Characterizing Blood Plasma Proteins from Organ Tissue in Rainbow Trout (*Oncorhynchus mykiss*) Using a Non-Targeted Proteomics Approach

by

Nancy Tannouri

A thesis submitted to the
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

Master of Science in Applied Bioscience

Faculty of Science

University of Ontario Institute of Technology (Ontario Tech University)

Oshawa, Ontario, Canada

December 2021

THESIS EXAMINATION INFORMATION

Submitted by: **Nancy Tannouri**

Master of Science in Applied Bioscience

Thesis title: Characterizing Blood Plasma Proteins from Organ Tissue in Rainbow Trout (*Oncorhynchus Mykiss*) Using a Non-Targeted Proteomics Approach

An oral defense of this thesis took place on December 02, 2021 in front of the following examining committee:

Examining Committee:

Chair of Examining Committee	Dr. Janice Strap
Research Supervisor	Dr. Denina Simmons
Examining Committee Member	Dr. Sean Forrester
Thesis Examiner	Dr. Dario Bonetta, Associate Professor (Ontario Tech University)

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

ABSTRACT

Protein expression patterns adapt to various cues to meet the needs of an organism. The dynamicity of an organism's proteome can therefore reveal information about an organism's health. Proteome databases contain limited information regarding organisms outside of medicinal biology. The Uniprot human and mouse proteomes are extensively reviewed and ~50% of both proteomes include tissue specificity, while >99% of the rainbow trout proteome lacks tissue specificity. This study aimed to expand knowledge on the rainbow trout proteome with a focus on blood plasma proteins. Blood, brain, heart, liver, kidney, and gills were collected from adult rainbow trout, plasma and tissue proteins were analyzed using liquid chromatography tandem mass spectrometry. Over 10,000 proteins were identified across all groups. Majority of the plasma proteome is shared with multiple tissue types, though 4-7% of the plasma proteome is uniquely shared with each tissue (gill > heart > liver > kidney > brain).

Keywords: proteomics; blood plasma; rainbow trout; fish;

AUTHOR'S DECLARATION

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology (Ontario Tech University) to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology (Ontario Tech 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 will be made electronically available to the public.

The research work in this thesis that was performed in compliance with the regulations of Research Ethics Board/Animal Care Committee under REB Certificate number/Animal care certificate file number 15759.

_____ NANCY TANNOURI

STATEMENT OF CONTRIBUTIONS

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication. I have used standard referencing practices to acknowledge ideas, research techniques, or other materials that belong to others. Furthermore, I hereby certify that I am the sole source of the creative works and/or inventive knowledge described in this thesis.

ACKNOWLEDGEMENTS

There are a whole lot of people I would like to thank for their constant encouragement throughout the course of completing my thesis. First, none of this would be possible without the support and guidance from my supervisor Dr. Denina Simmons. Your expertise paired with your humble and enthusiastic attitude has made my experience as a graduate student enjoyable, even in times of adversity. I will forever be thankful for your consideration and understanding over the past few years, especially in dealing with the loss of my father and adapting to a worldwide pandemic midway through my thesis. Your cheers and timely pushes truly helped get me here. I would also like to thank the Aquatic Omics Lab team for their help and advice. You all are some of the most genuine and intelligent people I have been so lucky to work alongside with and learn from. And to my committee member, Dr. Sean Forrester, for providing useful feedback and knowledge throughout the completion of my thesis.

To my mother Sanaa, and late father, Ahmad, thank you for being my biggest cheerleaders who never fail to tell the world just how proud you are of me. I am so happy to be able to show you that all the love and support you have poured into me has brought me to this point. None of this would be possible without the encouragement, love, care, and support you both have given me. To my two older sisters, Nour and Nagham, thank you for keeping my work-life balance in check and reminding me to live a little, you've made this journey so much more worthwhile. I am extremely grateful for my chosen family: Hatem, Nawal, Nadine, Nidal, Nader, and Tammy (+ the triplets & Maya) for their additional love and support. You are truly my go-to people who somehow always seem to show up and be there for me by lending a hand or continuously cheering me on through

all my life moments. Shout-out to my favourite few for always checking in on me and making my game-night dreams come to life: Alex, Danny, Jad, James, Linda, Rima, and Rochelle. The check-ins and small pushes from some of you meant a whole lot more than you may think. For the laughs and giggles: Amanda D, Amanda H, Dana, Hadi, Mareya, Mena, Rami, Ream, Roro, and Sammy. To my real-life science inspo, Sam (@science.sam), for the constant advice and mentorship, as well as editing and reading through important documents, such as my MSc application letter of intent.

Love and appreciate every one of you ☺

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure 8. Venn Diagram comparing plasma proteome to each tissue proteome. Full protein list for each sub-proteome was cross compared to the plasma proteome. The plasma proteome contained a total of 1,990 proteins. Overlapping proteins found amongst sub-proteomes.

Figure 9. Principal component analysis comparing all proteins detected in plasma to all proteins found in each tissue type. (A) PCA comparing liver versus plasma, (B) PCA comparing heart versus plasma, (C) PCA comparing gills versus plasma, (D) PCA comparing brain versus plasma, (E) PCA comparing kidney versus plasma. Distinct sub-proteomes were found amongst plasma and all five tissues analyzed. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

Figure 10. Plasma proteome at a glance. Tissue-specific information surrounding the rainbow trout plasma proteome. “Unique proteins” refers to proteins only found amongst plasma and each individual tissue.

Figure 11. Principal component analysis comparing proteins only found in plasma and brain. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

Figure 12. Protein expression heat-map of proteins unique to plasma and brain. Depicting all 77 proteins amongst plasma and brain; 22 and 28 proteins significantly increased and decreased in plasma, respectively (t-test). High levels of expression in red, low levels of expression in blue. Distance measure using Euclidian distance method with ward clustering on normalized data.

Figure 13. Figure. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and brain. Results intentionally highlight brain related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values were calculated using the Gene Ontology Panther Tool. Data only includes gene ontology IDs with $p < 0.05$.

Figure 14. Principal component analysis comparing proteins only found in plasma and heart. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

Figure 15. Protein expression heat-map of proteins unique to plasma and heart. Depicting top 100 proteins amongst plasma and heart; 49 and 48 proteins significantly increased and decreased in plasma, respectively (t-test). High levels of expression in red, low levels of expression in blue. Distance measure using Euclidian distance method with ward clustering on normalized data.

Figure 16. Figure. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and heart. Results intentionally highlight heart related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Figure 17. Principal component analysis comparing proteins only found in plasma and gill. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

Figure 18. Protein expression heat-map of proteins unique to plasma and gill. Depicting top 100 proteins amongst plasma and gill; 38 and 34 proteins significantly increased and decreased in plasma, respectively (t-test). High levels of expression in red, low levels of expression in blue. Distance measure using Euclidian distance method with ward clustering on normalized data.

Figure 19. Figure. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and gill. Results intentionally highlight gill related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than

yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Figure 20. Principal component analysis comparing proteins only found in plasma and kidney. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

Figure 21. Protein expression heat-map of proteins unique to plasma and kidney. Depicting all 98 proteins amongst plasma and kidney; 22 and 30 proteins significantly increased and decreased in plasma, respectively (t-test). High levels of expression in red, low levels of expression in blue. Distance measure using Euclidian distance method with ward clustering on normalized data.

Figure 22. Figure. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and kidney. Results intentionally highlight kidney related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Figure 23. Principal component analysis comparing proteins only found in plasma and liver. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

Figure 24. Protein expression heat-map of proteins unique to plasma and liver. Depicting top 100 proteins amongst plasma and liver; 43 and 43 proteins significantly increased and decreased in plasma, respectively (t-test). High levels of expression in red, low levels of expression in blue. Distance measure using Euclidian distance method with ward clustering on normalized data.

Figure 25. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and liver. Results intentionally highlight liver related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI

- GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Figure 26. REVIGO semantic similarity of gene ontology (GO) biological processes related to all proteins detected in plasma. Results intentionally highlight plasma related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Figure 27. REVIGO semantic similarity of gene ontology (GO) biological processes related to plasma-specific proteins, which are proteins only detected in plasma. Results intentionally highlight plasma related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

LIST OF ABBREVIATIONS AND SYMBOLS

ATP	Adenosine triphosphate
BSA	Bovine serum albumin
cAMP	Cyclic adenosine monophosphate
DDA	Data-dependent acquisition
DNA	Deoxyribonucleic acid
EBI	European bioinformatics institute
ER	Endoplasmic reticulum
F	Female
FDR	False discovery rate
GDP	Guanosine diphosphate
GLY	Glycine
GO	Gene Ontology
GOA	Gene ontology annotation
GTP	Guanosine triphosphate
HPLC	High-performance liquid chromatography
HnRNP	Heterogeneous nuclear ribonucleoproteins
H_2O	Dihydrogen monoxide
IAA	Iodoacetamide
ID	Identification
JNK	C-Jun N-terminal kinase
LEU	Leucine
M	Male
MAPK	Mitogen-activated protein kinase
MET	Methionine
MPP	Mass Profile Professional
mRNA	Messenger ribonucleic acid
MS/MS	Tandem mass spectrometry
MS-222	Tricaine methanesulfonate
MW	Molecular weight
NADP	Nicotinamide adenine dinucleotide phosphate
PCA	Principal component analysis
PHE	Phenylalanine
PI3K	Phosphoinositide 3-kinases
PLS-DS	Partial least squares discriminant analysis
P5C	1-pyrroline-5-carboxylic acid
Q-TOF	Quadrupole Time-of-Flight
RMP	Revolutions per minute
RNA	Ribonucleic acid
TEAB	Triethylammonium bicarbonate
TCEP	Tris(2-carboxyethyl) phosphine
TYR	Tyrosine
VAL	Valine
ºC	Degrees Celsius

1. Introduction

The focus of this study is to analyze the rainbow trout proteome with emphasis on the plasma proteome and the potential origin of such plasma proteins. The following chapter covers the function of plasma, general information surrounding proteomics, a review of the literature on the rainbow trout proteome, as well as non-lethal blood sampling.

1.1. Plasma

Blood plasma is a pale-yellow liquid component of whole blood that contains water, metabolites, electrolytes, clotting factors, hormones, dissolved gasses, sugars, and proteins (Mathew & Varacallo, 2020). Plasma proteins can be classified in three main groups; (1) proteins involved in plasma functionality, (2) signalling molecules, and (3) proteins released from tissue (Geyer, Holdt, Teupser, & Mann, 2017). A wide-range of proteins make their way into blood plasma as it circulates throughout an organism, rendering blood plasma an effective tool in monitoring overall organismal health. Specifically, blood plasma contains a gold-mine of potential protein biomarkers—such as growth factors, antibodies, and proteins contained in extracellular vesicles—that hold a substantial amount of information indicative of normal biological function, disease, or exposure response (Califf, 2018).

1.2. Proteomics

Proteins are multifunctional macromolecules involved in essentially all cellular processes. To understand the cellular and molecular processes of an organism one must understand protein structure and function, an area of research referred to as proteomics. Proteomics studies the entire set of proteins expressed by an organism, also known as the proteome, at *any* point in time (Consortium, 2020). The proteome is not static. Instead, protein

expression patterns fluctuate and adapt to internal or external cues – such as disease progression and environmental exposure – to meet the needs of an organism (López-Maury, Marguerat, & Bähler, 2008). The dynamicity of an organism's proteome and the study of proteomics can therefore reveal significant information about an organism's health, such as early detection of disease and environmental exposure.

Collaborative scientific efforts have resulted in an abundance of proteomic information regarding a wide-range of organisms. For example, the Uniprot database, which is online and publicly available, holds over 20,000 different reference proteomes, from bacteria to eukaryotes, that can be freely accessed (Consortium, 2020). Uniprot includes various tools and additional resources all in one hub, such as information pertaining to gene ontology and tissue specificity, ultimately facilitating research in large-scale proteomics and bioinformatics. The abundance of information, like tissue specificity, declines as we move away from organisms relevant to medical research (e.g. humans, mice, zebrafish) and move toward environmentally relevant species. For instance, the human proteome contains over 20,000 reviewed proteins, almost half of which include tissue specificity (Consortium, 2020; Zahn-Zabal et al., 2019). Mouse (*Mus musculus*) proteome has over 17,000 reviewed proteins, nearly half of which include tissue specificity. Lastly, zebrafish (*Danio rerio*), a fish species relevant to medical research, has over 3,000 reviewed proteins, with more than 10% having tissue specificity. Such information is limited in environmentally relevant species, reflecting the privation of research in environmental sciences. For instance, the rainbow trout (*Oncorhynchus* mykiss) proteome is classified as a complete reference proteome, though only five proteins have been reviewed, and only two proteins include tissue specificity – this means over 99% of the rainbow trout

reference proteome lacks information on tissue specification and origin. Similar trends are seen across various bioinformatics tools and databases. For example, Proteomic Identification Database results in greater than seven thousand results upon searching 'human', while searching 'rainbow trout' results in less than twenty hits (Vizcaíno et al., 2015).

1.3. Rainbow trout proteomics

Most research surrounding rainbow trout proteomics – which is scarce as is – looks at either blood plasma or tissue, though comparing the blood plasma proteome to multiple tissue proteomes is not available. Such studies have yet to be examined on any fish species, though a comprehensive study on zebrafish blood plasma and tissue specification using *in-silico* data-mining was done by Li et al. (2016). The preceding study used non-targeted proteomics to characterize blood plasma in zebrafish and found greater than 900 plasma proteins and no sex biases excluding abundant female yolk precursor proteins (e.g. Vitellogenin). Additionally, Li et al. (2016) cross compared their blood plasma dataset to databases containing protein tissue specificity and found most plasma proteins were expressed in whole-body or multiple organs, though liver was thought to be the predominant site of plasma protein production (C. Li, Tan, Lim, Lin, & Gong, 2016).

Current research surrounding rainbow trout proteomes include proteomic profiles of liver, spleen, head kidney, skeletal muscle, gills, blood plasma, and reproductive cells, typically exposed to various conditions. The proceeding information is summarized in Table 1. Causey et al. (2018) exposed rainbow trout to *Aeromonas salmonicida*, a bacterium that negatively impacts salmonids, and observed differences in liver proteomes between

control and exposed fish. The preceding study found proteins linked to immunity, stress, and vesicle-mediated transport were upregulated in the liver of exposed fish (Causey et al., 2018). Proteome analysis of rainbow trout exposed to Verapamil, a cardiovascular pharmaceutical found in aquatic environments, resulted in altered liver proteins compared to control; high concentrations of Verapamil resulted in upregulated proteins, while environmentally relevant concentrations resulted in downregulated proteins (Z. H. Li, Li, Sulc, Hulak, & Randak, 2012). Martin et al. (2001) analyzed the effect of short-term starvation on liver proteins and found abundance differences in 24 proteins between control and starved groups; some of which included proteins involved in protein degradation and energy metabolism (Martin, Cash, Blaney, & Houlihan, 2001). Baumgarner et al. (2013) also studied the impact of short-term starvation on rainbow trout and found differential expression of 40 proteins in intestinal epithelia involved in immunity and stress (Baumgarner, Bharadwaj, Inerowicz, Goodman, & Brown, 2013). The effects of *Yersinia ruckeri,* a bacterium causing enteric red mouth disease, on rainbow trout head kidney and spleen proteomes were analyzed by Kumar et al. (2018); the results indicated 34 and 85 differentially expressed proteins in head kidney and spleen, respectively. The preceding study found differentially expressed proteins to be involved in immune response, such as lysosomes and antioxidants (Kumar et al., 2018).

Aedo et al. (2019) analyzed the effects of cortisol-stimulated skeletal muscle in rainbow trout and found a total of 108 differentially expressed proteins amongst cortisol and cortisol-BSA treatments, close to 75% being involved in cellular processes and binding (Aedo, Fuentes-Valenzuela, Molina, & Valdés, 2019). Muscle exposed to hypoxic conditions resulted in 46 differentially expressed proteins compared to control in a time-

dependent manner; proteins identified in hypoxic groups were linked to homeostasis, energy metabolism, and muscle structure (Wulff, Jokumsen, Hojrup, & Jessen, 2012). Gills directly exposed to radiation demonstrated upregulation of annexin II, a cancer related protein, while gills receiving signals from irradiated cells demonstrated upregulation of protective proteins involved in oxidative damage (Smith, Wang, Bucking, Mothersill, & Seymour, 2007). Nynca et al. (2014) identified 206 sperm proteins in healthy rainbow trout and found that most proteins were involved in metabolic processes and transport, as well as catalytic activity and binding (Nynca, Arnold, Fröhlich, Otte, & Ciereszko, 2014). In a later study, Nynca et al. (2017) compared rainbow trout blood plasma proteins to seminal plasma proteins and found a total of 119 blood plasma proteins, most of which were acute phase proteins, as well as 54 differentially expressed proteins in seminal plasma involved in various signalling pathways (Nynca, Arnold, Fröhlich, & Ciereszko, 2017). Babaheydari et al. (2016) studied the effects of heat-shock on rainbow trout fertilized eggs and found 15 proteins decreased significantly in heat-shocked eggs, all of which were related to vitellogenin (Babaheydari, Keyvanshokooh, Dorafshan, & Johari, 2016).

1.4. Non-lethal blood sampling

Considering blood plasma is accessible and can be acquired non-lethally, using blood plasma proteins can serve as a promising method of environmental health monitoring. Aside from human biofluids, knowledge regarding environmentally relevant species and their respective biofluids are limited (Martyniuk & Simmons, 2016). Expanding our knowledge on blood plasma protein content and potential disease biomarkers will help push forth non-lethal environmental health monitoring programs in real-world settings,

ultimately reducing the need to sacrifice wildlife. Advancements related to non-lethal sampling considers the three Rs in animal research: replacement, refinement, and reduction, and is especially significant in research pertaining to endangered and at-risk species (Huang et al., 2016).

Table 1. Summary of rainbow trout proteomics found in literature.

TISSUE / EXPOSURE	MAIN FINDINGS	REFERENCE
liver / exposed to *Aeromonas salmonicida*	increase in abundance of proteins linked to immunity, stress, and vesicle mediated transport	Causey et al. (2018)
liver / Verapamil (cardiovascular pharmaceutical)	altered liver protein concentration: high concentrations resulted in upregulated proteins; environmentally relevant concentration resulted in downregulated proteins	Li et al. (2012)
liver / short-term starvation	24 proteins differed amongst control and treatment groups, some of which included proteins involved in protein degradation and energy metabolism	Martin et al. (2001)
intestinal epithelial / short-term starvation	40 proteins different amongst control and treatment groups, some of which included proteins involved in immunity and stress	Baumgarner et al. (2013)
head kidney and spleen / *Yersinia ruckeri*	34 and 85 differentially expressed proteins in head kidney and spleen, respectively. Some of which included proteins involved in immune response, such as lysosomes and antioxidants	Kumar et al. (2018)
skeletal muscle / cortisol	108 differentially expressed proteins amongst cortisol and cortisol-BSA treatments, close to 75% being involved in cellular processes and binding	Aedo et al. (2019)
skeletal muscle / hypoxic conditions	46 differentially expressed proteins compared to control in a time-dependent manner; proteins identified in hypoxic groups were linked to homeostasis, energy metabolism, and muscle structure	Wulff et al. (2012)

gills / radiation	Gills directly exposed to radiation demonstrated upregulation of annexin II. Gills receiving signals from irradiated cells demonstrated upregulation of protective proteins involved in oxidative damage	Smith et al. (2007)
sperm / control	identified 206 sperm proteins in healthy rainbow involved in metabolic processes and transport, as well as catalytic activity and binding	Nynca et al. (2014)
blood plasma and semen / control	119 blood plasma proteins and 54 differentially expressed seminal plasma proteins	Nynca et al. (2017)
fertilized eggs / heat stress	15 vitellogenin related proteins decreased in heat-shock eggs when compared to control	Babaheydari et al. (2016)

2. Rationale and research objectives

The goal of this study was to discover fundamental information regarding rainbow trout (*Oncorhynchus mykiss*) blood plasma using a systems-biology approach. Blood plasma proteins will be identified, and the plasma proteome will be compared to various tissue proteomes to assess potential tissue specific origin. Expanding our knowledge on blood plasma and tissue specificity could help drive future research in biomarker discovery, progress current proteomic databases, and improve current methods of environmental health monitoring.

3. Materials & Methods

3.1. Fish maintenance

Nineteen healthy adult rainbow trout (sex: 12F, 7M; weight: 245-460 g; fork length: 26-33 cm) were acquired from Linwood Acres Trout Farms LTD (ON, Canada). Fish were transferred to 1000 L flow-through aquarium (5 turnovers/24 hours) and acclimated for several weeks at 12ºC with a photoperiod of 16 h light and 8 h darkness. Fish were fed commercial trout pellets equivalent to 1.6% total body weight, once a day. Fish health and behaviour were monitored daily.

3.2. Blood sampling and tissue collection

Following the acclimation period, fish were anesthetized with 100 mg/L MS-222, weighed and measured for fork length. Blood was sampled using caudal vein puncture (posterior to the anal fin) with 21-gauge needle and heparinized vacutainers. Fish were euthanized by exsanguination. Separation of plasma from whole-blood was obtained after centrifugation at 2000 rpm for 20 minutes at 4ºC (Sorvall Legend RT Quick Set Refrigerated Benchtop Centrifuge). Plasma was flash frozen in liquid nitrogen then stored at -80ºC until further analysis. After plasma collection, various organ tissues were dissected; whole brain, heart, and kidney, as well as one gill, and a portion of liver were dissected from all biological samples. Samples were flash frozen in cryovials, and stored at -80ºC until further analysis.

