEST and X!TANDEM algorithms. Redundant use of MS/MS spectra were flirted out with the SQL Server system and the R statistical analysis system was used to perform Chi Square (X^2) analysis of frequency counts and ANOVA of the $\log_{10}$ precursor intensity results. Alpha-feto protein, fetal albumin, insulin, insulin like growth factors, platelet derived growth factors and proteins associated with HRAS/AKT growth pathway at the level of ligand, receptors, receptor associated enzyme and nucleic acid binding proteins including transcription factors were observed to be specifically enriched in fetal serum.

ACKNOWLEGMENTS

I would like to thank my supervisor Dr. John G. Marshall for giving me this opportunity to take my first steps into scientific research and allowing me to work in his lab. Thank you for your patience, understanding, and seeing potential in me as a scientist. I would also like to thank my committee members, Dr. Gupta and Dr. Sabatinos, who have taken their time to help me and provide useful suggestions for this project. Last but not least, I would like to thank my family and friends; I would not have been able to come this far without your support and understanding.

TABLE OF CONTENTS

LIST OF SELECTED ABBREVIATIONS

ABS: adult bovine serum

ACN: acetonitrile

AFP: alpha-feto protein

BSA: bovine serum of albumin

CID: collision-induced dissociation

DMEM: dolbecco modified eagle's medium

DMSO: dimethyl sulfoxide

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

FA: formic acid

FBS: fetal bovine serum

FGF: fibroblast growth factor

HAC: acetic acid

HSCs: haemopoietic stem cells

IgG: immunoglobulin

INS: insulin

LC-ESI-MS/MS: liquid chromatography –electrospray ionization – tandem mass spectrometry

MALDI-TOF: matrix-assisted laser desorption ionization –time of flight

m/z: mass to charge ratio

NBCS: newborn calf serum

PBS: phosphate buffered saline

PFA: para-formaldehyde

QA: quaternary amine

RBCs: red blood cells

RP –HPLC: reversed –phase high performance liquid chromatography

SDS-PAGE: sodium dodecyl sulfate –polyacrylamide gel electrophoresis

SQL Server: Structured Query Language

TGF: transforming growth factor

WBCs: white blood cells

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1.0 INTRODUCTION

1.1 Blood Proteins and Proteomics

Proteomics is the study of proteins and peptides' primary structure and quantification [1] . The proteome, which refers to all the possible proteins encoded by the genome, is neither as static nor as uniform as the genome [2]. The blood may contain the most diverse set of proteins since they communicate between all of the organs, tissues and cells. Blood is also a most significant source of therapeutic proteins, transfusion products and is commonly used for diagnostic testing [3]. Blood proteins serve many different functions such as lipid and hormone transporters, binding proteins and may regulate the function of the immune system. Similar blood proteins were observed in serum and plasma [4]. Serum is prepared after coagulation of blood cells by fibrin cleavage and polymerization. Plasma results when blood is collected with anti-coagulants such as EDTA and the cells are removed by low speed centrifugation without the activation of coagulation cascade [5]. The most effective separation for the selective purification of intact blood proteins prior to identification was partition chromatography which identified more proteins than electrophoresis [3]. The identification of proteins resolved by chromatography from tryptic peptides can be performed using electrospray ionization [6] and the peptide fragmentation spectra [7] from a linear quadrupole ion trap [8].

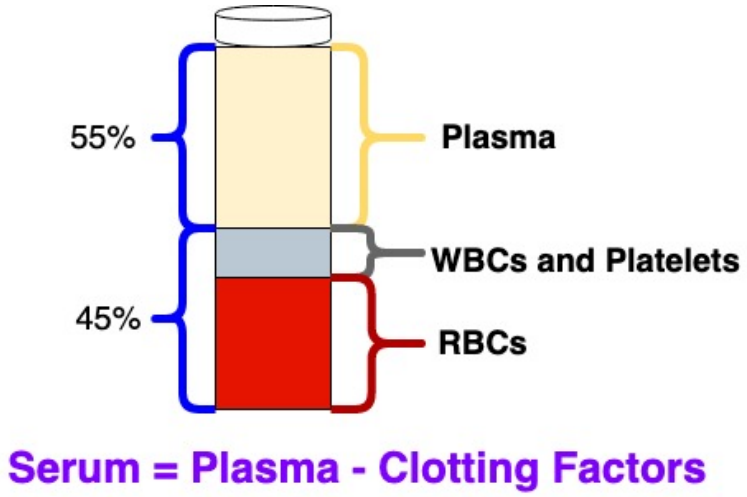


Figure 1.1. 1. Plasma versus serum of Human/Bovine blood. Plasma is treated with anti-coagulants and serum is treated after coagulation. Both plasma and serum do not contain WBCs, RBCs and platelets.

1.2 Adult Serum vs. Fetal Serum

Fetal serum is an essential supplementation of basal culture media for cell growth and proliferation [9]. Many animal cells can grow in a fetal serum-containing media such as Eagle's medium, while certain cell types require specific protein growth factors that are not present in serum [10]. The most widely used sera are from adult or fetal bovine origin [9], and the most widely used product manufactured from bovine blood is bovine serum for cell culture [11]. Bovine serum is the liquid fraction of clotted blood that contains a large number of nutritional factors as well as small molecules like amino acids, sugars, lipids, pyruvate and hormones that can increase cell growth [5].

Fetal serum is a mixture of thousands of protein factors needed for proliferation of cells in culture [10]. One of the major blood proteins is serum albumin, which accounts for 55% of blood proteins. Albumin is an important contributor to maintaining the osmotic pressure of plasma to

assist in the lipids and steroid hormones transportations [9]. Bovine serum of albumin (BSA) is the major component of the bovine serum, that together with transferrin is known to enhance growth in cultured cells [5]. Studies have shown that different serum constituents have distinct effects on cell cultures. Insulin alone, nerve growth factor, epidermal growth factor, insulin-like growth factor, fibroblast growth factor (FGF), platelet-derived growth factor, and transforming growth factor (TGF) [12-14] were found to increase proliferation but were inferior to effect of fetal serum [15] indicating there are as yet unknown growth factors in fetal serum. Thus, it is important to analyze the biological properties of adult bovine serum in comparison with fetal bovine serum for cell culture [11]. In general, FBS is an extremely complex mixture of a large number constituents, including both low and high molecular weight biomolecules with different growth-promoting and growth-regulating activities [9]. Previous studies showed that fetal serum contains more growth factors and has low immunoglobulin (IgG) content as well as lower levels of complements than adult serum (complement can interfere with immunoassays since they contain the undesired effect of lysing cells in culture). In addition, FBS provides essential compounds, such as hormones, binding and attachment factors that promote cell growth and maintain cellular activities [5,16].

1.2.1 Cell Growth in Adult and Fetal Serum

Within the tissues of intact animals, most cells tightly contact and interact specifically with other cells via various cellular junctions. In addition, the cells can contact the extracellular matrix, which a complex network that fills the gaps between cells with secreted proteins and carbohydrates. The extracellular matrix helps cells bind in tissue together while providing a lattice through which cells can move during early stage of animal differentiation. Most cultured

animal cells may require a surface to grow on in vivo when interacting with one another in an extracellular matrix surrounding. Many cell types can adhere to and grow on glass as they secret collagen and bind to the culture surface and function as a bridge between the cells [10]. For instance, the RAW 264.7 mouse macrophage was used in this experiment, which is a monocyte white blood cell that can differentiate into macrophage and adhere to the glass [17].

Cultured cells are useful models for studying cell growth and differentiation. When cells are removed from an embryo or an adult animal the adherent primary cells line may grow spontaneously in culture for a limited time before growth stops. Primary cultures may eventually die after many passages, even when provided with fetal serum [10]. For the past decades, FBS has been preferred for culturing cells over other bovine sera since it contains higher levels of nutrients, adhesion proteins and growth factors that promote cell growth [11], [18], [19]. Moreover, significantly lower levels of antibodies that reduce the risks posed by cross-reaction of antibodies in cell culture were found in FBS [11], [20].

1.3 The Importance of FBS

The variability of FBS results in the presence of unpredictable amounts of different bioactive components, including immunoglobulin, transcription factors and growth factors, which highly influence cell behavior [18]. Although there are several controversies about using animal serum in culture systems, the alternative use of other products for cell growth was not widely popularized. Most laboratories use FBS despite the high price and limited supply. The most common concern about FBS alternatives is the lack of solid evidence indicating that they exhibit equal performance as FBS. Studies have demonstrated that most cell lines were maintained in

FBS-containing serum while adult serum products have less effect on cell growth and proliferation. More importantly, the number of generations that certain cell lines had been cultured or whether these FBS alternatives serum would support long-term cell growth is still not known [21].

It remains possible that serum from different species may exhibit different effects on cell behavior in cultures [16]. These sera contain proteins that can adhere onto the material surface, which play crucial roles in subsequent cell behavior, including adhesion, spreading, and proliferation [22]. Some studies have shown that cells cultured in FBS exhibited a constrained morphology and did not extend as well as in adult bovine serum [21]. Cells in fetal serum tend to remain circular and spherical, which act like stem cells [16].

Fetal serum contains factors that can promote changes in gene expression [23], [24] and permit the indefinite culture of certain cell lines such as RAW 264.7 murine macrophages [25]. All of the factors in fetal serum that permit indefinite cell growth have yet to be determined [26]–[28]. Fetal serum or platelet lysates have also been shown to expand clonal cell lines and may have therapeutic applications towards human health [29]. There have been significant efforts to create cell growth media that contains the types of growth factors contained in natural serum for use in cell culture experiments [13]. There have been several attempts to compare fetal serum to adult serum from umbilical chords or other approaches [30]. The peptides and proteins specific to fetal serum may have therapeutic and basic biochemical importance [31]. The technology of identifying peptides from proteins using LC-ESI-MS/MS has been successfully applied to many cell and tissue types [32]–[35]. In this study, the application of LC-ESI-MS/MS to fetal versus

adult bovine serum was compared using partition chromatography of intact proteins and organic extraction of endogenous peptides.

1.4 Comparison of Step-wise Proteins versus Peptide Fraction

Pure substances can be measured with great sensitivity and log linearity using a mass spectrometer, and pre-fraction of proteins by chromatography shows even greater sensitivity results compared to crude serum. The use of partition chromatography of intact protein chains prior to digestion permitted the identification of the major blood proteins from tryptic peptides [3]. The analysis of serum or plasma by LC-ESI-MS/MS [36] requires that each sample must be partitioned into many fractions [37], [38], so that each enriched in a sub set of proteins or polypeptides prior to subsequent to nano-LC-ESI-MS/MS in order to achieve sufficient sensitivity to detect the many proteins that may be in pico-molar or lower concentrations [3], [39]. A comparison of intact proteins over many chromatography resins showed that quaternary amine resin identified the largest number of soluble proteins from intact serum proteins [39]. The use of quaternary amine ion exchange fractionation prior to digestion was effective for examining proteins released from cancer cells lines cultured in artificial media [37], [40]. A comparison of endogenous peptides purified using ultrafilters, ion exchange chromatography, C18, size exclusion chromatography and organic extraction showed that steps wise organic fractionation provided the best results [41], [42]. High signal to noise ratio of blood peptides is dependent on sample preparation to enrich certain proteins or peptides in each of the many sub-fractions, relieve competition and suppression of ionization, and thus achieve sensitivity [3], [38], [39].

1.5 Mass Spectrometry and Tandem Mass Spectrometry

Mass spectrometer (MS) is an ideal tool to quantify and identify the protein and complexes of blood in which abundant peptides may be detected with high intensity[39]. MS is also an analytical approach that is used to fractionate a complex mixture into fragments, measure its mass-to-charge ratio of ions [43]. The capacity of a MS to detect proteins at low concentrations in blood is critical to the discovery of new biomarkers of disease [39]. Furthermore, mass spectrometry is a diagnostic and cancer biomarker discovery tool; research has been done on discovering biomarkers for different types of cancers with this technology including ovarian, prostate, breast, bladder, renal etc [44]–[46].

A Tandem Mass Spectrometer (MS/MS) involves multiple steps of MS selection with fragmentation occurring in between the stages. In the first stage of a Tandem Mass Spectrometer, ions are formed in the ion source and then separated by mass to charge ratio (MS1-"Parent m/z") [43], [47]. Ions of a particular m/z, also known as precursor ions, are selected where CID (Collision-Induced Dissociation) process creates the fragment ions are separated and detected in the second stage of MS (MS2 or MS/MS- "Fragment m/z") [48]. Thus, precursor and product ions are both characterized independently by their m/z [49]. A single stage mass analysis allows for the identification of analytes of interest at which atmospheric pressure ionization produces mass spectra that can provide information of the molecular mass of the analytes. A two-stage mass (MS/MS) analysis allows for even more accurate compound identification, which it monitors how a parent ion fragments when exposed to an additional stage of ionization [50]. This is a very popular, reliable and well-established method of protein identification. Ion intensity is a measure of each m/z values in MS and then MS/MS fragments, which indicates how strong the

signal is to the detector [47]. Using MS/MS analysis, it is easier to quantify and identify target analytes in complex maxtices such us plants, animal tissues, soil, or human plasma [43], [47].

Mass Spectrometry has become invaluable to a broad range of fields and applications. For instance, mass spectrometry can be applied to proteomic studies, to determine protein's primary structure, function, folding and interactions [51]. MS can also identify peptides from the precursor and fragment masses and thus can be used to monitor enzyme reactions, chemical modifications and protein digestion [51], [52]. Furthermore, mass spectrometry can also be used for drug discovery, to determine significant structures of drugs and metabolites. Last but not least, mass spectrometry has a broad range of applications in genomic, environmental, geological, and forensic science studies. It has been suggested that MS can achieve much higher diagnostic sensitivity and specificity in comparison to the classical cancer biomarkers [44], [53].

1.5.1 RP-HPLC

Reversed-phase HPLC has also become an essential tool in the separation of proteins and peptides and it is able to separate peptides of nearly identical structure. Also, RP-HPLC is widely used in the biochemistry industry to characterize protein therapeutic products. The sensitivity of detection is a driving factor to develop HPLC columns of smaller diameter and for new types of stationary phases. The instrumental components of a HPLC are pumps, connecting tubing, an injection valve, and a detector. The pumping systems provide flow rates in the μL/min to nL/min range. There are usually two different solvents in the pumping systems, which are the aqueous buffer and the organic buffer. A reversed-phased chromatographic column is the most widely used separation technique for proteomic research. It is based on the interaction between a polar

hydrophilic mobile phase and the sample with a hydrophobic stationary phase [36]. HPLC is capable of carefully controlling the polarity via a gradient of the two solvents in order to further increase the selectivity of proteins that elute off the column [54].

1.5.2 Nano-Spray-HPLC-ESI-MS/MS

Mass spectrometry can be coupled with gas chromatography (GC) or high-pressure [55] liquid chromatography (HPLC) for identification and quantification of relatively small molecules. Nano-Spray ionization was the specific MS technique used in this project. A small volume (2μL) of the mobile phase containing the sample can be ionized and scanned at a time. Ionization occurs via an application of high voltage to the mobile phase and the ions are separated by an electromagnetic field generated in the MS [56]: lighter ions are deflected much more readily than heavier ones, as well as their charge are altered their direction of movement. This allows the MS to measure the m/z. The resulting ions can be detected based on this whereas the signal can be converted into a spectrum (chromatogram) [57].

1.6 Computation of Random and Independent Replicate Samples

Mathematical and statistical analysis of the many treatments, samples, replicates and fractions results in hundreds or thousands of LC-ESI-MS/MS runs and requires a robust computation and statistical strategy [46], [58]. The statistical validity of MS/MS correlation from the ion trap has been established for standards and human plasma [59]–[61] by multiple methods in agreement [37], [62]–[64]. Random and independent sampling of tryptic peptides from step-wise fractionation followed by LC-ESI-MS/MS [60] is a time and manual labor intensive approach that is sensitive, direct, and rests on few assumptions [37], [38]. Analysis of many experimental

treatments and controls with independent replicates samples [4], [46], [58] each fractioned into many sub-samples, requires large computing power to re-assemble the sub-fractions together, and organize and analyze the results across treatments [38], [65]. Extensive computation is required to match MS/MS spectra to peptide sequences using the X!TANDEM [28] and SEQUEST [66] algorithms and these results must be stored together with and related to the raw LC-ESI-MS/MS data in a SQL Sever database for statistical analysis. The large amount of mass spectra raw data and correlation results may be stored and related in SQL Server [67] for analysis by a generic statistical algorithm like those found in the R statistical system to summarize the results and compute error rates [65]. The combination of 64 bit SQL Server together with generic computation software such as the R package permits the use of a classical random and independent sampling of experimental treatments, and controls for subsequent analysis [65]. Here the peptides and proteins of bovine fetal serum were compared to adult serum by differential solubility in organic solvent versus quaternary amine chromatography prior to digestion to provide the identity and quantification of the factors unique to fetal serum. The combination of step wise preparative separation prior to analytical separation of the resulting peptides by nano LC-ESI-MS/MS with a linear ion trap was shown to be much more sensitive than the use of one step fractionation and micro electrospray, and has apparently detect many cellular proteins in plasma [4], [46].

2.0 HYPOTHESIS AND OBJECTIVES

The hypothesis of this study is that there was a difference in the proteins and peptides of FBS versus ABS that transform cell growth. The aim and objective was to observe the cell morphology and to compare the proteins and peptides in the FBS with ABS by random and independent sampling using the organic differential extraction and the step-wise salt extraction for partition chromatography over a quaternary amine resin. The proteins and peptides collected were then analyzed by a LC-ESI-MS/MS with a linear ion trap to enumerate and quantify the differences in peptide observation frequency by the Chi Square test or precursor intensity by ANOVA and the proteins observed to be specific to fetal serum may be subsequently assayed for their effect on cellular proliferation.

3.0 MATERIALS AND METHODS

3.1.Materials and Equipment

The model of MS used was the LTQ Tune Plus Finnigan Linear Ion Trap from Thermo Electron Corporation (Waltham, MA, USA). HPLC used was Agilent Technologies 1260 Infinity and 1200 Series. Ceramic QA resin was from BioRad. C18 ZipTips were obtained from Millipore (Bedford, MA). C18 HPLC resin (Zorbax 300 SB-C18 5-micron, 300 Angstrom) was from Agilent. HPLC grade solvents were obtained from Caledon Laboratories (Georgetown, Ontario, Canada). All other salts and reagents were obtained from Sigma-Aldrich-Fluka (St Louis, MO) except where indicated.

3.2.Blood Samples

FBS and ABS were supplied from Cell Grow, Sigma-Aldrich (Canada), Gibco by life technologies (New Zealand), MP Biomedical (MP, USA), Rocky Mountain Biologicals (USA), and Thermo Fisher. Blood samples were drawn into sera tubes. The samples were thawed, aliquot and re-frozen once before being used and discarded as described.

3.3.Cell Culture

Most cell lines can be grown using DMEM culture media with 5% Fetal Bovine Serum [68]. All media, blood samples, supplement and reagents were sterilized prior to cell culture. Micropipettors, pipette tips, micro cover glasses, and cell culture flasks/plates were autoclaved in order to create an aseptic environment. A volume of 25mL was added to a 500mL DMEM (5% serum). 0.5mL of raw cells 264.7 (white blood cells) were diluted with 3mL 5% FBS in a cell

culture flask and incubated in a humidified atmosphere of 5% CO_2 at 37^{o}C until confluence reached (the confluence of cells was observed under a light microscope). Media was changed after the first four hours since the raw cells contained DMSO when it was frozen down in the liquid nitrogen. Cells were passaged and scraped when they became 80% confluent (1:2 split in 1 to 2 days) [69].

3.4.Cell Proliferation Assay

The ability to culture cells under very low cell density is a stringent criterion of the quality of culture sera. Many experiments require growing cells at low density so that individual colonies can develop [21]. To compare the effects of each serum on cell growth, three independent controls of FBS and three independent treatments of ABS from different sources assay were performed and observed. Cells were seeded in 6-well plates with maximum 2mL of 5% (v/v) FBS and 5% (v/v) ABS in DMEM respectively. They were cultured in the incubator at 37^{o}C and sampled (passaged) every 48 hours. The rate of cell growth was measured by looking at the confluence of the cells under a light microscope and the data was recorded [11], [69].

3.5.Cell Fixation with Para-formaldehyde

Raw 264.7 Monocytic white blood cells can differentiate into macrophages, which adhere to cover slips [70]. 5% FBS or 5% ABS in media was removed from each well (cells were first cultured in FBS), and each cover slip with cells adhered to it was washed 3 times with 2mL 1X PBS solution. The 6-well cell culture plate was shake on a rocker plate for 30 minutes. 2mL of 2% Para-formaldehyde was added to each well in order to fix the cells. The cell culture plate was allowed to stay in Para-formaldehyde for no more than 2 hours (or 0.1% PFA overnight). The

Para-formaldehyde was removed and the cells were washed 3 times with 2 mL 1 x PBS [69], [71].

3.6.Cell Staining with Rhodamine Phalloidin

Cells used to measure length and shapes were quenched in 1X PBS with 10mM glycine for 15 minutes. Then the cells were stored in the 4^{o}C refrigerator for microscopic analysis. For cells that were used to stain with rhodamine phallodin, the cells were first permeabilized with 0.1% Triton X-100 in 1X PBS (used to dissolve the cell membranes). 2mL of 1/1000 dilution phalloidin in 1X PBS was added to the 6-well plate with fixed cells and stained for 30 minutes on a rocker plate [72]. The cells were then washed with 3 portions of 2mL 1X PBS. The cover slip that had the fixed cells on it was dried by holding the edge of it with a tweezers on a Kim-wipe and rotating the edge of the cover slip so that it could get dry faster. The cover slip was mounted onto a glass slide with Dako (mountant). It was dried and stored in the -20^{o}C freezer for microscopic imaging.

3.7.Salt Partition (Gravity) Chromatography Extraction of Proteins:

The preparative gravity chromatography was performed by adding 100µL scale columns in 3mL plastic transfer pipettes stoppered with glass wool. The tops of the plastic transfer pipettes were cut off with scissors in order to form an upper reservoir. The stoppered pipettes were placed within 15mL disposable falcon tubes to strike the column with the loading buffer. A 200µL volume of 50% slurry of quaternary amine (QA) resin was added to the top of the transfer pipette column and allowed to settle to the glass wool at the bottom by gravity. The column was then drained and equilibrated with 20mM Tricine 1xPBS loading buffer. A 25µL volume of serum

was diluted in 200µL of 20mM Tricine 1xPBS solution and then added to the column. Additional PBS loading buffer was added in order to bind the column and wash off excess sample. A flow-through elution (proteins that do not bind to the column) was collected into a 1.5mL eppendorf tube. The column was then loaded with a sequential gradient of salts (NaCl). All protein separations were performed by gravity-drip chromatography, and each protein fraction eluted was collected into a 1.5mL eppendorf tube. A column volume of 100-200µL can be run by gravity alone without the aid of frit or centrifugation. Given a binding capacity of 50mg/mL, such a column could bind up to 5mg of protein, which is more than sufficient for SDS-PAGE and HPLC-MS/MS [3], [39].

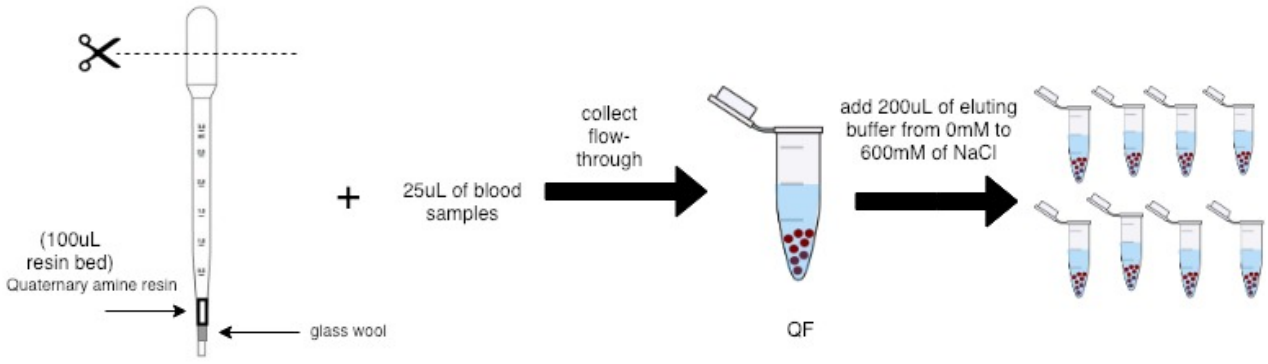


Figure 3.7. 1. Demonstration of step-wise salt gradient extraction of proteins over a QA resin. A total number of 16 fractions were collected for each individual ABS/FBS sample.

3.8.Trypsin Digestion of Intact Proteins

A 6M urea digestion buffer was prepared by dissolving 36.036g urea in a 100mL bottle with 50% ACN and 500mM Tris-HCl pH 8.8 (10X stock). 90µL of previously collected samples by the QA resin chromatography was mixed with 10µL of the 10X digestion buffer (Now the final digestion buffer is 1X). A 25µg of trypsin was dissolved and mixed in 50µL 1% acetic acid (HAC) on ice (0.5mg/mL). 1µL of dissolved trypsin was added to each sample (100µL mixture

of samples and digestion buffer). The samples were then incubated at 37^{O}C overnight (12 hours incubation). 1µL of 200mM fresh DTT was added to each sample (2mM in each sample). The samples were then heated at 50^{O}C for 5 minutes, another 1µL of trypsin was added and incubated for 2 hours at 37°C. The samples were quenched with 5% formic acid and freeze-dried for the MS analysis [73].

3.9.Organic Differential Extraction of Endogenous Peptides

The sample preparation was carried out by an organic differential extraction where 1800µL of 100% acetonitrile was added to the 200µL of FBS/ABS in the biological cleanroom. Disposable plastic 1.5mL sample tubes and plastic pipette tips were used to handle samples. FBS/ABS samples precipitation followed by the selective extraction of the pellet using a step gradient to achieve selectivity across sub-fractions and thus greater sensitivity. Sample mixture was vortexed (mixed) for 10 seconds, then centrifuged for 5 minute at 12,000 RCF. The supernatant (aqueous portion) was collected and the organic precipitate (pellet) that contains a much larger total amount of endogenous polypeptides was manually sequentially re-suspended in increasingly polar solvent (water content) of 70, 60, 50, 40, 30, 20, 10, 0% of acetonitrile, followed by water with 5% formic acid. After re-suspending the pellet in each organic solvent/water mixture, samples were centrifuged at 12,000 RCF for 5 minutes and the supernatant was collected. All samples were transferred to a fresh tube, dried in the rotary evaporator for 4 hours and kept at -80°C in aliquots for future uses [4], [46].