3.3. Plasma and tissue sample preparation

Plasma and tissue samples were removed from the -80ºC and placed on ice to gradually thaw. Consistent 100 mg portions of each tissue type were measured and mixed with 500

mL of 100 mM triethylammonium bicarbonate (TEAB) and two beads, then run on a ball mill for 1 minute at 20 Hz. Samples were centrifuged at 14,000 xg for 15 minutes at 4ºC. Tissue supernatants containing soluble proteins were collected and a portion was used to estimate protein concentration using a Qubit 4 Fluorometer. Approximately 1 mg of total protein concentration – ranging between 40-50 µL homogenate depending on tissue type – was transferred into low retention tubes for further sample preparation; 100 mM TEAB was added to homogenates with <50 µL to bring the final volume to 50 µL. Plasma and tissue homogenates were vortexed to obtain a uniform mixture, and 15 µL of each sample was transferred into low retention microcentrifuge tubes. Next, 35 µL of 100 mM TEAB buffer was added to each tube and vortexed. Proteins in plasma and homogenates were then reduced using 2.65 µL of 100 mM tris(2-carboxyethyl) phosphine (TCEP) in 100 mM TEAB, vortexed, and incubated at room temperature for 45 minutes. Then, 2.8 µL of alkylating solution (200 mM iodoacetamide (IAA) in 100 mM TEAB was added, vortexed, and incubated in the dark at room temperature for 45 minutes. Following incubation, 50 µL of chemical digestion solution (20% formic acid v/v) was added and vortexed gently. Each tube was lid-locked and incubated at 115ºC for 30 minutes using a VWR 96 heating block. Next, samples were evaporated to ~20 µL using a centrifugal evaporator (Genevac miVac Quattro Concentrator) for ~40 minutes. Samples were re-suspended in 20 µL high-performance liquid chromatography (HPLC) buffer (95% H_2O, 5% acetonitrile, 0.1% formic acid), and gently vortexed until dried pellets were completely reconstituted. Reconstituted samples were centrifuged at 10,000 xg for 10 minutes to precipitate any debris. Next, 20 µL of the sample supernatant and 2 µL of internal peptide

standard (Sigma Aldrich, HPLC peptide standard mixture, H2016) were added to 2 mL screw thread HPLC vials containing 250 µL pp bottom spring inserts. Peptide solutions were stored at 4ºC until instrumental analysis.

3.4. Proteomics

Reverse phase separation of each sample was completed using an Agilent 1260 Infinity Binary LC and Zorbax, 300SB-C18, 1.0 × 50 mm 3.5 µm column (Agilent Technologies Canada Inc., Mississauga, ON). Specifically, 2 µL of peptide solution from each sample was injected into the instrument and separated using reverse phase chromatography. Liquid chromatography was coupled to The Agilent 6545 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) to detect and identify peptides. Refer to Appendix A for detailed instrumental methods. Each run included a solvent blank and a BSA digest standard (Agilent Technologies Canada Inc., Mississauga, ON), as well as an external standard (Sigma Aldrich, HPLC peptide standard mixture, H2016), which were injected between every 10 samples to monitor baseline, carry-over, drift, sensitivity, and overall instrumental variation during the runtime. Tissue and plasma samples were injected twice per each individual fish.

Spectrum Mill Software (Version B.06.00, Agilent Technologies) was used to analyze the spectral files of each sample to identify proteins. Specifically, proteins were identified by cross-comparing detected peptides against known proteins found in the Uniprot Reference Proteome for rainbow trout (Proteome ID#UP000193380) in January 2020. Proteins were manually accepted when the following criteria were met, (1) peptide score (quality match between the observed spectrum and theoretical spectrum) greater than 6

in at least one peptide, and (2) a %SPI (percent of the spectral intensity accounted for by the theoretical fragments) greater than 70%, both of which are recommended for validating results using an Agilent Q-TOF mass spectrometer. Once peptides were identified and sequenced using Spectrum Mill at the MS/MS level, missing values at the MS1 level were quantified using a data-dependent acquisition (DDA) workflow in Skyline 20.2 (MacCoss Lab Software) with a cut-off score of 0.9, 5-minute retention time window, and 5 missed cleavages with transition settings for TOF (Pino et al., 2020). The list of Uniprot protein identifiers and their respective spectral intensities were exported from Skyline and uploaded into Mass Profiler Professional 15.1 (Agilent Technologies) to normalize data (discussed below) and obtain data files compatible with Excel. The list of identified rainbow trout protein accession numbers was blasted against the human Uniprot Reference Proteome (Proteome ID#UP000005640) to obtain human protein orthologs based on sequence similarity.

3.5. Normalization using internal standards

The standards used to account for instrumental variation contained five peptides, three of which were stable peptides used to obtain scale values for normalization. The peptides of interest were VAL-TYR-VAL (MW = 379.5; Sigma Aldrich H2016, V 8376), TYR-GLY-GLY-PHE-MET (MW= 573.7; Sigma Aldrich H2016, M 6638), and TYR-GLY-GLY-PHE-LEU (MW= 555.6; Sigma Aldrich H2016, L 9133). MassHunter Qualitative software (Agilent Technologies) was used to integrate standard peaks and obtain peak area detected by specific mass. Initially, the external standard peak area was to be used to scale experimental data, though due to poor external standard peptide detection, we

opted to use internal standards – spiked into each individual sample – to normalize the data. For each individual sample, peak area for the three peptide standards were averaged – call this 'value A' (e.g., the three detected peptide standards of interest [VAL-TYR-VAL, TYR-GLY-GLY-PHE-MET, and TYR-GLY-GLY-PHE-LEU] in sample #1 were averaged to get 'value A'). Then, the average peak area of each run (1 run = 10 samples) was calculated – call this 'value B' (e.g., run 1 included sample #1 to sample #10, 'value A' for all ten samples were averaged to obtain 'value B'). To get the scale value for each individual sample, value B (pooled run average) was divided by value A (average peak area in each individual sample). This value was typically close to 1. Data files were imported into Mass Profiler Professional (MPP) and normalized based on the defined scale value (multiply protein abundance values by the calculated scale value). The resulting normalized data can be visualized in Figure 1. Data was exported from MPP into Excel, where further data cleanup was conducted. Technical replicates were consolidated using the maximum protein abundance found amongst the two iterative runs per sample, and all proteins containing less than 5000 mean intensity values were removed.

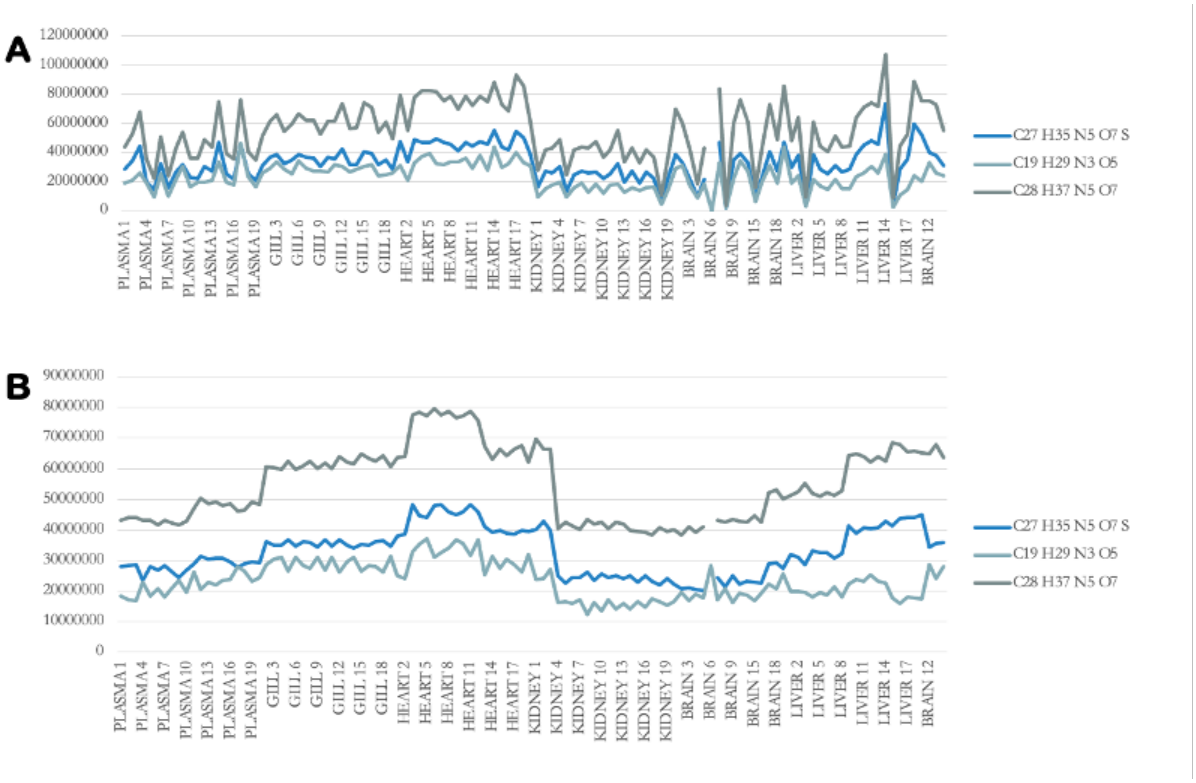


Figure 1. Instrumental Variation Corrected Using Internal Standards. Standard mix contained five peptides, 3 of which were used for normalization. 2uL of internal standards were spiked into each sample. Peak integration and area were found using MassHunter Qual. (A) Peak intensity before normalization, and (B) Peak intensity after normalization.

3.6. Statistical data analysis

Proteomics statistical analysis was conducted using MetaboAnalyst 5.0 where data was normalized by the median, log transformed, and scaled using pareto scaling (Pang et al., 2021). Principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA), and heat maps were also carried out using MetaboAnalyst 5.0. Any statistical analysis and significance test using MetaboAnalyst contained an FDR corrected p-value cut-off of 0.05. Gene Cards and Uniprot (Consortium, 2020; Stelzer et al., 2016) were used to investigate function and tissue specific information surrounding proteins of interest. REVIGO was used to visualize semantic similarities of the gene ontology biological processes linked to the identified proteins (Supek, Bošnjak, Škunca, & Šmuc, 2011). REVIGO includes a user-provided p-value option, which in our experiment was acquired using Gene Ontology Panther (P. D. Thomas et al., 2003; Paul D. Thomas et al., 2006); p-values were based on the number of proteins linked to a given gene ontology ID, where for each gene ontology ID, the number of proteins in the input dataset are compared to what would be expected. The preceding comparison is based on the Homo Sapien reference list and statistically analyzed using Fisher's Exact Test; a test used to determine non-random representation in categorical data. GraphPad Prism software (Version 9.1.2) was used to create bar graphs in Figure 6.

4. Results

4.1. Overview

The overview section includes full proteomic profiles and comparisons of all six sample types analyzed; brain, gill, heart, kidney, liver, and plasma.

4.1.1. Whole-body proteomic analysis

A total of 10,564 proteins were identified across all six groups analyzed, almost half of which were non-redundant (Figure 2). The gill proteome made up 20.18% of the +10,000 proteins identified, followed by plasma (18.84%), heart (17.61%), liver (16.02%), kidney (13.76%), and brain (13.59%). Whole-body comparison showed distinct sub-proteomes across the 6 groups analyzed, depicted by both qualitative and quantitative results (Figure 2, 3 & 4). Qualitatively, each sample group contained a sub-set of unique proteins, ranging from 200 to over 500 proteins; the brain (233) and liver (372) had the lowest number of unique proteins, while gill (499) and plasma (576) had the largest number of unique proteins (Figure 2). Quantitatively, both PCA and heat-map analysis showed distinct clustering amongst all six groups (Figure 3 & 4). Specifically, PCA analysis showed the closest relationship amongst the brain and liver proteomes, while the plasma and gill proteomes were the most contrasting (Figure 3). Additionally, the heat-map clearly showed differences in protein abundance levels in all six groups (Figure 4).

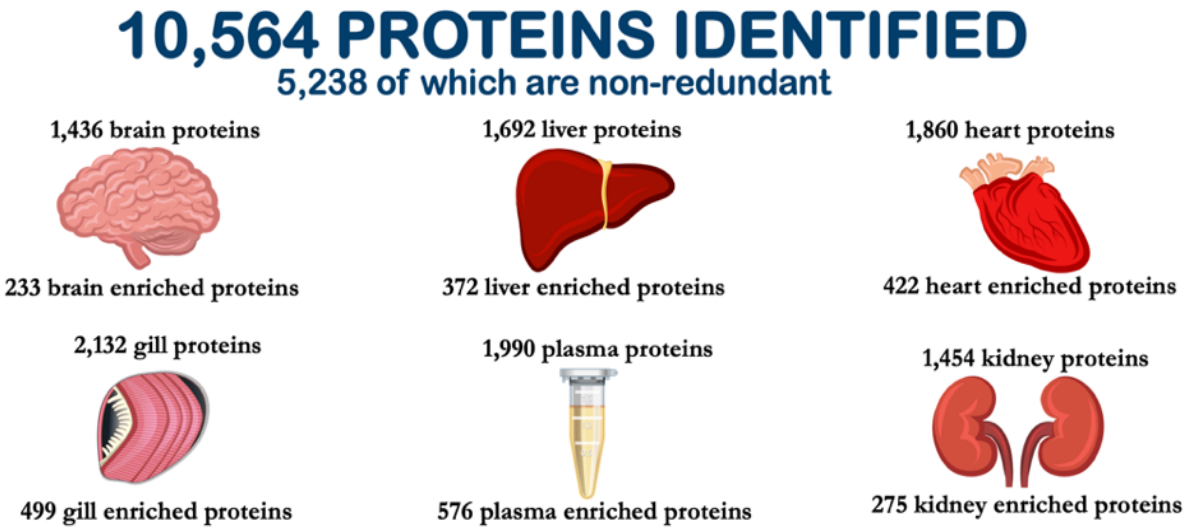


Figure 2. Overview of the proteins identified in brain, liver, heart, gills, plasma, and kidney. “Non-redundant proteins” refers to total number of unique proteins. “Enriched proteins” refers to proteins distinct to only one tissue type.

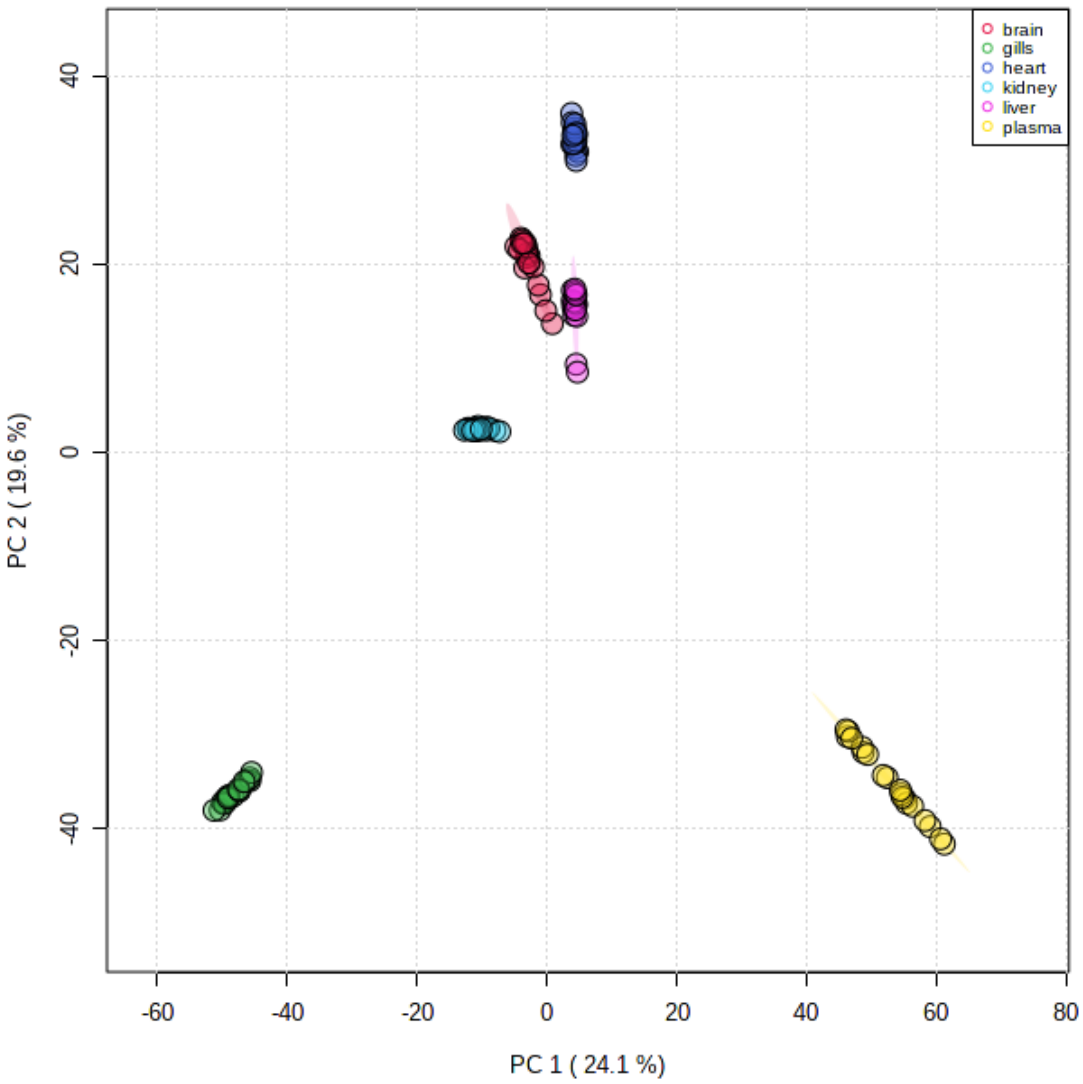


Figure 3. Principal Component Analysis (PCA) was preformed to compare whole-body rainbow trout proteome, specific to brain, gills, heart, kidney, liver, and plasma proteomes. Nineteen biological samples were used for all six sample types, represented by a single point on the graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

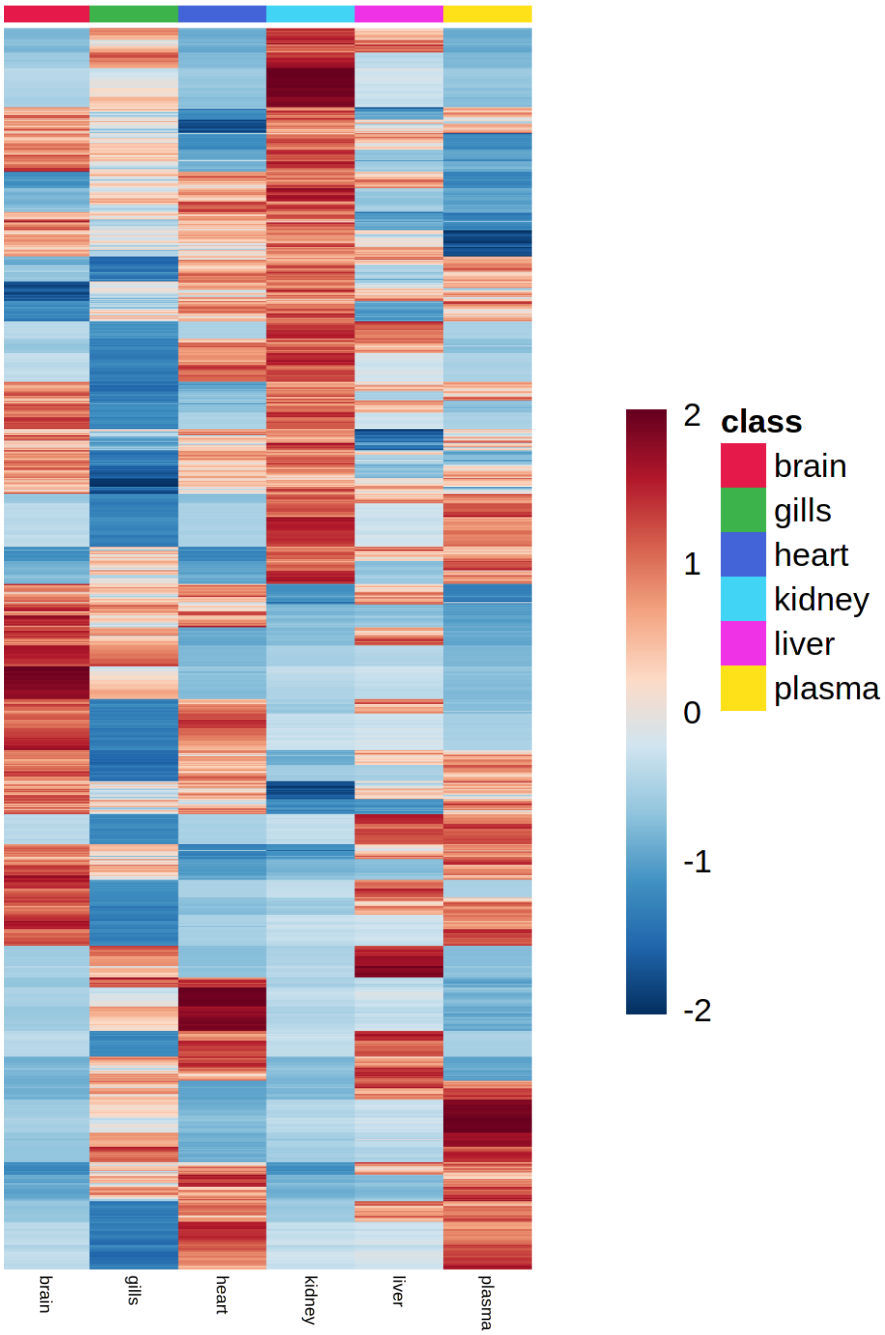


Figure 4. Heat-map of protein abundance in brain, gill, heart, kidney, liver, and plasma. High levels of expression in red, low levels of expression in blue. Distance measure using Euclidian distance method. The heat-map represents grouped averages of each sample type; such that one row demonstrates the average of each protein in all 19 samples, per group.

4.1.2. Sex differences

Clustering in the PCA plot appeared to be tissue specific and unrelated to sex (Figure 5A). Further modelling of the data using PLS-DA showed clustering was driven by tissue type, though low R^2 and negative Q2 values indicated that the PLS-DA was a poor fit and not a predictive model of the original dataset (Figure 5B). Additionally, all accession numbers linked to the egg-yolk protein (vitellogenin; Vtg) in rainbow trout were mapped to the APOB-100 protein in humans; APOB-100 protein was identified in all male and female biological samples in all six sample groups analyzed (Figure 6B). Again, no sex specific differences were apparent, though differences amongst tissue types were present; e.g., APOB-100 was predominantly found in the brain (Figure 6A). Sex difference analysis was also conducted in each individual tissue type, though no specific differences were seen (Figure 7). PCA plots showed overlapping clusters across all six groups analyzed (Figure 7A-7F).

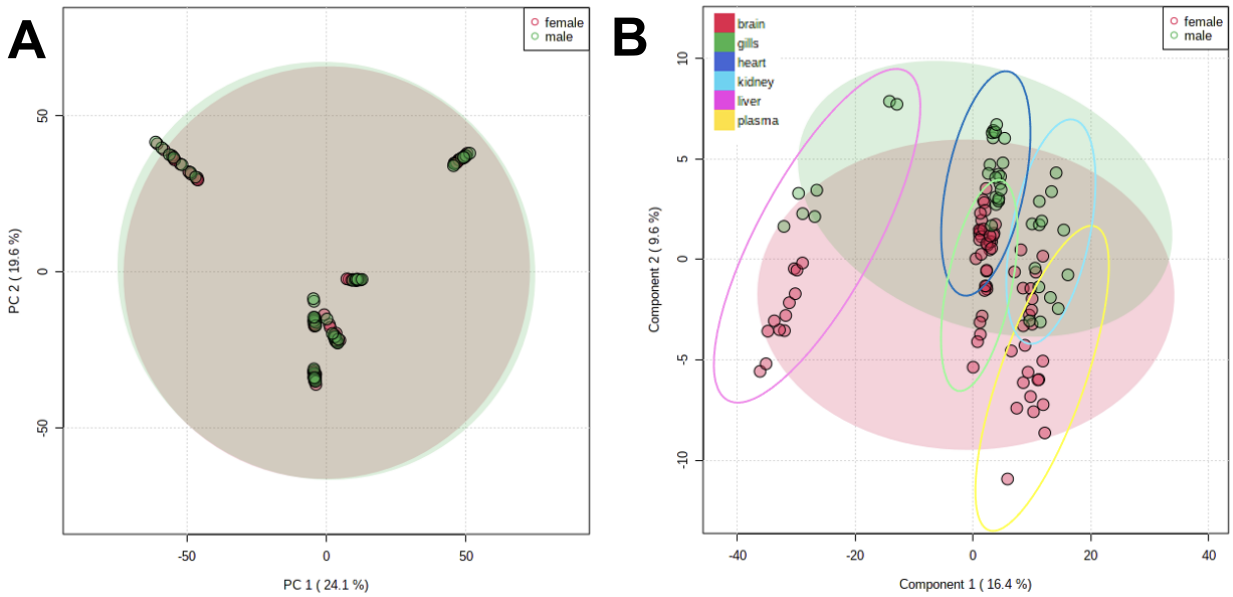


Figure 5. Whole-body sex difference comparison in rainbow trout. Nineteen biological samples total; twelve females (red dots) and seven males (green dots). (A) Principal Component Analysis (PCA) comparing male versus female proteomes in all tissues, (B) Partial Least Squares Discriminant Analysis (PLS-DA) comparing male versus female proteomes in all tissues. PLS-DA resulted in negative Q2 values and low R^2 values, indicating poor fit and not a predictive model of original data. Slight clustering in PCA and PLS-DA are sample-specific, and not related to sex differences.

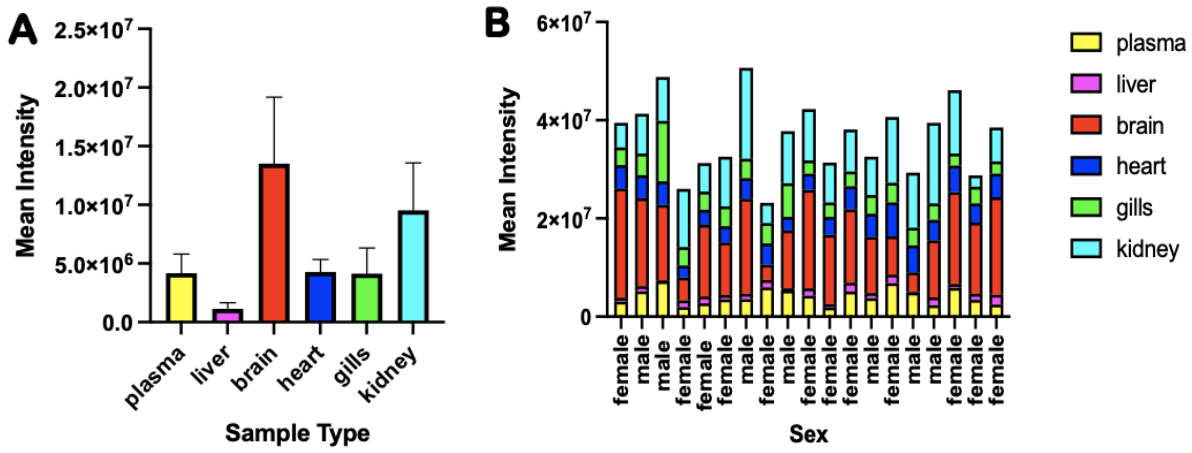
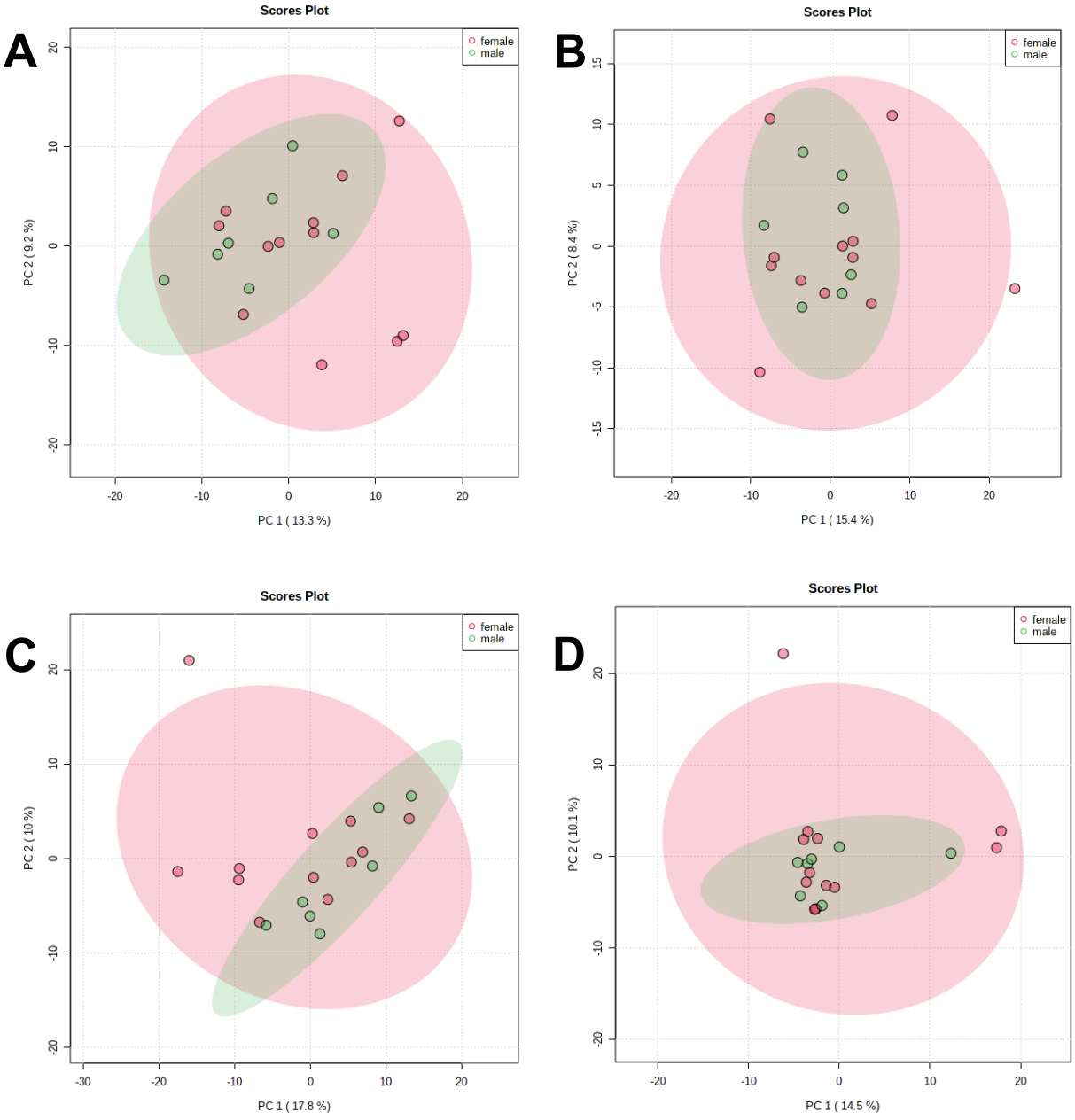


Figure 6. APOB-100 protein abundance patterns in male and female samples. Mean protein abundance levels are plotted on the y-axis, and sample type (A) or sex (B) are plotted on the x-axis. Four rainbow trout protein accession numbers (A0A060Z709, A0A060Y552, A0A060W7E7, A0A060W754) were identified as vitellogenin proteins, all of which were mapped to the human protein ortholog, APOB-100.


A
Scores Plot
female
male
PC 2 (9.2 %)
PC 1 (13.3 %)
B
Scores Plot
female
male
PC 2 (8.4 %)
PC 1 (15.4 %)
C
Scores Plot
female
male
PC 2 (10 %)
PC 1 (17.8 %)
D
Scores Plot
female
male
PC 2 (10.1 %)
PC 1 (14.5 %)

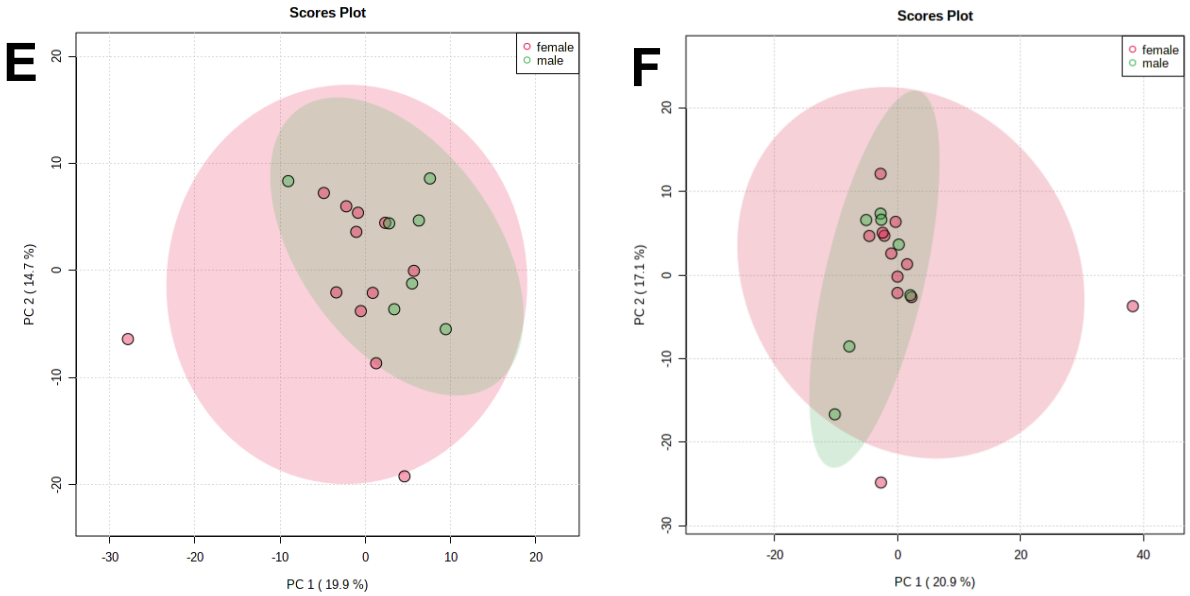


Figure 7. Principal Component Analysis (PCA) comparing male versus female proteomes in each sample type; (A) liver, (B), heart, (C) gills, (D) brain, (E) kidney, (F) plasma. Nineteen biological samples total; twelve females (red dots) and seven males (green dots).

4.2. Comparing plasma proteome to each tissue proteome

To analyze tissue specificity in the plasma proteome, the full plasma proteome was cross-compared to each individual tissue proteome. Qualitative observations using Venn diagrams showed at least 500 overlapping proteins between plasma and all five tissues analyzed (Figure 8). Specifically, 539 proteins were shared amongst the plasma and kidney proteome, followed by brain (540), liver (611), heart (660), and gill (713). Note that there are over-lapping proteins found across all groups; e.g., a number of proteins shared between plasma and liver could be present in any or all of the other proteomes. Statistical analysis (PCA) revealed that although the plasma proteome contained shared proteins with each tissue type, isolated clustering was still apparent between the plasma proteome and each individual tissue (Figure 9A-9E).

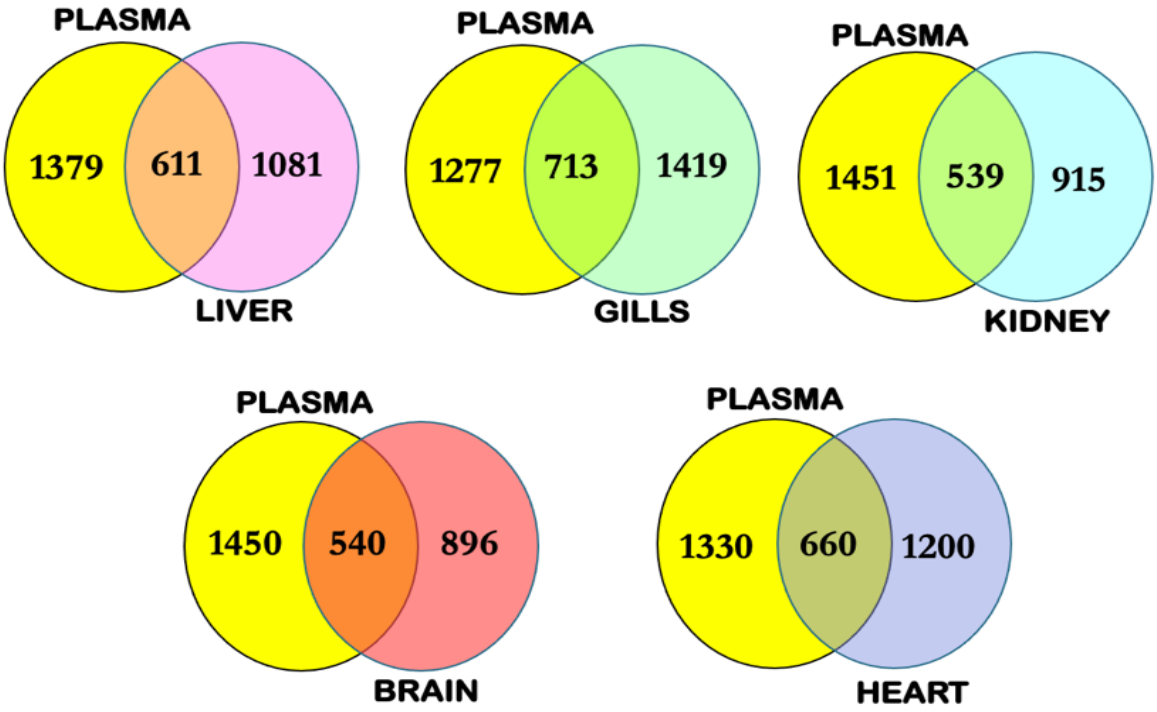
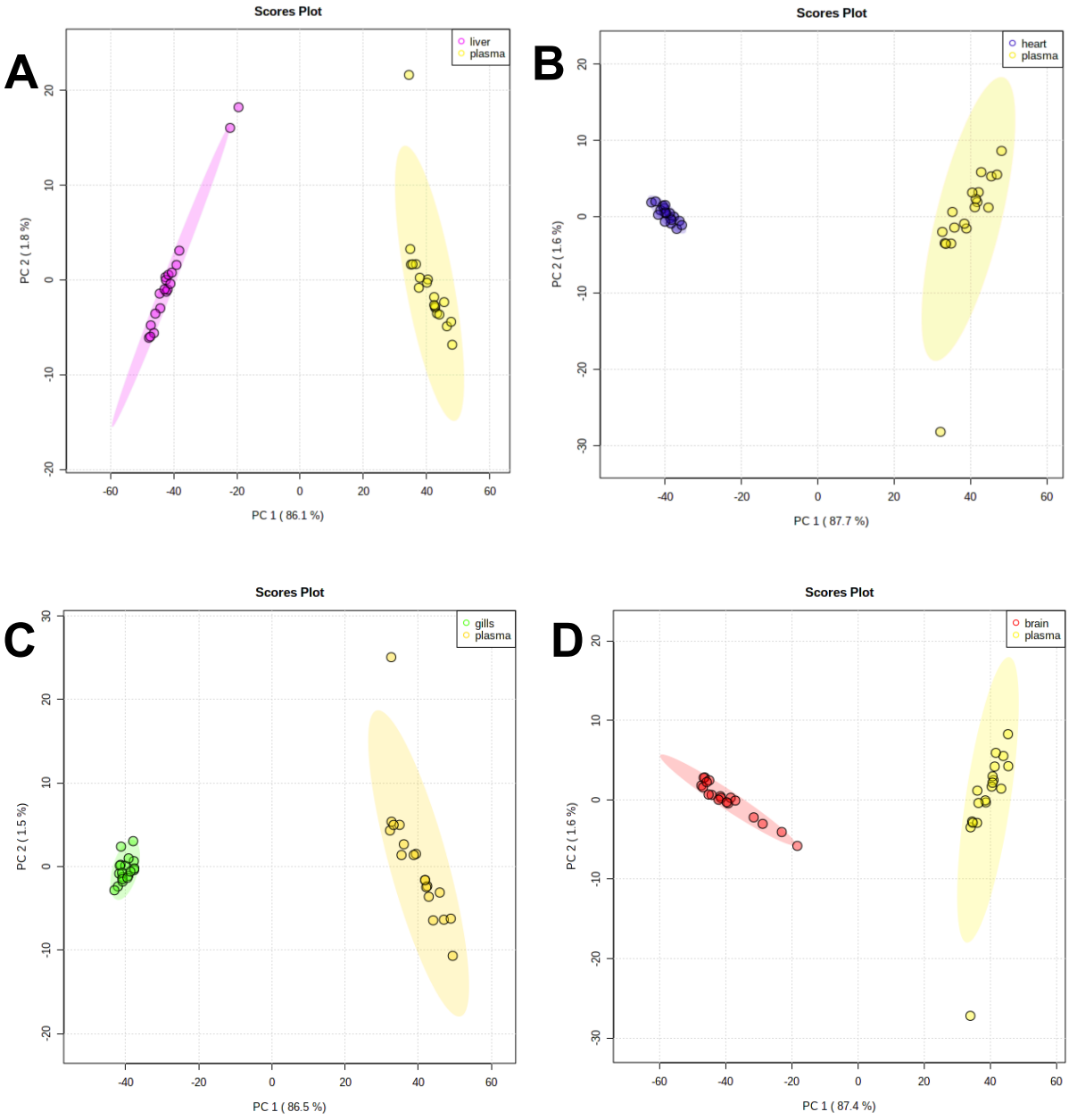


Figure 8. Venn Diagram comparing plasma proteome to each tissue proteome. Full protein list for each sub-proteome was cross compared to the plasma proteome. The plasma proteome contained a total of 1,990 proteins. Overlapping proteins found amongst sub-proteomes.