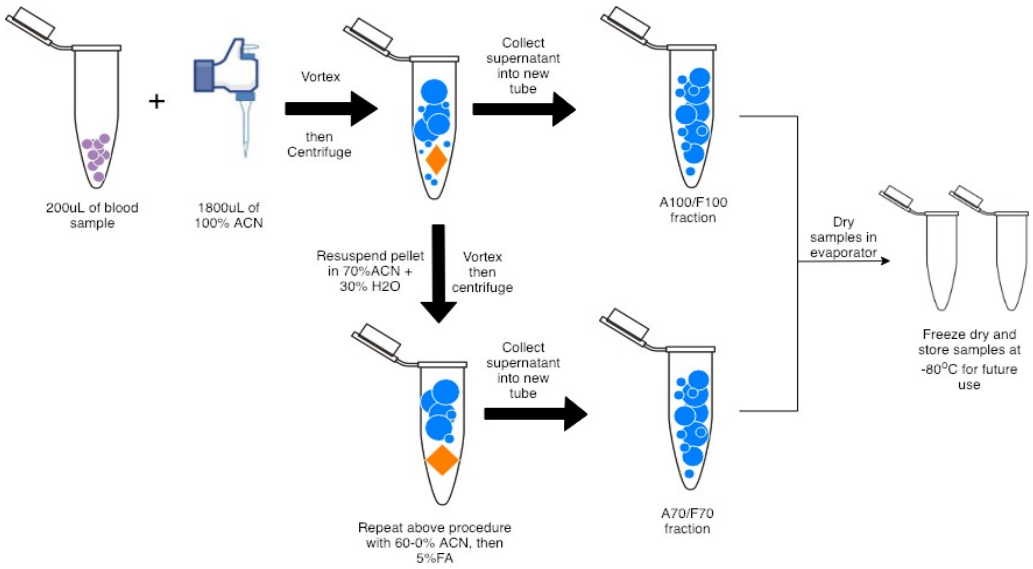


Figure 3.8. 1. Demonstration of step-wise organic extraction of peptides with acetonitrile/water solvents. A total of 10 fractions were collected for each individual ABS/FBS sample.

3.10.Dumbroff Protein Assay

The concentration of proteins in each fraction was determined by dot blotting method of Dumbroff on a Whatman number 1 filter paper alongside BSA standards (range from 0.0mg/mL to 5.0mg/mL). The proteins were mixed in a 1:1 ratio with 2X SDS sample loading buffer, and 2µL of volume was dotted on Whatman number 1 filter paper along with BSA standards at the bottom in the same detergent buffer before drying. The dot-blot was rinsed in methanol and then stained with CBBR in water, methanol and acetic acid. It was allowed to stain for 15 minutes. The blots were de-stained in water, methanol and acetic acid. The Dumbroff method has an advantage prior to SDS-PAGE: the sample content can be measured in detergent and then directly loaded on the gel [3], [39].

3.11.Protein Resolved by SDS-PAGE:

Tricine-SDS-PAGE was used to resolve the protein fractions for the comparison selectivity [3], [39]. The protein fractions in 1:1 2X SDS were resolved on polyacrylamide gels. Polyacrylamide gels with 9% monomer resolving gels with Tris-HCl pH 8.8 and 4% monomer stacking gels with Tris-HCl pH 6.8 were run in tricine anode and cathode buffers. Protein fractions were resolved and compared by SDS-PAGE alongside protein molecular weight standards [74]. The gels were then stained with CBBR in methanol and acetic acid [39].

3.12.C18 Column Packing

A C18 reversed phase HPLC column was packed with a column guard. For every sample, the C18 resin solvent was prepared by mixing a C18 Zorbax 5 micron 300 Angstrom with chloroform. C18 packing was performed with a high-pressure liquid chromatography (HPLC) pump set at 100% acetonitrile in 0.1% of acetic acid (100% organic buffer) over a 15cm capillary silica column (150μm inner diameter) at a flow-rate of 200μL/min (pump was set to normal mode –range from 20-200μL/min). In this case, the column guard was filled with 2/3 of the mixture (Zorbax and chloroform) and attached to the pump. The C18 resin migrated down to the column by high-pressure pump. If the pressure of the pump went above 200 bars, the pump had to be stopped and the flow rate was set to 0μL/min. Once the entire silica column was fully packed, it was sealed and dated and the resin type (C18) was labeled.

3.13.Blank Runs

A blank run was where there was no samples injected to the column and it was recorded on a chromatogram over a 30-minute gradient from 5% to 65% of acetonitrile using Nano-LC-ESI-

MS/MS. Each new packed column was washed at least 3 times with both water and acetonitrile and recorded as a blank on a chromatogram. The blanks were run in order to wash off the contaminants of the column before column conditioning and sample injections to the mass spectrometer. After the blank run was finished, the column was then equilibrated by flushing the C18 column with a high organic buffer of 50% acetonitrile for 20 minutes, and then 5% acetonitrile for 10 minute.

3.14.Sample Introduction into the MS Detector

The samples were introduced from a syringe into the loop via a Rhodynne injector before injecting onto the analytical column, and then into the LC flow using a divert valve that enters the MS detector [50]. 20µL mixture of protein sample and 5%FA were introduced each time.

3.15.Column Conditioning with TRP3

The purpose of column conditioning was to coat the C18 column with a digested TRP3 protein. A TRP3 standard was used to confirm the sensitivity and mass accuracy of the system, which made of 5µL of each three non-human proteins (rabbit muscle glycogen phosphorylase B, yeast alcohol dehydrogenase, and bovine cytochrome c-like-protein) in 1485µL of 1% of acetic acid. The mass spectrometer was cleaned and calibrated. The HPLC pump was set to 50% acetonitrile for 20 minutes and then 5% acetonitrile for 10 minutes manually in order to clean the column. The column conditioning was performed by the TRP3 peptides coated onto the C18 reverse phase column in a 120-minute tune method. A sample injection valve was used for HPLC transferring sample from a syringe to a sample loop. The sample loop was connected via a change in valve configuration, to a high-pressure mobile phase stream, which carried the sample

into the column. A 20µL of TRP3 was introduced using a syringe into the sample loop via a port on the injection valve on LOAD. The sample injector was then switched to INJECT and the column was attached to the mass spectrometer ion trap source. The HPLC pump was then set to 50% acetonitrile and sample injector was switched to LOAD when the first TRP3 peak appeared on the chromatogram. A second TRP3 peak appeared at 50% acetonitrile. This procedure was repeated three times until a total of six peaks were collected [4], [46], [60].

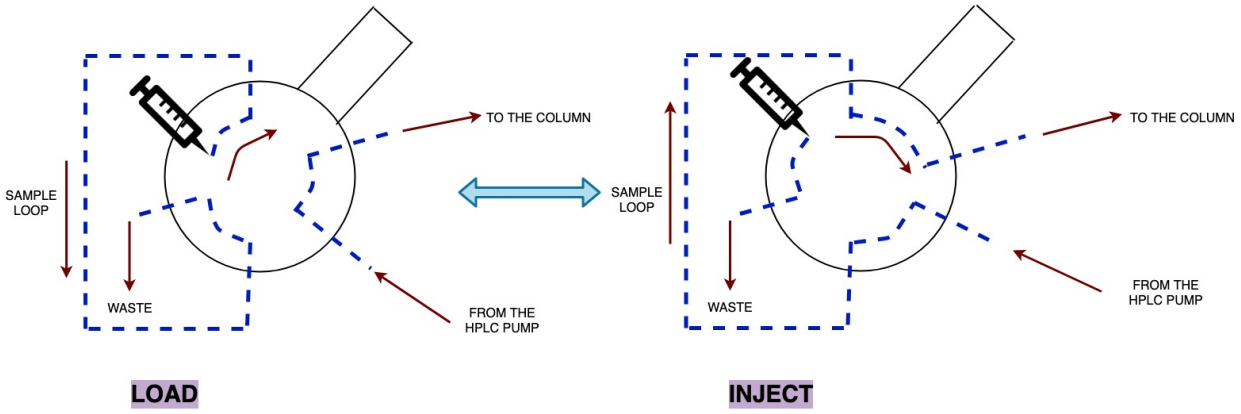


Figure 3.15. 1. HPLC manual injection. Sample was introduced with a syringe and stayed in the sample loop when it was on LOAD. Solvents picked up samples from the sample loop and pushed it down to the column when it was on INJECT.

3.16.Sample Zip-tipping

Each of the previously extracted protein fractions was re-suspended by adding 200µL of 5% formic acid and vortexed to mix well. Sample was then zip-tipped in 65% acetonitrile (organic solvent) and 5% formic acid (aqueous solvent). 5 volumes of 1mL solutions were prepared in 5 eppendorf tubes labeled as "B1 A1 S A2 B2", where B was 65% acetonitrile organic solvent, A was 5% formic acid, and S was the extracted protein fraction. The organic solvent was used to concentrate and precipitate the peptides from the sample onto the resin pipette tip, whereas the

aqueous solvent was used to wash off other compounds from the environment and desalt. A P20 micropipettor was set to 20μL, and a C18 reversed phase pipette zip-tip was placed on the pipettor. 20μL of B1 solution was pipetted slowly and aspirated to avoid introducing air into the packing materials and to maximize binding (the C18 resin could not be dried out once it was hydrated). The solution was then slowly expelled back into the tube. The solution was pipetted and expelled 10 times into the same tube. The buffer A was the wash solution; 20μL of A1 solution was aspirated and expelled back into the tube for 5 times. The pipette tip was clean and placed to the sample tube. 20μL of the sample solution was aspirated and expelled back into the eppendorf tube for 20 times. 20μL of A2 solution was aspirated one time only, and then the solution was expelled into a waste container. At last, 2μL aliquot of B2 solution was aspirated (the 2μL mark was indicated on the pipette tip) and expelled into a new eppendorf tube. The micropipettor was set to 18μL, and then 18μL of 5% formic acid was added to the tube with the 2μL buffer B solution. The zip-tipped protein sample was kept in ice [4], [46].

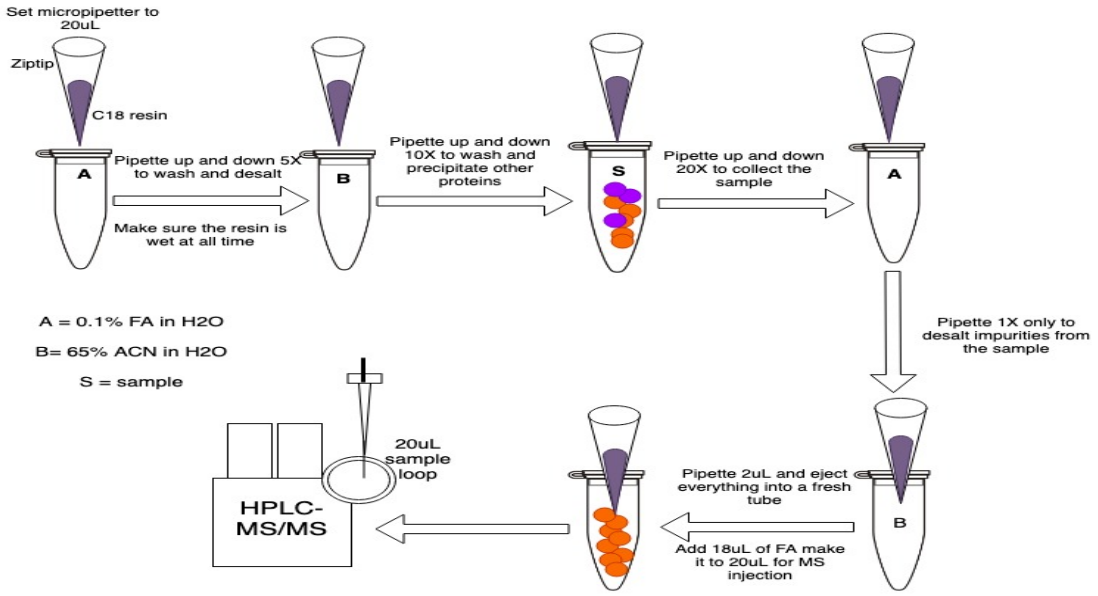


Figure 3.16. 1. Demonstration of C18 sample zip-tipping. The C18 zip-tip was washed and desalted before collecting the samples for MS detection.

3.17.Analytical LC-ESI-MS/MS

All preparative peptides were collected by zip-tipping prior to injections as described above. Fractionated proteins were resolved by C18 reversed phase chromatography (150µm ID, 15cm column). The sample was analyzed over an 85 minutes gradient sequence, which started at 5%, ramped to 12% and then increased to 65% for 5 minutes, decreased to 50% for 15 minutes then declined to a final proportion of 5% of acetonitrile at a flow rate of 5µL/min with an Agilent 1200 series capillary pump via an emitter needle Nano-spray into an ion trap. The MS/MS data chromatograms were searched against a library of proteins, cDNAs, and genomic DNA using SEQUEST or X!TANDEM algorithms [3].

3.18.Column Equilibration

The column was equilibrated to the initial aqueous/organic solvents composition before performing another analysis. The purpose of the equilibration was to wash off all the previous peptides that stuck to the column. A column equilibration was performed by flushing the C18 column with a high organic buffer (50% ACN) for 20 minutes. The alternative column equilibration was to make the first run a blank run before the real analysis [39], [41].

3.19.Correlation Analysis

In this study 1,554,347 precursor ions from fetal versus adult MS/MS spectra were recorded by nano-LC-ESI-MS/MS. Correlation analysis of ion trap data was performed with X!TANDEM [7] and SEQUEST [66] algorithms to match tandem mass spectra to peptide sequences from a library of 209,111 bovine proteins that differ by at least one amino acid from RIKEN, IMAGE, RefSeq, NCBI, Swiss Prot, TrEMBLE, ENSEMBL, UNIPROT and UNIPARC along with available Gene Symbols, all previous accession numbers, description fields and any other available annotation that was rendered non-redundant by protein sequence in SQL Server last assembled in 2015 [38], [60]. Endogenous peptides with precursors greater than 10,000 (E4) arbitrary counts were fit as fully tryptic peptides and/or tryptic phosphor-peptides on separate servers for each algorithm and the results combined, and compared in SQL Server/R. The X!TANDEM default ion trap data settings of ±3 m/z from precursors peptides considered from 300 to 2000 m/z with a tolerance of 0.5 Da error in the fragments were used [7], [37], [39], [63], [64], [75]. The best fit peptide of the MS/MS spectra to fully tryptic and/or tryptic phospho peptides at charge states of +2 versus +3 were accepted with additional acetylation, or oxidation of methionine and with possible loss of water or ammonia. The resulting accession numbers,

actual and estimated masses, correlated peptide sequences, peptide and protein scores, resulting protein sequences and other associated data were captured and assembled together in an SQL Server relational database [65].

3.20.Data Sampling, Sorting, Transformation and Visualization

The linear quadrupole ion trap provided the precursor ion intensity values and the peptide fragment MS/MS spectra that were correlated to specific tryptic peptide of STYP phosphopeptides by the X!TANDEM and SEQUEST algorithms. The MS and MS/MS spectra together with the results of the X!TANDEM and SEQUEST algorithms were parsed into a SQL Server database and redundant fits of MS/MS spectra were filtered to the best fit and the observation frequency was corrected on the basis of total MS/MS spectra before statistical and graphical analysis with the generic R data system [37], [62], [65], [76], [77]. The peptide to protein correlation counts for each gene symbol were compared for fetal versus adult serum using the Chi Square test using equation #1:

i)

$$\frac{[(Fetal - Adult)\text{^}2]}{Adult + 1}$$

EQN#1

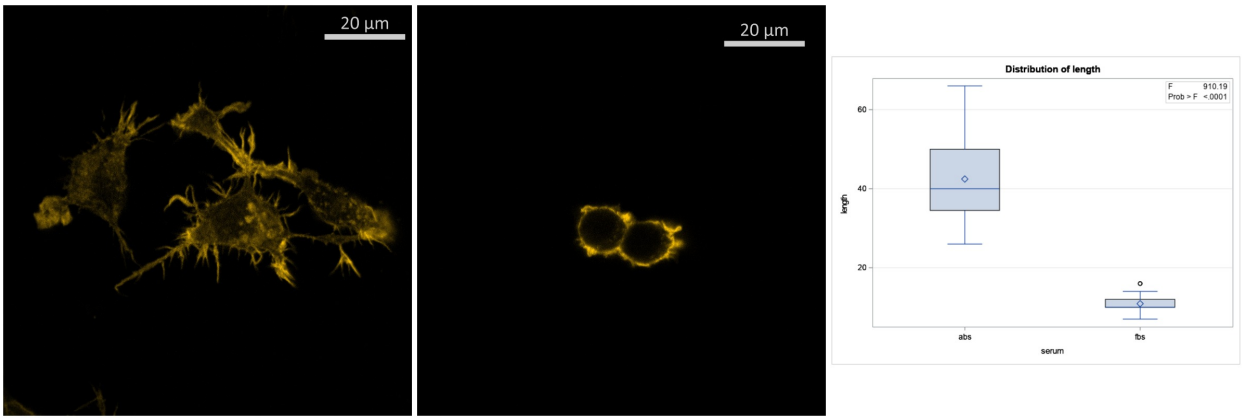
The precursor intensity data for MS/MS spectra were $\log_{10}$ transformed, tested for normality and analyzed across institution/study and diseases versus controls by means, standard errors and ANOVA [37], [62], [77]. The entirely independent analysis of the precursor intensity using the rigorous ANOVA with Tukey Kramer HSD test versus multiple controls was achieved using a 64-bit R server.

4.0 RESULTS

There was ~99% agreement on the proteins independently identified by peptidomics versus proteomics. Exogenous tryptic peptide sequences showed low levels of overlap with endogenous peptide sequences (23%). Both sets of independent peptide sequences mapped to the same group of protein accessions (99%), which was an unambiguous and clear demonstration that LC-ESI-MS/MS with a sensitive ion trap shows low error rates of peptide and protein identification. LC-ESI-MS/MS of blood with a simple quadrupole ion trap was a sensitive and reliable technology that permits the accurate identification and relative quantification of observation frequency or precursor intensity from tens of thousands of blood proteins by random and independent sampling. It appears that the combination of step wise fractionation of serum proteins over disposable quaternary amine resin prior to tryptic digestion, or organic extraction of endogenous peptides, and LC-ESI-MS/MS will provide a sensitive and direct method to compare the expression of proteins over different physiological states such as fetal versus adult serum. Many proteins associated with cell growth pathways and DNA or RNA binding factors were observed to differ between the fetal versus adult blood. The analysis of the results by STRING [78] has reinforced the previous conclusion from human serum that blood contains a large number of cellular proteins that apparently exist as protein complexes in circulation and at least some of these must come from blood cells.

4.1.Cell morphology

The cells grown in fetal serum divided rapidly forming small spherical cells that frequently forms spherical foci organized into a globe of cells that extended vertically upwards piled onto each other. In contrast, cells cultured in adult serum frequency formed a monolayer of elongated rhomboids with dendritic extensions that divided slowly and often died after a few passages. Staining with rhodamine phalliodin showed changes in the morphology of the cells (see Appendix). Cells cultured in FBS that formed in rounded symmetrical shape was measured to be average 10 microns in length while cells cultured in ABS with dendritic shape was measured to be average 40 microns in length. One-way ANOVA analysis showed statistically significant measurement of the cell lengths between adult versus fetal serum (Fig. 1)(see Appendix for more cell images).



A B C

Figure 1. The effect of fetal versus adult serum on the morphology of RAW 264.7 murine macrophages cultured in DMEM plus 5% bovine fetal or adult serum. The cells were stained with rhodamine pholloidin and the images were taken under a confocal microscope. Panel A: Cells grown in Adult bovine serum; B, Cells grown in Fetal bovine serum; C, the distribution of cell length in Adult (ABS) versus Fetal (FBS) serum. Ten cells were measured from 3 spots on 3 cover slips from 3 independent batches of fetal versus adult serum. The longest axis of the cell was measured.

One-way ANOVA Statistical Test on the means between the cell lengths of ABS versus FBS:

Estimated Std. Error t value Pr(>|t|)
(Intercept) 46.0438 2.1458 21.457 <2e-16 *
Slide -1.1375 0.6852 -1.660 0.0989 .
Replicate -0.7792 0.6852 -1.137 0.2572
Treatment[T.FBS] -31.5375 1.0409 -30.299 <2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 6.583 on 156 degrees of freedom
Multiple R-squared: 0.8552, Adjusted R-squared: 0.8524
F-statistic: 307.1 on 3 and 156 DF, p-value: < 2.2e-16**

4.2.Protein partition chromatography

The proteins of ABS and FBS were separated over ceramic quaternary amine resin [37], [39] and the fractions tested for protein content by the Dumbroff method [79] prior to resolving the proteins by tricine SDS-PAGE that showed selectivity over the course of the salt step gradient (Fig. 2A &3).

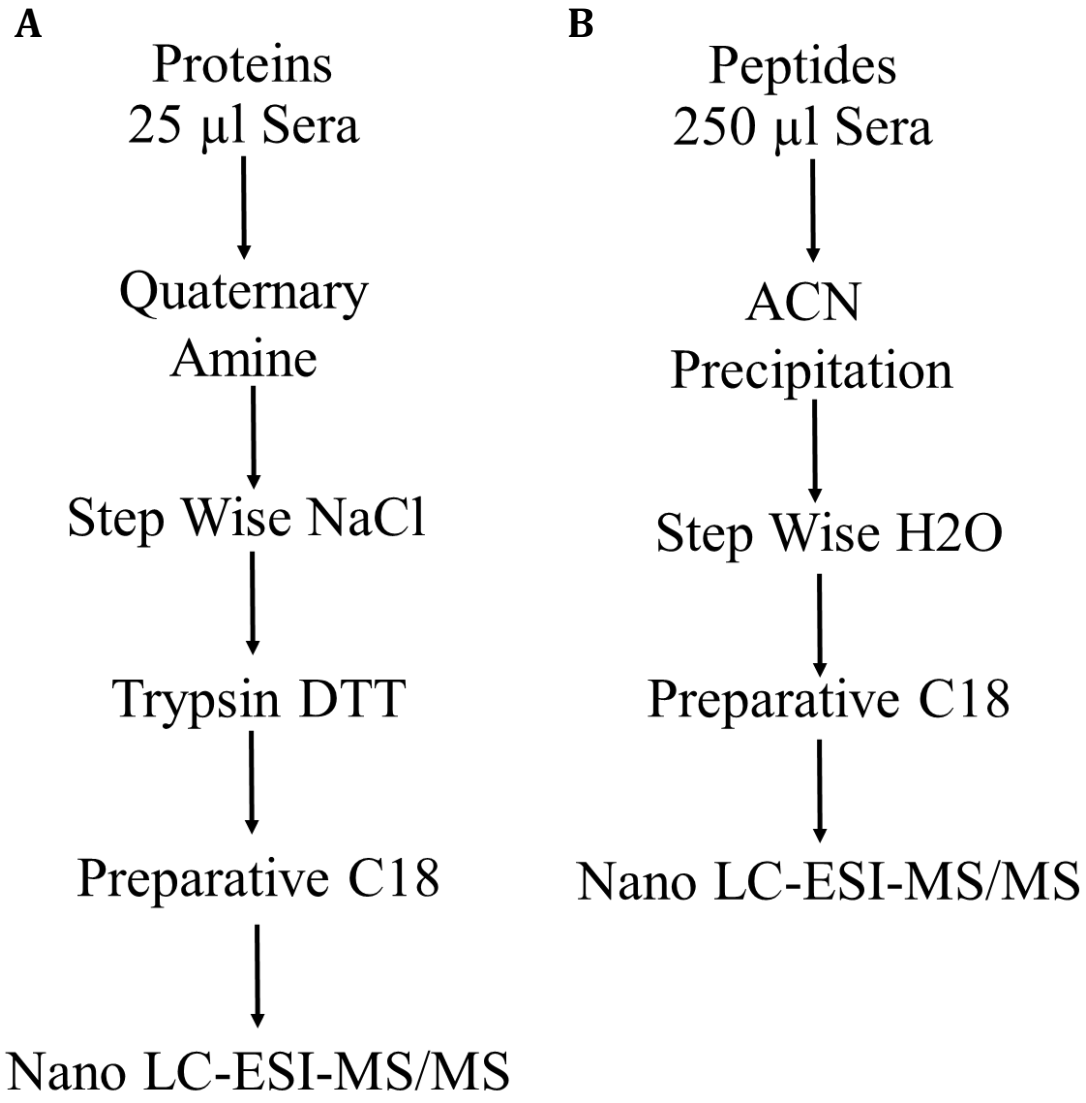


Figure 2. Flowchart of the intact proteins and endogenous peptides isolation analysis steps. A: Steps of intact proteins by step-gradient salt extraction over a QA resin partition chromatography. B: Steps of endogenous peptides by step-wise acetonitrile water extraction.

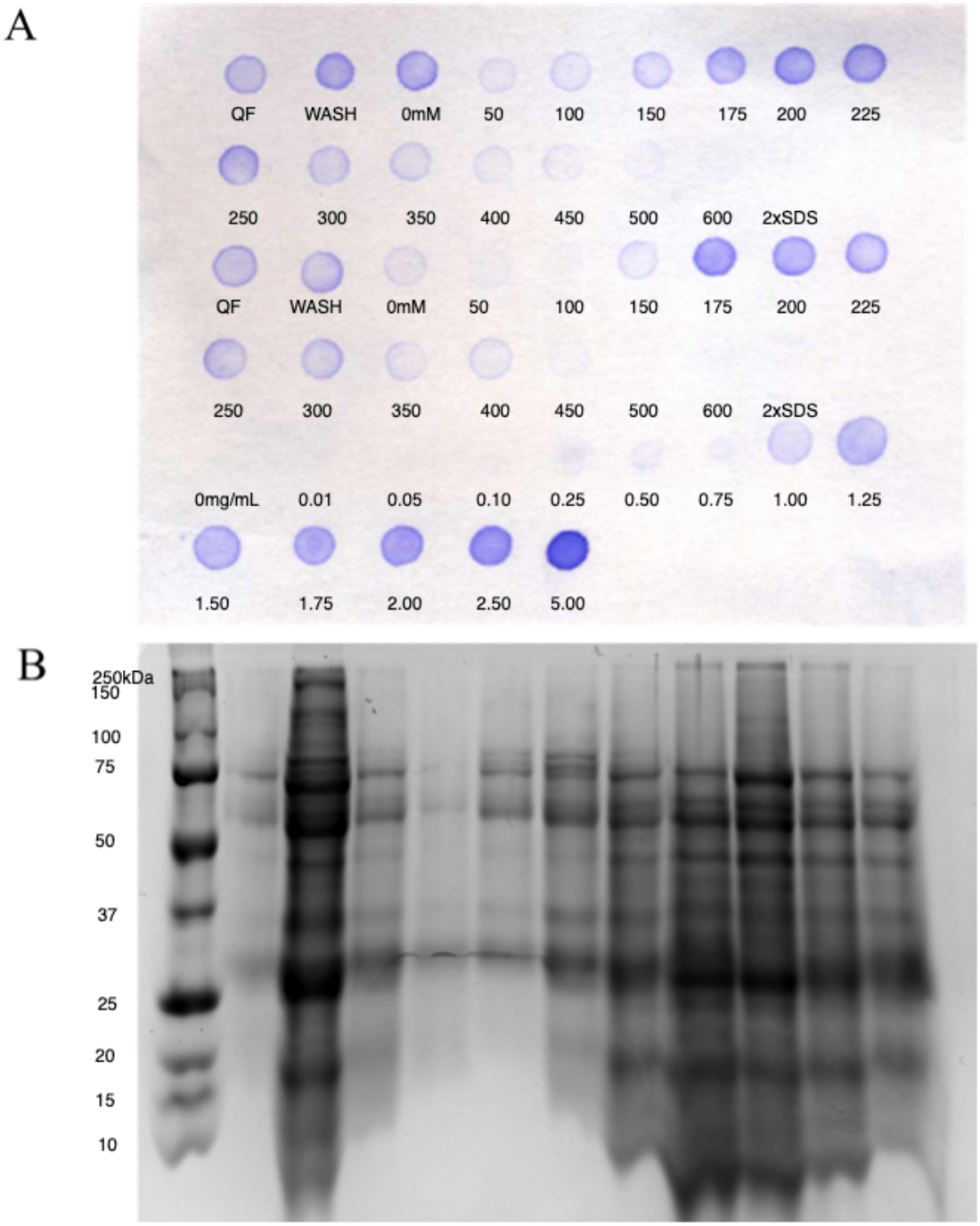


Figure 3. The fractionation of FBS by a step gradient of NaCl in 20 mM Tris pH 8.85 buffer. Panels A: Dumbroff protein assay shows the protein potent in each sample fraction; Panel B: Tricine SDS-PAGE shows the selectivity between sample fractions. Lanes: 1, Molecular Weight Standard; 2, Flow-through; 3, Wash1; 4, Wash2; 5, 0mM NaCl; 6, 50mM NaCl; 7, 100mM; 8, 150mM; 9, 200mM; 10, 250mM; 11, 300mM; 12, 350mM.