A
Scores Plot
liver
plasma
PC 2 (1.8 %)
PC 1 (86.1 %)
B
Scores Plot
heart
plasma
PC 2 (1.6 %)
PC 1 (87.7 %)
C
Scores Plot
gills
plasma
PC 2 (1.5 %)
PC 1 (86.5 %)
D
Scores Plot
brain
plasma
PC 2 (1.6 %)
PC 1 (87.4 %)

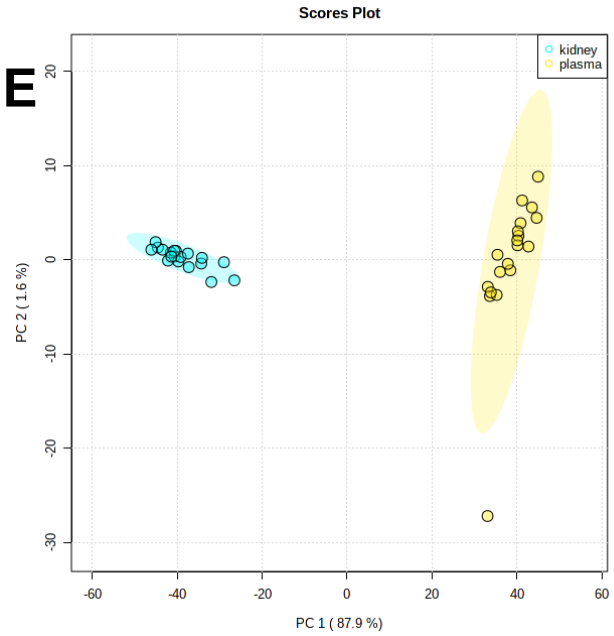


Figure 9. Principal component analysis comparing all proteins detected in plasma to all proteins found in each tissue type. (A) PCA comparing liver versus plasma, (B) PCA comparing heart versus plasma, (C) PCA comparing gills versus plasma, (D) PCA comparing brain versus plasma, (E) PCA comparing kidney versus plasma. Distinct sub-proteomes were found amongst plasma and all five tissues analyzed. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

4.3. Comparing proteins unique to plasma and each individual tissue type

Upon comparing the plasma proteome to the full proteome of each tissue type, further statistical analysis (e.g., heat-map analysis) produced results that highlighted the proteins that were not similar across the plasma and individual tissue proteomes, e.g., proteins that were only present in plasma, or only present in the individual tissues. To better understand the proteins shared between plasma and each tissue, we opted to compare the subgroup of proteins that were unique to plasma and each individual tissue type. Qualitative results of the plasma proteome showed almost half of the proteins found in plasma (43%) have multi-organ origins (found in 2+ tissues), 30% are plasma-specific proteins (only found in plasma), 7% are gill-specific, 6% are heart-specific, 6% are liver-specific, 5% are kidney specific, and 4% are brain specific (Figure 10).

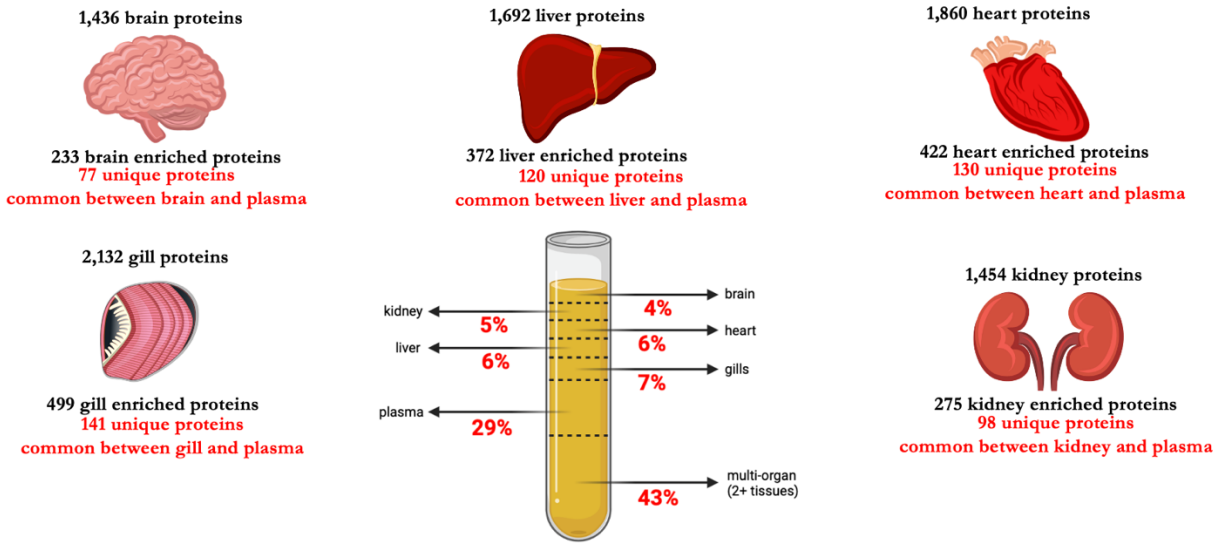


Figure 10. Plasma proteome at a glance. Tissue-specific information surrounding the rainbow trout plasma proteome. “Unique proteins” refers to proteins only found amongst plasma and each individual tissue.

4.3.1. Plasma versus brain proteome

The proceeding analysis compared the 77 proteins that were only shared amongst the plasma and brain proteomes. Although all 77 proteins were found across brain and plasma, PCA analysis showed that protein abundance levels differed and still resulted in significantly different sub-proteomes (Figure 11). Of the 77 proteins, 22 proteins were significantly increased in plasma, and 28 proteins were significantly decreased in plasma when compared to brain (FDR corrected p-value < 0.05; Figure 12). The function and expression annotations for the preceding significant proteins are summarized in Table 1. Gene Card and Uniprot tissue-expression data showed that 8 out the 50 significant proteins associated with brain and plasma are predominantly expressed in the human brain: GDI1, ACBD7, RIMBP2, ESRRB, SOX11, KCNG1, RASGRF2, and BBOX1. The functions of the preceding 8 proteins include binding and transport, catalytic activity, synaptic transmission, and exchange factors (Table 1). Gene ontology analysis of the 77 proteins associated with plasma and brain contained many biological processes specific to brain function, most of which could be summarized into (1) regulatory functions, (2) signaling, and (3) behaviour (Figure 13).

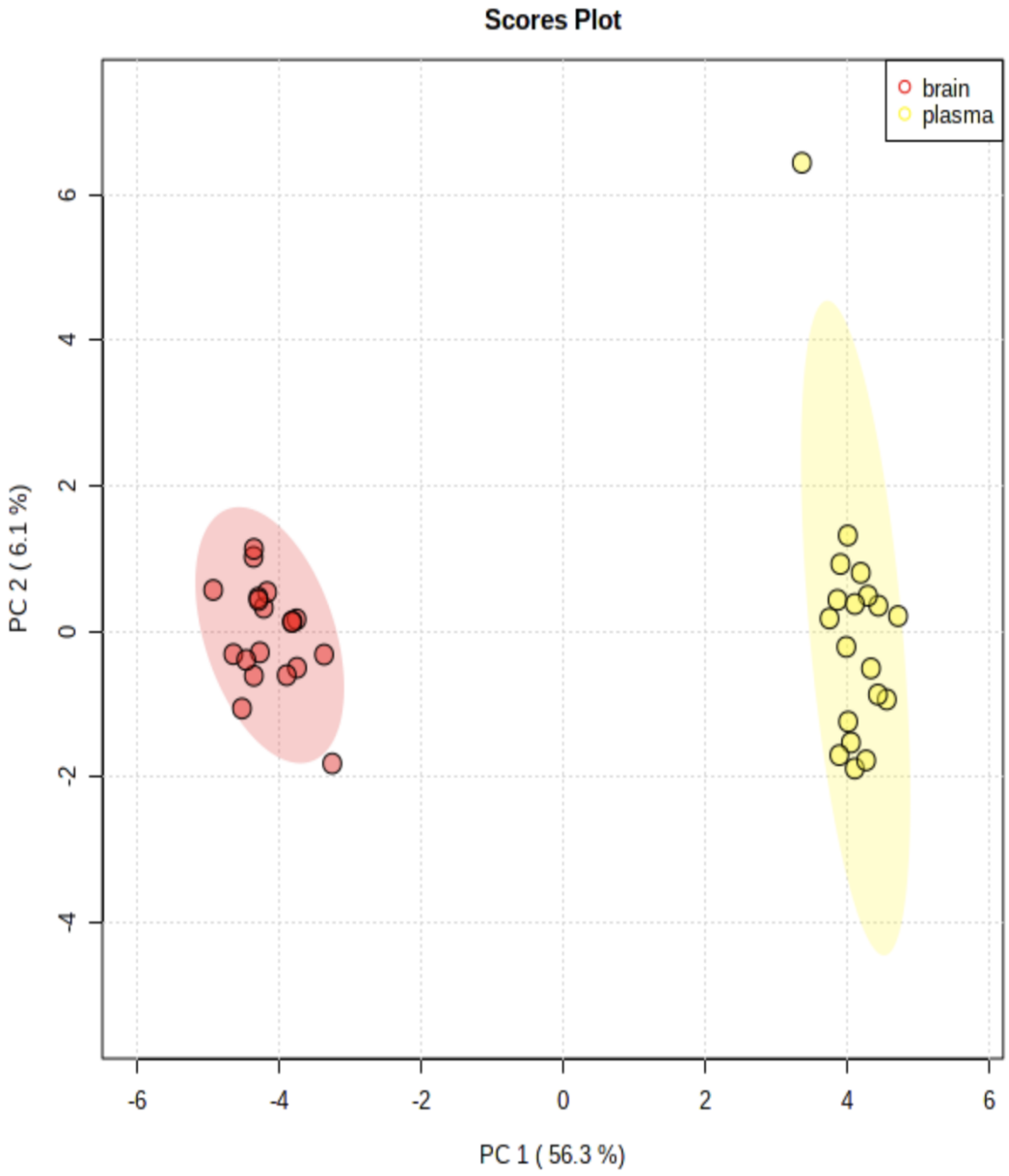


Figure 11. Principal component analysis comparing proteins only found in plasma and brain. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

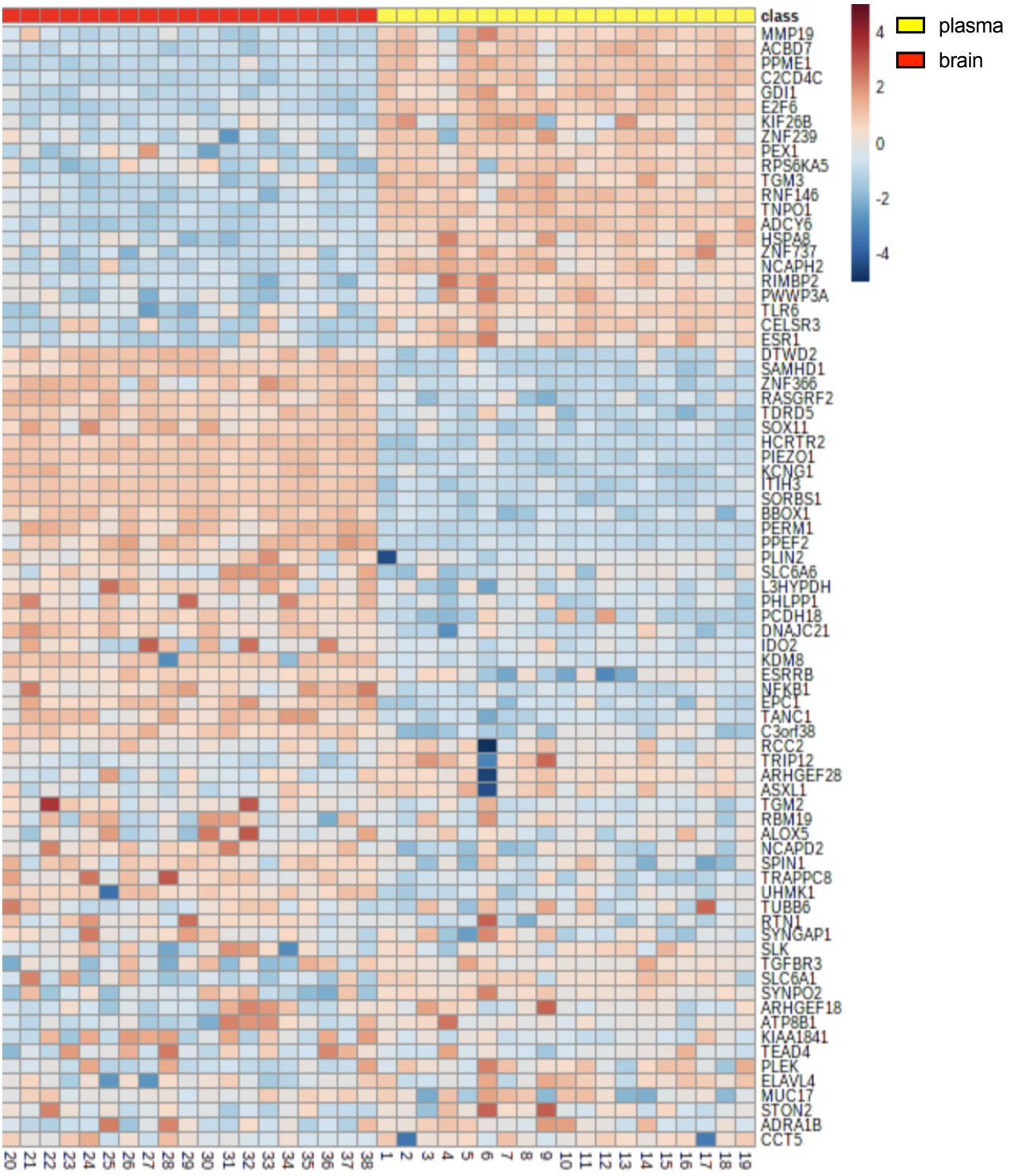


Figure 12. Protein expression heat-map of proteins unique to plasma and brain. Depicting all 77 proteins amongst plasma and brain; 22 and 28 proteins significantly increased and decreased in plasma, respectively (t-test).

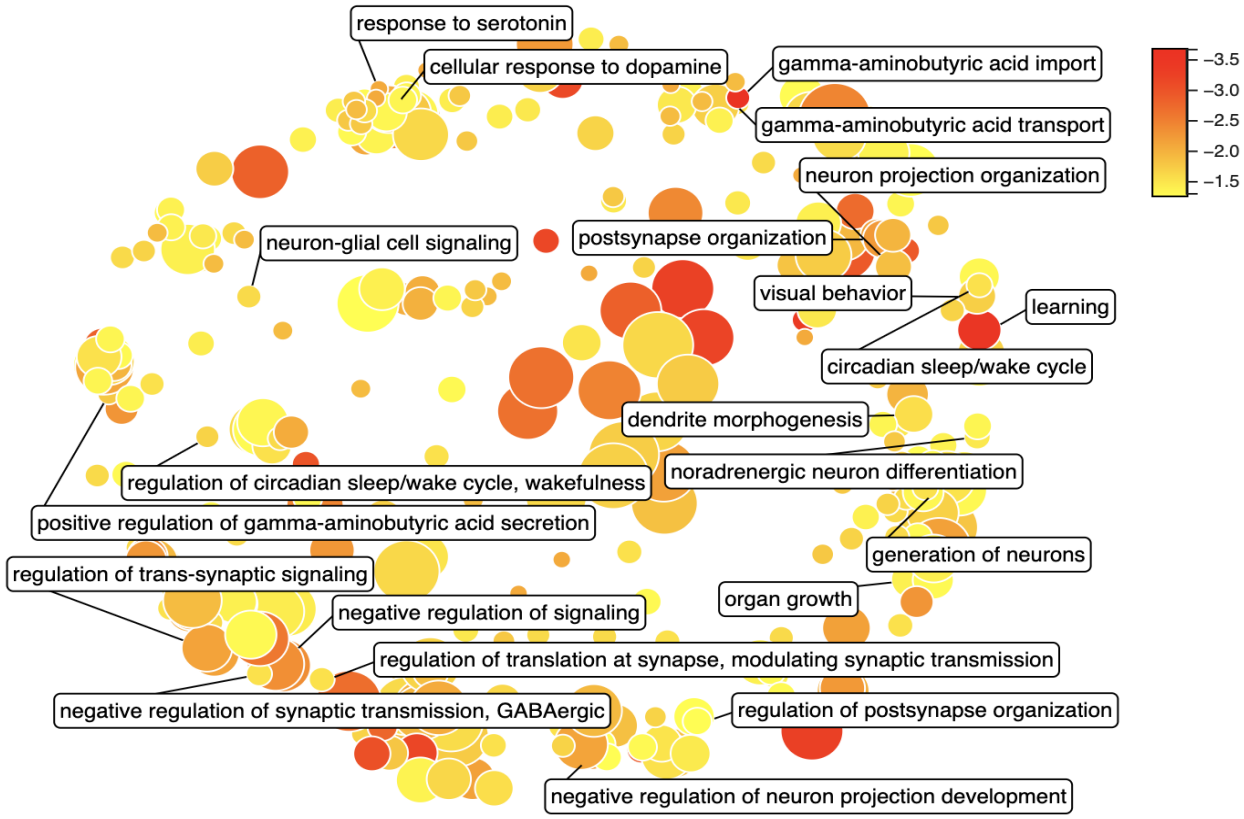


Figure 13. Figure. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and brain. Results intentionally highlight brain related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values were calculated using the Gene Ontology Panther Tool. Data only includes gene ontology IDs with $p < 0.05$.

Table 2. Function and expression of proteins uniquely detected in plasma and brain. Protein list based on proteins that were significantly increased (red font) or decreased (blue font) in plasma when compared to the brain.

HUMAN PROTEIN SYMBOL ORTHOLOGS[1]	FUNCTION[2]	EXPRESSION[2]
PPME1 Protein Phosphatase Methylesterase 1	demethylates proteins	widely expressed
TNPO1 Transportin 1	nuclear transport receptor	widely expressed
E2F6 E2F Transcription Factor 6	inhibits E2F-dependent transcription	widely expressed
GDI1 GDP Dissociation Inhibitor 1	regulates GDP/GTP exchange	widely expressed; highest levels in the brain
ACBD7 Acyl-CoA Binding Domain Containing 7	binds medium and long acyl-CoA esters	predominantly expressed in the nervous system
HSPA8 Heat Shock Protein Family A (Hsp70) Member 8	molecular chaperon	widely expressed
ADCY6 Adenylate Cyclase 6	catalyzes the formation of cAMP	peripheral blood mononuclear leukocytes
CELSR3 Cadherin EGF LAG Seven-Pass G-Type Receptor 3	cell signaling during nervous system formation	widely expressed
MMP19 Matrix Metallopeptidase 19	endopeptidase	widely expressed; not detected in brain or blood
RNF146 Ring Finger Protein 146	E3 ubiquitin-protein ligase	widely expressed; high levels in brain of Alzheimer patients
ESR1 Estrogen Receptor 1	nuclear hormone receptor	widely expressed
RIMBP2 RIMS Binding Protein 2	role in synaptic transmission	brain
TLR6 Toll Like Receptor 6	innate immune response to gram-positive bacteria/fungi	blood and immune, stomach
NCAPH2 Non-SMC Condensin II Complex Subunit H2	regulatory subunit of the condensin-2 complex	widely expressed

RPS6KA5 Ribosomal Protein S6 Kinase A5	serine/threonine-protein kinase	widely expressed
TGM3 Transglutaminase 3	catalytic activity	widely expressed
PWWP3A PWWP Domain Containing 3A, DNA Repair Factor	DNA damage response pathway	N/A
TRIP12 Thyroid Hormone Receptor Interactor 12	E3 ubiquitin-protein ligase	blood and lymph nodes
PEX1 Peroxisomal Biogenesis Factor 1	protein import into peroxisome	widely expressed
KIF26B Kinesin Family Member 26B	essential for embryonic kidney development	widely expressed
PHLPP1 PH Domain and Leucine Rich Repeat Protein Phosphatase 1	protein phosphatase	widely expressed
ESRRB Estrogen Related Receptor Beta	transcription factor	brain, liver, gallbladder
PLIN2 Perilipin 2	development and maintenance of adipose tissue	widely expressed
TDRD5 Tudor Domain Containing 5	required during spermiogenesis	widely expressed
L3HYPDH Trans-L-3-Hydroxyproline Dehydratase	catalytic activity	widely expressed
IDO2 Indoleamine 2,3-Dioxygenase 2	catalytic activity	widely expressed
SOX11 SRY-Box Transcription Factor 11	transcription factor	predominantly expressed in brain and heart
EPC1 Enhancer Of Polycomb Homolog 1	component of the NuA4 histone acetyltransferase complex	blood and heart

NCAPD2 Non-SMC Condensin I Complex Subunit D2	regulatory subunit of the condensin complex	widely expressed
UHMK1 U2AF Homology Motif Kinase 1	phosphorylation	widely expressed at the mRNA level
KCNG1 Potassium Voltage-Gated Channel Modifier Subfamily G Member 1	potassium channel subunit	predominantly expressed in brain and placenta
PPEF2 Protein Phosphatase With EF-Hand Domain	phototransduction	retina
ZNF366 Zinc Finger Protein 366	transcriptional repression activity	t-lymphocyte, heart, adipocyte
RASGRF2 Ras Protein Specific Guanine Nucleotide Releasing Factor 2	calcium-regulated nucleotide exchange factor	widely expressed; predominantly expressed in brain and plasma
TANC1 Tetratricopeptide Repeat, Ankyrin Repeat and Coiled-Coil Containing 1	scaffold component in postsynaptic density	widely expressed
NFKB1 Nuclear Factor Kappa B Subunit 1	pleiotropic transcription factor	widely expressed
TRAPPC8 Trafficking Protein Particle Complex Subunit 8	endoplasmic reticulum to Golgi apparatus trafficking	widely expressed
KDM8 Lysine Demethylase 8	enzymatic activity	blood, heart, pancreas, prostate, ovaries
PIEZO1 Piezo Type Mechanosensitive Ion Channel Component 1	pore-forming subunit of non-specific cation-channels	widely expressed; link to Alzheimer and Parkinson disease
ITIH3 Inter-Alpha-Trypsin Inhibitor Heavy Chain 3	hyaluronan carrier or binding protein	widely expressed
DTWD2 DTW Domain Containing 2	catalytic activity	widely expressed
PERM1 PPARGC1 And ESRR Induced Regulator, Muscle 1	glucose and lipid metabolism	heart, muscle

BBOX1 Gamma-Butyrobetaine Hydroxylase 1	catalytic activity	predominantly expressed in kidney, liver and brain
C3orf38 Chromosome 3 Open Reading Frame 38	apoptosis regulation	widely expressed
SAMHD1 SAM And HD Domain Containing Deoxynucleoside Triphosphate Triphosphohydrolase 1	defense response, DNA end resection at stalled replication forks	widely expressed
SORBS1 Sorbin And SH3 Domain Containing 1	tyrosine phosphorylation	widely expressed

Note: 46 out of 50 significant proteins listed. Proteins removed if information on function and expression were incomplete.