4.3.Peptide organic extraction

The organic precipitate of endogenous peptides from ABS and FBS were separated by a water acetonitrile step gradient and differential centrifugation, tested for protein content by the Dumbroff method before resolving the polypeptides on tricine SDS-PAGE that showed selectivity for low molecular mass polypeptides (Fig. 2B &4).

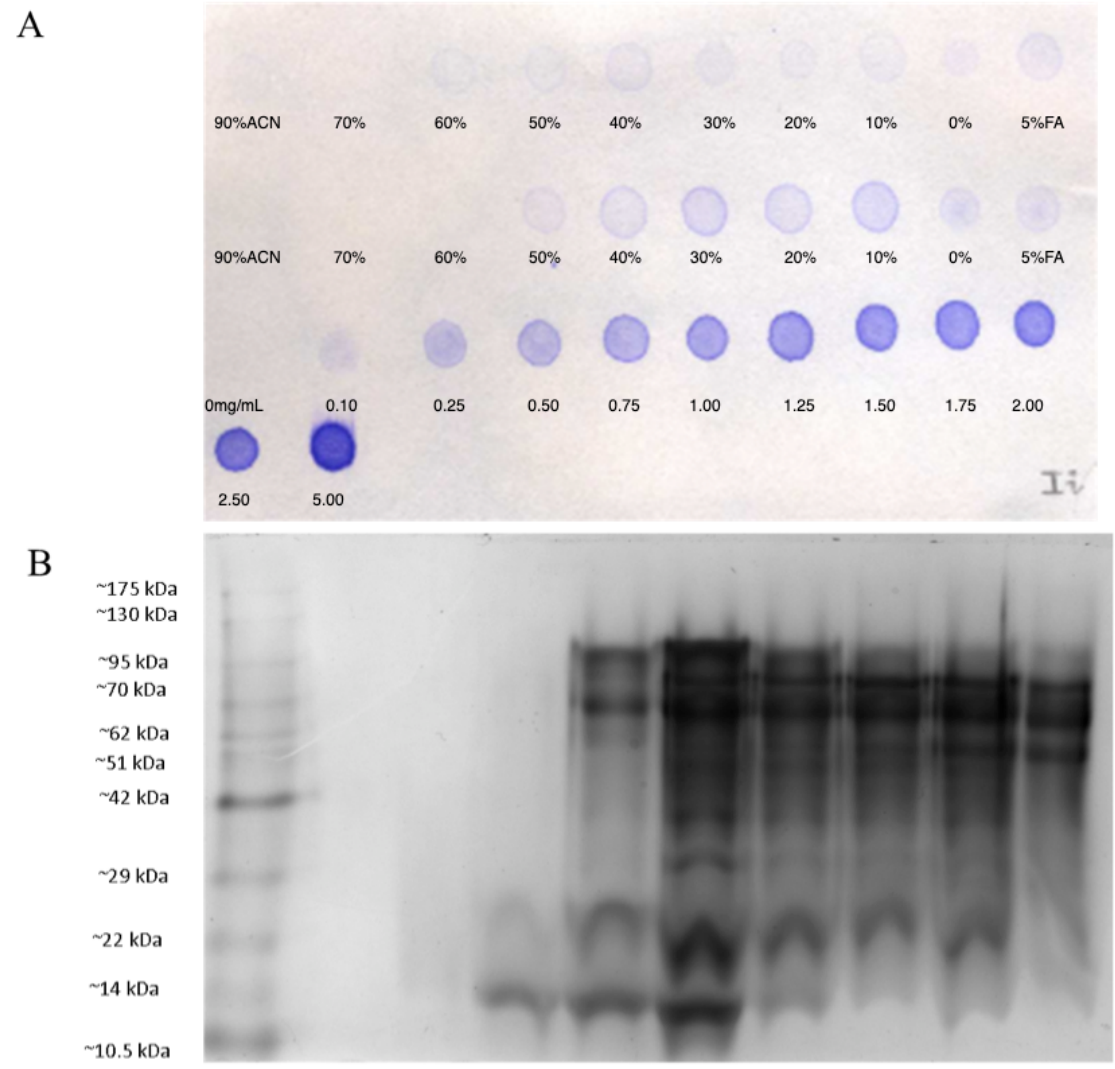


Figure 4. The sequential extraction of low molecular mass polypeptides from serum using a step-wise solublization in organic/water. Panels A: Dumbroff protein assay; Panel B: Tricine SDS-PAGE shows the selectivity between sample fractions. Lanes 1, Molecular Weight Standard; 2, 90% acetonitrile supernatant; 3, 70% ACN supernatant, 4, 60% ACN

supernatant; 5, 50% ACN supernatant; 6, 40% ACN supernatant; 7, 30% ACN supernatant; 8, 20% ACN supernatant; 9, 10% ACN supernatant; 10, 0% ACN supernatant.

4.4.Normality and variation across treatments and replicates

The $\log_{10}$ precursor intensity values from all treatments and replicates together approached a linear and Gaussian distribution (Fig. 5A). The average results of the 6 experiments (2 treatments x 3 replicates) were comparable in terms of the intensity of the precursor peptides obtained with and without the consideration of phosphorylation (Fig. 5B).

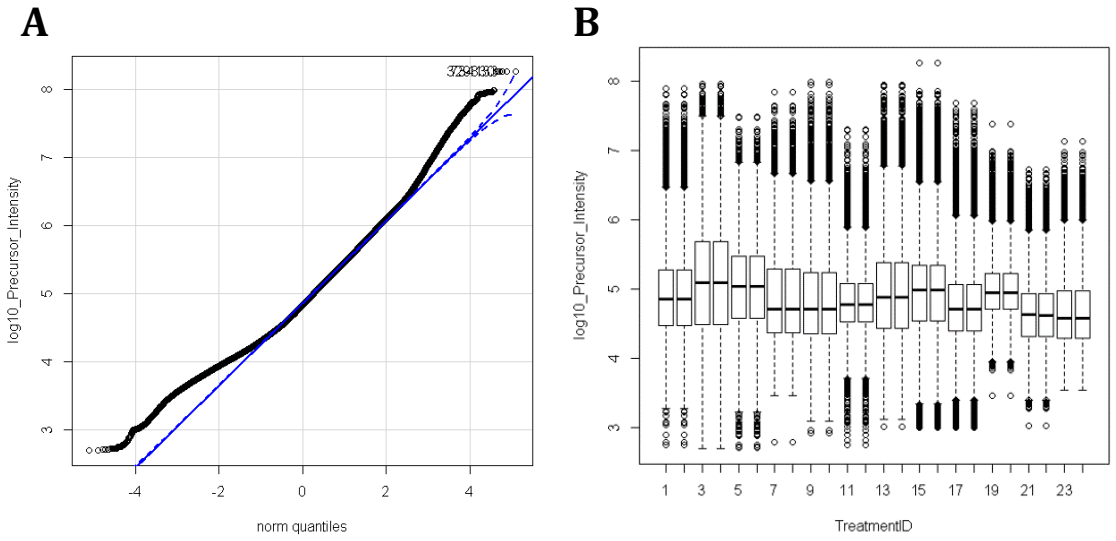


Figure 5. The reproducibility of the fetal versus adult serum samples. Panels: A, the quantile Q-Q plot showing the normality of the log10 transformed intensity values; B, the box plot showing the average $\log_{10}$ intensity and variation in log10 and 99% confidence interval is shown for the adult versus fetal serum replicates. (ID# 1-12:ABS treatments; #13-24: FBS treatments)

4.5.SQL Server Filtering

The pool of tryptic peptides from proteins and endogenous tryptic (TRYP) peptides and/or tryptic phosphor-peptides (STYP) were randomly and independently sampled without replacement by liquid chromatography, nano-electrospray ionization and tandem mass spectrometry (LC-ESI-MS/MS) [60] from FBS versus ABS. The raw correlations from precursors ≥ E4 intensity counts were filtered to retain only the best fit by charge state and peptide sequence in SQL Server to entirely avoid re-use of the same MS/MS spectra. The LC-ESI-MS/MS of adult serum resulted in 1,553,347 MS/MS spectra that spectra resulting in 61,152 best fit correlations tryptic by the X!TANDEM (Table 1).

Table 1. The filtering of the MS/MS spectra and the resulting correlations to peptides to ensure that only the best hit was accepted. The total number of MS/MS spectra of greater than E4 counts collected in this study was 526,870 MS/MS spectra from organic extraction and 1,027,477 from quaternary amine fractionation of proteins followed by digestion resulting in a total of 1,554,347 MS/MS spectra.

	Redundant MS/MS						
	PSTY Mod	Total Spectra Count	Redundant MS/MS Spectra	Distinct MS/MS Spectra	Protein Identifications	Distinct Protein Identifications	Distinct Peptide Identifications
TRYP	0	2616468	36560334	2353004	36560334	204332	2777338
TRYP-PSTY	1	2616468	36297856	2351922	36297856	201220	2598348
	Best Hit Spectra						
		Total Spectra Count	Redundant MS/MS Spectra	Distinct MS/MS Spectra	Protein Identifications	Distinct Protein Identifications	Distinct Peptide Identifications
	No Mods	2616468	7782701	2353004	7782701	179653	1063402
	PSTY	2616468	7730807	2351922	7730807	176176	1040077
	TOTAL	5232936					
	Best Hit by Engine	isPSTYMOD	Redundant MS/MS Spectra	Distinct MS/MS Spectra	Protein Identifications	Distinct Protein Identifications	Distinct Peptide Identifications
	SearchEngineID						
X!TANDEM	2	0	383535	127250	383535	57308	61152
X!TANDEM	2	1	342083	116827	342083	44146	57435
SEQUEST	9	0	7399166	2344666	7399166	178895	1021440
SEQUEST	9	1	7388724	2345002	7388724	176133	998966

4.6.Peptide and protein identification

Organic extraction identified a set of 58,200 proteins compared to 59,799 proteins identified from protein separation over salt partition chromatography. A set of 142,197 peptides was only found in organic extraction compared to a set of 348,543 peptides found only in salt partition chromatography showed a 99% confidence on the protein identity between ABS versus FBS (Table 2). The optimal organic solvent composition was from 40 to 60%of ACN while the optimal salt elution was from 100 to 175 mM NaCl (Table 3). The average results of the 6 experiments were comparable in terms of the intensity of the precursor peptides obtained (Fig.5). The filtered results were then analyzed by the generic R statistical system that showed more than 55,000 protein accession and more than 20,000 proteins gene symbols were detected by the sum of the X!TANDEM and SEQUEST results with at least 5 peptides (Fig. 6).

Table 2. Comparison of organic extraction of peptides versus separation of proteins over quaternary amine chromatography (See Fig.2). The redundant and distinct proteins and peptides and the overlap between treatments were computed in SQL Server.

Total Distinct Proteins	Organic Distinct Proteins	Salt Distinct Proteins	Common Distinct Proteins	Found Only in Organic	Found Only in Salt
60233	58200	59799	57766	434	2033
Total Distinct peptides	Organic Distinct peptides	Salt Distinct peptides	Common Distinct Peptides	Found only in Organic	Found Only in Salt
573640	225097	431443	82900	142197	348543

Table 3. Proteins separation over the salt fractions versus endogenous peptide separation over organic water on the proteins and tryptic peptides identified (See Fig.2).

Peptides				
FractionNumber	Concentration (%)	Redundant Protein	Distinct Protein	Distinct Peptide
1	0%AcN	223397	44397	48853
2	5%AcN	249899	46250	55977
3	10%AcN	250107	45728	53953
4	20%AcN	236723	44892	49899
5	30%AcN	249846	46750	58077
6	40%AcN	263797	48017	63652
7	50%AcN	271287	47568	62042
8	60%AcN	266640	47248	62155
9	70%AcN	278624	46975	59369
10	90%AcN	274823	46150	54845
			463975	568822
Proteins				
FractionNumber	Concentration (mM)	Redundant Protein	Distinct Protein	Distinct Peptide
1	0mM	430748	51459	87107
2	50mM	393872	50332	79094
3	100mM	517203	52308	91084
4	150mM	478902	52205	93276
5	175mM	530335	53440	109606
6	200mM	379348	50820	83834
7	225mM	324976	50150	78085
8	250mM	278455	48049	64995
9	300mM	382052	50746	81509
10	350mM	351982	49386	73052
11	400mM	326096	48878	69684
12	450mM	261822	46261	57871
13	500mM	404293	50257	79141
14	600mM	322479	48271	67073
15	BdDigest	211181	43258	46105
16	QF	399748	50099	77835
			795919	1239351

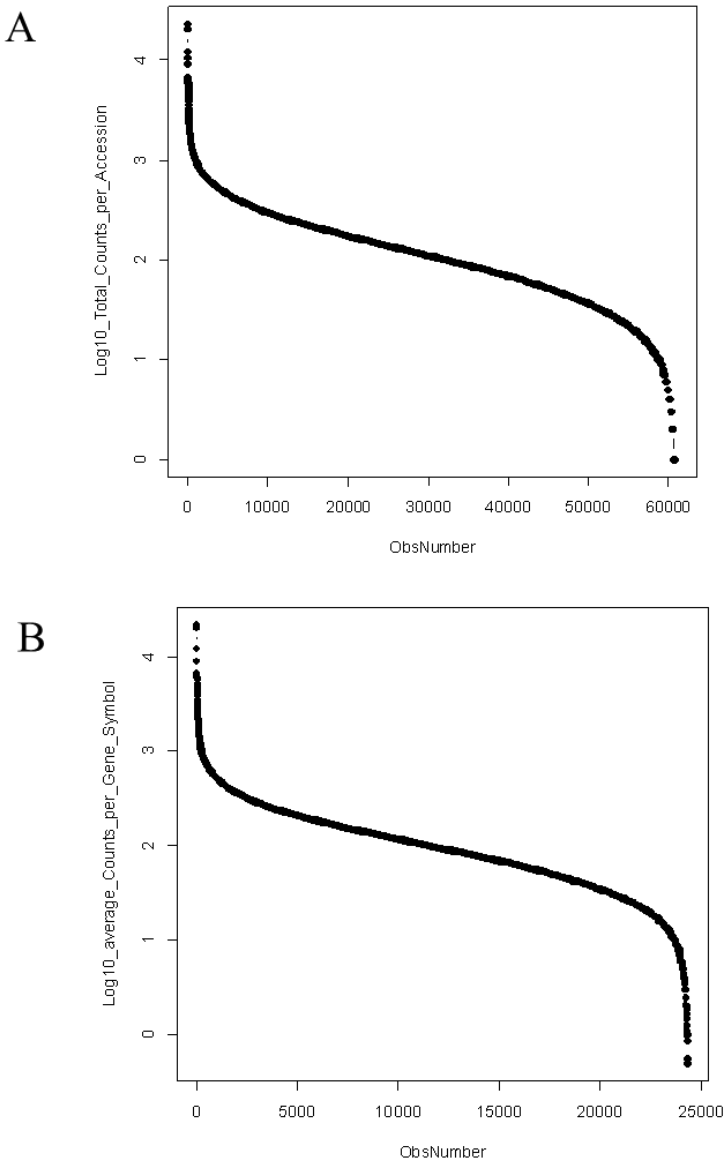


Figure 6. The total number of protein accessions and Gene symbols from the Bovine protein library that were correlated by the X!TANDEM and SEQUEST algorithms combined. Panels: A, redundant protein accessions (after filtration of the redundant MS/MS); B, best correlation per Gene symbol.

4.7.X!TANDEM versus SEQUEST

The SEQUEST algorithm identified some 58,000 proteins with at least 5 peptides and X!TANDEM identified identified some 12,000 protein gene symbols with multiple peptides (Fig. 7). X!TANDEM was used to identified the proteins by gene symbols whereas SEQUEST was used to count the relative abundance per protein accession.

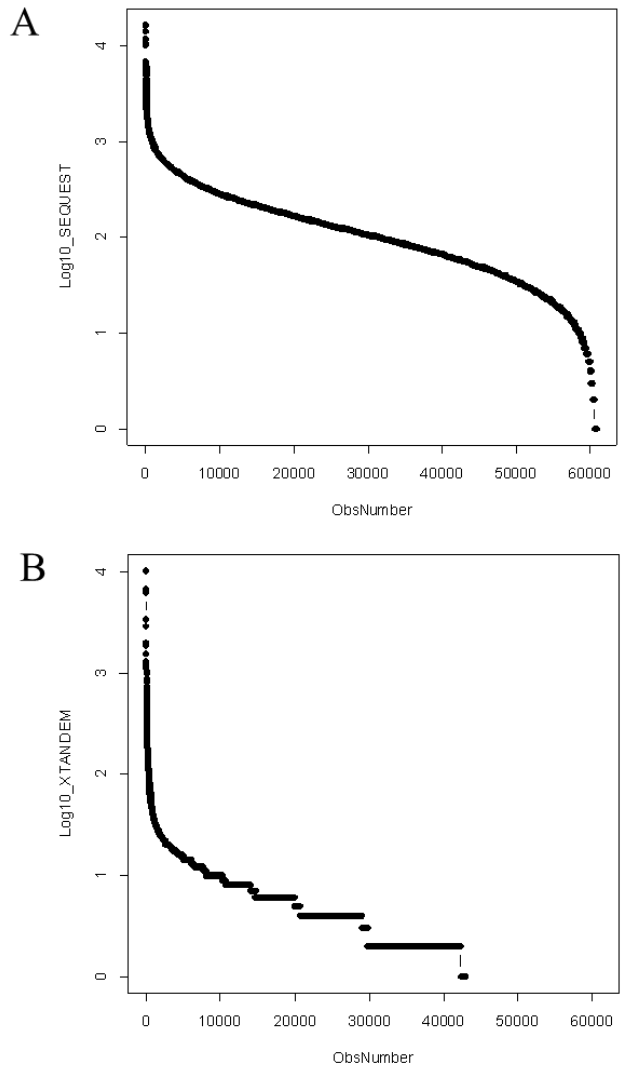


Figure 7. The protein accessions identified by the SEQUEST and X!TANDEM algorithms separately. Panels: A, Proteins identified using SEQUEST showed some 57,308 protein accessions; B, Proteins identified using X!TANDEM showed some 44,146 protein accessions.

4.8.X!TANDEM p-value and FDR corrected q-value

Computation of cumulative p-values and FDR corrected q-values for each gene symbol by the method of Benjamini and Hochberg [55] that showed some 12,000 proteins had a q-value of 0.01 or less (Fig. 8).

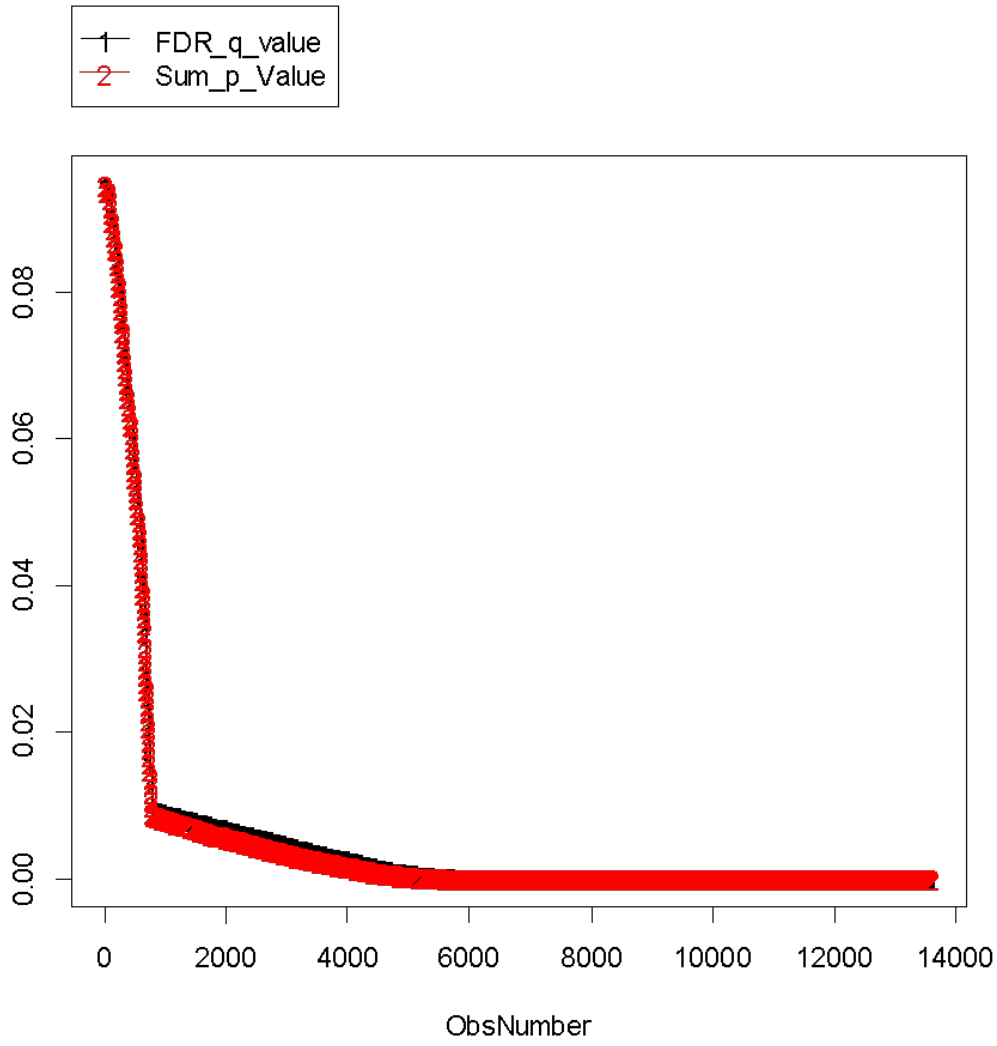


Figure 8. The cumulative p-value and FDR corrected q-value of the non-redundant peptides per gene symbol computed from the X!TANDEM results using the R statistical system.

4.9.Chi Square analysis

Unbiased Chi Square analysis of fetal proteins that showed $\chi 2$ values greater that 60 (DF=1) compared to adult serum after correcting observation frequency by the number of MS/MS spectra collected (Table 4). The observed frequency difference plot passed through the 0 point (no difference in observed frequency) at the 0 quantile point (mean of difference distribution) clearly indicating the observation frequency values were proportionally corrected prior to Chi Square comparison. Proteins that showed highly significant increases in observation frequency in the fetal serum included alpha-fetoprotein that is known to be expressed in the fetal liver. Other proteins raised in the fetal serum included, alpha-2-HS-glycoprotein precursor, serpin peptidase inhibitor clade A, fetuin B, gamma globin, inter-alpha-trypsin inhibitor heavy chain H3, collagen and calcium-binding EGF domain- protein, pyruvate carboxylase, adenylate kinase, hemoglobin subunit beta, kallikrein K, thrombospondin 4 and many others. The Chi Square analysis showed some proteins with $\chi 2$ values that were apparently far too large ($\chi 2 \geq 60$, $p<0.0001$, DF=1) to all have resulted from random sampling error (Fig. 9).

Table 4. Fetal serum specific proteins detected by fully tryptic peptides and fully tryptic phosphopeptides from X!TANDEM and/or SEQUEST or that show a Chi Square ($\chi 2$) value of ≥ 60.

Total Protein Identification Count	GeneSymbol	DescriptionFasta	GoFunction	GoProcess	TANDEM	PD_MSF
13062	AHSG	Alpha-2-HS-glycoprotein			7364	5698
6313	SERPINA1	Alpha-1-antiproteinase			3602	2711
1256	FETUB	Fetuin-B			575	681
1603	ITIH3	inter-alpha-trypsin inhibitor heavy chain H3 isofo			738	865
1714	ITIH3	Inter-alpha-trypsin inhibitor heavy chain H3			774	940
1673	ITIH3	inter-alpha (globulin) inhibitor H3			758	915
2472	AFP	Alpha-fetoprotein			1222	1250
6646	LOC104974567	translation initiation factor IF-2-like			4	6642
6731	PDCD4	programmed cell death protein 4			2	6729
433	RRAGB	ras-related GTP-binding protein B			6	427
10513	DGAT1	Truncated diacylglycerol acyltransferase 1 (Fragment)			6	10507
5829	SH2B1	SH2B1 (Fragment)			6	5823
5841	SH2B1	SH2B1 (Fragment)			8	5833
2625		Uncharacterized protein KIAA2012 homolog			9	2616
2625	KIAA2012	Uncharacterized protein KIAA2012 homolog			9	2616
5913	SH2B1	SH2B adapter protein 1			12	5901
5914	SH2B1	SH2B adapter protein 1 isoform X4			14	5900
5922	SH2B1	SH2B adapter protein 1 isoform X3			12	5910
5933	SH2B1	SH2B adapter protein 1 isoform X2			12	5921
2836	SPRY4	protein sprouty homolog 4 isoform X1			14	2822
429	RRAGB	ras-related GTP-binding protein B isoform X1			6	423
434	RRAGB	Uncharacterized protein			6	428
5216	KIRREL3	Uncharacterized protein			0	5216
5223	KIRREL3	kin of IRRE-like protein 3 precursor			0	5223
3615	FBLL1	Uncharacterized protein			125	3490
6016	CCDC9	coiled-coil domain-containing protein 9 isoform X1			20	5996
6012	CCDC9	coiled-coil domain-containing protein 9			20	5992
6417	RBMX	RNA-binding motif protein, X chromosome			33	6384
5805	NFATC2IP	NFATC2-interacting protein			18	5787
6172	GAR1	H/ACA ribonucleoprotein complex subunit 1			30	6142
1493	IL13RA1	Uncharacterized protein			0	1493
5850	SFPQ	splicing factor, proline- and glutamine-rich			34	5816
5724	HMX1	homeobox protein HMX1			20	5704
2991	L2HGDH	L-2-hydroxyglutarate dehydrogenase, mitochondrial			0	2991

5733	HMX1	Uncharacterized protein	18	5715
5825	SFPQ	splicing factor, proline- and glutamine-rich isofo	28	5797
540	GNPTG	N-acetylglucosamine-1-phosphotransferase subunit g	12	528
6021	BRWD3	Uncharacterized protein	10	6011
589	GNPTG	N-acetylglucosamine-1-phosphotransferase subunit g	18	571
597	GNPTG	N-acetylglucosamine-1-phosphotransferase subunit g	18	579
414	FGA	Fibrinogen alpha chain	159	255
6115	BRWD3	bromodomain and WD repeat-containing protein 3 iso	10	6105
6115	BRWD3	bromodomain and WD repeat-containing protein 3 iso	10	6105
412	FGA	Fibrinogen alpha chain	159	253
4396	ADGRV1	Uncharacterized protein	26	4370
4548	PLA2G2A	phospholipase A2, group IIA (platelets, synovial f	4	4544
4545	PLA2G2A	Phospholipase A2, membrane associated	4	4541
262	LOC101902853	homeobox protein ESX1-like	0	262
4549	LOC100125947	uncharacterized protein LOC100125947 precursor	4	4545
4459	ADGRV1	G-protein coupled receptor 98	26	4433
6132	CTR9	Uncharacterized protein	28	6104
6148	CTR9	RNA polymerase-associated protein CTR9 homolog	26	6122
3191	MORF4L1	mortality factor 4-like protein 1 isoform X1	2	3189
3174	MORF4L1	mortality factor 4-like protein 1 isoform X3	2	3172
558	FGA	fibrinogen alpha chain isoform X1	163	395
2162	NXNL2	nucleoredoxin-like protein 2 isoform X1	4	2158
3102			8	3094
2581	H2AFX	histone H2A.x	6	2575
596	LOC100138633	N-acetylglucosamine-1-phosphotransferase subunit g	8	588
3081	MORF4L1	Uncharacterized protein	2	3079
2062	NXNL2	nucleoredoxin-like protein 2 isoform X2	4	2058
3139		Hemoglobin fetal subunit beta	1671	1468
1651	ETFA	Electron transfer flavoprotein subunit alpha, mito	2	1649
1652	ETFA	Electron transfer flavoprotein subunit alpha, mito	2	1650
348	PRDM11	Uncharacterized protein	213	135
382		fibrinogen A-alpha chain	152	230
1216	ZRSR2	U2 small nuclear ribonucleoprotein auxiliary facto	8	1208
1234	ZRSR2	Uncharacterized protein	8	1226
2121	PPP1R12B	Uncharacterized protein	10	2111
258	LOC787241	Uncharacterized protein	0	258
122	SUSD4	sushi domain-containing protein 4 isoform X1	0	122
122	SUSD4	Uncharacterized protein	0	122
338	HSPB8	Heat shock protein beta-8	2	336
1592	NFKBIL1	NF-kappa-B inhibitor-like protein 1 isoform X1	10	1582
1947	EME2	Uncharacterized protein	4	1943
6439	ITIH2	Inter-alpha-trypsin inhibitor heavy chain H2	3303	3136
6242	ITIH2	inter-alpha-trypsin inhibitor heavy chain H2 precu	3256	2986

1605	NFKBIL1	NF-kappa-B inhibitor-like protein 1	16	1589
1700		Uncharacterized protein	931	769
2239		Uncharacterized protein	6	2233
396	LUM	Lumican	168	228
1477	BDKRB1	B1 bradykinin receptor	2	1475
1476		Bradykinin B1 receptor (Fragment)	2	1474
441	LOC101902766	vegetative cell wall protein gp1-like	0	441
2310	AKT1	RAC-alpha serine/threonine-protein kinase	8	2302
2270	METTL25	methyltransferase like 25	4	2266
195	ACOT6	putative acyl-coenzyme A thioesterase 6 isoform X2	128	67
235	BAG4	BAG family molecular chaperone regulator 4	0	235
553	FBLN1	fibulin-1 precursor	86	467
554	FBLN1	Fibulin-1	86	468
643		Uncharacterized protein	0	643
1085	TTLL10	Uncharacterized protein	2	1083
605	TMEM143	Transmembrane protein 143	4	601
1475	CCDC124	Coiled-coil domain-containing protein 124	0	1475
2856			1495	1361
375	MCAT	MCAT protein (Fragment)	203	172
586	IK	protein Red	218	368
613	TMEM143	transmembrane protein 143 isoform X5	4	609
604	FLJ10922	hypothetical protein LOC55260	4	600
605	TMEM143	transmembrane protein 143 isoform X3	4	601
747		Uncharacterized protein C11orf71 homolog	2	745
615	TMEM143	transmembrane protein 143 isoform X2	4	611
1276	KEAP1	kelch-like ECH-associated protein 1	2	1274
613	TMEM143	transmembrane protein 143 isoform X4	4	609
1252	KEAP1	Kelch-like ECH-associated protein 1 (Fragment)	4	1248
1435	MRVI1	protein MRVI1 isoform X1	10	1425
1435	MRVI1	protein MRVI1 isoform a	10	1425
615	TMEM143	transmembrane protein 143 isoform X1	4	611
839	TMEM114	transmembrane protein 114 isoform X1	10	829
1437	MRVI1	Protein MRVI1	8	1429
1145	TTLL10	inactive polyglycylase TTLL10 isoform X2	4	1141
3222	LEO1	RNA polymerase-associated protein LEO1 isoform X3	2	3220
2161		Uncharacterized protein	1170	991
218	FUT2	galactoside 2-alpha-L-fucosyltransferase 2 isoform	0	218
438	LGMN	Legumain	7	431
814	TMEM114	transmembrane protein 114	6	808
1142	TTLL10	inactive polyglycylase TTLL10 isoform X1	10	1132
1867	GRK5	G protein-coupled receptor kinase 5	14	1853
439	LGMN	legumain isoform X1	8	431
123	SUSD4	sushi domain-containing protein 4	0	123

264	HRAS	GTPase HRas isoform 2	0	264
5051	AHDC1	AT-hook DNA-binding motif-containing protein 1	43	5008
1045	AMBP	Protein AMBP	403	642
2766	NTN3	netrin-3 precursor	4	2762
717	TMEM181	LOW QUALITY PROTEIN: transmembrane protein 181	30	687
1058	AMBP	Protein AMBP	403	655
905	UBN1	Uncharacterized protein	8	897
553	PRDM11	zinc finger protein 862 isoform X1	219	334
300	THBS4	Thrombospondin-4	67	233
821	URI1	Unconventional prefoldin RPB5 interactor	2	819
822	URI1	Unconventional prefoldin RPB5 interactor	2	820
828	C19orf2	RPB5-mediating protein	2	826
237	ACOT6	Uncharacterized protein	120	117
921	CNTD2	cyclin N-terminal domain-containing protein 2 isof	2	919
489	PRDM11	zinc finger protein 862 isoform X3	217	272
1040	CCBE1	Uncharacterized protein	10	1030
1127	CCBE1	collagen and calcium-binding EGF domain-containing	20	1107
218	FUT2	Galactoside 2-alpha-L-fucosyltransferase 2	0	218
194	KLNK	Kallikrein K	105	89
286	HRAS	GTPase HRas isoform 1	0	286
780	TMEM220	transmembrane protein 220	0	780
958	UBN1	ubinuclein-1 isoform X1	10	948
1784	GRK5	G protein-coupled receptor kinase 5 isoform X1	14	1770
811	SSPN	sarcospan	4	807
953			8	945
996	JADE2	protein Jade-2 isoform X4	8	988
602	PFDN6	Prefoldin subunit 6	4	598
942	CNTD2	cyclin N-terminal domain-containing protein 2 isof	2	940
620	LOC515676	keratin-associated protein 10-6 isoform X3	4	616
590	PFDN6	prefoldin subunit 6 isoform X3	4	586
1133	LOC101906155	uncharacterized protein LOC101906155	10	1123
984	JADE2	protein Jade-2 isoform X1	8	976

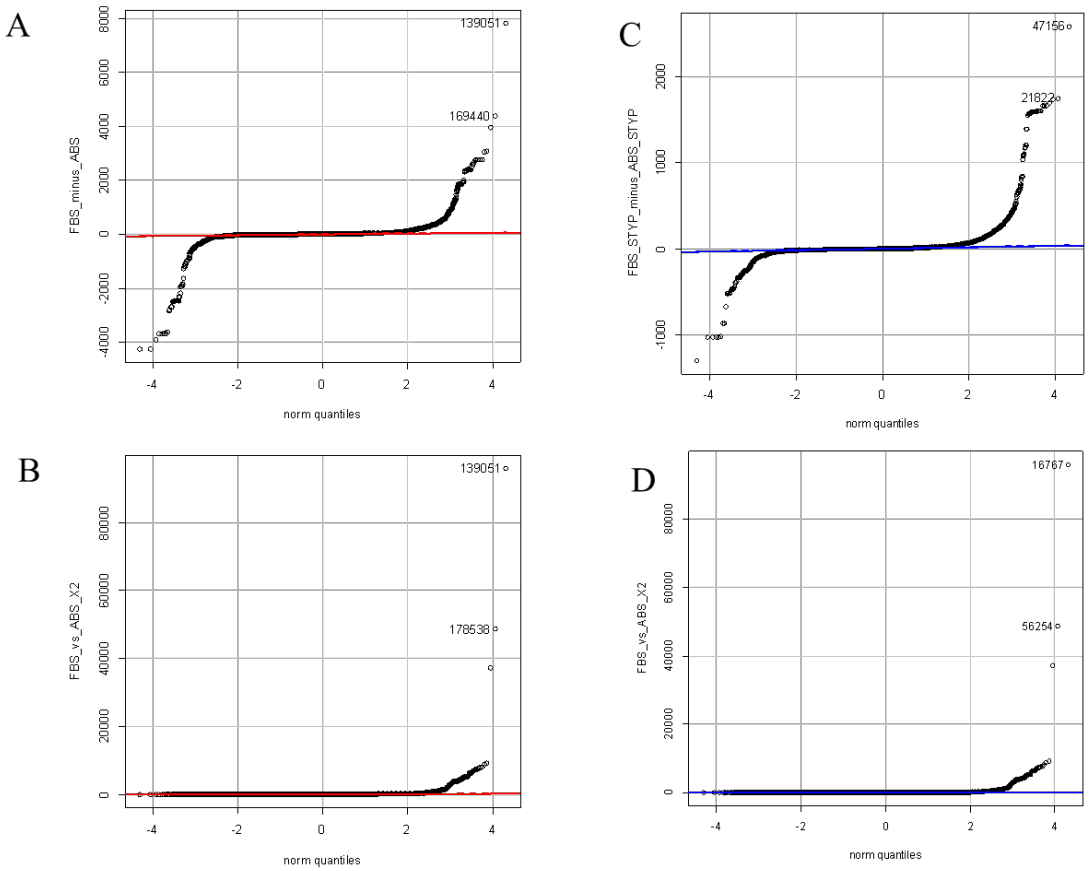


Figure 9. Quantile plots of the corrected difference in observation frequency (Delta) and Chi Square values of the fetal versus adult serum. The difference in observation frequency of fetal serum versus adult serum using the Quantile plot that tended to zero (see red line). Panels: A, Tryptic peptide corrected difference (delta) in observation frequency; B, tryptic peptide Chi Square χ2; C, Tryptic and/or phosphor-tryptic peptides after the corrected difference (delta) in observation frequency; D, tryptic and/or STYP peptide Chi Square χ2.

4.10.Erythrocytes

Hemoglobin fetal subunit beta, Hemoglobin subunit beta that were observed in the fetal versus adult serum thousands of times in close proportions, similar to the ratio of sampled MS/MS spectra from fetal versus adult serum but few other known red blood cells proteins were observed. Some, but not all proteins annotated from erythrocyte or red blood cells were observed in the fetal and/or adults serum.

4.11.Granulocytes, macrophages, monocytes, neutrophils, platelets or rbcs

It remains possible that the cellular proteins observed in serum were released from white blood cells during sample collection. However only a small minority of all known macrophage, granulocyte, neutrophil, monocyte proteins were observed in plasma by X!TANDEM that typically showed little trend to differences between fetal and adult plasma. For example, macrophage receptor MARCO isoform, macrophage scavenger receptors macrophage-stimulating protein receptor, natural resistance-associated macrophage protein, macrophage colony-stimulating factor, macrophage mannose receptor, phospholipase A2, group IIA and other were observed in serum. Similarly, a small subset of all known granulocyte, neutrophil, monocyte or mast cell proteins were observed in plasma. Specific platelet associated proteins including platelet glycoprotein IV (CD36) or platelet endothelial cell adhesion molecule isoform, platelet glycoprotein Ib alpha chain, Platelet-derived growth factor D isoform X1, Platelet-activating factor receptor, platelet-activating factor receptor isoform X1 and platelet factor 4 (PF4) were observed in fetal serum, however most of the known platelet associated proteins were not strongly detected.

4.12.Hormones, ligands, growth factors, chemokines cytokines, necrosis factors and interleukins

A biased search was performed for tryptic and/or phosphor-peptides from factors identified by X!TANDEM and counted with SEQUEST in fetal versus adult serum that might play regulatory roles in signal reception, gene expression related to cell growth and development that resulted in a set of gene symbols showed Chi Square ($\chi 2$) values of ≥ 9. The SQL Server database was queried for peptides and/or phosphor-peptides for ligands, receptors, growth factors, cytokines, chemokines, interleukins, and necrosis factors along with their receptors, kinases, phosphatases, g-proteins, cyclases, phosphodiesterases, transcription factors, nucleic acid binding proteins identified by X!TANDEM that showed greater frequency in fetal serum (Table 5). Growth or regulatory hormones that showed increased observation frequency in fetal serum included parathyroid hormone/parathyroid hormone-related peptide receptor, hormone-sensitive lipase-like, Lutropin-choriogonadotropic hormone receptor, parathyroid hormone 2 receptor precursor, thyrotropin-releasing hormone, thyrotropin-releasing hormone receptor, glycoprotein hormone beta-5, steroid hormone receptor ERR1, anti-Muellerian hormone type-2 receptor precursor. Platelet associated molecules that were enriched in fetal serum by Chi Square analysis included Platelet factor 4, platelet endothelial cell adhesion molecule, platelet-derived growth factor C, platelet glycoprotein 4, Platelet-derived growth factor receptor-like detected by X!TANDEM and SEQUEST but some other platelet derived growth factors were only detected by SEQUEST. Multiple growth factors and associated receptors or binding proteins showed greater observation frequency in fetal serum including macrophage colony-stimulating factor 1, multiple epidermal growth factor-like domains, lens epithelium-derived growth factor, transforming growth factor beta-2, pro-epidermal growth factor, Insulin-like growth factor-binding protein, latent-

transforming growth factor beta-binding protein, Insulin-like growth factor-binding protein, hepatoma-derived growth factor-related protein 3, fibroblast growth factor 7 precursor, Fibroblast growth factor receptor, Fibroblast growth factor, Fibroblast growth factor 18, insulin-like growth factor II isoform X, Connective tissue growth factor and others. C-X-C motif chemokine 6, C-X-C chemokine receptor type 4, C-C motif chemokine 21 precursor, Chemokine (C-C motif) ligand 21, C-C motif chemokine 8, cardiotrophin-like cytokine factor 1, cytokine receptor-like factor 1 isoform X1, Complement C1q and tumor necrosis factor-related protein, Tumor necrosis factor alpha-induced protein 8-like, tumor necrosis factor (ligand) superfamily members. Interleukins or associated factors were observed including toll/interleukin-1 receptor domain-containing adaptor protein (TIRAP), pro-interleukin-16 isoform X1 (IL16), interleukin-17D (IL17), interleukin 8 receptor alpha (CXCR1), interleukin-27 subunit alpha precursor (IL27) and interleukin-18 receptor accessory protein (IL18RAP). Additional factors that were observed included nuclear receptor corepressor 2 isoform X2, cytokine receptor-like factor 1 isoform X1, Atrial natriuretic peptide receptor 2, scavenger receptor class A member 3, ALK tyrosine kinase receptor and others. Specific examples of fetal specific proteins include G-protein coupled receptor 98, Leucine-rich repeat-containing G-protein-coupled receptor, G-protein-signaling modulator 1, adhesion G-protein coupled receptor G7, serine/threonine-protein phosphatase 2B catalytic, receptor-type tyrosine-protein phosphatase mu, ectonucleotide pyrophosphatase/phosphodiesterase, Protein phosphatase 2, G protein-coupled receptor kinase 5, adenylate kinase 8, Serine/threonine-protein kinase 33, Mitogen-activated protein kinase 13, RNA-binding motif protein, AT-hook DNA-binding motif-containing protein, transcription factor 15, doublesex- and mab-3-related transcription factor, transcriptional enhancer factor TEF-3, pre-B-cell leukemia transcription factor 2, chromodomain-helicase-DNA-binding protein 1,

zinc finger protein 862 and others. The zinc finger and BTB domain-containing protein 20 and many nuclear receptors including ROR-alpha were observed in both fetal and adult serum. Many proteins that were enriched by counting tryptic peptides showed even greater observation frequency in fetal serum when phosphorylation was taken into account, for example alpha feta protein (AFP).

Table 5. The proteins shown were identified by X!TANDEM with at least three independent peptide identifications and showed significant Chi Square values (≥9) from both tryptic (TRYP) and phosphor-tryptic peptides.

GeneSymb	DescriptionFasta	Fetal_TRYP_X2	Fetal_TRYP_Delta	Fetal_STYP_X2	Fetal_STYP_Delta
AFP	Alpha-fetoprotein	5348.602116	1113.946	1556.325992	449.8026
NULL	Hemoglobin fetal subunit beta	854.2719438	676.044	176.7079574	272.4286
PDGFC	platelet-derived growth factor C isoform X2	12.3596	11.66	0.636712967	3.8268
FGF13	fibroblast growth factor 13 isoform X2	9.058456333	10.426	2.512436338	4.7552
TNF	Tumor necrosis factor	17.34245633	14.426	1.982581813	2.4388
HDGFRP3	hepatoma-derived growth factor-related protein 3	25.42848711	15.128	0.217444516	1.4746
INS	insulin isoform X1	22.74426021	20.788	44.60723627	34.7044
NULL	Insulin (Fragment)	24.265008	8.532	12.80995681	7.1582
ADGRL3	Adhesion G protein-coupled receptor L3	9.326156108	18.576	9.950830602	-15.4538
ADGRV1	G-protein coupled receptor 98	1112.372984	626.632	2195.764741	1061.3328
ADORA3	adenosine receptor A3 isoform 2 precursor	28.86375625	21.49	55.10568452	23.4746
ADRA2B	Alpha-2B adrenergic receptor (Fragment)	19.21192467	28.406	67.9008967	64.8834
ALK	ALK tyrosine kinase receptor	55.88393086	67.28	13.32430466	-33.851
CRLF1	cytokine receptor-like factor 1 isoform X1	31.61250625	22.49	41.1135674	27.9492
GPR68	ovarian cancer G-protein coupled receptor 1 isoform X1	33.4643805	16.362	15.35385856	11.7552
GPR84	G-protein coupled receptor 84 isoform X1	19.4724135	21.618	29.74095992	33.6178
GRID1	glutamate receptor ionotropic, delta-1	71.7184188	70.854	314.3117279	189.2922
GRIN3B	glutamate receptor ionotropic, NMDA 3B isoform X1	194.6561805	118.386	10.33909278	23.4088
GRK4	G protein-coupled receptor kinase 4 isoform X13	32.68163462	29.15	17.17261096	31.2864
GRK5	G protein-coupled receptor kinase 5	373.5386091	215.218	495.8210284	340.6202
NPR2	Atrial natriuretic peptide receptor 2	24.68672653	27.214	57.73592009	48.0566
NULL	Ig-like receptor A3	12.3596	11.66	85.10406817	22.597
OR10C1	olfactory receptor 10C1	62.13958017	38.618	162.6927626	69.8626
PRLR	Prolactin receptor (Fragment)	47.71521424	40.278	9.557999923	14.8268
SCARA3	scavenger receptor class A member 3	26.52385225	54.26	366.8439746	311.2022
TRPM3	LOW QUALITY PROTEIN: transient receptor potential cation chan	20.1097216	41.344	69.6317658	85.0982
TRPM7	transient receptor potential cation channel subfamily M member	16.2294087	48.006	25.73137734	56.4862
LGR4	Leucine-rich repeat-containing G-protein-coupled receptor 4	36.60403776	35.278	16.09956841	34.5162
ARAF	serine/threonine-protein kinase A-Raf	14.6906317	19.916	9.134377013	15.7044
ARHGAP22	rho GTPase-activating protein 22 isoform X6	33.83385306	60.728	66.80683781	115.88
BAZ1B	tyrosine-protein kinase BAZ1B isoform X1	13.58472512	42.346	38.24375118	72.119
CHKB	choline/ethanolamine kinase isoform X2	77.8075236	27.894	21.0356952	27.8984
DYRK1B	dual specificity tyrosine-phosphorylation-regulated kinase 1B isof	15.45208278	18.852	39.50967953	43.0924
ENPP1	ectonucleotide pyrophosphatase/phosphodiesterase family mem	46.70131284	29.788	28.42717701	27.7044
ERMARD	endoplasmic reticulum membrane-associated RNA degradation p	14.83352337	16.788	13.15999009	-22.9434
GUK1	guanylate kinase isoform X6	36.51382044	18.128	20.94301009	-28.9434
INPP5J	phosphatidylinositol 4,5-bisphosphate 5-phosphatase A isoform X	17.52325008	29.002	13.86270096	18.985
IRS2	LOW QUALITY PROTEIN: insulin receptor substrate 2	12.34738713	16.852	29.67798006	21.791
JAK2	Tyrosine-protein kinase	16.37276166	33.854	54.48583008	70.0266
LOC10713	Guanylate cyclase	34.60903153	24.256	71.9202806	41.5462
LOC10713	hormone-sensitive lipase-like	17.51585506	17.256	10.46691381	18.3014
LRRK2	leucine-rich repeat serine/threonine-protein kinase 2 isoform X1	231.156482	169.984	207.2239401	180.9458
MAP3K13	Mitogen-activated protein kinase kinase kinase 13	124.2969135	54.618	9.821326633	-18.5404
MAP3K5	mitogen-activated protein kinase kinase kinase 5	17.8233082	32.428	13.33753	-31.2032
MARS	methionine--tRNA ligase, cytoplasmic isoform X2	22.86991506	38.258	162.9276714	104.4804
MYLK3	myosin light chain kinase 3 isoform X1	13.03970312	25.534	18.77508114	-32.7136
NULL	guanylate cyclase	43.62402689	28.022	84.90731635	46.985
PLK5	inactive serine/threonine-protein kinase PLK5 isoform X1	64.47198084	76.174	94.94012583	111.522
RALGAPA1	ral GTPase-activating protein subunit alpha-1 isoform X6	15.97455031	41.728	15.64689987	33.7968
SRPK1	LOW QUALITY PROTEIN: SRSF protein kinase 1	13.14351607	27.13	17.51624315	43.6952
STK33	serine/threonine-protein kinase 33 isoform X3	161.9011727	94.364	242.5024154	123.6028
STK33	serine/threonine-protein kinase 33 isoform X1	179.512809	99.364	229.140802	122.0416
TTBK2	tau-tubulin kinase 2	56.72732498	54.832	23.59860779	30.3372
UPP1	uridine phosphorylase 1 isoform X1	9.058456333	10.426	12.3853946	14.5104
LOC52950(	aryl hydrocarbon receptor repressor	13.916126	13.958	41.9056751	25.0716

4.13.Growth factors

It has been previously demonstrated that fetal plasma contains insulin, insulin like growth factors and other transforming growth factors [80]. Multiple growth factors and associated receptors or binding proteins showed greater observation frequency in fetal serum including macrophage colony-stimulating factor 1, multiple epidermal growth factor-like domains, lens epithelium-derived growth factor, transforming growth factor beta-2, pro-epidermal growth factor, Insulin-like growth factor-binding protein, latent-transforming growth factor beta-binding protein, Insulin-like growth factor-binding protein, hepatoma-derived growth factor-related protein 3, fibroblast growth factor 7 precursor, Fibroblast growth factor receptor, Fibroblast growth factor, Fibroblast growth factor 18, insulin-like growth factor II isoform X, connective tissue growth factor and others (Table 6).

Table 6. The growth factors that differ between fetal versus adult serum by Chi Square analysis of the tryptic (TRYP) and or serine, threonine or tyrosine phosphorylated tryptic peptides (STYP) identified by SEQUEST and X!TANDEM.