[1]Human protein symbol ortholog: the list of identified rainbow trout protein accession numbers was blasted against the Human Uniprot Reference Proteome (Proteome ID#UP000005640) to obtain human protein orthologs based on sequence similarity.

[2]Function and expression annotations were interpreted from Gene Cards and Uniprot and are specific to humans; 'widely expressed' refers to proteins that have been detected ubiquitously, or in numerous tissues or cell types.

4.3.2. Plasma versus heart proteome

The proceeding analysis compared the 130 proteins that were only shared amongst the plasma and heart proteomes. Although all 130 proteins were found across heart and plasma, PCA analysis showed that protein abundance levels differed and still resulted in significantly different sub-proteomes (Figure 14). Of the 130 proteins, 49 proteins were significantly increased in plasma, and 48 proteins were significantly decreased in plasma when compared to heart (FDR corrected p-value < 0.05; Figure 15). The function and expression annotations for the preceding significant proteins are summarized in Table 2. Gene Card and Uniprot tissue-expression data showed that 18 out the 97 significant proteins associated with heart and plasma are predominantly expressed in the human heart: PDE4B, NTN4, CTDSPL2, ADGRF3, MYO10, MYBPC3, SHH, FRRS1, ZNF423, LRP4, SLC2A4, PNKP, TRIM27, HES1, NYAP1, SETBP1, ZNF280D, and MYO19. The functions of the preceding 18 proteins include enzymatic activity, development, receptor and transport, transcription, and DNA damage repair (Table 2). Gene ontology analysis of the 130 proteins associated with plasma and heart contained many biological processes specific to heart function, most of which could be summarized into (1) muscle activity, and (2) heart and vasculature development (Figure 16).

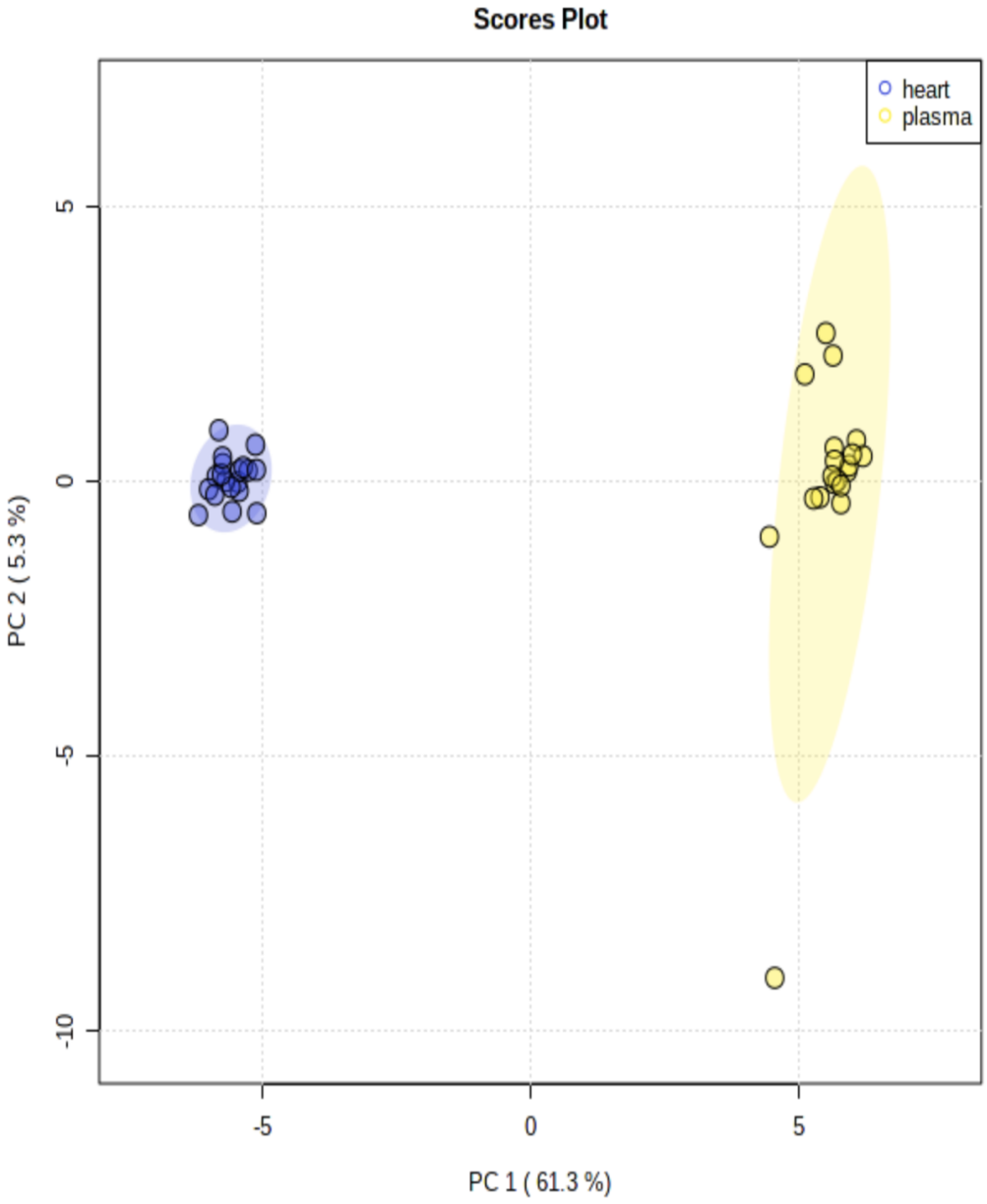


Figure 14. Principal component analysis comparing proteins only found in plasma and heart. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

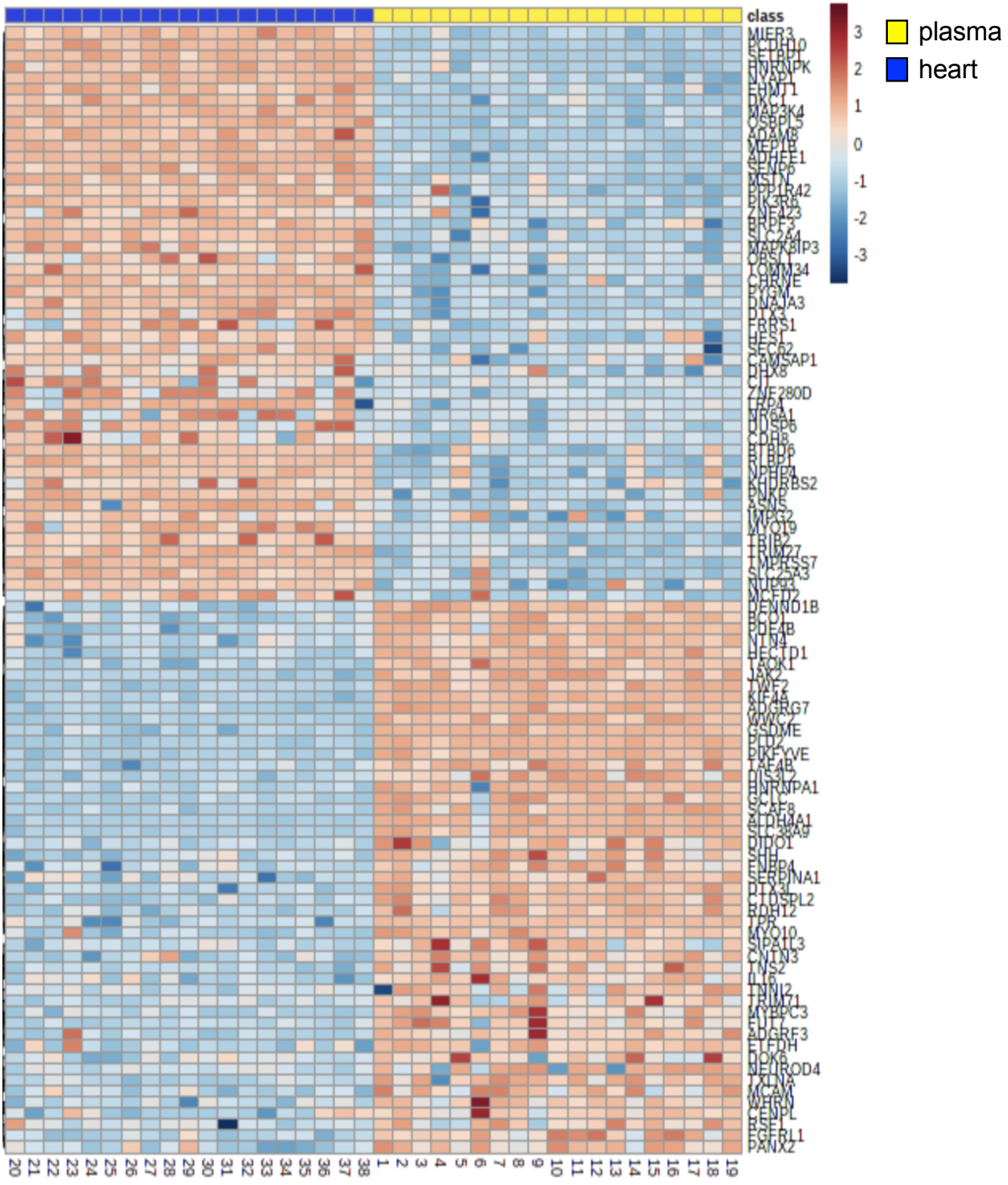


Figure 15. Protein expression heat-map of proteins unique to plasma and heart. Depicting top 100 proteins amongst plasma and heart; 49 and 48 proteins significantly increased and decreased in plasma, respectively (t-test).

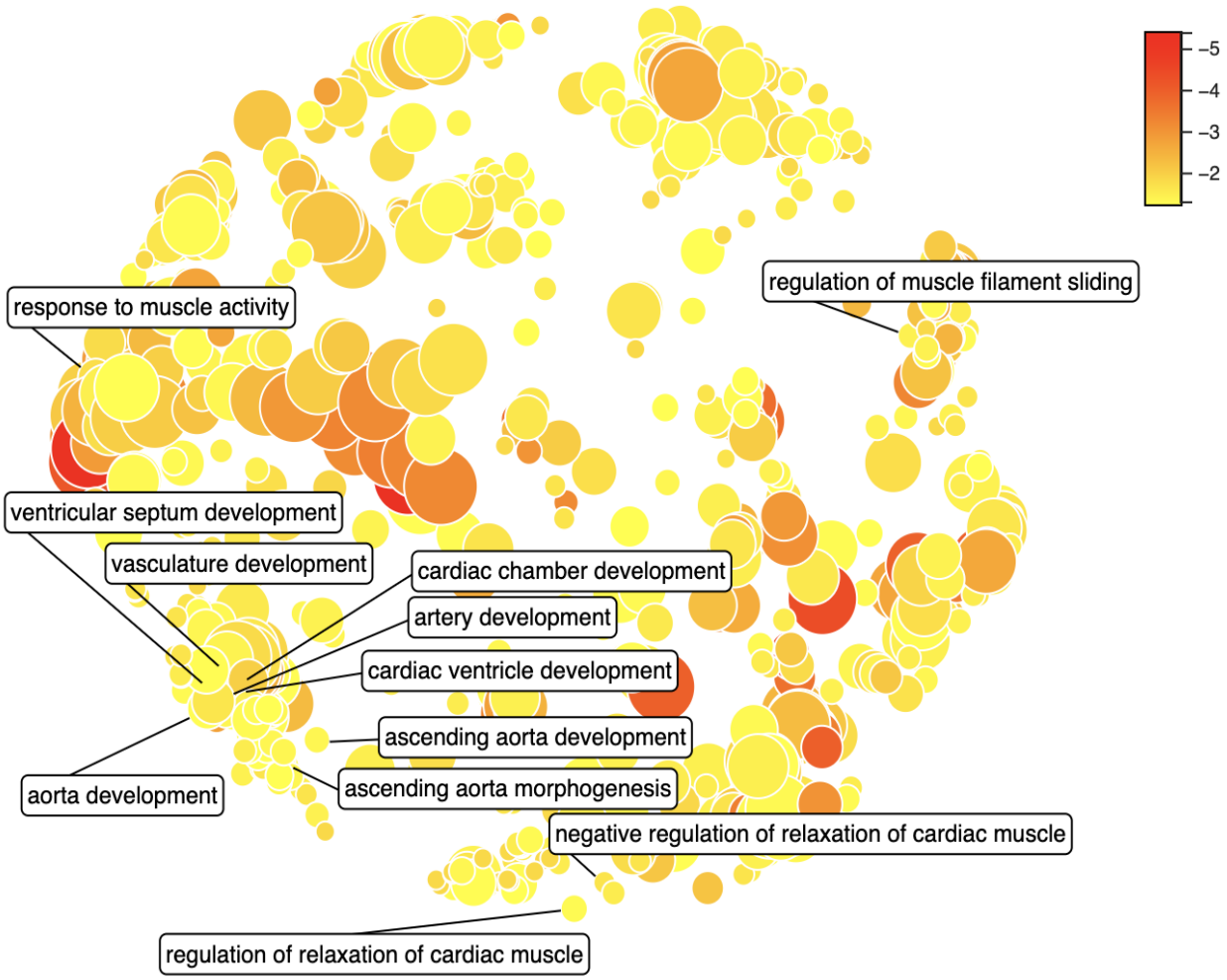


Figure 16. Figure. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and heart. Results intentionally highlight heart related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Table 3. Function and expression of proteins uniquely detected in plasma and heart. Protein list based on proteins that were significantly increased (red font) or decreased (blue font) in plasma when compared to heart.

HUMAN PROTEIN SYMBOL ORTHOLOGS[1]	FUNCTION[2]	EXPRESSION[2]
JAK2 Janus Kinase 2	non-receptor tyrosine kinase	widely expressed
GSDME Gasdermin E	precursor of a pore-forming protein	widely expressed
PLD2 Phospholipase D2	phospholipase for phosphatidylcholine	widely expressed
ALDH4A1 Aldehyde Dehydrogenase 4 Family Member A1	irreversible conversion of P5C to glutamate	widely expressed
ADGRG7 Adhesion G Protein-Coupled Receptor G7	orphan receptor	liver
TPR Translocated Promoter Region Protein	component of the nuclear pore complex	widely expressed
SLC38A9 Solute Carrier Family 38 Member 9	lysosomal amino acid transporter	placenta
TAF4B TATA-Box Binding Protein Associated Factor 4b	transcription factor	blood, colon muscle, gallbladder, testis
NEUROD4 Neuronal Differentiation 4	mediates neuronal differentiation	serum and plasma
DENND1B DENN Domain Containing 1B	guanine nucleotide exchange factor	immune cells, pancreas
TNS2 Tensin 2	tyrosine-protein phosphatase	widely expressed
HECTD1 HECT Domain E3 Ubiquitin Protein Ligase 1	E3 ubiquitin-protein ligase	widely expressed
PIKFYVE Phosphoinositide Kinase, FYVE-Type Zinc Finger Containing	dual specificity kinase	widely expressed
TXLNA Taxilin Alpha	intracellular vesicle traffic	widely expressed

KIF4A Kinesin Family Member 4A	iron-sulfur cluster binding motor protein	widely expressed
SCAF8 SR-Related CTD Associated Factor 8	anti-terminator protein required to prevent early mRNA termination during transcription	widely expressed
HNRNPA1 Heterogeneous Nuclear Ribonucleoprotein A1	involved in the packaging of pre-mRNA into hnRNP particles	widely expressed
PDE4B Phosphodiesterase 4B	hydrolyzes cAMP	brain, heart, lung, skeletal muscle
TAOK1 TAO Kinase 1	serine/threonine-protein kinase	widely expressed
BCO1 Beta-Carotene Oxygenase 1	involved in vitamin A metabolism	intestine, eye, liver
TNNI2 Troponin I2, Fast Skeletal Type	inhibitory subunit of troponin	tonsil, brain, liver, lung, testis
DTX3L Deltex E3 Ubiquitin Ligase 3L	E3 ubiquitin-protein ligase	widely expressed
MCAM Melanoma Cell Adhesion Molecule	plays a role in cell adhesion	widely expressed
TCP1 T-Complex 1	assists the folding of proteins upon ATP hydrolysis	widely expressed
NTN4 Netrin-4	neural, kidney, and vascular development	heart, kidney, lung, secretory organs, testis
WHRN Whirlin	hearing and vision	N/A
CTDSPL2 CTD Small Phosphatase Like 2	probable phosphatase	blood, retina, heart, gut, secretory and reproductive organs
GCLC Glutamate-Cysteine Ligase Catalytic Subunit	glutamate-cysteine ligase	widely expressed

TRIM71 Tripartite Motif Containing 71	E3 ubiquitin-protein ligase	liver, pancreatic juice, reproductive organs
TWF2 Twinfilin Actin Binding Protein 2	actin-binding protein	widely expressed
FUT7 Fucosyltransferase 7	catalyzes L-fucose	spleen, adipocyte
ADGRF3 Adhesion G Protein-Coupled Receptor F3	orphan receptor	plasma, platelet and heart
RDH12 Retinol Dehydrogenase 12	dehydrogenase/reductase with a preference for NADP	predominantly in the retina
MYO10 Myosin X	actin-based motor molecule	plasma, retina, heart, lung, secretory and reproductive organs
MYBPC3 Myosin Binding Protein C3	thick filament-associated protein in striated muscles	predominantly in the heart
SHH Sonic Hedgehog Signaling Molecule	early embryo development patterning	blood, heart, liver, placenta
SIPA1L3 Signal Induced Proliferation Associated 1 Like 3	epithelial cell morphogenesis, polarity, adhesion and cytoskeleton organization	blood, retina, liver, adipocyte, secretory organs
CENPL Centromere Protein L	assembly of kinetochore proteins	nasal respiratory epithelium
FNBP4 Formin Binding Protein 4	role in regulating cytoskeletal dynamics during cell division and migration	widely expressed
FGFRL1 Fibroblast Growth Factor Receptor Like 1	negative effect on cell proliferation	widely expressed
IL16 Interleukin 16	immune response	predominantly in blood and immune cells
ETFDH Electron Transfer Flavoprotein Dehydrogenase	essential for electron transfer	widely expressed

DOK6 Docking Protein 6	intracellular adaptor	blood, brain, kidney, spinal cord, reproductive organs
DIDO1 Death Inducer-Obliterator 1	putative transcription factor	widely expressed
RSF1 Remodeling And Spacing Factor 1	nuclear protein	widely expressed
KIF21A Kinesin Family Member 21A	microtubule-binding motor protein	widely expressed
CNTN3 Contactin 3	mediates cell surface interactions	widely expressed; predominantly in serum and plasma
CDH8 Cadherin 8	calcium-dependent cell adhesion proteins	predominantly in brain, placenta, and testis
PPP1R42 Protein Phosphatase 1 Regulatory Subunit 42	plays a role in the control of centrosome integrity	testis
NUP93 Nucleoporin 93	plays a role in nuclear pore complex assembly and maintenance	widely expressed
SEC62 SEC62 Homolog, Preprotein Translocation Factor	mediates post-translational transport of precursor polypeptides across ER	widely expressed
MCFD2 Multiple Coagulation Factor Deficiency 2, ER Cargo Receptor Complex Subunit	plays a role in the secretion of coagulation factors	widely expressed
KHDRBS2 KH RNA Binding Domain Containing, Signal Transduction Associated 2	RNA-binding protein, regulation of splicing	predominantly in blood and brain
SPI1 Spi-1 Proto-Oncogene	transcriptional activator	predominantly in blood and immune cells
CHRNE Cholinergic Receptor Nicotinic Epsilon Subunit	acetylcholine receptor	monocyte, b-lymphocyte, gut
FRRS1 Ferric Chelate Reductase 1	reduces ferric to ferrous iron	plasma, heart, secretory and reproductive organs

CIT Citron Rho-Interacting Serine/Threonine Kinase	plays a role in cytokinesis	widely expressed
EHMT1 Euchromatic Histone Lysine Methyltransferase 1	histone methyltransferase	widely expressed
ZNF423 Zinc Finger Protein 423	transcription factor	widely expressed; predominantly in bone marrow, brain, and heart
BRPF3 Bromodomain And PHD Finger Containing 3	scaffold subunit	N/A
MAPK8IP3 Mitogen-Activated Protein Kinase 8 Interacting Protein 3	scaffold proteins selectively mediates JNK signaling	widely expressed
SLC25A3 Solute Carrier Family 25 Member 3	involved in phosphate transport	widely expressed
DTX3 Deltex E3 Ubiquitin Ligase 3	E3 ubiquitin ligase	widely expressed
LRP4 LDL Receptor Related Protein	low-density lipoprotein receptor-related proteins	plasma, heart, bone, brain
SLC2A4 Solute Carrier Family 2 Member 4	glucose transporter	widely expressed; predominantly in heart and adipocytes
NPHP4 Nephrocystin 4	involved in the organization of apical junctions	widely expressed
NR6A1 Nuclear Receptor Subfamily 6 Group A Member 1	orphan nuclear receptor	platelet, stomach, and testis
PIK3R6 Phosphoinositide-3-Kinase Regulatory Subunit 6	regulatory subunit of the PI3K gamma complex	blood, liver
HNRNPK Heterogeneous Nuclear Ribonucleoprotein K	RNA-binding protein	widely expressed
MEP1B Meprin A Subunit Beta	membrane metallopeptidase	pancreatic juice, cervix
PNKP Polynucleotide Kinase 3'-Phosphatase	DNA damage repair	widely expressed; predominantly in heart and blood

DNAJA3 DnaJ Heat Shock Protein Family (Hsp40) Member A3	apoptotic signal transduction	widely expressed
TRIM27 Tripartite Motif Containing 27	E3 ubiquitin-protein ligase, transcriptional repressor activity	immune cells, heart, liver, secretory and reproductive organs
PYGM Glycogen Phosphorylase, Muscle Associated	carbohydrate metabolism	widely expressed; predominantly in blood
HES1 Hes Family BHLH Transcription Factor 1	transcriptional repressor	B-lymphocytes, heart
ADAM8 ADAM Metallopeptidase Domain 8	cell-cell and cell-matrix interactions	blood and immune cells
DUSP6 Dual Specificity Phosphatase 6	inactivate kinases	pancreas, keratinocytes
NYAP1 Neuronal Tyrosine Phosphorylated Phosphoinositide-3-Kinase Adaptor 1	activates PI3K	brain, heart, lung, breast, pancreas
TMPRSS7 Transmembrane Serine Protease 7	serine protease	brain, ovary, lung, pancreas
RLBP1 Retinaldehyde Binding Protein 1	soluble retinoid carrier	brain, retina
DKC1 Dyskerin Pseudouridine Synthase 1	ribosome biogenesis and telomere maintenance	widely expressed
SENP6 SUMO Specific Peptidase 6	protease	blood and immune cells, brain, placenta, reproductive organs
BTBD6 BTB Domain Containing 6	adapter protein for E3 ubiquitin-protein ligase complex	lens
SETBP1 SET Binding Protein 1	protein containing several structural motifs	plasma, cerebrospinal fluid, heart, adipocyte
PCDH10 Protocadherin 10	calcium-dependent cell adhesion protein	predominantly in brain

TOMM34 Translocase Of Outer Mitochondrial Membrane 34	import precursor proteins into mitochondria	widely expressed
ZNF280D Zinc Finger Protein 280D	transcription factor	plasma, B-lymphocyte, brain, heart, pancreas, placenta, testis
ADHFE1 Alcohol Dehydrogenase Iron Containing 1	catalytic activity	widely expressed
TRIB2 Tribbles Pseudokinase 2	interacts with MAPK kinases	peripheral blood leukocytes
OSBPL5 Oxysterol Binding Protein Like 5	lipid transporter	widely expressed
MSTN Myostatin	negative regulation of skeletal muscle cell proliferation	vitreous humor
MAP3K4 Mitogen-Activated Protein Kinase Kinase Kinase 4	phosphorylates and activates MAP2K4 and MAP2K6	widely expressed
ASNS Asparagine Synthetase	involved in asparagine synthesis	widely expressed
MIER3 MIER Family Member 3	transcriptional repressor	B-lymphocytes, frontal cortex, placenta, reproductive organs
MYO19 Myosin XIX	actin-based motor molecule	heart, lung, pancreas, placenta

Note: 95 out of 97 significant proteins listed. Proteins removed if information on function and expression were insufficient.