TotalProteinIde	GeneSymbol	DescriptionFasta	TANDEM	PD_MSF	FBS_minus_ABS	FBS_minus_ABS_STYP	FBS_minus_ABS_X2	FBS_minus_ABS_STYP_X2
532	EGF	pro-epidermal growth factor	6	526	78.302291	7.01224416	62.56376302	1.49004752
257	MEGF6	multiple epidermal growth factor-like domains prot	6	251	63.086063	-9.11122532	180.9023339	2.243633158
260	MEGF6	multiple epidermal growth factor-like domains prot	4	256	62.32006	-8.11122532	184.9423752	1.778161519
478	Ledgf	lens epithelium-derived growth factor	4	474	61.810105	97.08010098	106.1246967	200.5222555
478	Ledgf	lens epithelium-derived growth factor	4	474	61.810105	97.08010098	106.1246967	200.5222555
357	MEGF6	multiple epidermal growth factor-like domains prot	8	349	60.448084	-21.69464431	125.9989951	7.354024871
172	bovine IGF2	Insulin-like growth factor 2 (Fragment)	6	166	46.384072	10.62959156	86.05928541	8.691401287
517	MEGF8	multiple epidermal growth factor-like domains prot	18	499	34.534147	-3.78418538	23.85214618	0.19093412
181	IGFBP-4	Insulin-like growth factor-binding protein 4 (Frag	4	177	33.086063	-1.74081688	49.75852567	0.121217736
178	MEGF11	multiple epidermal growth factor-like domains prot	2	176	32.576108	-1.08954107	28.68115709	0.098924979
199	IGFBP4	Insulin-like growth factor-binding protein 4	7	192	31.618069	-5.4260211	41.65426197	0.949732419
192	MEGF11	multiple epidermal growth factor-like domains prot	2	190	28.874117	-2.37040844	20.84286581	0.432218167
192	MEGF11	multiple epidermal growth factor-like domains prot	2	190	28.874117	1.47219367	20.84286581	0.21673542
511	LTBP4	latent-transforming growth factor beta-binding pro	4	507	28.384072	-5.63622758	32.22622173	0.235311565
246	FGFR2	Fibroblast growth factor receptor	6	240	19.32006	-31.00944009	17.7745104	16.57905818
246	FGFR2	Fibroblast growth factor receptor	6	240	19.32006	-31.00944009	17.7745104	16.57905818
243	FGFR2	fibroblast growth factor receptor 2 isoform X3	6	237	17.086063	-32.00944009	13.26970677	17.6655906
243	FGFR2	fibroblast growth factor receptor 2 isoform X3	6	237	17.086063	-32.00944009	13.26970677	17.6655906
243	FGFR2	fibroblast growth factor receptor 2 isoform X5	6	237	16.086063	-32.00944009	11.76188286	17.6655906
243	FGFR2	fibroblast growth factor receptor 2 isoform X5	6	237	16.086063	-32.00944009	11.76188286	17.6655906
84	HDGFRP3	hepatoma-derived growth factor-related protein 3	13	71	15.128024	1.47219367	25.42856779	0.21673542
251	FGFR2	fibroblast growth factor receptor 2 isoform X2	8	243	14.618069	-31.00944009	8.90366422	16.57905818
230	FGFR	Fibroblast growth factor receptor	8	222	14.32006	-25.88597061	9.764958019	12.40895323
213	FGFR	Fibroblast growth factor receptor	6	207	14.022051	-26.60510324	10.92321746	13.35531167
251	FGFR2	fibroblast growth factor receptor 2 isoform X6	8	243	13.618069	-31.00944009	7.72715847	16.57905818
500	LTBP3	latent-transforming growth factor beta-binding pro	6	494	13.30015	40.40714092	3.468509608	19.20867103
503	LTBP3	latent-transforming growth factor beta-binding pro	6	497	13.30015	40.12627355	3.468509608	18.72230034
503	LTBP3	latent-transforming growth factor beta-binding pro	6	497	13.30015	40.12627355	3.468509608	18.72230034
504	LTBP3	latent-transforming growth factor beta-binding pro	6	498	13.30015	38.84540618	3.468509608	17.34443197
137	FGF18	Fibroblast growth factor	2	135	13.192036	35.75306104	13.38690876	142.0312637
137	FGF18	Fibroblast growth factor	2	135	13.192036	35.75306104	13.38690876	142.0312637
233	FGFR2	fibroblast growth factor receptor 2	6	227	12.852066	-27.88597061	7.181547846	14.40050661
207		Fibroblast growth factor receptor 2 variant 6	6	201	12.32006	-25.76250113	7.227803734	13.27412929
204	FGFR2	fibroblast growth factor receptor 2 isoform X11	6	198	12.022051	-28.32423587	8.029428347	15.42812188
138	FGF18	Fibroblast growth factor 18	2	136	11.958039	35.75306104	10.21390691	142.0312637
138	FGF18	Fibroblast growth factor 18	2	136	11.958039	35.75306104	10.21390691	142.0312637
154	IGF2BP3	insulin-like growth factor 2 mRNA-binding protein	10	144	11.724042	-3.14515373	9.163544055	0.329733066
141	PDGFC	platelet-derived growth factor C isoform X2	10	131	11.66003	3.82091786	12.3596636	0.6347571
141	PDGFC	platelet-derived growth factor C isoform X1	10	131	11.66003	3.82091786	12.3596636	0.6347571
215	FGFR2	fibroblast growth factor receptor 2 isoform X8	8	207	11.554057	-27.32423587	6.674811658	14.35795896
213		Fibroblast growth factor receptor	8	205	11.554057	-23.76250113	6.674811658	11.2931292
215	FGFR2	fibroblast growth factor receptor 2 isoform X12	8	207	10.554057	-27.32423587	5.569405958	14.35795896
88	FGF13	fibroblast growth factor 13 isoform X2	8	80	10.426033	4.75306104	9.058513676	2.510176583
241	FGFR2	fibroblast growth factor receptor 2 isoform X1	8	233	10.384072	-26.88597061	4.313158052	13.3862114
212	FGFR2	fibroblast growth factor receptor 2 isoform X10	8	204	9.554057	-27.32423587	4.564000258	14.35795896
241	FGFR2	fibroblast growth factor receptor 2 isoform X4	8	233	9.384072	-26.88597061	3.522432292	13.3862114
205	FGFR2	fibroblast growth factor receptor 2 isoform X7	8	197	7.32006	-23.20076639	2.551584686	11.21407419
202	FGFR2	fibroblast growth factor receptor 2 isoform X9	8	194	5.32006	-23.20076639	1.347763734	11.21407419
90	IGF2	insulin-like growth factor II isoform X1	8	82	4.192036	13.75306104	1.351781987	21.01629866
102	CTGF	Connective tissue growth factor	8	94	3.958039	12.75306104	1.119005195	18.07117399
100	NTRK2	BDNF/NT-3 growth factors receptor isoform X3	8	92	3.724042	-14.21301055	0.924565921	12.62560431
197	IGF1R	Insulin-like growth factor 1 receptor (Fragment)	22	175	3.682081	15.3826526	0.484204303	11.26790481
30		insulin-like growth factor-binding protein 5	2	28	3.064012	3	1.877633907	9
91	GRB10	growth factor receptor-bound protein 10 isoform X2	13	78	1.66003	10.1913263	0.250518146	9.442102887
36	TGFBR3L	transforming growth factor-beta receptor type 3-li	4	32	0.596018	6.71913263	0.050748208	22.57337165
436	EGFR	epidermal growth factor receptor	10	426	-6.997859	26.99055991	1.02020897	12.560178
112	FIBP	acidic fibroblast growth factor intracellular-bind	2	110	-7.381931	13.75306104	2.27053772	21.01629866
112	FIBP	acidic fibroblast growth factor intracellular-bind	2	110	-7.381931	13.75306104	2.27053772	21.01629866
109	PDGFRL	Platelet-derived growth factor receptor-like prote	2	107	-7.977949	15.1913263	3.53598168	20.97967225
115	PDGFRL	platelet-derived growth factor receptor-like isofo	4	111	-11.211946	12.34872419	6.616196479	10.89221351
119	PDGFRL	Platelet-derived growth factor receptor-like prote	4	115	-13.445943	11.06785682	9.039669158	8.166496973

4.14.Pathway and gene ontology analysis using the STRING algorithm

Many tryptic peptides (TYRP) and/or phospho-tryptic peptides (STYP) from growth or regulatory pathways were identified by both X!TANDEM and SEQUEST showed greater observation frequency by Chi Square in fetal compared to adult serum (Fig. 10). In a computationally independent method, the fetal proteins that showed a higher observation frequency from both tryptic and phosphorylated tryptic peptides were analyzed by the STRING protein-protein interaction network. The STRING analysis showed a significant p-value that so many previously established protein-protein interactions were just randomly sampled from the genome (Fig. 10&11). The distribution of the cellular locations, molecular functions, and biological processes (Table 7) were significant with respect to the category distribution of the human genome. The proteins that were observed to differ between the fetal and adult serum by unbiased Chi Square analysis were organized at least in part around the PI3K, AKT, HRAS growth and proliferation pathways [81] that may include hormones and ligands, receptors, receptor associated enzymes such as kinases, phosphatases, cyclases and g-proteins or their associated factors, as well as transcription factors, DNA and RNA binding proteins and specific histones or histone modification enzymes. Calmodulin-independent adenylate cyclase, calmodulin-binding transcription activator, calcium/calmodulin-dependent protein kinase type 1, Calcium/calmodulin-dependent 3',5'-cyclic nucleotide phosphodiesterase showed significantly increased observation frequency (Fig. 10&11).

Table 7. The summary of STRING analysis of the fetal serum enriched proteins with respect to a random sampling of the human Genome for Gene Symbols showed the top 400 proteins with the most significant Chi Square ($\chi 2$) value (See Fig.11). The Gene Symbols from unbiased Table IV and biased SQL Queries V were combined for analysis by the STRING algorithm that showed strong enrichment and protein-protein interactions in the fetal specific factors.

Network Stats	
Number of nodes:	292
Number of edges:	461
Average node degree:	3.16
Avg. local clustering coefficient:	0.447
Expected number of edges:	331
PPI enrichment p-value:	9.60E-12
Your network has significantly more interactions than expected (what does that mean?)	
Functional enrichments in your network	

Biological Process (GO)	
Pathway ID	pathway description
GO:0051346	negative regulation of hydrolase activity
GO:0072376	protein activation cascade
GO:0010466	negative regulation of peptidase activity
GO:2000257	regulation of protein activation cascade
GO:0045861	negative regulation of proteolysis
GO:0010951	negative regulation of endopeptidase activity
GO:0010605	negative regulation of macromolecule metabolic process
GO:0030449	regulation of complement activation
GO:0043086	negative regulation of catalytic activity
GO:0051248	negative regulation of protein metabolic process
GO:0044092	negative regulation of molecular function
GO:0031324	negative regulation of cellular metabolic process
GO:0002920	regulation of humoral immune response
GO:0009892	negative regulation of metabolic process
GO:0051336	regulation of hydrolase activity
GO:0006953	acute-phase response
GO:0052547	regulation of peptidase activity
GO:0006956	complement activation
GO:0032269	negative regulation of cellular protein metabolic process
GO:0002526	acute inflammatory response
GO:0070613	regulation of protein processing
GO:0030162	regulation of proteolysis
GO:0006959	humoral immune response

GO:0052548	regulation of endopeptidase activity
GO:0002673	regulation of acute inflammatory response
GO:0072378	blood coagulation, fibrin clot formation
GO:0006957	complement activation, alternative pathway
GO:0030212	hyaluronan metabolic process
GO:0044710	single-organism metabolic process
GO:0048585	negative regulation of response to stimulus
GO:0050994	regulation of lipid catabolic process
GO:1903034	regulation of response to wounding
GO:0006958	complement activation, classical pathway
GO:0006103	2-oxoglutarate metabolic process
GO:0051004	regulation of lipoprotein lipase activity
GO:0034377	plasma lipoprotein particle assembly
GO:0048523	negative regulation of cellular process
GO:0071827	plasma lipoprotein particle organization
GO:0090207	regulation of triglyceride metabolic process
GO:0002455	humoral immune response mediated by circulating immunoglobulin
GO:2001234	negative regulation of apoptotic signaling pathway
GO:0016584	nucleosome positioning
(less ...)	

Molecular Function (GO)	
Pathway ID	pathway description
GO:0061134	peptidase regulator activity
GO:0030414	peptidase inhibitor activity
GO:0061135	endopeptidase regulator activity
GO:0004866	endopeptidase inhibitor activity
GO:0004857	enzyme inhibitor activity
GO:0030234	enzyme regulator activity
(less ...)	

Cellular Component (GO)	
Pathway ID	pathway description
GO:0072562	blood microparticle
GO:0005615	extracellular space
GO:0044421	extracellular region part
GO:0034364	high-density lipoprotein particle
GO:0031982	vesicle
GO:0031988	membrane-bounded vesicle
GO:0070062	extracellular exosome
GO:0005576	extracellular region
GO:0060205	cytoplasmic membrane-bounded vesicle lumen

GO:0071682	endocytic vesicle lumen
GO:0043227	membrane-bounded organelle
GO:0043226	organelle
GO:0034366	spherical high-density lipoprotein particle
GO:0031012	extracellular matrix
GO:0030496	midbody
(less ...)	

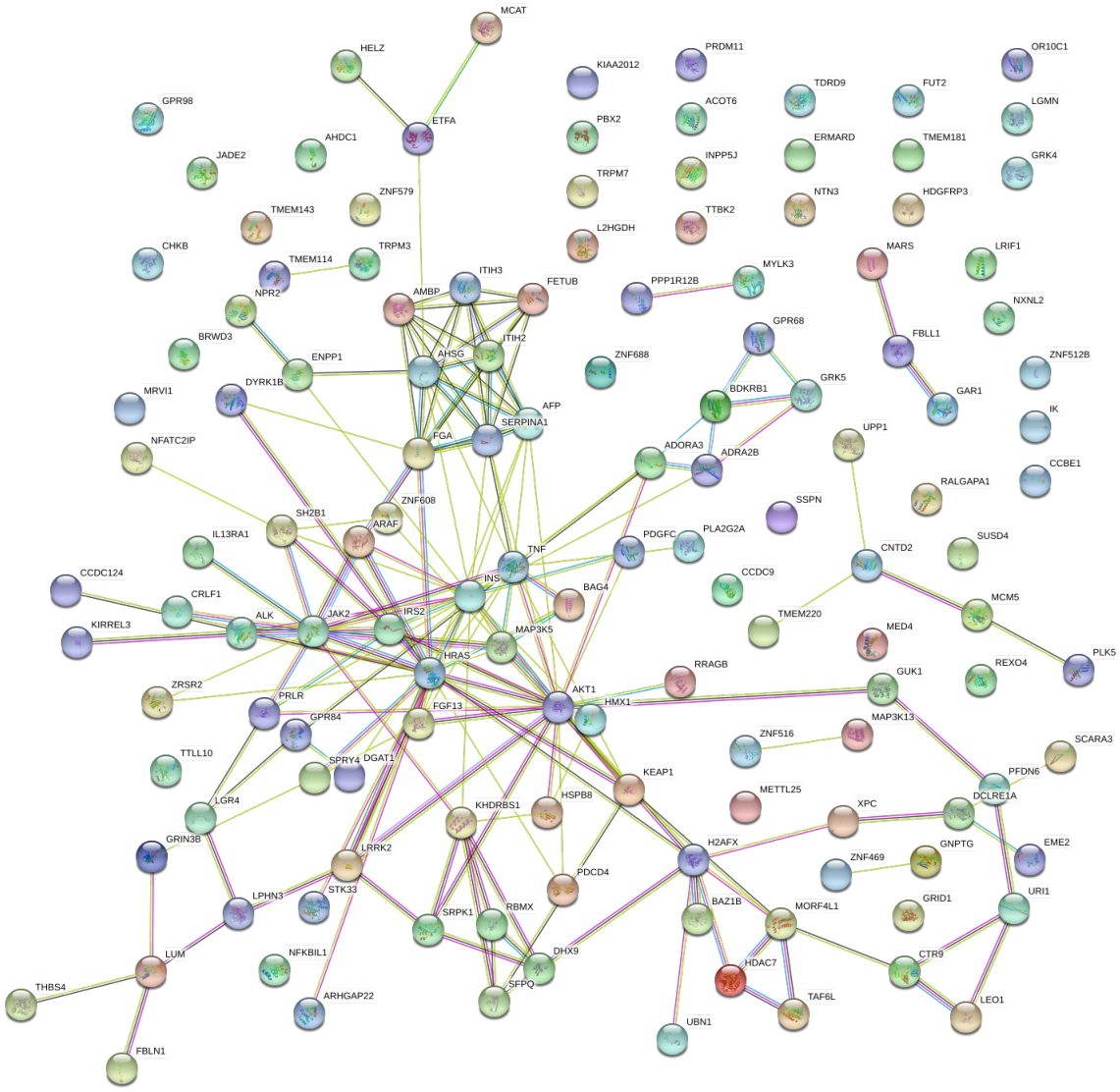


Figure 10. The Fetal serum STRING network where Chi Square $\chi 2 \geq 15$ from fully tryptic peptides. The network shown is where fetal serum observations exceeded those of adult serum by tryptic peptide frequency difference with $\chi 2$ value greater than 15 at degrees of freedom of 1 ($p < 0.0001$).

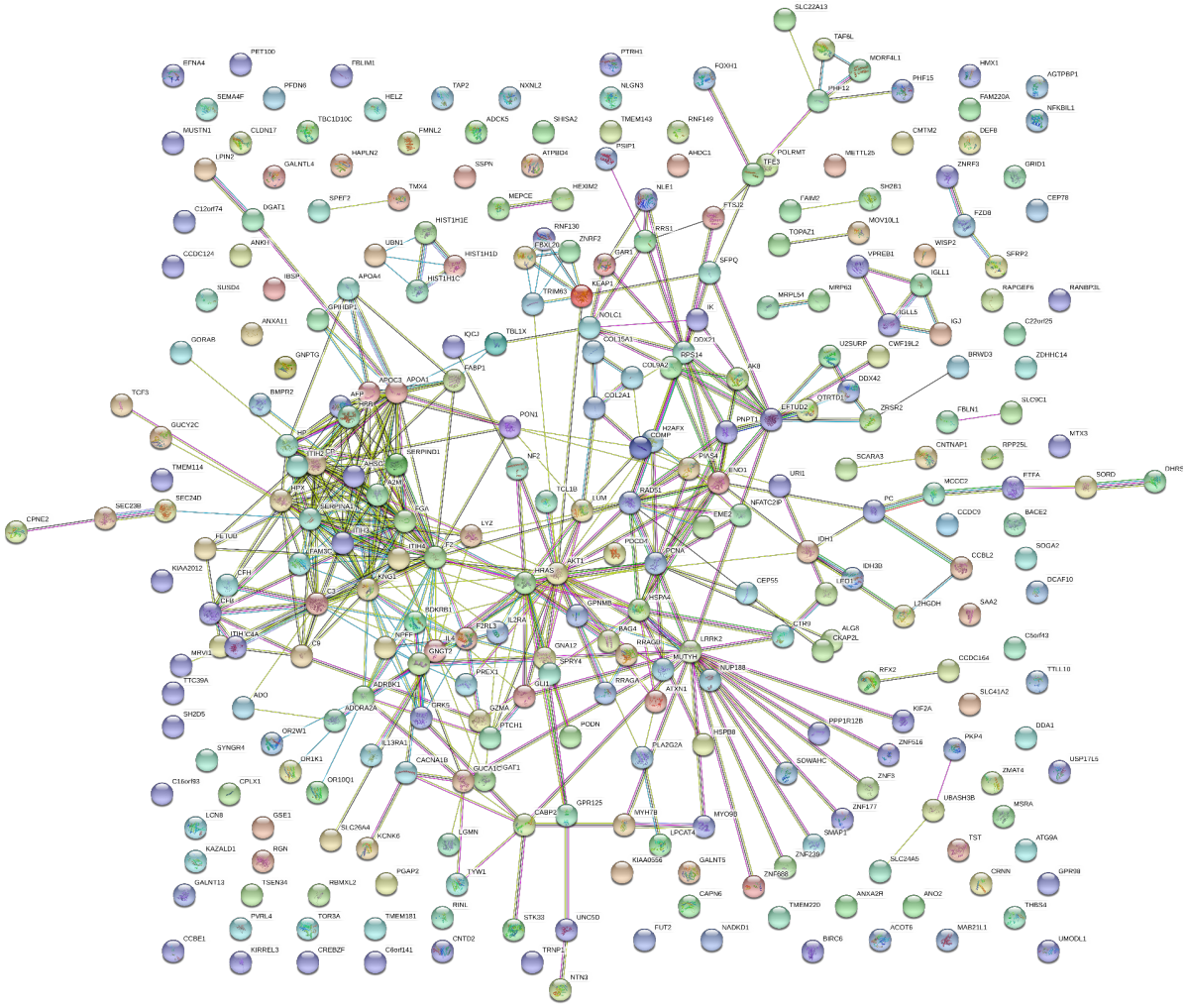


Figure 11. The STRING network shows the top 400 proteins with the most significant chi-square value in the fetal serum. Many AKT/HRAS proliferation pathways as well as AFP and Insulin like growth factors were observed.

4.15.ANOVA quantification of alpha feto protein (AFP)

AFP was observed to be specific to the fetal serum and was markedly increased in the observation frequency compared to adult serum. Two independent computations returned similar results with the same protein AFP implicated. Exemplary MS/MS spectra were observed from Alpha feto protein (Fig. 12). Analysis of log10 transformed precursor intensity from MS/MS spectra correlated by X!TANDEM showed significant increased in Alpha-feto protein in all three of the fetal serum replicates both with and without phosphorylation (Table 8). Some peptides were found to be specific to AFP and highly observed in fetal serum. Variations of log10 transformed precursor intensity of peptides in AFP indicated that different peptides were observed in fetal versus adult serum but they are all specific to AFP, which was shown by the normality distribution (Fig. 13). The peptides LGEYYLQNAFLVAYTK, KAPQLTSPELMALTR, FLGDRDFNQLSSR, and QKPQITEEQLETVVADFSGLLEK were examples of sequences found to be unique to AFP and not Albumin (searched against BLAST), and they were only observed in fetal serum with enriched amounts compared to other sequences (Table 9 & Fig. 14).

Table 8. The mean and standard deviation of the Log10 precursor intensity values for treatments between adult versus fetal serum. Alpha-feto protein was highly enriched in FBS over the partition chromatography extraction (See appendix Table 10).

	Mean Intensity	Standard deviation	AFP Counts	Treatment Name
1	5.478582	0.58540897	30	ABS-NaCl Extraction
2	5.210961	0.45857083	20	ABS- NaCl Extraction PSTY
3	6.066390	0.80549511	90	ABS- NaCl Extraction
4	6.123193	0.72567329	44	ABS- NaCl Extraction PSTY
5	5.450042	0.54669255	40	ABS- NaCl Extraction
6	5.393417	0.54968106	14	ABS- NaCl Extraction PSTY
7	NA	NA	NA	ABS-ACN Extraction
8	4.756725	0.22025096	2	ABS-ACN Extraction PSTY
9	4.875267	0.55411930	9	ABS- ACN Extraction
10	4.942789	0.41958090	7	ABS- ACN Extraction PSTY
11	4.591639	0.38176752	4	ABS- ACN Extraction
12	4.813441	0.74986619	5	ABS- ACN Extraction PSTY
13	5.323178	0.45386238	105	FBS-NaCl Extraction
14	5.310902	0.43212468	48	FBS- NaCl Extraction PSTY
15	5.294939	0.57956018	873	FBS- NaCl Extraction
16	5.359196	0.55721166	344	FBS- NaCl Extraction PSTY
17	5.027520	0.41713324	307	FBS-NaCl Extraction
18	5.028439	0.44348405	153	FBS- NaCl Extraction PSTY
19	5.307961	NA	1	FBS-ACN Extraction
20	4.715307	0.11598031	2	FBS-ACN Extraction PSTY
21	4.250918	0.26384156	6	FBS-ACN Extraction
22	4.410122	0.36034454	18	FBS-ACN Extraction PSYT
23	4.637643	0.24639235	4	FBS-ACN Extraction
24	5.005919	0.09273483	2	FBS-ACN Extraction PSTY

Table 9. Peptide sequences specific to AFP. FLGDRDFNQLSSR, LGEYYLQNAFLVAYTK, QKPQITEEQLETVVADFSGLLEK, and KAPQLTSPELMALTR were examples of peptides that were specific to AFP but not albumin with enriched amounts (searched against BLAST) and found in fetal serum only.

Peptide Sequences	Intensity Values	Counts
AAPASIPPFPVLEPVTSCK	5.178533	15
AENATECFETK	4.741233	2
AENATECFETKVTSITK	5.174305	1
AFLVAYTKKAPQLTSPELMALTR	5.12508	11
AITVTK	5.030537	1
APQLTSPELMALTR	5.163889	11
CCQSQEQEVCFTEEGPALISK	5.261659	6
DFNQLSSR	4.803041	10
DFNQLSSRDKDLSMARFTYEYSR	5.229139	1
DKDLSMAR	5.027066	32
DKDLSMARFTYEYSRR	5.274222	1
DLCQVQGVPLQTMK	4.958126	10
DLCQVQGVPLQTMKQQFLINLVK	4.563741	1
DLSMARFTYEYSRRHTK	4.643211	1
DVLTIIEKPTGSK	5.168966	46
EIPEKYGLSDCCSR	4.979192	3
EIPEKYGLSDCCSRTGEERHDCFLAHK	5.007572	1
EKEIPEKYGLSDCCSR	4.983226	4
EKEIPEKYGLSDCCSRTGEER	4.824199	2
ELFINR	5.035353	1
ELFINRYIYEIAR	4.903483	11
ELRESSLLNQHICAVMGKFGPRTFR	5.687433	1
ELRESSLLNQHICAVMGKFGPRTFRAITVTKVSQK	4.472155	1
ESSLLNQHICAVMGK	5.554158	10
ESSLLNQHICAVMGKFGPRTFR	6.135839	1
FGPRTFR	4.474459	1
FIFHKDLCQVQGVPLQTMK	5.403825	1
FIFHKDLCQVQGVPLQTMKQQFLINLVK	5.105651	1
FLGDRDFNQLSSR	5.531424	156
FLGDRDFNQLSSRDKDLSMAR	4.230859	1
FPKANFTEIQK	5.307961	1
FTYEYSR	5.239719	57
FTYEYSRR	5.095873	7
GNVLECLQDGER	5.066557	33
GYQELLEK	5.173775	92
HDCFLAHKK	4.860238	8
HPVLYAPTILSVANQYNK	5.073591	8
HTKLAVPIILR	5.127088	19
IIPHCCKAENATECFETKVTSITK	5.833767	1
KAAPASIPPFPVLEPVTSCK	5.206178	8

KAAPASIPPFPVLEPVTSCKSYK	5.370925	1
KAAPASIPPFPVLEPVTSCKSYKENR	5.552489	2
KAPQLTSPELMALTR	5.292359	181
KAPQLTSPELMALTRK	5.547794	13
KMANAGAICCHLSEDKQLACGEGVADLIIGHLCIR	5.431735	2
LAVPIILR	5.215772	77
LAVPIILRVAK	5.423892	2
LGEYYLQNAFLVAYTK	5.056435	186
LGEYYLQNAFLVAYTKK	5.229229	6
LPTTLELGHCIIHAENDDKPEGLSPNVNR	4.572126	3
LVMDVAHIHEECCK	4.94861	4
LVMDVAHIHEECCKGNVLECLQDGER	5.369604	2
MKWVVSFFLLFLLNFSDSR	4.637449	1
MKWVVSFFLLFLLNFSDSRTMHK	4.657245	1
QKPQITEEQLETVVADFSGLLEK	5.428821	153
QLACGEGVADLIIGHLCIR	5.252301	2
QPAGCLENQVSAFLEEICR	4.637113	5
QPAGCLENQVSAFLEEICREK	5.563586	3
QQFLINLVK	5.542871	39
RHPVLYAPTILSVANQYNK	5.240158	36
RHPVLYAPTILSVANQYNKIIPHCCK	4.727681	1
RHTKLAVPIILR	4.038676	2
RPCFSSLVVDETYVPPPFSDDK	4.614878	2
RSCGLFQK	4.990376	14
SCGLFQK	4.814539	17
SYKENRELFINR	5.237495	13
TFRAITVTKVSQKFPK	4.77845	4
TGEERHDCFLAHK	4.473828	12
TGEERHDCFLAHKK	4.823098	33
TRAALGV	4.90512	38
VAKGYQELLEK	5.046219	33
VMSYICSQQDILSR	5.203705	20
VMSYICSQQDILSRQIAECCK	5.309928	1
VSQKFPK	4.880662	70
VTSITKELR	4.787943	4
VTSITKELRESSLLNQHICAVMGK	4.828695	2
VTSITKELRESSLLNQHICAVMGKFGPR	5.010108	1
YGLSDCCSR	4.891239	22
YIQESQALAK	5.82608	52
YIQESQALAKR	4.607971	3
YIYEIAR	5.544451	248
YIYEIARR	5.46179	59
YLYEIAR	5.64936	129
YLYEIARR	5.451315	51

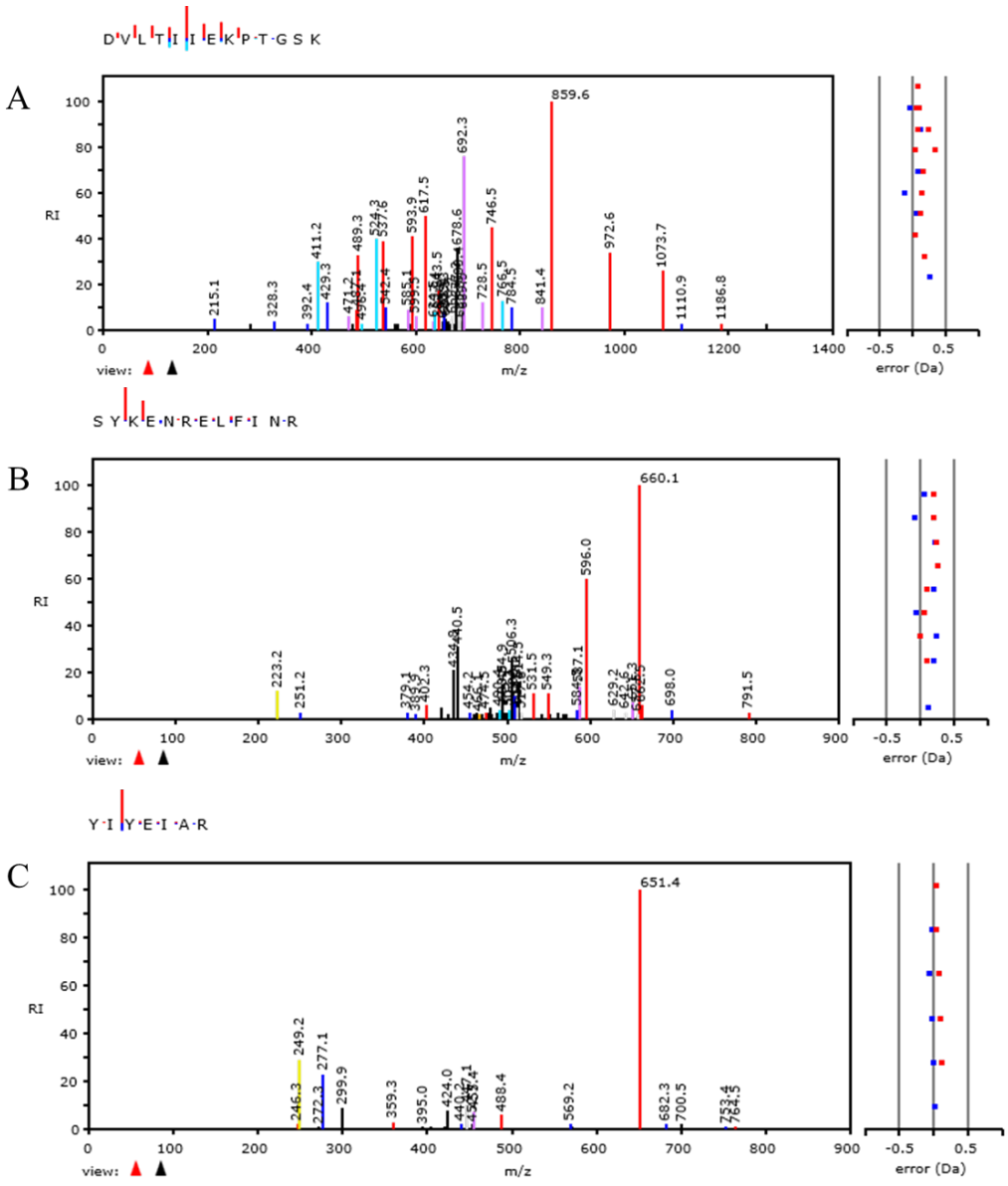


Figure 12. Example MS/MS spectra from alpha feto protein that was specific to fetal serum from the X!TANDEM algorithm. Panels: A, the peptide DVLTIIEKPTGSK;, B, the peptide SYKENRELFINR; C, the peptide YIYEIAR.