[1]Human protein symbol ortholog: the list of identified rainbow trout protein accession numbers was blasted against the Human Uniprot Reference Proteome (Proteome ID#UP000005640) to obtain human protein orthologs based on sequence similarity.

[2]Function and expression annotations were interpreted from Gene Cards and Uniprot and are specific to humans; 'widely expressed' refers to proteins that have been detected ubiquitously, or in numerous tissues or cell types.

4.3.3. Plasma versus gill proteome

The proceeding analysis compared the 141 proteins that were only shared amongst the plasma and gill proteomes. Although all 141 proteins were found across gill and plasma, PCA analysis showed that protein abundance levels differed and still resulted in significantly different sub-proteomes (Figure 17). Of the 141 proteins, 38 proteins were significantly increased in plasma, and 34 proteins were significantly decreased in plasma when compared to gill (FDR corrected p-value < 0.05; Figure 18). The function and expression annotations for the preceding significant proteins are summarized in Table 3. Since gills are not found in humans, we opted to data mine tissue-specific information (Gene Card and Uniprot) on proteins found in lungs considering gill and lung function are most comparable between fish and humans. Of note, 3 out of the 72 significant proteins are only or predominantly expressed in the lungs: BNIPL, KIF23, and PTGER4. The functions of the preceding 3 proteins include bridge molecule, cytokinesis, and receptor activity (Table 3). Gene ontology analysis of the 141 proteins associated with plasma and gills contained many biological processes that could be related to gill function, most of which could be summarized into (1) response mechanism, (2) immunity, (3) xenobiotic transport, and (4) cartilage development (Figure 19).

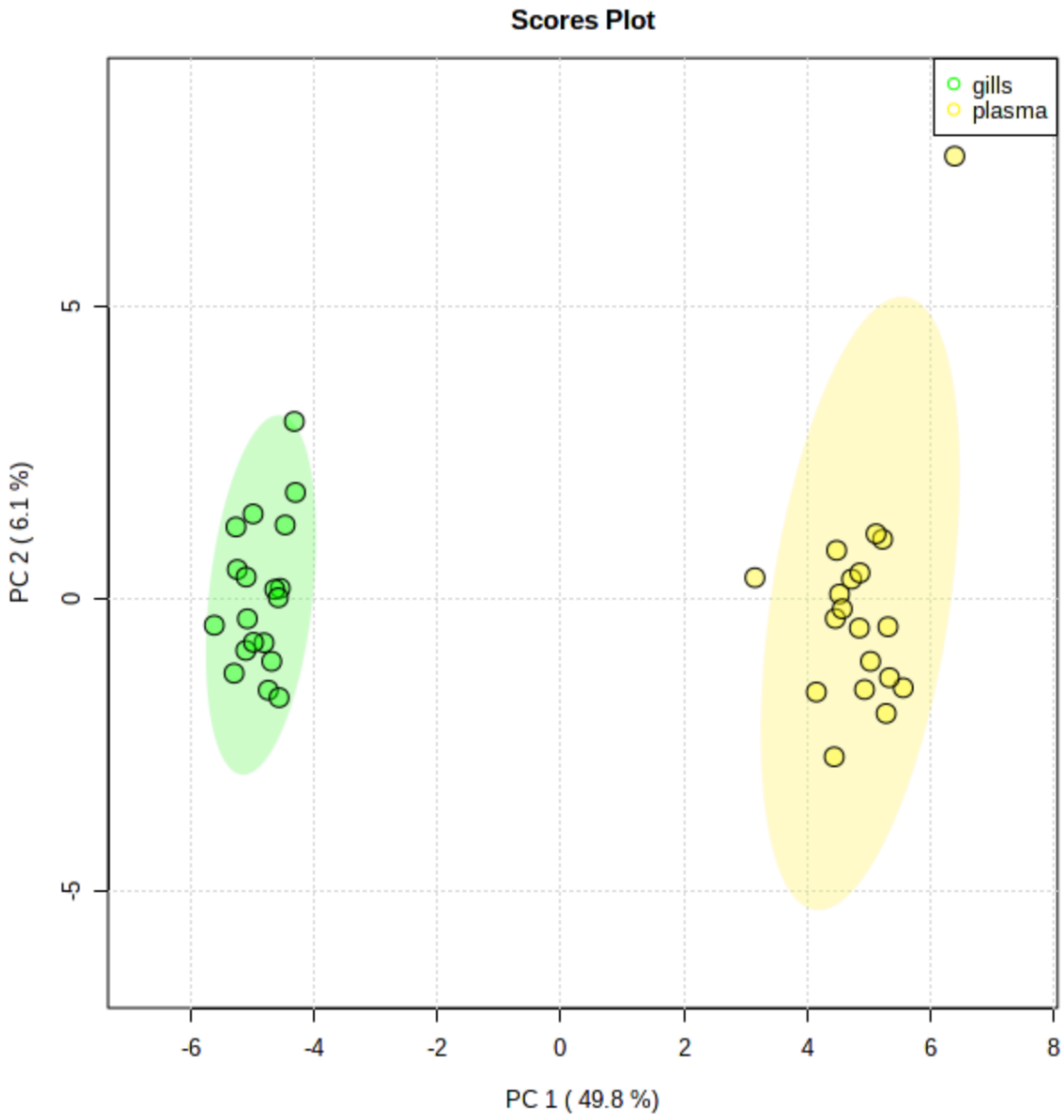


Figure 17. Principal component analysis comparing proteins only found in plasma and gill. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

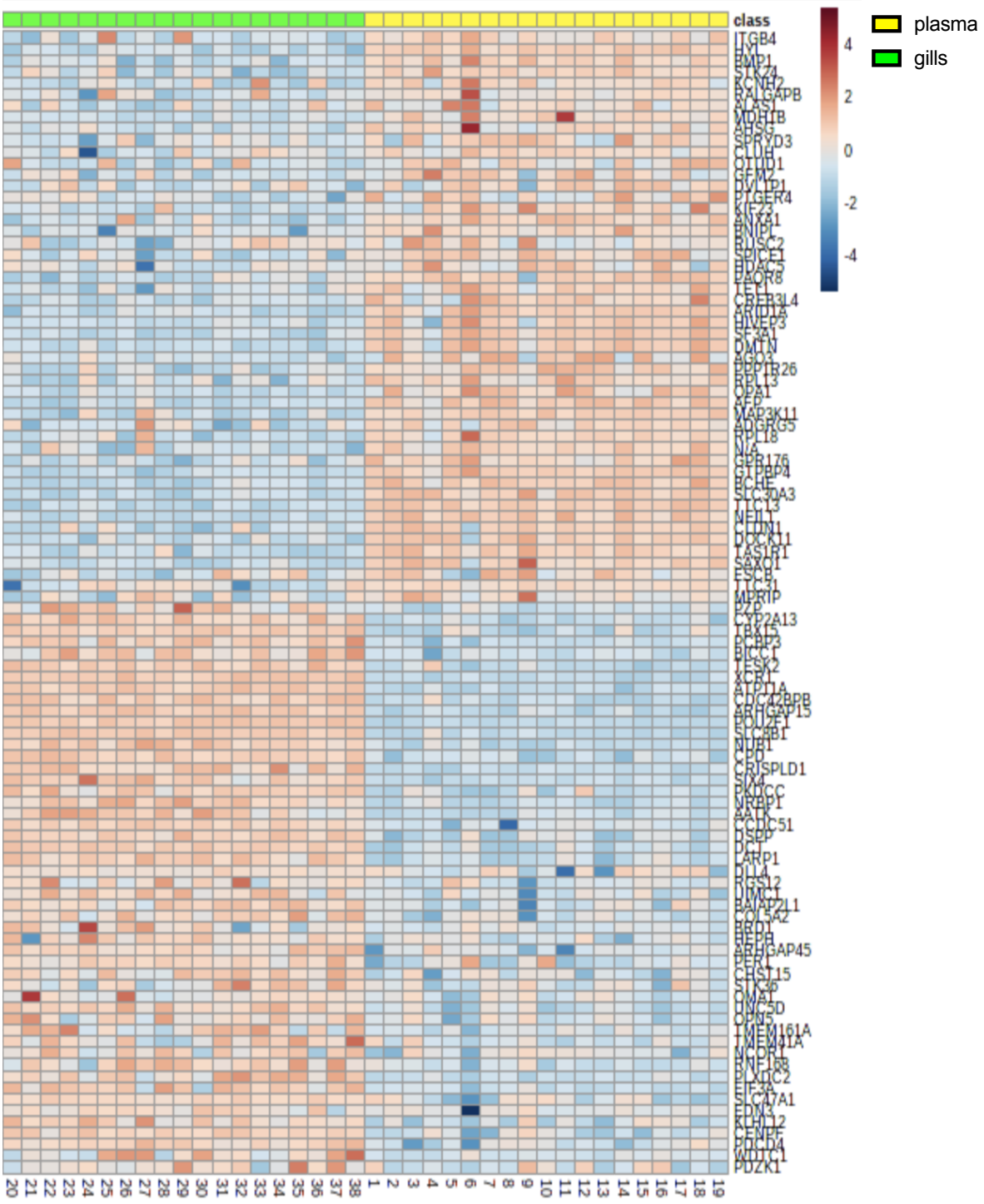


Figure 18. Protein expression heat-map of proteins unique to plasma and gill. Depicting top 100 proteins amongst plasma and gill; 38 and 34 proteins significantly increased and decreased in plasma, respectively (t-test).

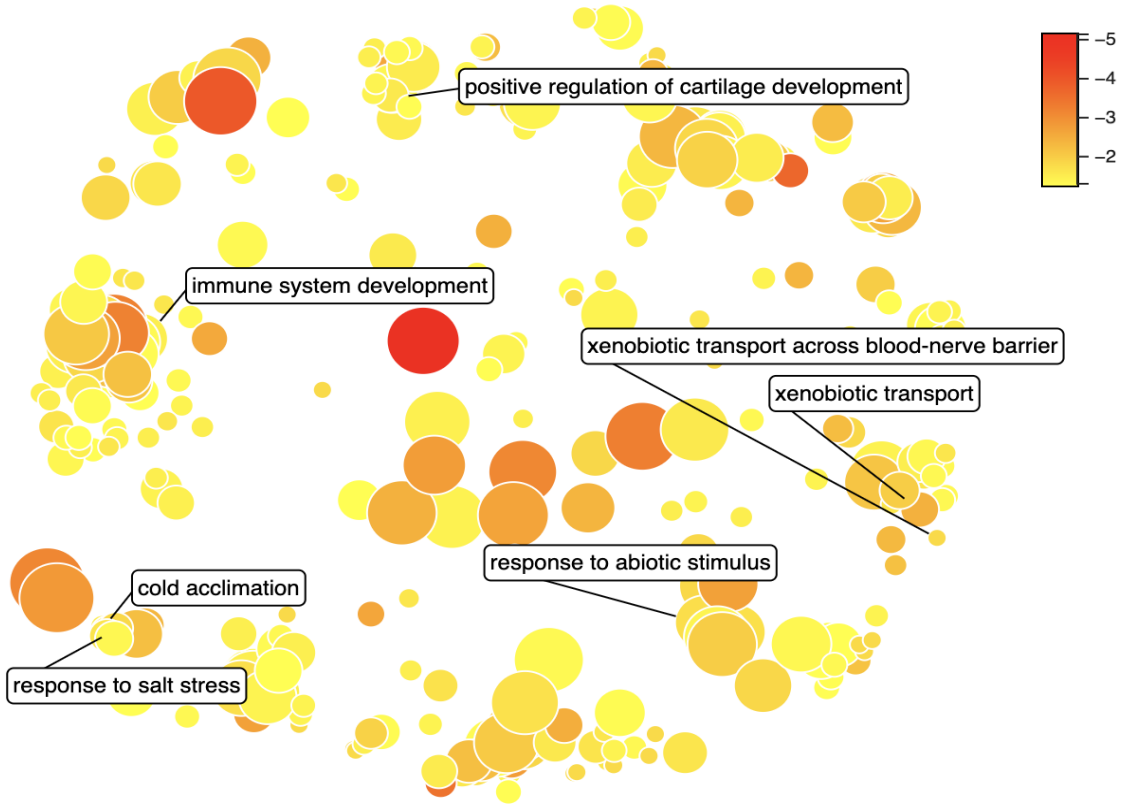


Figure 19. Figure. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and gill. Results intentionally highlight gill related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Table 4. Function and expression of proteins uniquely detected in plasma and gill. Protein list based on proteins that were significantly increased (red font) or decreased (blue font) in plasma when compared to gill.

HUMAN PROTEIN SYMBOL ORTHOLOGS[1]	FUNCTION[2]	EXPRESSION[2]
HYI Putative Hydroxypyruvate Isomerase	carbohydrate transport and metabolism	widely expressed
OPA1 OPA1 Mitochondrial Dynamin Like GTPase	dynamin-related GTPase	widely expressed
ARID1A AT-Rich Interaction Domain 1A	transcription of certain genes by altering chromatin structure	widely expressed
GTPBP4 GTP Binding Protein 4	GTP binding protein involved in 60S ribosomal subunit biogenesis	widely expressed
NEIL1 Nei Like DNA Glycosylase 1	DNA damage repair	blood and heart
SF3A1 Splicing Factor 3a Subunit 1	pre-mRNA splicing	widely expressed
DMTN Dematin Actin Binding Protein	structural role in erythrocytes	widely expressed
CLDN1 Claudin 1	major constituents of tight junctions	widely expressed
RPL13 Ribosomal Protein L13	60S subunit of the ribosomal protein	widely expressed
BMP1 Bone Morphogenetic Protein 1	induces formation of cartilage in vivo	widely expressed
CREB3L4 CAMP Responsive Element Binding Protein 3 Like 4	transcriptional activator, role in the unfolded protein response	widely expressed
BCHE Butyrylcholinesterase	cholinesterase	widely expressed
RPL18 Ribosomal Protein L18	60S subunit of the ribosomal protein	widely expressed
AGO3 Argonaute RISC Catalytic Component 3	RNA-mediated gene silencing	widely expressed

DOCK11 Dedicator Of Cytokinesis 11	activates small G proteins	widely expressed
TAS1R1 Taste 1 Receptor Member 1	responds to umami taste	heart
SLC30A3 Solute Carrier Family 30 Member 3	involved in accumulation of zinc in synaptic vesicles	brain, bone, testis
AFP Alpha Fetoprotein	major plasma protein, binding	widely expressed
PAQR8 Progestin And AdipoQ Receptor Family Member 8	binding activity	brain, spinal cord, kidney, prostate, testis
BNIPL BCL2 Interacting Protein Like	bridge molecule	brain, breast, urine, placenta, lung
SAXO1 Stabilizer Of Axonemal Microtubules 1	microtubule binding	widely expressed
BCL7B BAF Chromatin Remodeling Complex Subunit BCL7B	positive regulator of apoptosis	widely expressed
HIVEP3 HIVEP Zinc Finger 3	transcription factor	predominantly in blood and immune cells
STK24 Serine/Threonine Kinase 24	serine/threonine protein kinase	widely expressed
MAP3K11 Mitogen-Activated Protein Kinase Kinase Kinase 11	serine/threonine protein kinase	widely expressed
KIF23 Kinesin Family Member 23	involved in cell cycle cytokinesis	brain, liver, lung, reproductive organs
TET1 Tet Methylcytosine Dioxygenase 1	DNA demethylation	widely expressed
PTGER4 Prostaglandin E Receptor 4	receptor activity	immune cells, lung
SPICE1 Spindle And Centriole Associated Protein 1	regulator involved in mitosis	widely expressed
ANXA1 Annexin A1	innate immune response	widely expressed
PPP1R26 Protein Phosphatase 1 Regulatory Subunit 26	inhibits phosphatase activity	widely expressed

ALAS1 5'-Aminolevulinate Synthase 1	involved in heme biosynthesis	liver, reproductive organs
AHSG Alpha 2-HS Glycoprotein	serum glycoprotein	widely expressed
HEPH Hephaestin	iron and copper transport	widely expressed
BAIAP2L1 BAR/IMD Domain Containing Adaptor Protein 2 Like 1	actin cytoskeleton remodeling	stomach, liver, secretory and reproductive organs
CENPF Centromere Protein F	involved in cell division and proliferation	widely expressed
NCOR1 Nuclear Receptor Corepressor 1	transcriptional activity	widely expressed
PZP Pregnancy Zone Protein	inhibits proteinase activity	widely expressed
UNC5D Unc-5 Netrin Receptor D	cell-cell adhesion and cell guidance	brain, liver, placenta
PKDCC Protein Kinase Domain Containing, Cytoplasmic	tyrosine-protein kinase activity	heart, liver
BICC1 BicC Family RNA Binding Protein 1	RNA-binding protein	plasma, monocyte, pancreas
SLC47A1 Solute Carrier Family 47 Member 1	solute transporter	widely expressed
EIF3A Eukaryotic Translation Initiation Factor 3 Subunit A	RNA-binding component	widely expressed
PLXDC2 Plexin Domain Containing 2	role in tumor angiogenesis	widely expressed
WDTC1 WD And Tetratricopeptide Repeats 1	substrate receptor	widely expressed
DSPP Dentin Sialophosphoprotein	calcium ion and collagen binding	expressed in teeth
PDCD4 Programmed Cell Death 4	inhibits translation	widely expressed
PCBP3 Poly(RC) Binding Protein 3	RNA-binding protein	widely expressed

AATK Apoptosis Associated Tyrosine Kinase	tyrosine-protein kinase activity	brain
NRBP1 Nuclear Receptor Binding Protein 1	binding activity	widely expressed
CCDC51 Coiled-Coil Domain Containing 51	potassium channel activity	widely expressed
TESK2 Testis Associated Actin Remodelling Kinase 2	serine/threonine protein kinase	B-lymphocyte, prostate, placenta, testis
LARP1 La Ribonucleoprotein 1, Translational Regulator	RNA-binding protein	widely expressed
ARHGAP15 Rho GTPase Activating Protein 15	regulation of GTPase activity	predominantly in blood and immune cells
NUB1 Negative Regulator of Ubiquitin Like Proteins 1	protein ubiquitination	widely expressed
ATP11A ATPase Phospholipid Transporting 11A	binding activity	widely expressed
CDC42BPB CDC42 Binding Protein Kinase Beta	serine/threonine protein kinase	widely expressed
CYP2A13 Cytochrome P450 Family 2 Subfamily A Member 13	enzymatic activity	widely expressed
SLC8B1 Solute Carrier Family 8 Member B1	sodium/calcium antiporter	pancreas, adrenal, placenta, testis
POU2F1 POU Class 2 Homeobox 1	transcription factor	widely expressed

Note: 60 out of 72 significant proteins listed. Proteins removed if information on function and expression were insufficient.
[1]Human protein symbol ortholog: the list of identified rainbow trout protein accession numbers was blasted against the Human Uniprot Reference Proteome (Proteome ID#UP000005640) to obtain human protein orthologs based on sequence similarity.
[2]Function and expression annotations were interpreted from Gene Cards and Uniprot and are specific to humans; 'widely expressed' refers to proteins that have been detected ubiquitously, or in numerous tissues or cell types.