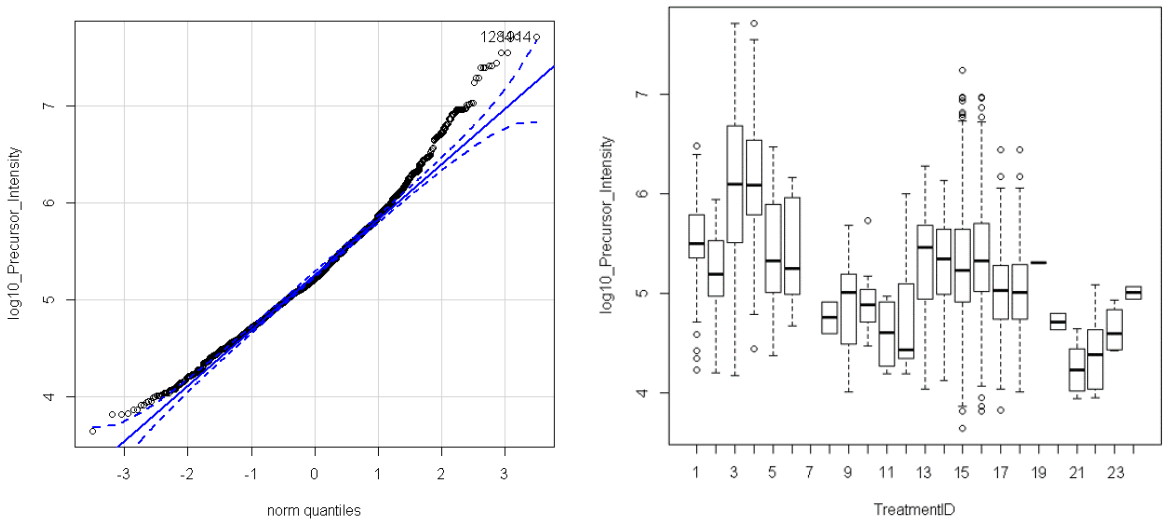


Figure 13. AFP in fetal and adult bovine serum between treatments and replicates. Treatments and replicates from salt partition chromatography and organic extraction with and without the consideration of phosphorylation of AFP were shown here. Panels: A, the quantile Q-Q plot showing the normality of the log10 transformed intensity values; B, the box plot showing the average log10 intensity and variation in log10 for the adult versus fetal serum replicates.

ANOVA analysis showed a probability of ~2e-16 that the AFP between fetal versus adult serum samples were the statistically significant:

Response: log10_Precursor_Intensity

	Sum Sq	Df	F value	Pr(>F)
Peptide	**138.92**	**82**	**7.2359**	**< 2.2e-16 *****
TreatmentID	**133.02**	**21**	**27.0545**	**< 2.2e-16 *****
Peptide:TreatmentID	**60.05**	**175**	**1.4655**	**0.0001405 *****
Residuals	**432.68**	**1848**		

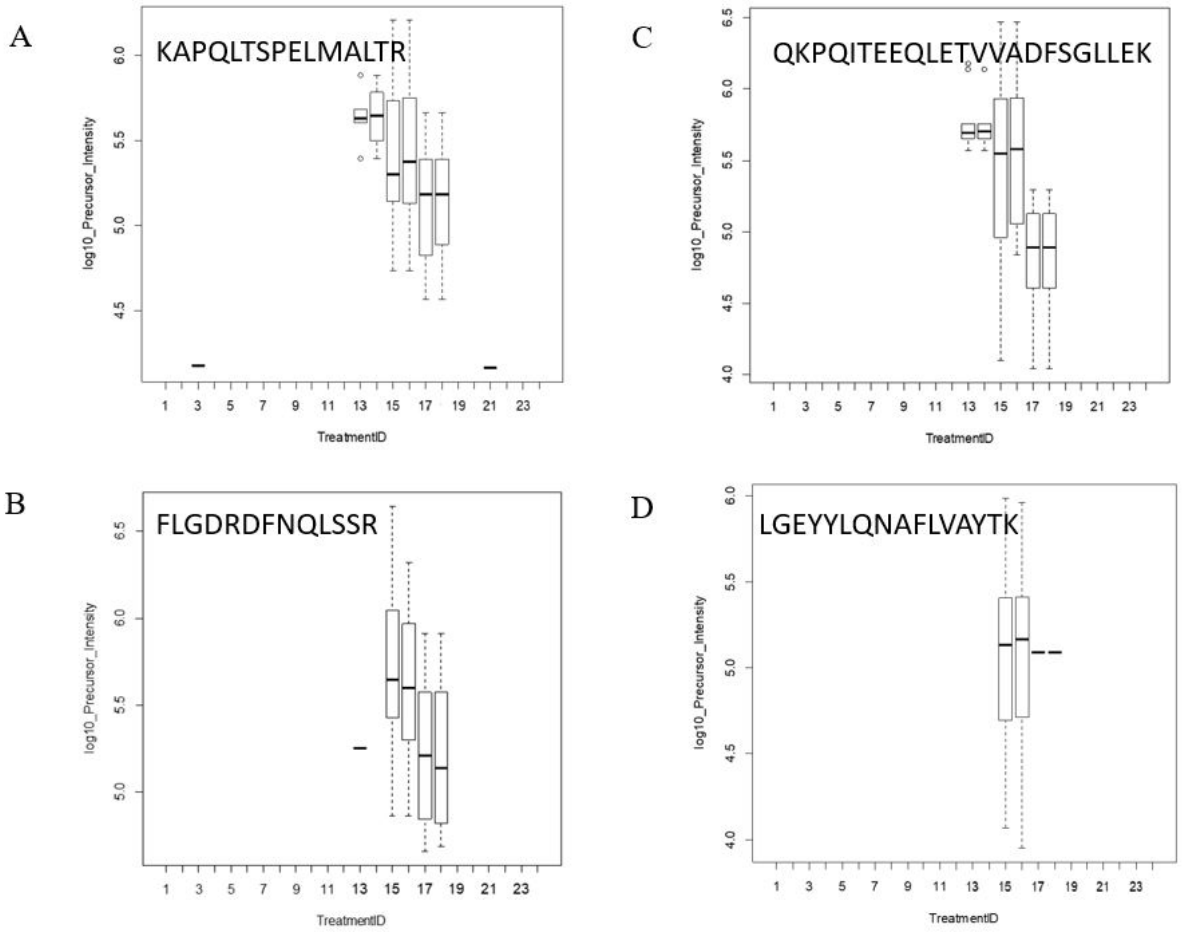


Figure 14. Some peptides were found to be specific to AFP and highly observed in fetal serum. The four peptides shown above (Panels A, B, C, and D) were found to be unique to AFP but not Albumin, which were only found in fetal serum. These peptides from AFP were only observed from partition chromatography (See Table 8). Panels: A, the peptide KAPQLSPELMALTR; B, the peptide FLGDRDFNQLSSR; C, QKPQITEEQLETVVADFSGLLEK; D, the peptide LGEYYLQNAFLVAYTK.

ID# from 1-12: ABS treatments; ID# from 13-24: FBS treatments (See Table 8)

4.16. Insulin in FBS and Addition of Insulin in ABS cell culture

Insulin was another example of a protein found with increased observation frequency in FBS. Exemplary MS/MS spectra were observed from Insulin (Fig. 15). The variations of INS peptides in fetal versus adult serum indicated that different peptides were observed (Fig. 16). Addition of insulin to Raw cell macrophages in ABS cell culture resulted in rounded cells similar to FBS controls. Different amounts of insulin from 0.2μg, 2μg, and 20μg added to ABS cell culture were compared with 0μg of insulin in ABS culture and FBS culture as controls. The addition of 2μg of insulin seemed to have the optimal results to the cultured cells in ABS where 0.2μg of insulin did not seem to have any effect and 20μg of insulin seemed to be too saturated (overdose) so the cells started to die. (Fig. 17&18)

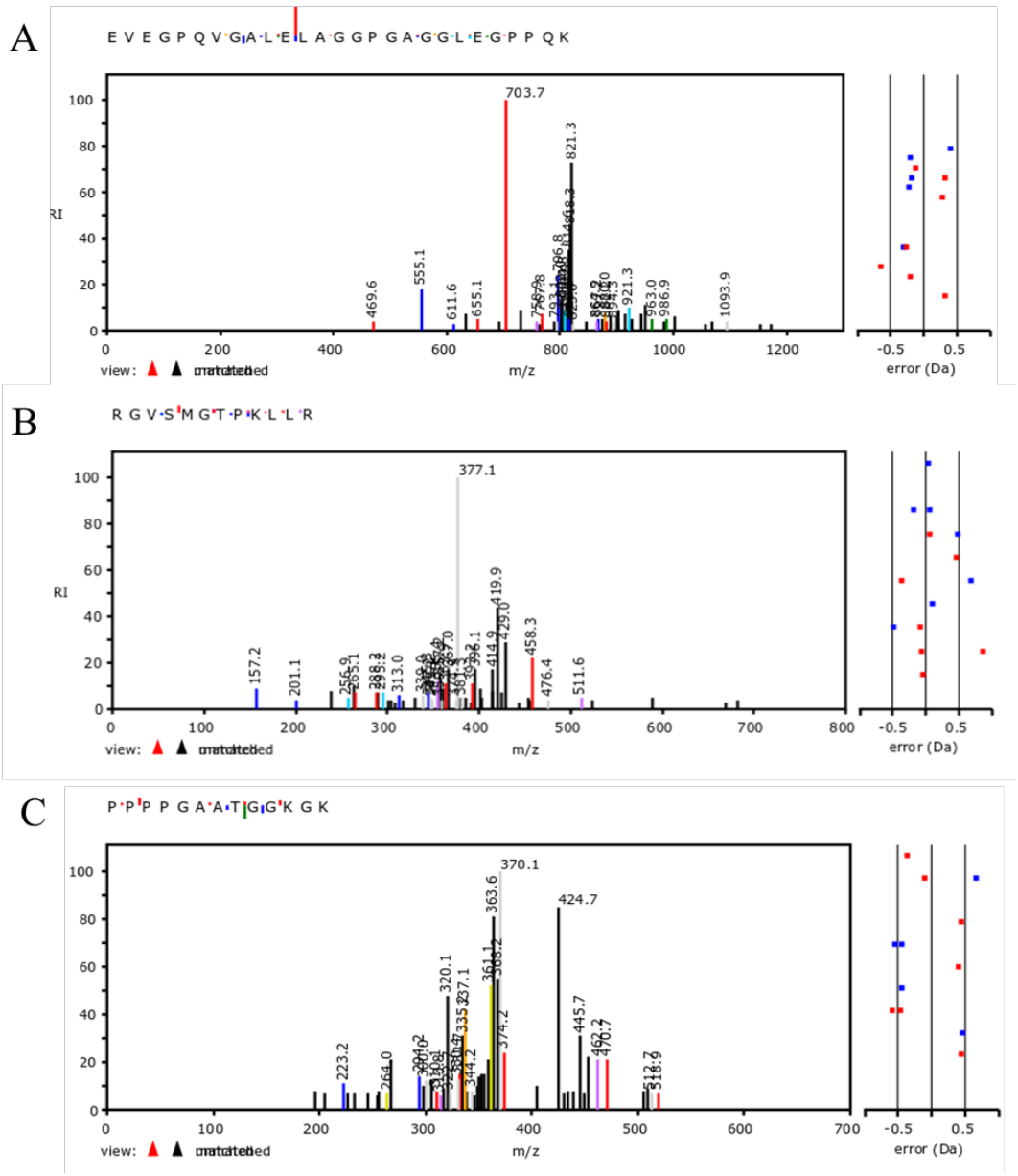


Figure 15. Example MS/MS spectra from high abundance insulin that was specific to fetal serum from X!TANDEM algorithm. Panels: A, the peptide EVEGPQVGALEKLAGGPPGAGGLEGPPQK; B, the peptide RGVSMGTPKLLR; C, the peptide PPPPGAATGGKGK.

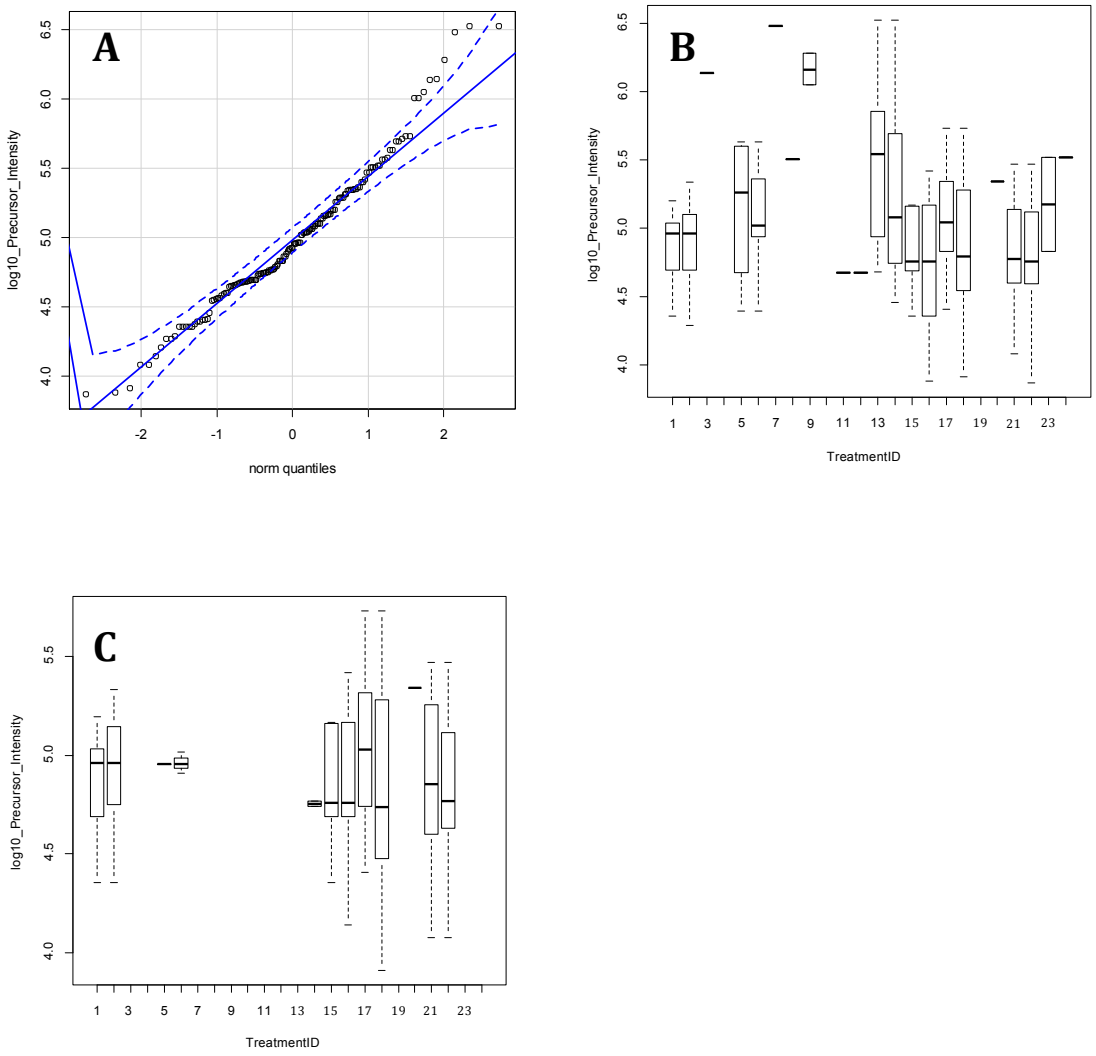


Figure 16. The variation in INS peptides in the fetal versus adult serum samples. Panels: A, the QQ plot showing the normality of the log10 transformed intensity values; B, all of the peptides from insulin (INS); C, the peptide MGPGRAR from insulin. The box plot showing the average log10 intensity and variation in log10 and 99% confidence interval is shown for the adult versus fetal serum replicates ANOVA analysis show a significant effect of treatment and peptides both with a p value of ~2e-16.

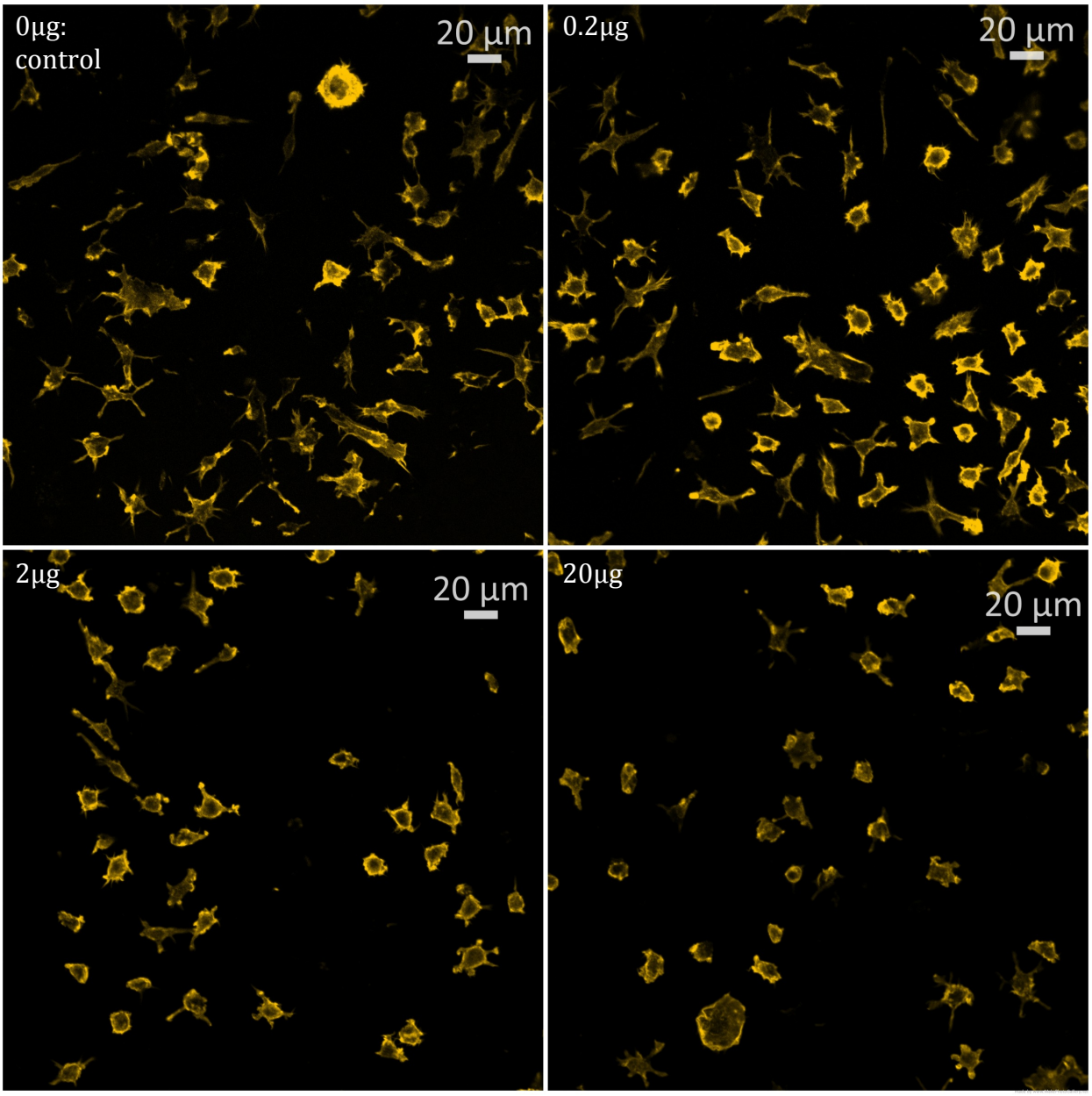


Figure 17. Raw cells cultured in 5% ABS and stained with rhodamine phalloidin. The images were taken under a 20X magnification confocal microscope. Addition of 0µg, 0.2µg, 2µg, and 20µg of insulin to Raw cell 264.7 in ABS cell culture showed different cell morphologies. Different amounts of INS were compared to the control (no addition of INS), where 2µg of INS gave the similar results to FBS, 0.2µg of INS did not seem to have any results to the cell and 20µg of INS seemed to be saturated.

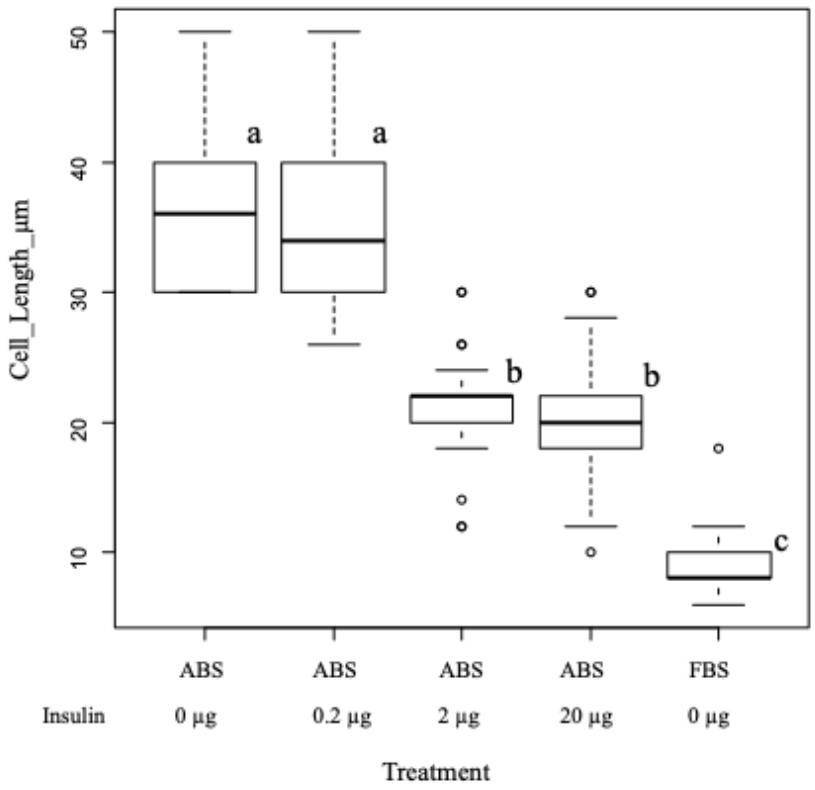


Figure 18. ANOAVA analysis shows the means of cell length between the levels of insulin addition to the ABS cell culture. Ten cells were measured from 3 spots of each treatment. The longest axis of the cell was measured.

```
> Summary(AnovaModel.1)

             Df   Sum Sq  Mean Sq   F value Pr(>F)
Treatment    4    16064   4016       157.3 <2e-16 ***
Residuals    145  3703    26
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> with(Dataset, numSummary(Cell_Length_um, groups=Treatment,
+   statistics=c("mean", "sd")))
                        mean            sd       data:n
ABS                    36.400000   5.834499     30
ABS_0.2_ug_Insulin     35.533333   7.040441     30
ABS_2_ug_Insulin       21.133333   4.125057     30
ABS_20_ug_Insulin      20.200000   4.708027     30
FBS                    8.933333    2.211776     30
```

5.0 DISCUSSION

The aim and objective of this study was to identify proteins and peptides specific to fetal versus adult serum by random and independent sampling of peptides from digested proteins (proteomics) or endogenous peptides (peptidomics) using a LC-ESI-MS/MS with a linear quadrupole ion trap to reveal the factors that may account for proliferative phenotype observed. There was near perfect (99%) agreement on the proteins of bovine serum between the independent proteomic and peptidomic methods clearly and unambiguously demonstrated that the type I error of protein identification was low. However, there was large variation in the observation frequency and/or intensity of the peptides and proteins identified between fetal versus adult serum. Here for the first time all of the factors that differ between fetal versus adult serum are revealed.

5.1. Agreement between independent methods

Multiple independent lines of evidence including the effect of fetal versus adult serum on cells, the results of LC-ESI-MS/MS with both the X!TANDEM and SEQUEST algorithm (computed independently with and without phosphorylation), Chi square analysis of observation frequency [82], and previously observed structural and functional relationships from STRING analysis [78], all agreed that there were significant statistical and biological differences in the proteins of fetal versus adult serum. However, the source of the many cellular proteins observed in sera is not clear and at least some of them can be accounted for from red blood cells or white blood cells that may occur during sampling.

5.2.Release of cellular proteins from blood cells

Many unexpected cellular proteins have been observed in serum of plasma by liquid chromatography and tandem mass spectrometry by multiple groups in agreement [76], [83]–[85]. At present it is not clear if the many receptors observed in serum were cleaved from the cell surfaces by proteases. Neither it is clear if the cellular proteins, other than those from blood cell lysis, were secreted in solution, released in membrane bound vesicles like exosomes, were bound to circulating nucleic acids [4], or otherwise released from cells. It remains possible that some of the cellular factors observed were released from white blood cells like platelets or immune cells and it will be important to rule out cells as a source of error in the serum or plasma proteome. Some, but not all, the proteins that are known to be derived from platelets, monocytes, macrophages, neutrophils, or granulocytes were observed in the serum. The fraction of all known white or red blood cells proteins observed, and the even smaller subset that showed significantly greater observation frequency in fetal serum, were not consistent with contamination from white, or red blood cells as the largest source of the differences between fetal versus adult serum. However, the observation that only a small fraction of all known white blood cells proteins were observed and did not show significantly greater observation frequency in fetal serum, are not consistent with contamination from white blood as the source of the fetal serum specific proteins. White blood cells are known to secrete phospholipase A2 [86] that was also observed in serum and plasma by LC-ESI-MS/MS in human and bovine samples. The observed hemoglobin confirms that some red blood cell lysis occurred but it is difficult to account for the transcription factor, histones and DNA binding proteins observed from enucleated red blood cells. The platelet proteome is known to possess at least ~1,100 proteins [87] and platelet derived factors have effects similar to fetal serum on cell cultures [29]. Platelet lysates are known to permit the

expansion of cultured cell lines in a manner similar to fetal serum [13]. Some platelet derived factors such as platelet associated factor 4 (chemokine (C-X-C motif) ligand 4) observed in fetal serum have been associated with the biology of cancer [88].

5.3.Quaternary amine chromatography versus organic extraction

There was previously observed good agreement between the proteins that resulted from partition chromatography and endogenous peptides extracted from serum in single step extractions analyzed by micro chromatography [3]. Subsequently the use of step wise salt elution of partition chromatography resins followed by micro electrospray identified 4,396 distinct proteins by X!TANDEM [39]. Step wise organic fractionation with micro electrospray previously identified 3,463 protein from serum [38] and the combination with nano-spray identified some 14,000 proteins by X!TANDEM from the aggregation of hundreds of LC-ESI-MS/MS experiments [4]. Here the use of partition chromatography and organic extraction together by nano LC-ESI-MS/MS results in more than 12,000 proteins identified with multiple peptides by X!TANDEM that shows good agreement with the number of proteins reported from the HuPO (Human Proteome Organization) plasma proteome consortium [65]. Taken together the results indicate that more than 12,000 proteins can be reliably monitored by step-wise quaternary amine pre-fractionation prior to digestion.

5.4.Fetal proteins

The specific identification of the alpha feto protein [89] and fetal hemoglobin in the fetal serum seems to indicate that the LC-ESI-MS/MS system was capable of faithfully identifying the proteins that differ between fetal and adult serum [90]. Alpha feto protein (AFP) stands out as an

example of a protein that was markedly different between adult versus fetal serum that is strongly suspected of directly mediating effects on cell growth: AFP is known to stimulate proliferation and inhibit apoptosis [91], [92], binds to a DNA-binding zinc finger transcription factor of the nuclear receptors and may be an important regulator of embryonic development [93].

The example of AFP also showed that the proteins were stable and did not degrade easily during the extraction over partition chromatography and organic extraction. This was indicated by looking at the accession protein counts that there was only a few of the AFP observed in the endogenous peptides extraction. (Table 8).

5.5.Cellular and multi cellular growth pathways

Unbiased chi square analysis of protein frequencies between fetal versus adult serum showed proto oncogenes from the cellular growth pathway such as AKT and HRAS were observed in the fetal serum consistent with a role for the pathway in the rapid growth of developing cells. The AKT/HRAS pathway is connected to many of the receptor and receptor enzyme proteins discovered specific to fetal serum. Increased frequency of adenylate cyclase, calmodulin-binding transcription activator, calcium/calmodulin-dependent protein kinase type 1, Calcium/calmodulin-dependent 3',5'-cyclic nucleotide phosphodiesterase links the PI3K, AKT, HRAS proliferation pathways to the cyclic AMP (cAMP) responsive element binding CREB/ATF bZIP transcription factor.

5.6.Potential Transforming factors from sera

The specific increase in certain peptide hormones, growth factors or platelet derived proteins might be the source of the transformative effects of the fetal serum or platelet lysates. Peptide hormones and other potent ligands such as growth factor or interleukins observed present in relatively low levels present the greatest challenge to detection by LC-ESI-MS/MS. Formerly, insulin like growth factor II was barely detectable [41], [42], but after step wise fractionation together with nano scale LC-ESI-MS/MS insulin, insulin like growth factor II, insulin binding protein, and insulinoma-associated protein were all observed tens or hundreds of times and by both X!TANDEM and SEQUEST. A large variety of soluble protein or peptide factors that may contribute to the capacity of fetal serum to drive cell growth indefinitely were observed such as platelet-derived growth factor. The interleukins were undetectable in previous studies using single step protein or peptides extractions with micro electrospray ionization. However, interleukin-16 isoform X1 (IL16), interleukin-17D IL were strongly detected here with multiple step-wise fractionation and nano LC-ESI-MS/MS. Since fetal serum or platelet derived factors may be used to expand cultured cells, the specific increase in some growth factor proteins, including in platelet derived growth factors, was consistent with the transformative effects of the fetal serum. The apparent increase in growth factors, cytokines, chemokines, interleukin, or tumor necrosis factors may be consistent with the existence of factors in the fetal serum that drive the immortal replication of RAW 264.7 and other cells lines [12] and may permit the perfection of synthetic cell media that will prevent the contamination of secreted cellular proteins from culture media components in cell culture models of cancer [28], [29].

5.7.Receptors and receptor signaling associated cellular and nuclear factors

G protein receptors and receptor kinases as well as receptor associated enzymes, monomeric G proteins were observed in serum and might still play a role as co-receptors via protein-protein interactions outside of the cell [94]. Peptides mapping to overlapping tryptic peptides and sequences observed to multiple unique GPCR 98 variant that did not contain the transmembrane or cytosolic region may indicate the extracellular region of the receptor was shed from the cell surface by proteolytic activity and so the soluble receptor might not be functional in the conventional sense. While receptor fragments might be shed from the cell surface by proteolytic activity this pathway cannot explain the presence of intracellular receptor associated signaling factors. Cellular signaling proteins have been observed by many groups in human serum and plasma, and here are also observed in bovine serum and cannot be released by proteolytic action alone. It has suggested that secretion, or exosomes, may be responsible for the presence of apparently high concentrations of unexpected cellular proteins and nucleic acids in blood serum or plasma [85], [95].

5.8.DNA and RNA binding proteins

H2A Histone Family Member X (H2AFX) is associated with the Nijmegen Breakage Syndrome and Ataxia-Telangiectasia and the proliferation of cells in cancer [96]. Proteins such as bromodomains, PHD and Zinc finger domains [42] are known to bind DNA and/or RNA molecules that may influence the maintenance of the chromosomes and gene expression [97], [98].

The nuclear receptor for alpha-feto protein is a zinc finger DNA binding factor [99]. Cell free DNA exists in serum or plasma at ng/ml concentrations [95], [100] and may remain bound to proteins such as histones, zinc fingers and transcription factors. The enriched presence of the CREB transcription factor links the MAPK and CAMK with AKT/HRAS growth pathway. It seems unlikely that such nuclear factors should be released by reticulocytes or enucleated red blood cells directly into a blood and so the source of these nucleic acid binding proteins in plasma is unclear.

6.0 CONCLUSION

The overwhelming agreement between the entirely independent methods of protein fractionating by partition chromatography and digestion versus endogenous peptide extraction using organic solvents indicates that the level of type I error for protein identification by LC-ESI-MS/MS with a linear quadrupole ion trap is low. It appears that some of the proteins or peptides in blood might be cleaved from the surface of, or released from, white blood or red blood cells potentially as an artifact of collection. Fetal serum also contained high levels of alfa-feto protein, and fetal albumin as well as insulin, platelet derived and transforming growth factors that were consistent with the capacity to support indefinite cell proliferation. Unexpectedly, fetal serum also contained proteins associated with the AKT/HRAS growth pathways including cellular receptors, receptor associated enzymes and other signal transduction factors as well as transcription factors and DNA/RNA binding proteins were observed in fetal serum. The source of the apparently reproducible set of cellular proteins was not clear [101], but STRING analysis seems to indicate that the proteins may be participating in extracellular protein-protein or protein-nucleic acid interactions in circulation [85]. The combination of classical sample partition, random and independent sampling by LC-ESI-MS/MS with a sensitive, economical and robust linear ion trap, peptide identification from MS/MS spectra using SEQUEST or X!TANDEM with data filtering and storage in 64 bit SQL Server for statistical analysis with open source R was apparently sufficient to identify and compare some 12,000 different protein gene symbols from blood serum in a robust and inexpensive approach. It appeared that some as yet unknown fetal cells have released many proteins related to the AKT/HRAS growth pathway in the serum where their possible residual function(s), if any, were unclear.

APPENDICES

Table A.1. Experimental treatment table. The data of all QA column extractions, organic extractions from ABS and FBS, and the data of all blank runs and column conditioning were collected by the Tandem Mass Spectrometry, and computed in the SQL server.

Treatment	FileCount	TreatmentName
0	2	ProteinLibrary
1	64	BB_ABS_NaClExtraction_BOVINE_rep1
2	64	BB_ABS_NaClExtraction_BOVINE_rep1_PSTY
3	64	BB_ABS_NaClExtraction_BOVINE_rep2
4	64	BB_ABS_NaClExtraction_BOVINE_rep2_PSTY
5	64	BB_ABS_NaClExtraction_BOVINE_rep3
6	64	BB_ABS_NaClExtraction_BOVINE_rep3_PSTY
7	40	BB_ABS_OrganicExtraction_BOVINE_rep1
8	40	BB_ABS_OrganicExtraction_BOVINE_rep1_PSTY
9	40	BB_ABS_OrganicExtraction_BOVINE_REP2
10	40	BB_ABS_OrganicExtraction_BOVINE_REP2_PSTY
11	40	BB_ABS_OrganicExtraction_BOVINE_rep3
12	40	BB_ABS_OrganicExtraction_BOVINE_rep3_PSTY
13	20	BB_CTRL_BLANKS_BOVINE_REP1
14	20	BB_CTRL_Blanks_BOVINE_REP1_PSTY
15	20	BB_CTRL_BLANKS_HUMAN_REP1
16	15	BB_CTRL_BLANKS_HUMAN_REP1_PSTY
17	56	BB_CTRL_CC_BOVINE_REP1
18	56	BB_CTRL_CC_BOVINE_REP1_PSTY
19	56	BB_CTRL_CC_HUMAN_REP1
20	42	BB_CTRL_CC_HUMAN_REP1_PSTY
21	3500	BB_CTRL_RNG_BOVINE_REP1
22	3500	BB_CTRL_RNG_BOVINE_REP1_PSTY
23	72	BB_CTRL_TRP3_HUMAN_REP1
24	54	BB_CTRL_TRP3_HUMAN_REP1_PSTY
25	64	BB_FBS_NaClExtraction_BOVINE_rep1
26	64	BB_FBS_NaClExtraction_BOVINE_rep1_PSTY
27	64	BB_FBS_NaClExtraction_BOVINE_rep2
28	64	BB_FBS_NaClExtraction_BOVINE_rep2_PSTY
29	64	BB_FBS_NaClExtraction_BOVINE_rep3
30	64	BB_FBS_NaClExtraction_BOVINE_rep3_PSTY
31	40	BB_FBS_OrganicExtraction_BOVINE_rep1
32	40	BB_FBS_OrganicExtraction_BOVINE_rep1_PSTY
33	40	BB_FBS_OrganicExtraction_BOVINE_REP2
34	40	BB_FBS_OrganicExtraction_BOVINE_REP2_PSTY
35	40	BB_FBS_OrganicExtraction_BOVINE_rep3
36	40	BB_FBS_OrganicExtraction_BOVINE_rep3_PSTY

Table A.2. Precursor intensity of INS in log_{10} between treatments of ABS and FBS.

	log10_Precursor_Intensity	Sd(mean)	Mean
1	4.853978	0.09069283	9
2	4.884109	0.07600723	17
3	6.137435	NA	1
5	5.135398	0.29102713	4
6	5.089139	0.15941548	7
7	6.480922	NA	1
8	5.505216	NA	1
9	6.163147	0.11384392	2
11	4.670567	NA	1
12	4.670567	NA	1
25	5.48388	0.22131844	8
26	5.264823	0.18671255	13
27	4.812131	0.12651844	6
28	4.75955	0.16224421	10
29	5.070476	0.12763632	10
30	4.839104	0.10069634	21
32	5.339887	NA	1
33	4.850236	0.09479241	17
34	4.804412	0.08244537	24
35	5.173349	0.34666499	2
36	5.520014	NA	1

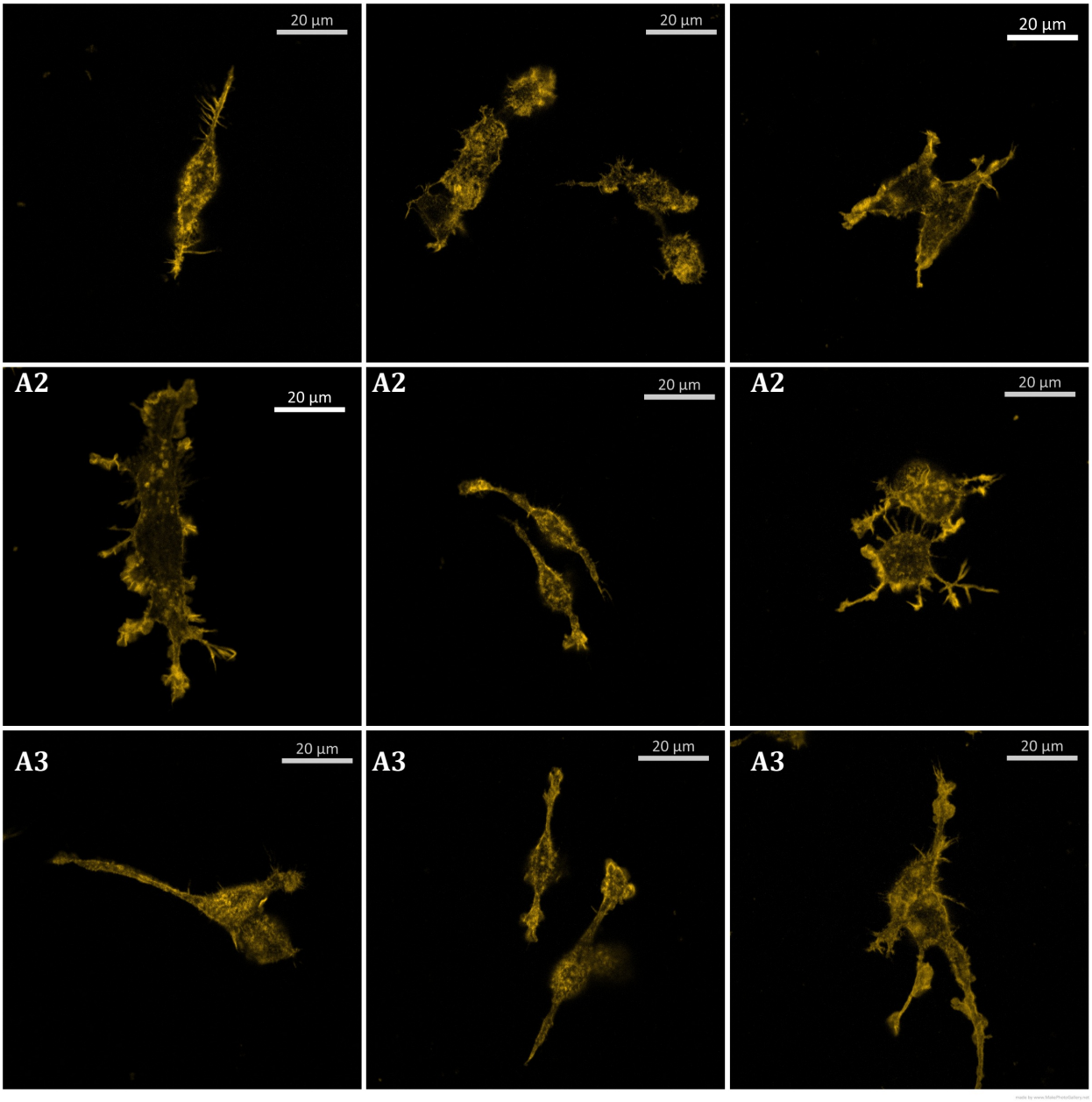


Figure B.1. The effect on the morphology of RAW cell 264.7 murine macrophages cultured in DMEM plus 5% ABS and stained with rhodamine phallodin. The cells were cultured in three independent batches of ABS and observed to have dendritic shapes under a 63X magnification confocal microscope. Panels: A1, 3 different individual experiments from ABS batch 1; A2, 3 different individual experiments from ABS batch 2; A3, 3 different individual experiments from ABS batch 3.

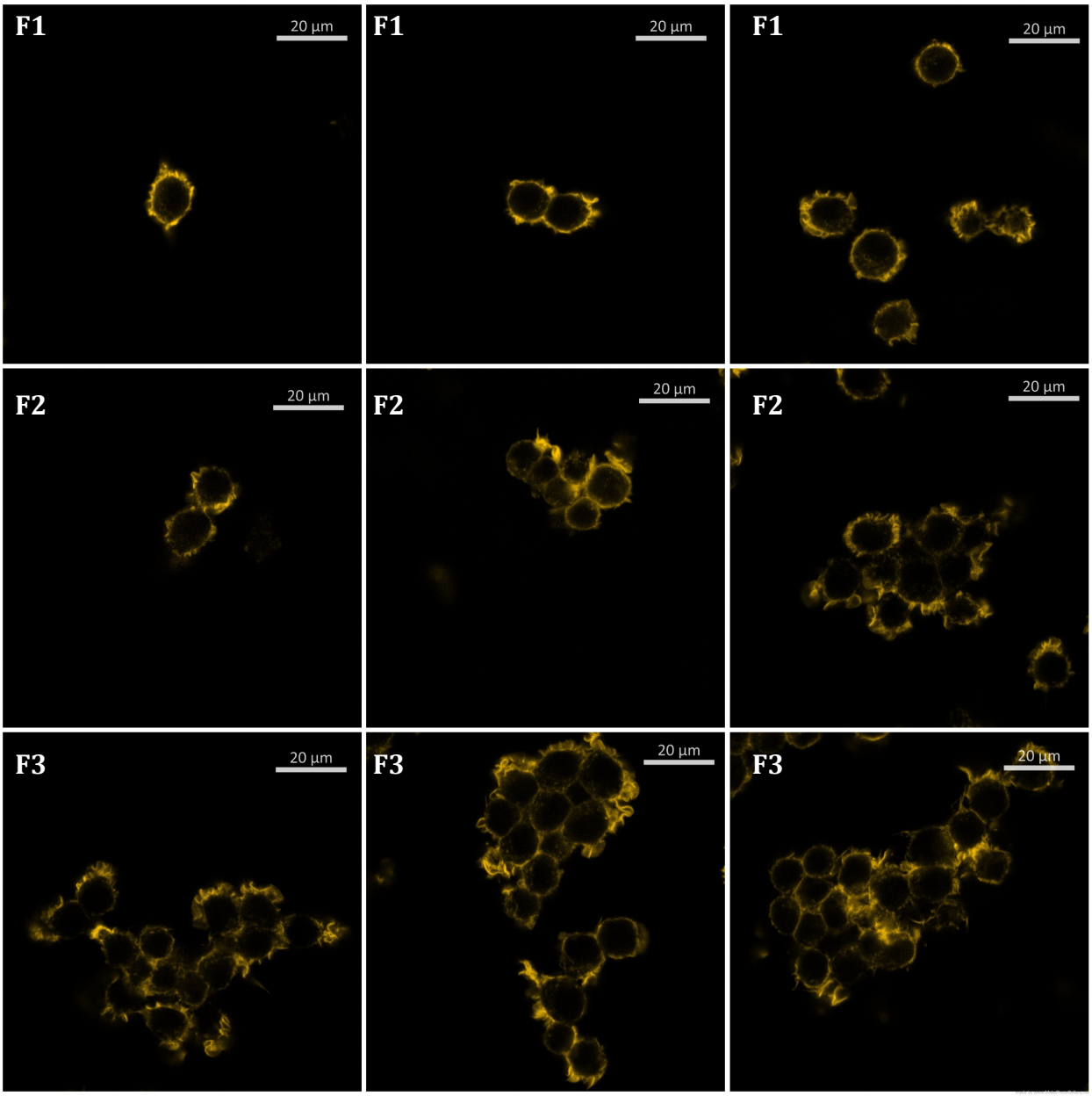


Figure B.2. The effect on the morphology of RAW cell 264.7 murine macrophages cultured in DMEM plus 5% FBS and stained with rhodamine phallodin. The cells were cultured in three independent batches of FBS and observed to have rounded shapes under a 63X magnification confocal microscope. Panels: F1, 3 different individual experiments from FBS batch 1; F2, 3 different individual experiments from FBS batch 2; F3, 3 different individual experiments from FBS batch 3.

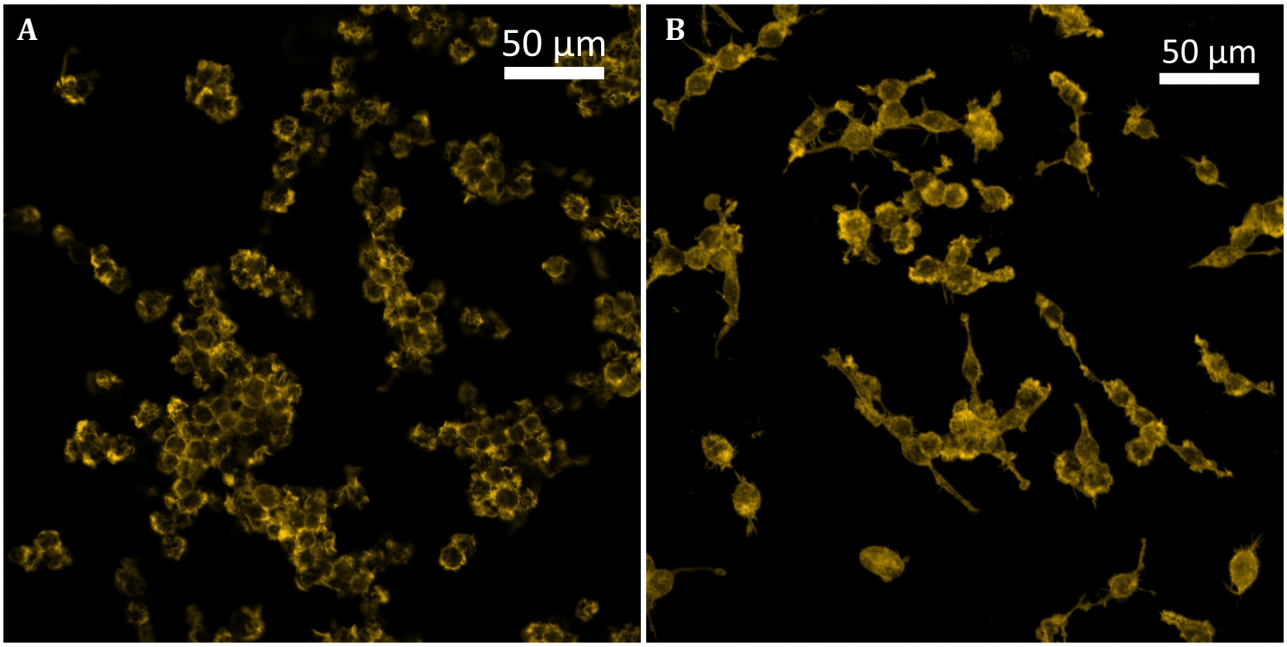


Figure B.3. The RAW cells cultured in the ABS/FBS were fixed with para-formaldehyde and stained with rhodamine phalloidin. The images were taken under a 20X magnification confocal microscope. Panels: A, cells grown in fetal serum that divided rapidly and piled onto each other; B, cells grown in adult serum that divided slowly and often died.

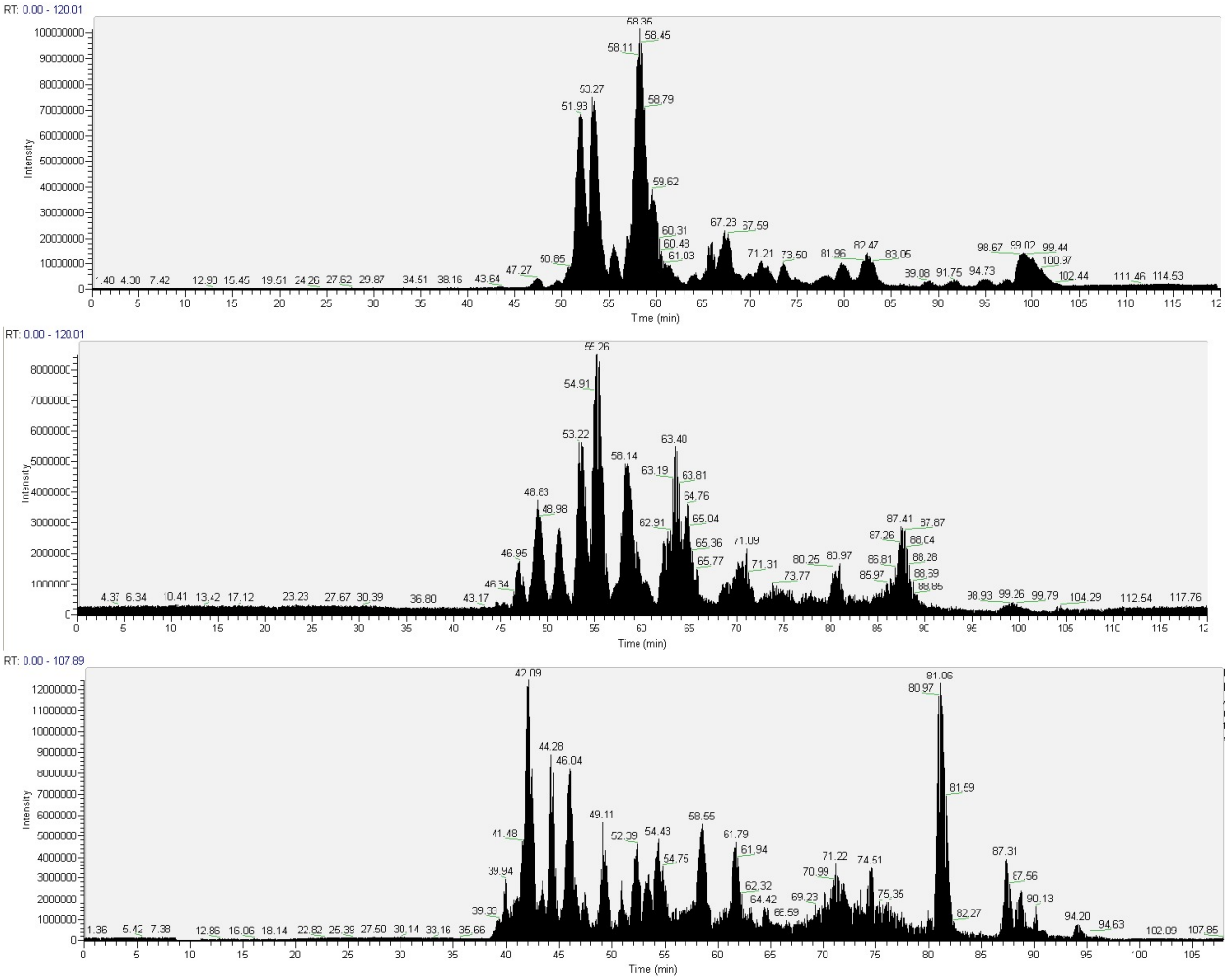


Figure B.4. Base-peak chromatograms of ABS sample collected over QA resin partition chromatography with 225mM NaCl in tricine eluting buffer between 3 individual replicates.