4.3.4. Plasma versus kidney proteome

The proceeding analysis compared the 98 proteins that were only shared amongst the plasma and kidney proteomes. Although all 98 proteins were found across kidney and plasma, PCA analysis showed that protein abundance levels differed and still resulted in significantly different sub-proteomes (Figure 20). Of the 98 proteins, 22 proteins were significantly increased in plasma, and 30 proteins were significantly decreased in plasma when compared to kidney (FDR corrected p-value < 0.05; Figure 21). The function and expression annotations for the preceding significant proteins are summarized in Table 4. Gene Card and Uniprot tissue-expression data showed that only 1 out the 52 significant proteins associated with kidney and plasma are predominantly expressed in the human kidney: CYP27B1, which plays a role in vitamin D metabolism (Table 4). Gene ontology analysis of the 98 proteins associated with plasma and kidney contained many biological processes specific to kidney function, most of which could be summarized into (1) regulation of vitamin D, (2) nitrogen and amino acid metabolism, (3) aldosterone secretion, and (4) urine volume (Figure 22).

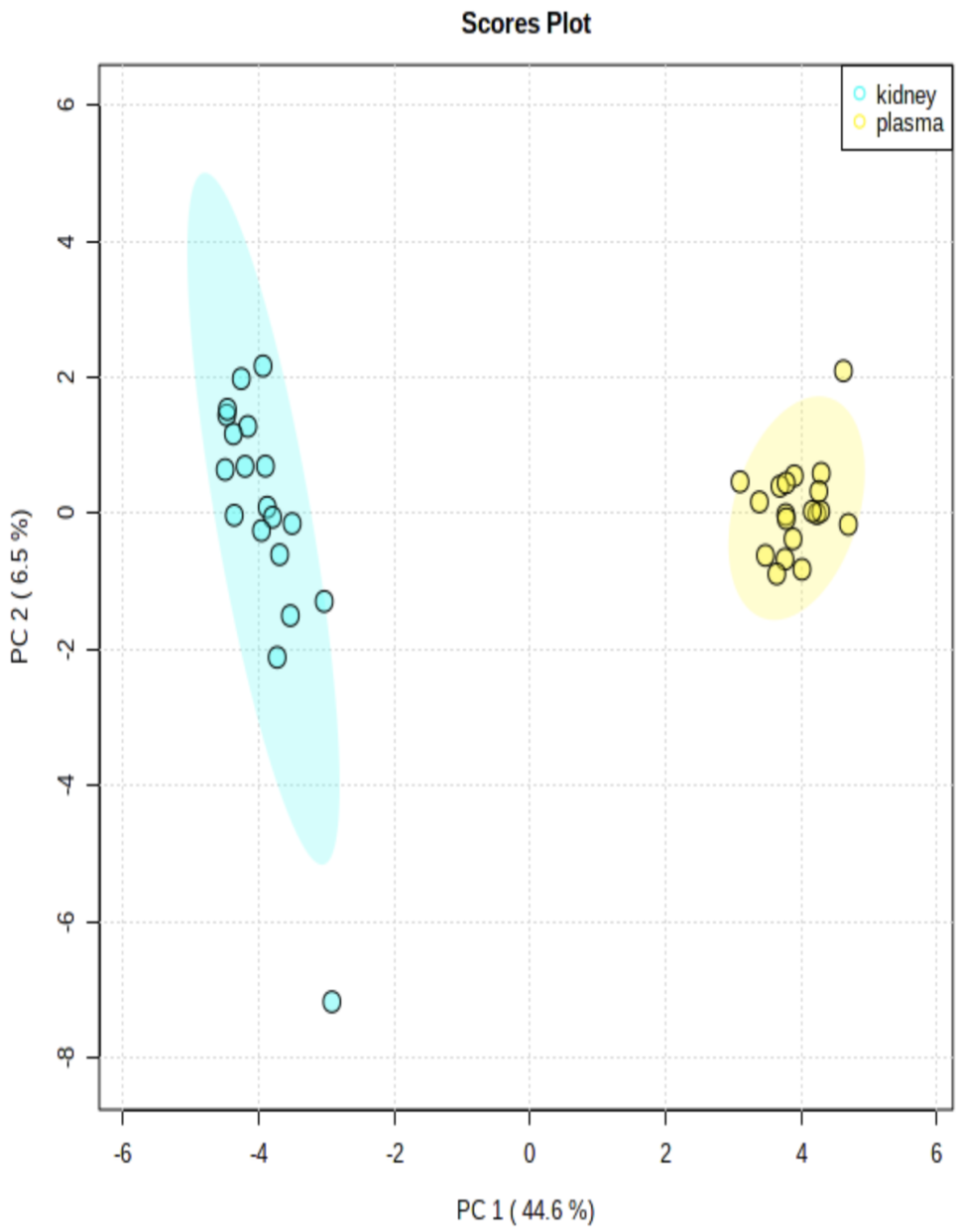


Figure 20. Principal component analysis comparing proteins only found in plasma and kidney. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

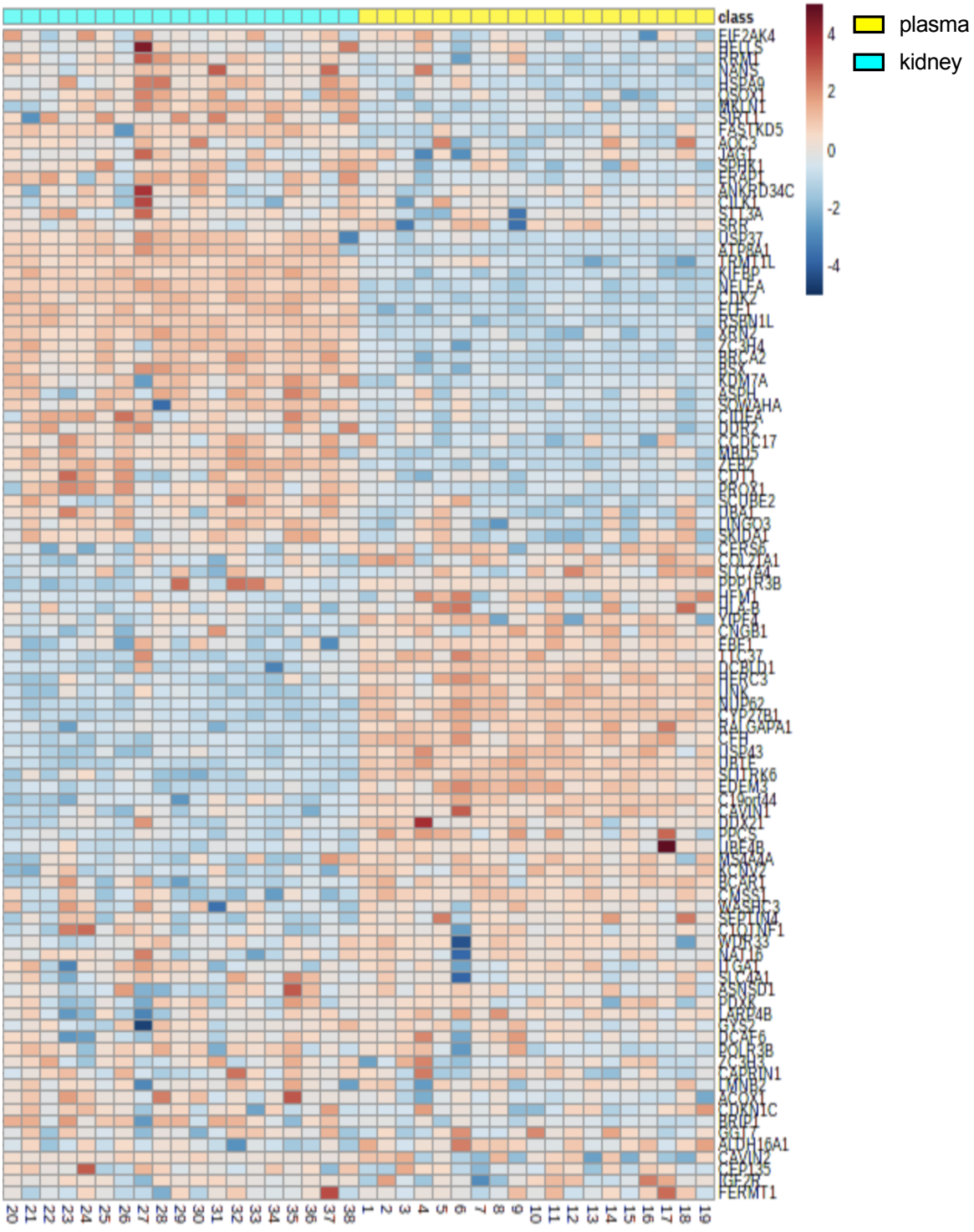


Figure 21. Protein expression heat-map of proteins unique to plasma and kidney. Depicting all 98 proteins amongst plasma and kidney; 22 and 30 proteins significantly increased and decreased in plasma, respectively (t-test).

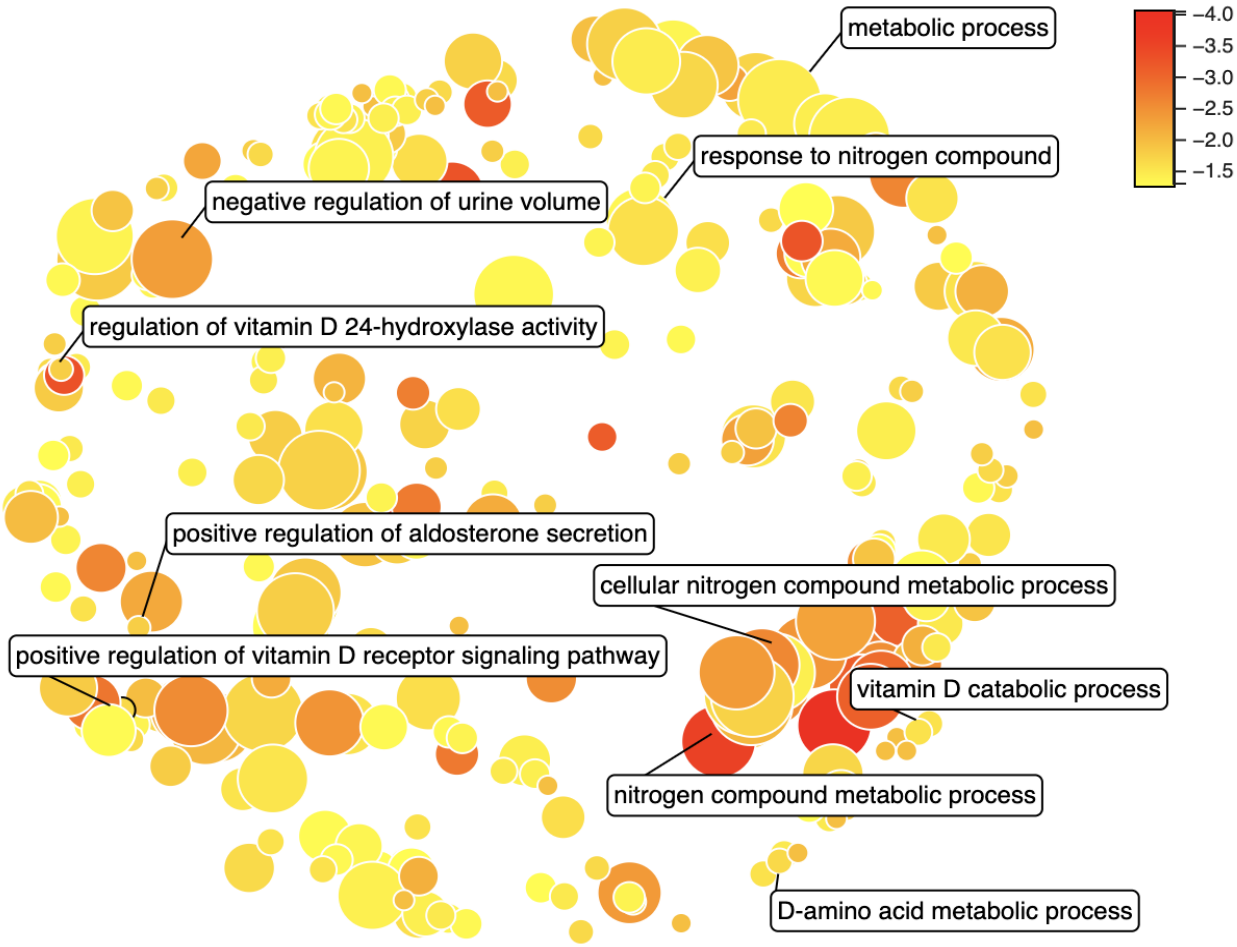


Figure 22. Figure. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and kidney. Results intentionally highlight kidney related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Table 5. Function and expression of proteins uniquely detected in plasma and kidney. Protein list based on proteins that were significantly increased (red font) or decreased (blue font) in plasma when compared to kidney.

HUMAN PROTEIN SYMBOL ORTHOLOG[1]	FUNCTION[2]	EXPRESSION[2]
CYP27B1 Cytochrome P450 Family 27 Subfamily B Member 1	vitamin D metabolism	plasma and kidney
UBTF Upstream Binding Transcription Factor	DNA binding	widely expressed
CFH Complement Factor H	protein binding	widely expressed
UBE4B Ubiquitination Factor E4B	ubiquitin-protein ligase activity	widely expressed
NUP62 Nucleoporin 62	nuclear binding activity	widely expressed
EDEM3 ER Degradation Enhancing Alpha-Mannosidase Like Protein 3	endoplasmic reticulum-associated degradation	widely expressed
RALGAPA1 Ral GTPase Activating Protein Catalytic Subunit Alpha 1	GTPase activator activity	widely expressed
PPCS Phosphopantothenoylcysteine Synthetase	biosynthesis of coenzyme A from vitamin B5	widely expressed
YIPF4 Yip1 Domain Family Member 4	Golgi structure maintenance	widely expressed

HERC3 HECT And RLD Domain Containing E3 Ubiquitin Protein Ligase 3	E3 ubiquitin-protein ligase	blood and immune cells, bone, liver
SLITRK6 SLIT And NTRK Like Family Member 6	regulator of neurite outgrowth	brain, lung, liver
HFM1 Helicase For Meiosis 1	involved in meiosis	blood, heart, pancreas, reproductive organs
UNK Unk Zinc Finger	RNA-binding protein	widely expressed
SEPTIN4 Septin 4	cytoskeletal GTPase	widely expressed
TTC37 Tetratricopeptide Repeat Domain 37	may mediate protein-protein interactions	widely expressed
DDX21 DExD-Box Helicase 21	RNA helicase	widely expressed
HLA-B Major Histocompatibility Complex, Class I, B	immune response	widely expressed
STT3A STT3 Oligosaccharyltransferase Complex Catalytic Subunit A	transferase activity	widely expressed
DDR2 Discoidin Domain Receptor Tyrosine Kinase 2	tyrosine kinase involved in tissue remodeling	widely expressed
NANS N-Acetylneuraminate Synthase	enzymatic activity, biosynthesis of sialic acids	widely expressed
RRM1 Ribonucleotide Reductase Catalytic Subunit M1	DNA synthesis	widely expressed

UBA1 Ubiquitin Like Modifier Activating Enzyme 1	ubiquitin activating enzyme activity	widely expressed
ERAP1 Endoplasmic Reticulum Aminopeptidase 1	aminopeptidase, immune response	widely expressed
MBD5 Methyl-CpG Binding Domain Protein 5	chromatin binding	widely expressed
CDT1 Chromatin Licensing and DNA Replication Factor 1	DNA replication	B-lymphocyte, heart
PROX1 Prospero Homeobox 1	transcription	widely expressed
ZC3H4 Zinc Finger CCCH-Type Containing 4	RNA binding	widely expressed
MKLN1 Muskelin 1	E3 ubiquitin-protein ligase, mediates cell spreading	widely expressed
PPP1R3B Protein Phosphatase 1 Regulatory Subunit 3B	regulates glycogen synthesis	liver, skeletal muscle, heart
TRMT1L TRNA Methyltransferase 1 Like	motor coordination, behaviour	widely expressed
USP37 Ubiquitin Specific Peptidase 37	DNA replication	brain, prostate, heart
FASTKD5 FAST Kinase Domains 5	RNA binding	widely expressed
BRCA2 BRCA2 DNA Repair Associated	DNA repair	breast, thymus

HSPA9 Heat Shock Protein Family A (Hsp70) Member 9	chaperone protein	widely expressed
CDK2 Cyclin Dependent Kinase 2	serine/threonine-protein kinase	widely expressed
SIRT1 Sirtuin 1	transcription	widely expressed
ATP8A1 ATPase Phospholipid Transporting 8A1	cation-transporting ATPase activity	widely expressed
NELFA Negative Elongation Factor Complex Member A	transcriptional regulation	widely expressed
ELF1 E74 Like ETS Transcription Factor 1	transcription factor	widely expressed
ZEB2 Zinc Finger E-Box Binding Homeobox 2	transcriptional regulation	widely expressed
KIFBP Kinesin Family Binding Protein	binding activity	widely expressed
XRN2 5'-3' Exoribonuclease 2	transcriptional regulation	widely expressed
RSBN1L Round Spermatid Basic Protein 1 Like	lysine-specific demethylase	widely expressed

Note: 43 out of 52 significant proteins listed. Proteins removed if information on function and expression were insufficient.
[1]Human protein symbol ortholog: the list of identified rainbow trout protein accession numbers was blasted against the Human Uniprot Reference Proteome (Proteome ID#UP000005640) to obtain human protein orthologs based on sequence similarity.
[2]Function and expression annotations were interpreted from Gene Cards and Uniprot and are specific to humans; 'widely expressed' refers to proteins that have been detected ubiquitously, or in numerous tissues or cell types.

4.3.5. Plasma versus liver proteome

The proceeding analysis compared the 120 proteins that were only shared amongst the plasma and liver proteomes. Although all 120 proteins were found across liver and plasma, PCA analysis showed that protein abundance levels differed and still resulted in significantly different sub-proteomes (Figure 23). Of the 120 proteins, 43 proteins were significantly increased in plasma, and 43 proteins were significantly decreased in plasma when compared to liver (FDR corrected p-value < 0.05; Figure 24). The function and expression annotations for the preceding significant proteins are summarized in Table 5. Gene Card and Uniprot tissue-expression data showed that 9 out the 86 significant proteins associated with liver and plasma are predominantly expressed in the human liver: CISH, COL20A1, FAM200A, CYP24A1, MYLK3, MGAT5B, ERBB2, CES2, and CYP2C8. The functions of the preceding 9 proteins include cytokine signaling, metabolism, phosphorylation, enzymatic and receptor activity, and detoxification (Table 5). Gene ontology analysis of the 120 proteins associated with plasma and liver contained many biological processes specific to liver function, most of which could be summarized into (1) metabolism and biosynthesis, (2) lipid and cholesterol transport, (3) fatty-acid oxidation, and (4) lipoprotein particle clearance and remodeling (Figure 25).

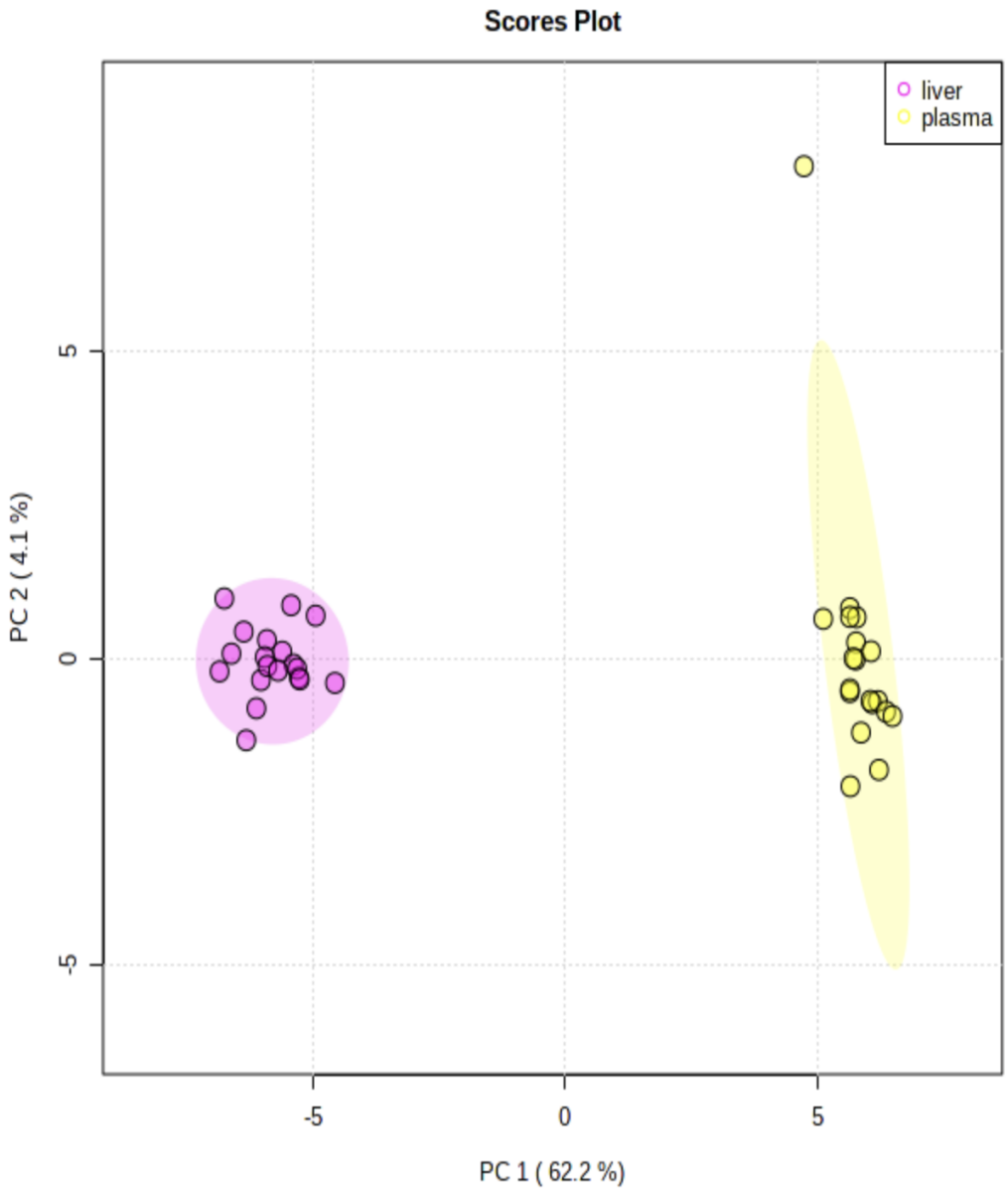


Figure 23. Principal component analysis comparing proteins only found in plasma and liver. Nineteen biological samples were used for all six sample types, represented by a single point on each graph. Each point depicts the list of proteins and their respective intensities per biological sample. Metaboanalyst was used to normalize, transform, scale, and analyze the data. Data was normalized by the median, log transformed, and scaled using pareto scaling.