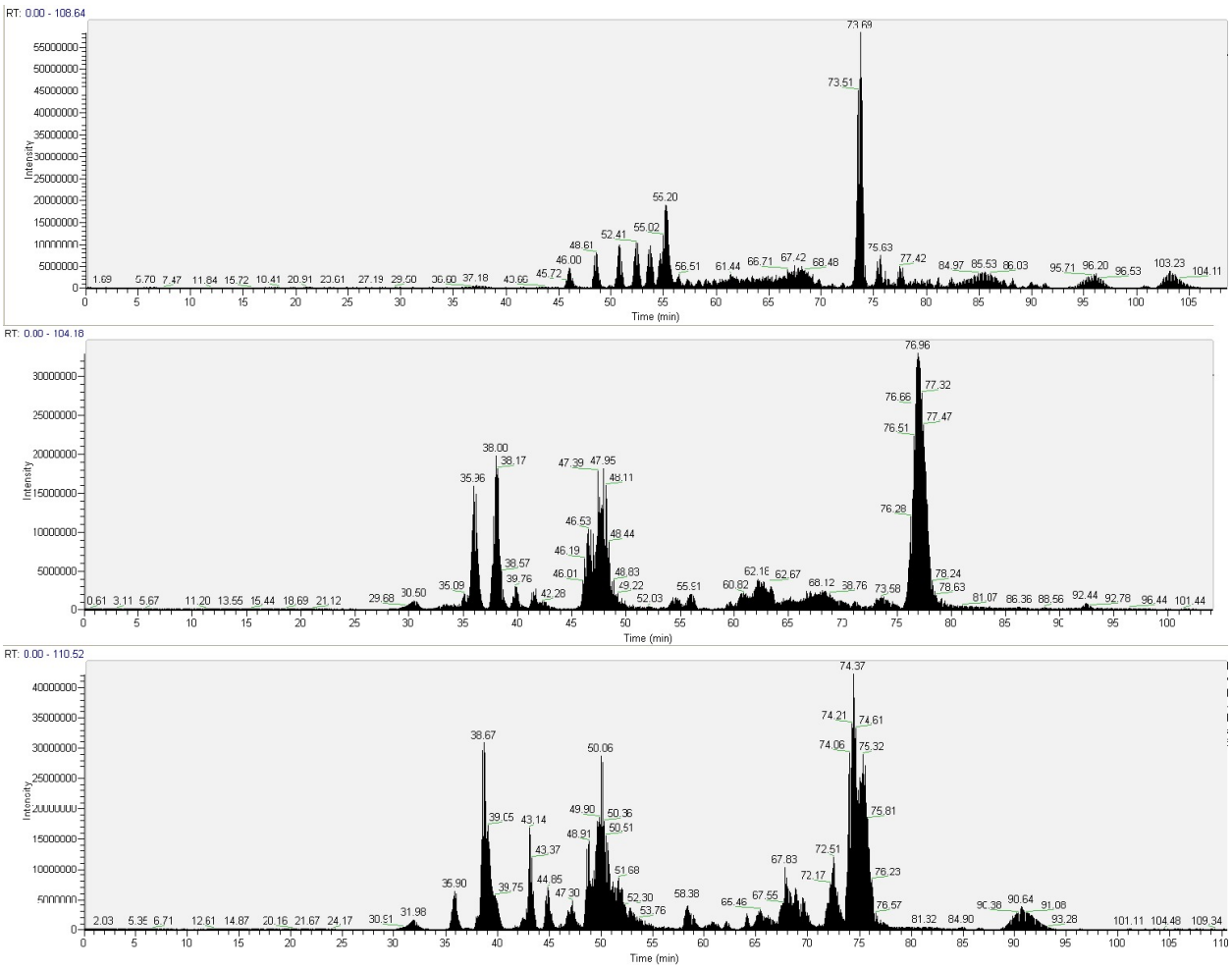


Figure B.5. Base-peak chromatograms of ABS sample collected over 40% acetonitrile organic extraction between 3 individual replicates.

REFERENCES

[1] J. Adams, "The Proteome: Discovering the Structure and Function of Proteins," *Nat. Educ.*, vol. 1, no. 3, p. 6, 2008.

[2] H. B. Vickery and C. L. A. Schmidt, "The history of the discovery of the amino acids," *Chem. Rev.*, 1931.

[3] J. Marshall *et al.*, "Human serum proteins preseparated by electrophoresis or chromatography followed by tandem mass spectrometry," *J. Proteome Res.*, 2004.

[4] J. Dufresne *et al.*, "The plasma peptidome.," *Clin. Proteomics*, vol. 15, p. 39, 2018.

[5] M. Johnson, "Fetal Bovine Serum," *Mater. Methods*, 2012.

[6] J. Fenn, M. Mann, C. Meng, S. Wong, and C. Whitehouse, "Electrospray ionization for mass spectrometry of large biomolecules," *Science (80-.).*, vol. 246, no. 4926, pp. 64–71, Oct. 1989.

[7] R. Craig and R. C. Beavis, "TANDEM: matching proteins with tandem mass spectra," *Bioinformatics*, vol. 20, no. 9, pp. 1466–1467, Jun. 2004.

[8] J. C. Schwartz, M. W. Senko, and J. E. . Syka, "A two-dimensional quadrupole ion trap mass spectrometer," *J. Am. Soc. Mass Spectrom.*, vol. 13, no. 6, pp. 659–669, Jun. 2002.

[9] G. Gstraunthaler, "Alternatives to the use of fetal bovine serum: serum-free cell culture.," *ALTEX*, 2003.

[10] H. Lodish, B. Arnold, S. L. Zipursky, P. Matsudaira, D. Baltimore, and J. E. Darnell, "Growth of Animal Cells in Culture," in *Molecular Cell Biology*, 2000.

[11] J. E. Yu *et al.*, "Potential use of adult bovine serum obtained during the slaughtering process as a biological reagent," *Animal Cells Syst. (Seoul).*, 2013.

[12] J. I. Jones, A. J. D'Ercole, C. Camacho-Hubner, and D. R. Clemmons, "Phosphorylation of insulin-like growth factor (IGF)-binding protein 1 in cell culture and in vivo: effects on affinity for IGF-I.," *Proc. Natl. Acad. Sci.*, vol. 88, no. 17, pp. 7481–7485, Sep. 1991.

[13] L. Johansson, J. Klinth, O. Holmqvist, and S. Ohlson, "Platelet lysate: A replacement for fetal bovine serum in animal cell culture?," *Cytotechnology*, 2003.

[14] S. G. Ball, C. A. Shuttleworth, and C. M. Kielty, "Vascular endothelial growth factor can signal through platelet-derived growth factor receptors.," *J. Cell Biol.*, vol. 177, no. 3, pp. 489–500, May 2007.

[15] T. J. Anderson, C. A. Lapp, M. A. Billman, and G. S. Schuster, "Effects of transforming growth factor-beta and platelet-derived growth factor on human gingival fibroblasts grown in serum-containing and serum-free medium.," *J. Clin. Periodontol.*, vol. 25, no. 1, pp. 48–55, Jan. 1998.

[16] J. I. Heger *et al.*, "Human serum alters cell culture behavior and improves spheroid formation in comparison to fetal bovine serum," *Exp. Cell Res.*, 2018.

[17] S. Marino, J. G. Logan, D. Mellis, and M. Capulli, "Generation and culture of osteoclasts.," *Bonekey Rep.*, vol. 3, p. 570, 2014.

[18] E. Bauman, P. L. Granja, and C. C. Barrias, "Fetal bovine serum-free culture of endothelial progenitor cells—progress and challenges," *Journal of Tissue Engineering and Regenerative Medicine*. 2018.

[19] A. Honegger and R. E. Humbel, "Insulin-like growth factors I and II in fetal and adult bovine serum. Purification, primary structures, and immunological cross-reactivities," *J.*

Biol. Chem., 1986.
[20] W. A. Ellis, E. F. Logan, and J. J. O'Brien, "Serum immunoglobulins in aborted and non-aborted bovine foetuses," *Clin. Exp. Immunol.*, 1978.
[21] C. Y. Fang, C. C. Wu, C. L. Fang, W. Y. Chen, and C. L. Chen, "Long-term growth comparison studies of FBS and FBS alternatives in six head and neck cell lines," *PLoS One*, 2017.
[22] A. Kikuchi, H. Taira, T. Tsuruta, M. Hayashi, and K. Kataoka, "Adsorbed serum protein mediated adhesion and growth behavior of bovine aortic endothelial cells on polyamine graft copolymer surfaces," *J. Biomater. Sci. Polym. Ed.*, 1996.
[23] K. O'Donoghue and N. M. Fisk, "Fetal stem cells," *Best Practice and Research: Clinical Obstetrics and Gynaecology*. 2004.
[24] R. Prywes, A. Dutta, J. A. Cromlish, and R. G. Roeder, "Phosphorylation of serum response factor, a factor that binds to the serum response element of the c-FOS enhancer.," *Proc. Natl. Acad. Sci.*, 2006.
[25] J. M. MacSween, K. Rajaraman, R. Rajaraman, and R. A. Fox, "Macrophage migration inhibition factor (MIF): reducing the variables.," *J. Immunol. Methods*, vol. 52, no. 1, pp. 127–36, 1982.
[26] T. T. PUCK, S. J. CIECIURA, and A. ROBINSON, "Genetics of somatic mammalian cells. III. Long-term cultivation of euploid cells from human and animal subjects.," *J. Exp. Med.*, 1958.
[27] D. Iyer *et al.*, "Serum response factor MADS box serine -162 phosphorylation switches proliferation and myogenic gene programs," *Proc. Natl. Acad. Sci.*, 2006.
[28] D. Kwon *et al.*, "The effect of fetal bovine serum (FBS) on efficacy of cellular reprogramming for induced pluripotent stem cell (iPSC) generation," *Cell Transplant.*, 2016.
[29] T. Burnouf, H. A. Goubran, T. M. Chen, K. L. Ou, M. El-Ekiaby, and M. Radosevic, "Blood-derived biomaterials and platelet growth factors in regenerative medicine," *Blood Rev.*, 2013.
[30] C. Li, F. Wang, and L. Liu, "[Ultra-high performance liquid chromatography-mass spectrometry for analysis of newborn and fetal bovine serum components].," *Nan Fang Yi Ke Da Xue Xue Bao*, vol. 34, no. 5, pp. 751–3, May 2014.
[31] J. M. Castellano *et al.*, "Human umbilical cord plasma proteins revitalize hippocampal function in aged mice," *Nature*, 2017.
[32] D. F. Hunt, J. R. Yates, J. Shabanowitz, S. Winston, and C. R. Hauer, "Protein sequencing by tandem mass spectrometry.," *Proc. Natl. Acad. Sci. U. S. A.*, 1986.
[33] A. L. McCormack *et al.*, "Direct Analysis and Identification of Proteins in Mixtures by LC/MS/MS and Database Searching at the Low-Femtomole Level," *Anal. Chem.*, 1997.
[34] M. P. Washburn, D. Wolters, and J. R. Yates, "Large-scale analysis of the yeast proteome by multidimensional protein identification technology," *Nat. Biotechnol.*, 2001.
[35] J. Link, " Direct analysis of large protein complexes using mass spectrometry.," *Nat. Biotechnol.*, 1999.
[36] S. . F. Tsunoda, M.; Nonaka, "Determination of methylated arginines by column-switching high-performance liquid chromatography –fluorescence detection," *Analyst*, vol. 130, no. 10, pp. 1410–1413, 2005.
[37] A. K. Florentinus, P. Bowden, G. Sardana, E. P. Diamandis, and J. G. Marshall, "Identification and quantification of peptides and proteins secreted from prostate epithelial

cells by unbiased liquid chromatography tandem mass spectrometry using goodness of fit and analysis of variance," *J. Proteomics*, vol. 75, no. 4, pp. 1303–1317, Feb. 2012.
[38] J. Dufresne, A. Florentinus-Mefailoski, P. Bowden, and J. G. Marshall, "A method for the extraction of the endogenous tryptic peptides (peptidome) from human EDTA plasma," *Anal. Biochem.*, vol. 549, pp. 188–196, May 2018.
[39] M. Tucholska *et al.*, "Human serum proteins fractionated by preparative partition chromatography prior to LC-ESI-MS/MS," *J. Proteome Res.*, 2009.
[40] G. Sardana, J. Marshall, and E. P. Diamandis, "Discovery of candidate tumor markers for prostate cancer via proteomic analysis of cell culture-conditioned medium.," *Clin. Chem.*, vol. 53, no. 3, pp. 429–37, Mar. 2007.
[41] M. Tucholska *et al.*, "Endogenous peptides from biophysical and biochemical fractionation of serum analyzed by matrix-assisted laser desorption/ionization and electrospray ionization hybrid quadrupole time-of-flight," *Anal. Biochem.*, vol. 370, no. 2, pp. 228–245, Nov. 2007.
[42] M. Tucholska, A. Florentinus, D. Williams, and J. G. Marshall, "The endogenous peptides of normal human serum extracted from the acetonitrile-insoluble precipitate using modified aqueous buffer with analysis by LC–ESI–Paul ion trap and Qq-TOF," *J. Proteomics*, vol. 73, no. 6, pp. 1254–1269, Apr. 2010.
[43] X. Han, A. Aslanian, J. R. Yates, and III, "Mass spectrometry for proteomics.," *Curr. Opin. Chem. Biol.*, vol. 12, no. 5, pp. 483–90, Oct. 2008.
[44] E. P. Diamandis, "Mass Spectrometry as a Diagnostic and a Cancer Biomarker Discovery Tool," *Mol. Cell. Proteomics*, 2004.
[45] M. Lu *et al.*, "Proteomics and mass spectrometry for cancer biomarker discovery.," *Biomark. Insights*, vol. 2, pp. 347–60, Oct. 2007.
[46] J. Dufresne *et al.*, "The plasma peptides of ovarian cancer," *Clin. Proteomics*, vol. 15, no. 1, p. 41, Dec. 2018.
[47] Y. V Karpievitch, A. D. Polpitiya, G. A. Anderson, R. D. Smith, and A. R. Dabney, "Liquid Chromatography Mass Spectrometry-Based Proteomics: Biological and Technological Aspects.," *Ann. Appl. Stat.*, vol. 4, no. 4, pp. 1797–1823, 2010.
[48] W. M. A. Niessen and D. Falck, "Introduction to Mass Spectrometry, a Tutorial," 2015.
[49] L. Sleno and D. A. Volmer, "Ion activation methods for tandem mass spectrometry," *Journal of Mass Spectrometry*. 2004.
[50] "Finnigan ™ LTQ ™ Getting Started 97055-97012 Revision A," 2003.
[51] T. E. Angel *et al.*, "Mass spectrometry-based proteomics: existing capabilities and future directions.," *Chem. Soc. Rev.*, vol. 41, no. 10, pp. 3912–28, May 2012.
[52] P. Lössl, M. van de Waterbeemd, and A. J. Heck, "The diverse and expanding role of mass spectrometry in structural and molecular biology.," *EMBO J.*, vol. 35, no. 24, pp. 2634–2657, 2016.
[53] G. Mitulović and K. Mechtler, "HPLC techniques for proteomics analysis - A short overview of latest developments," *Briefings in Functional Genomics and Proteomics*. 2006.
[54] C. T. Mant, Z. Jiang, B. E. Boyes, and R. S. Hodges, "An improved approach to hydrophilic interaction chromatography of peptides: Salt gradients in the presence of high isocratic acetonitrile concentrations," *J. Chromatogr. A*, 2013.
[55] O. Chertov *et al.*, "Organic solvent extraction of proteins and peptides from serum as an effective sample preparation for detection and identification of biomarkers by mass

spectrometry," *Proteomics*, 2004.
[56] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse, "Electrospray ionization-principles and practice," *Mass Spectrom. Rev.*, vol. 9, no. 1, pp. 37–70, Jan. 1990.
[57] A. I. Nesvizhskii, A. Keller, E. Kolker, and R. Aebersold, "A statistical model for identifying proteins by tandem mass spectrometry.," *Anal. Chem.*, 2003.
[58] J. C. Howard, A. Florentinus-Mefailoski, P. Bowden, W. Trimble, S. Grinstein, and J. G. Marshall, "OxLDL receptor chromatography from live human U937 cells identifies SYK(L) that regulates phagocytosis of oxLDL," *Anal. Biochem.*, 2016.
[59] A. Florentinus-Mefailoski and J. G. Marshall, "Pyridoxamine-5-phosphate enzyme-linked immune mass spectrometric assay substrate for linear absolute quantification of alkaline phosphatase to the yoctomole range applied to prostate specific antigen," *Anal. Chem.*, 2014.
[60] J. Dufresne, A. Florentinus-Mefailoski, J. Ajambo, A. Ferwa, P. Bowden, and J. Marshall, "Random and independent sampling of endogenous tryptic peptides from normal human EDTA plasma by liquid chromatography micro electrospray ionization and tandem mass spectrometry," *Clin. Proteomics*, vol. 14, no. 1, p. 41, Dec. 2017.
[61] J. Dufresne, A. Florentinus-Mefailoski, P.-H. Zhu, P. Bowden, and J. G. Marshall, "Re-evaluation of the rabbit myosin protein standard used to create the empirical statistical model for decoy library searching," *Anal. Biochem.*, vol. 560, pp. 39–49, Nov. 2018.
[62] P. Bowden, T. Thavarajah, P. Zhu, M. McDonell, H. Thiele, and J. G. Marshall, "Quantitative Statistical Analysis of Standard and Human Blood Proteins from Liquid Chromatography, Electrospray Ionization, and Tandem Mass Spectrometry," *J. Proteome Res.*, vol. 11, no. 4, pp. 2032–2047, Apr. 2012.
[63] P. Zhu, P. Bowden, M. Tucholska, and J. G. Marshall, "Chi-square comparison of tryptic peptide-to-protein distributions of tandem mass spectrometry from blood with those of random expectation," *Anal. Biochem.*, vol. 409, no. 2, pp. 189–194, Feb. 2011.
[64] P. Zhu, P. Bowden, M. Tucholska, D. Zhang, and J. G. Marshall, "Peptide-to-protein distribution versus a competition for significance to estimate error rate in blood protein identification," *Anal. Biochem.*, vol. 411, no. 2, pp. 241–253, Apr. 2011.
[65] P. Bowden, R. Beavis, and J. Marshall, "Tandem mass spectrometry of human tryptic blood peptides calculated by a statistical algorithm and captured by a relational database with exploration by a general statistical analysis system.," *J. Proteomics*, vol. 73, no. 1, pp. 103–11, Nov. 2009.
[66] J. R. Yates, J. K. Eng, A. L. McCormack, and D. Schieltz, "Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database.," *Anal. Chem.*, vol. 67, no. 8, pp. 1426–36, Apr. 1995.
[67] E. F. Codd, "A relational model of data for large shared data banks. 1970.," *MD. Comput.*, vol. 15, no. 3, pp. 162–6.
[68] O. Uysal, T. Sevimli, M. Sevimli, S. Gunes, and A. Eker Sariboyaci, "Cell and Tissue Culture: The Base of Biotechnology," *Omi. Technol. Bio-Engineering*, pp. 391–429, Jan. 2018.
[69] A. R. Baydoun, *Cell culture techniques*. 2005.
[70] B. Taciak *et al.*, "Evaluation of phenotypic and functional stability of RAW 264.7 cell line through serial passages," *PLoS One*, vol. 13, no. 6, p. e0198943, Jun. 2018.
[71] G. Melino, E. Candi, and P. M. Steinert, "Assays for Transglutaminases in Cell Death,"

Methods Enzymol., vol. 322, pp. 433–472, Jan. 2000.
[72] H. Frickmann, E. Schröpfer, and G. Dobler, "Actin assessment in addition to specific immuno-fluorescence staining to demonstrate rickettsial growth in cell culture.," *Eur. J. Microbiol. Immunol. (Bp).*, vol. 3, no. 3, pp. 198–203, Sep. 2013.
[73] A. J. Link and J. LaBaer, "Solution protein digest.," *Cold Spring Harb. Protoc.*, vol. 2011, no. 2, p. pdb.prot5569, Feb. 2011.
[74] H. Schägger and G. von Jagow, "Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa," *Anal. Biochem.*, vol. 166, no. 2, pp. 368–379, Nov. 1987.
[75] J. M. Chick *et al.*, "A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides," *Nat. Biotechnol.*, vol. 33, no. 7, pp. 743–749, Jul. 2015.
[76] P. Bowden, V. Pendrak, P. Zhu, and J. G. Marshall, "Meta sequence analysis of human blood peptides and their parent proteins.," *J. Proteomics*, vol. 73, no. 6, pp. 1163–75, Apr. 2010.
[77] A. K. Florentinus, A. Jankowski, V. Petrenko, P. Bowden, and J. G. Marshall, "The Fc receptor-cytoskeleton complex from human neutrophils," *J. Proteomics*, vol. 75, no. 2, pp. 450–468, Dec. 2011.
[78] C. von Mering *et al.*, "STRING: known and predicted protein-protein associations, integrated and transferred across organisms," *Nucleic Acids Res.*, vol. 33, no. Database issue, pp. D433–D437, Dec. 2004.
[79] S. Ghosh, S. Gepstein, J. J. Heikkila, and E. B. Dumbroff, "Use of a scanning densitometer or an ELISA plate reader for measurement of nanogram amounts of protein in crude extracts from biological tissues," *Anal. Biochem.*, vol. 169, no. 2, pp. 227–233, Mar. 1988.
[80] T. Yao and Y. Asayama, "Animal-cell culture media: History, characteristics, and current issues," *Reprod. Med. Biol.*, vol. 16, no. 2, pp. 99–117, Apr. 2017.
[81] B. Xin *et al.*, "Critical role of Myc activation in mouse hepatocarcinogenesis induced by the activation of AKT and RAS pathways," *Oncogene*, vol. 36, no. 36, pp. 5087–5097, Sep. 2017.
[82] P. Zhu *et al.*, "Comparison of protein expression lists from mass spectrometry of human blood fluids using exact peptide sequences versus BLAST," *Clin. Proteomics*, vol. 2, no. 3–4, pp. 185–203, Dec. 2006.
[83] J. N. Adkins *et al.*, "Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry.," *Mol. Cell. Proteomics*, vol. 1, no. 12, pp. 947–55, Dec. 2002.
[84] G. S. Omenn *et al.*, "Overview of the HUPO Plasma Proteome Project: Results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database," *Proteomics*, vol. 5, no. 13, pp. 3226–3245, Aug. 2005.
[85] J. Marshall, P. Bowden, J. C. Schmit, and F. Betsou, "Creation of a federated database of blood proteins: a powerful new tool for finding and characterizing biomarkers in serum," *Clin. Proteomics*, vol. 11, no. 1, p. 3, Dec. 2014.
[86] J. Marshall *et al.*, "Involvement of Cytosolic Phospholipase A_2 and Secretory Phospholipase A_2 in Arachidonic Acid Release from Human Neutrophils," *J. Immunol.*, vol. 164, no. 4, pp. 2084–2091, Feb. 2000.

[87] L. Senzel, D. V Gnatenko, and W. F. Bahou, “The platelet proteome.,” *Curr. Opin. Hematol.*, vol. 16, no. 5, pp. 329–33, Sep. 2009.
[88] D. Cervi *et al.*, “Platelet-associated PF-4 as a biomarker of early tumor growth,” *Blood*, vol. 111, no. 3, pp. 1201–1207, Oct. 2007.
[89] L. R. Purves, M. Macnab, and I. Bersohn, “Serum alpha-feto-protein. I. Immunodiffusion and immuno-assay results in cases of primary cancer of the liver.,” *S. Afr. Med. J.*, vol. 42, no. 42, pp. 1138–41, Nov. 1968.
[90] A. Adinolfi, M. Adinolfi, and Lessof, “Alpha-feto-protein during development and in disease.,” *J. Med. Genet.*, vol. 12, no. 2, pp. 138–51, Jun. 1975.
[91] G. J. Mizejewski, “The alpha-fetoprotein third domain receptor binding fragment: in search of scavenger and associated receptor targets,” *J. Drug Target.*, vol. 23, no. 6, pp. 538–551, Jul. 2015.
[92] M. P. Laderoute and L. M. Pilarski, “The inhibition of apoptosis by alpha-fetoprotein (AFP) and the role of AFP receptors in anti-cellular senescence.,” *Anticancer Res.*, vol. 14, no. 6B, pp. 2429–38.
[93] S. Wang, Z. Zou, X. Luo, Y. Mi, H. Chang, and D. Xing, “LRH1 enhances cell resistance to chemotherapy by transcriptionally activating MDC1 expression and attenuating DNA damage in human breast cancer,” *Oncogene*, vol. 37, no. 24, pp. 3243–3259, Jun. 2018.
[94] M. J. Jimenez-Dalmaroni *et al.*, “Soluble CD36 Ectodomain Binds Negatively Charged Diacylglycerol Ligands and Acts as a Co-Receptor for TLR2,” *PLoS One*, vol. 4, no. 10, p. e7411, Oct. 2009.
[95] J. Aucamp, A. J. Bronkhorst, C. P. S. Badenhorst, and P. J. Pretorius, “The diverse origins of circulating cell-free DNA in the human body: a critical re-evaluation of the literature,” *Biol. Rev.*, vol. 93, no. 3, pp. 1649–1683, Aug. 2018.
[96] X. Liu *et al.*, “Insulin-like growth factor-1 receptor knockdown enhances radiosensitivity via the HIF-1α pathway and attenuates ATM/H2AX/53BP1 DNA repair activation in human lung squamous carcinoma cells,” *Oncol. Lett.*, vol. 16, no. 1, pp. 1332–1340, May 2018.
[97] F. M. J. Jacobs *et al.*, “An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons,” *Nature*, vol. 516, no. 7530, pp. 242–245, Dec. 2014.
[98] M. Unoki, J. Okutsu, and Y. Nakamura, “Identification of a novel human gene, ZFP91, involved in acute myelogenous leukemia.,” *Int. J. Oncol.*, vol. 22, no. 6, pp. 1217–23, Jun. 2003.
[99] Z. Xie *et al.*, “Zinc finger protein ZBTB20 is a key repressor of alpha-fetoprotein gene transcription in liver,” *Proc. Natl. Acad. Sci.*, vol. 105, no. 31, pp. 10859–10864, Aug. 2008.
[100] Y. M. D. Lo *et al.*, “Maternal Plasma DNA Sequencing Reveals the Genome-Wide Genetic and Mutational Profile of the Fetus,” *Sci. Transl. Med.*, vol. 2, no. 61, pp. 61ra91-61ra91, Dec. 2010.
[101] C. Théry, L. Zitvogel, and S. Amigorena, “Exosomes: composition, biogenesis and function,” *Nat. Rev. Immunol.*, vol. 2, no. 8, pp. 569–579, Aug. 2002.