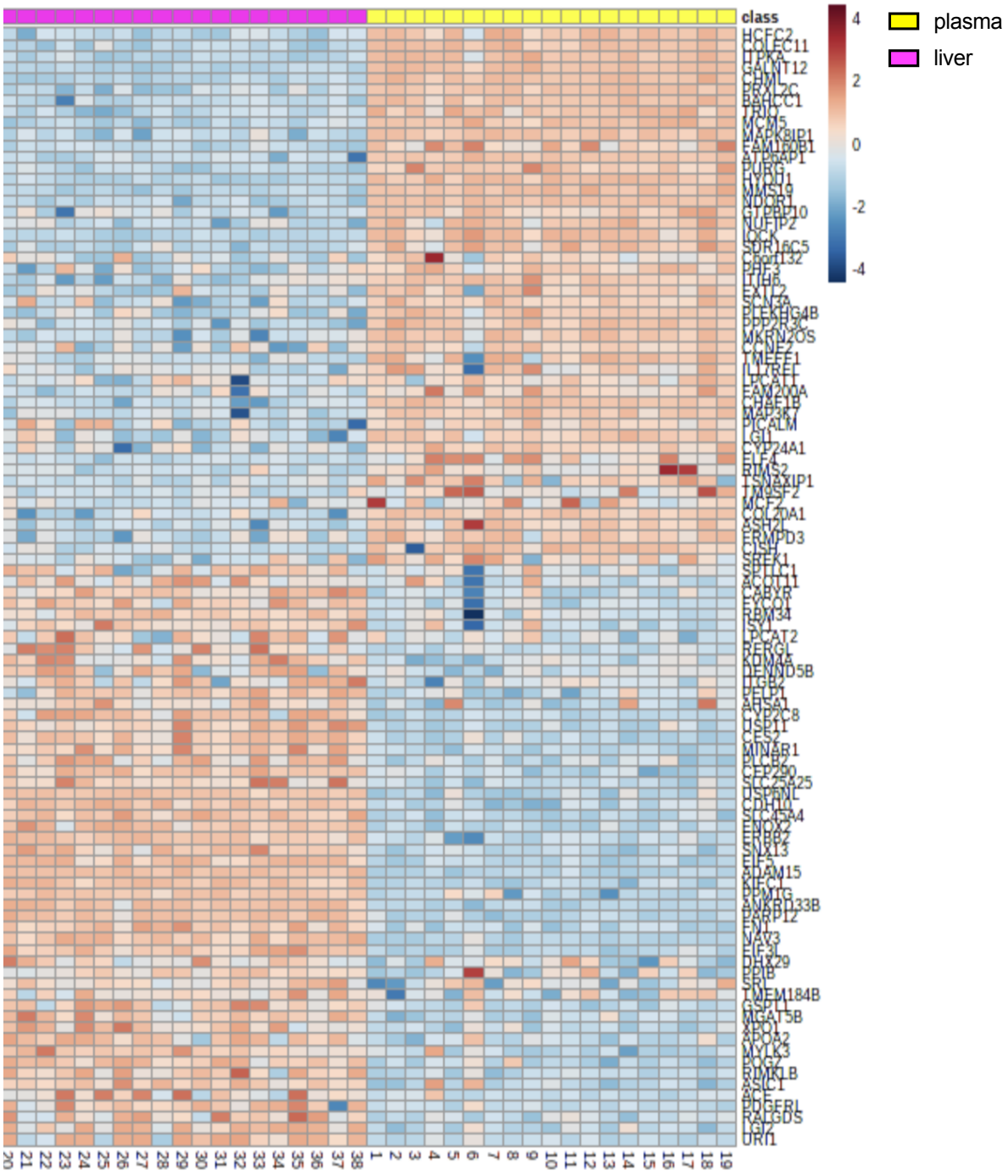


Figure 24. Protein expression heat-map of proteins unique to plasma and liver. Depicting top 100 proteins amongst plasma and liver; 43 and 43 proteins significantly increased and decreased in plasma, respectively (t-test).

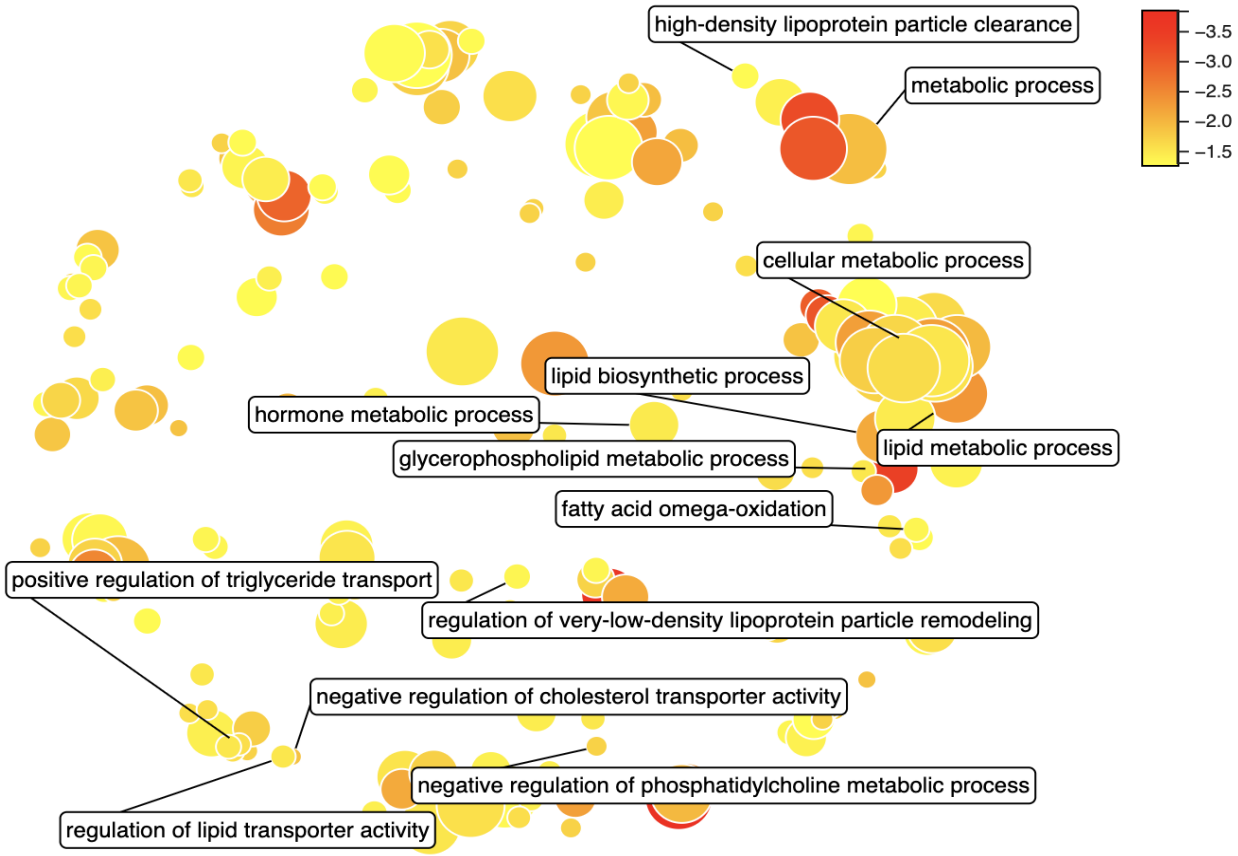


Figure 25. REVIGO semantic similarity of gene ontology (GO) biological processes related to proteins unique to plasma and liver. Results intentionally highlight liver related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Table 6. Function and expression of proteins uniquely detected in plasma and liver. Protein list based on proteins that were significantly increased (red font) or decreased (blue font) in plasma when compared to liver.

HUMAN PROTEIN SYMBOL ORTHOLOG[1]	FUNCTION[2]	EXPRESSION[2]
ELF4 E74 Like ETS Transcription Factor 4	transcriptional activator	widely expressed
GALNT12 Polypeptide N-Acetylgalactosaminyltransferase 12	transferase activity	widely expressed
NDOR1 NADPH Dependent Diflavin Oxidoreductase 1	electron transfer	widely expressed
HYOU1 Hypoxia Up-Regulated 1	cytoprotective cellular mechanism	widely expressed
CHML CHM Like Rab Escort Protein	binding activity	brain
CHAF1B Chromatin Assembly Factor 1 Subunit B	mediate chromatin assembly in DNA replication	lung, reproductive organs
ATP6AP1 ATPase H+ Transporting Accessory Protein 1	mediates organelle acidification	widely expressed
MMS19 MMS19 Homolog, Cytosolic Iron-Sulfur Assembly Component	iron-sulfur assembly	widely expressed
ITPKA Inositol-Trisphosphate 3-Kinase A	regulates inositol phosphate metabolism	blood and immune cells, brain

MCM5 Minichromosome Maintenance Complex Component 5	chromatin binding, DNA replication	widely expressed
COLEC11 Collectin Subfamily Member 11	innate immunity	widely expressed
CISH Cytokine Inducible SH2 Containing Protein	cytokine signaling suppressor	liver, kidney, adipocyte, epithelial tissue
FAM160B1 FHF Complex Subunit HOOK Interacting Protein 2A	required for proper functioning of the nervous system	predominantly expressed in brain
RIMS2 Regulating Synaptic Membrane Exocytosis 2	presynaptic protein	widely expressed
MAPK8IP1 Mitogen-Activated Protein Kinase 8 Interacting Protein 1	protein kinase inhibition	predominantly expressed in brain and immune cells
MAP3K7 Mitogen-Activated Protein Kinase Kinase Kinase 7	serine/threonine kinase	widely expressed
TRIO Trio Rho Guanine Nucleotide Exchange Factor	actin cytoskeleton remodeling	widely expressed
COL20A1 Collagen Type XX Alpha 1 Chain	probable collagen protein	blood, heart, and liver
PPP2R3C Protein Phosphatase 2 Regulatory Subunit B"Gamma	serine/threonine phosphatase	widely expressed

NUFIP2 Nuclear FMR1 Interacting Protein 2	RNA binding	widely expressed
SDR16C5 Short Chain Dehydrogenase/Reductase Family 16C Member 5	oxidoreductase	widely expressed
FAM200A Family With Sequence Similarity 200 Member A	unknown, similar to transposase proteins	platelet, heart, liver, testis
ASH2L ASH2 Like, Histone Lysine Methyltransferase Complex Subunit	transcriptional activator	widely expressed
CYP24A1 Cytochrome P450 Family 24 Subfamily A Member 1	vitamin D3 metabolism	heart, liver
FRMPD3 FERM And PDZ Domain Containing 3	signal transduction	predominantly expressed in blood and immune cells
MCF2 MCF.2 Cell Line Derived Transforming Sequence	guanine nucleotide exchange factor	predominantly expressed in brain
EXTL2 Exostosin Like Glycosyltransferase 2	biosynthesis of heparan-sulfate	widely expressed
PHF3 PHD Finger Protein 3	transcription factor	widely expressed
LGI1 Leucine Rich Glioma Inactivated 1	regulates voltage-gated potassium channels	predominantly expressed in brain
CCNE2 Cyclin E2	cell cycle regulation	widely expressed

TM9SF2 Transmembrane 9 Superfamily Member 2	small molecule transport or ion channel	widely expressed
SCN3A Sodium Voltage-Gated Channel Alpha Subunit 3	mediates voltage-dependent sodium ion permeability	brain, breast, colon, small bowel
FYCO1 FYVE And Coiled-Coil Domain Autophagy Adaptor 1	microtubule transport	widely expressed
ITGB2 Integrin Subunit Beta 2	cell adhesion and signaling	widely expressed
ASIC1 Acid Sensing Ion Channel Subunit 1	proton gated channel	brain
ISY1 ISY1 Splicing Factor Homolog	RNA binding	widely expressed
MYLK3 Myosin Light Chain Kinase 3	phosphorylates myosin heavy and light chains	predominantly expressed in heart, liver, and testis
CABYR Calcium Binding Tyrosine Phosphorylation Regulated	calcium ion binding activity	ovaries, testis, brain, pancreas
RBM34 RNA Binding Motif Protein 34	RNA binding	widely expressed
MGAT5B Alpha-1,6-Mannosylglycoprotein 6-Beta-N-Acetylglucosaminyltransferase B	glycosyltransferase	liver, brain
PELP1 Proline, Glutamate And Leucine Rich Protein 1	coactivator of estrogen receptor mediated transcription	widely expressed

XPO1 Exportin 1	cell cycle regulation	widely expressed
POGZ Pogo Transposable Element Derived With ZNF Domain	mitotic cell cycle progression	widely expressed
RALGDS Ral Guanine Nucleotide Dissociation Stimulator	GTPase regulator activity	brain, heart, adipocyte
FN1 Fibronectin 1	binds fibronectin	widely expressed
ACE Angiotensin I Converting Enzyme	angiotensin I conversion	widely expressed
PPM1G Protein Phosphatase, Mg2+/Mn2+ Dependent 1G	dephosphorylation of pre-mRNA splicing factors	widely expressed
USP11 Ubiquitin Specific Peptidase 11	deubiquitinating enzyme	widely expressed
ERBB2 Erb-B2 Receptor Tyrosine Kinase 2	cell surface receptor	heart, liver, lung, secretory and reproductive organs
SLC25A25 Solute Carrier Family 25 Member 25	calcium-binding mitochondrial carrier	widely expressed
KDM4A Lysine Demethylase 4A	histone demethylase activity	widely expressed
SNX13 Sorting Nexin 13	intracellular trafficking	widely expressed
PARP12 Poly(ADP-Ribose) Polymerase Family Member 12	ribosyl-transferase activity	widely expressed
KIFC1 Kinesin Family Member C1	microtubule motor activity	widely expressed

CES2 Carboxylesterase 2	detoxification of xenobiotics	predominantly expressed in liver and stomach
CDH10 Cadherin 10	calcium-dependent cell adhesion protein	brain
GSPT1 G1 To S Phase Transition 1	translation termination	widely expressed
EIF3L Eukaryotic Translation Initiation Factor 3 Subunit L	translation initiation	widely expressed
URI1 URI1 Prefoldin Like Chaperone	transcription regulation	widely expressed
PLCB2 Phospholipase C Beta 2	phosphoric diester hydrolase activity	predominantly expressed in blood and immune cells
CYP2C8 Cytochrome P450 Family 2 Subfamily C Member 8	metabolism	predominantly in liver and gallbladder
CEP290 Centrosomal Protein 290	cilia formation	widely expressed
USP6NL USP6 N-Terminal Like	GTPase activator activity	widely expressed
EIF5 Eukaryotic Translation Initiation Factor 5	translation	widely expressed
NAV3 Neuron Navigator 3	nervous system development	widely expressed; highly expressed in brain
ADAM15 ADAM Metallopeptidase Domain 15	cell adhesion and signaling	monocytes, colon, breast, placenta, ovaries

Note: 66 out of 86 significant proteins listed. Proteins removed if information on function and expression were insufficient.

[1]Human protein symbol ortholog: the list of identified rainbow trout protein accession numbers was blasted against the Human Uniprot Reference Proteome (Proteome ID#UP000005640) to obtain human protein orthologs based on sequence similarity.

[2]Function and expression annotations were interpreted rom Gene Cards and Uniprot and are specific to humans; 'widely expressed' refers to proteins that have been detected ubiquitously, or in numerous tissues or cell types.

4.4. Exploring plasma-specific proteins

A total of 1,990 proteins were identified in plasma, 576 of which were only detected in plasma and no other sample group. Gene ontology analysis of the proteins associated with plasma contained many biological processes specific to a wide range of functions, some of which include (1) immunity, (2) transport, (3) cell communication and adhesion, (4) regulation and response, (5) secretion, and (6) signaling (Figure 26, Figure 27). The top 20 most abundant proteins detected in the full plasma proteome are listed in Table 6. The functions pertaining to the top 20 proteins generally include transport, motor protein activity, enzymatic activity, transcription, binding, and Golgi function. The top 20 most abundant plasma-specific proteins, which are proteins uniquely identified in the plasma proteome, are listed in table 7. The functions pertaining to the top 20 plasma-specific proteins generally include transport, binding, enzymatic activity, metabolism, transcription, Golgi disassembly, endocytosis, and apoptosis. COPB1, NXF1, LYL1, AK9, and GDF5 human ortholog proteins were amongst the top 20 most abundant proteins in both the full plasma proteome and the plasma-specific proteome.

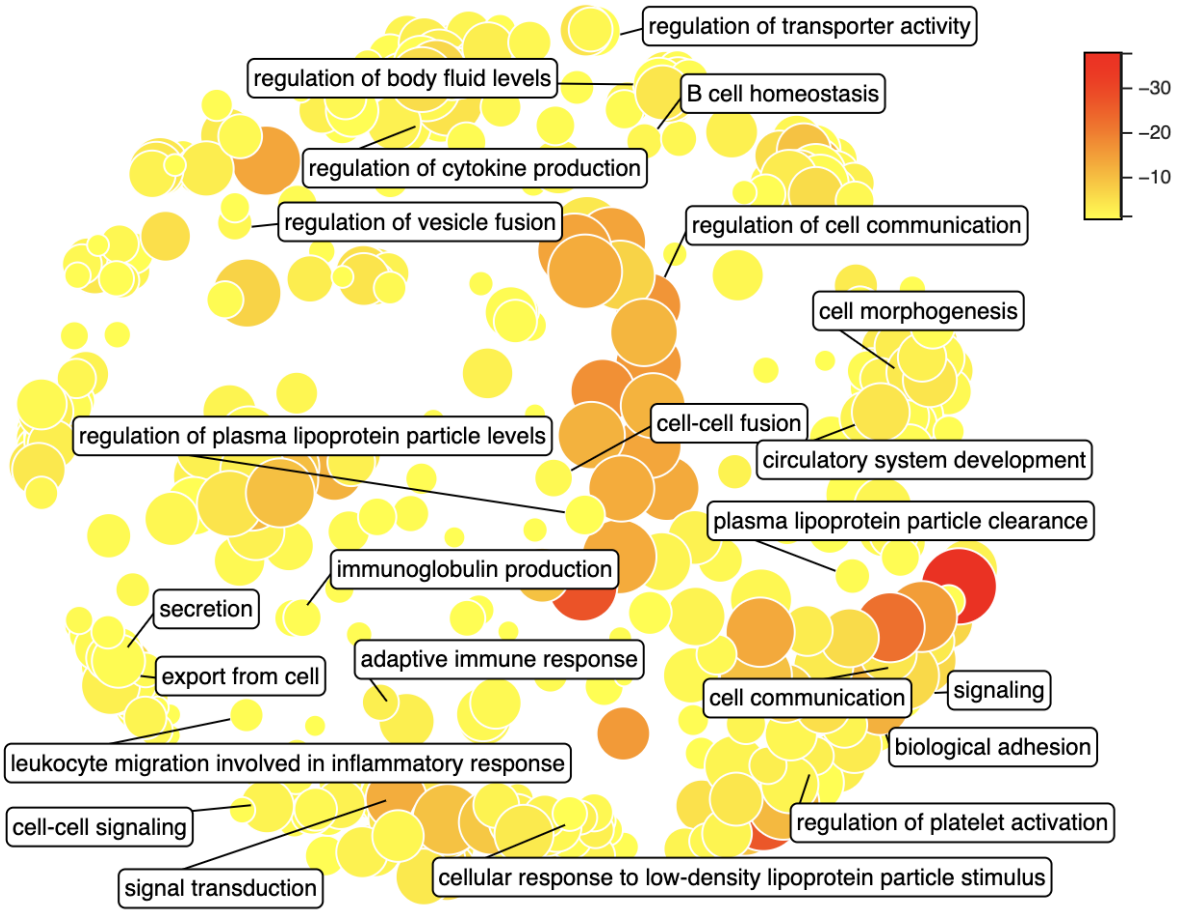


Figure 26. REVIGO semantic similarity of gene ontology (GO) biological processes related to all proteins detected in plasma. Results intentionally highlight plasma related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

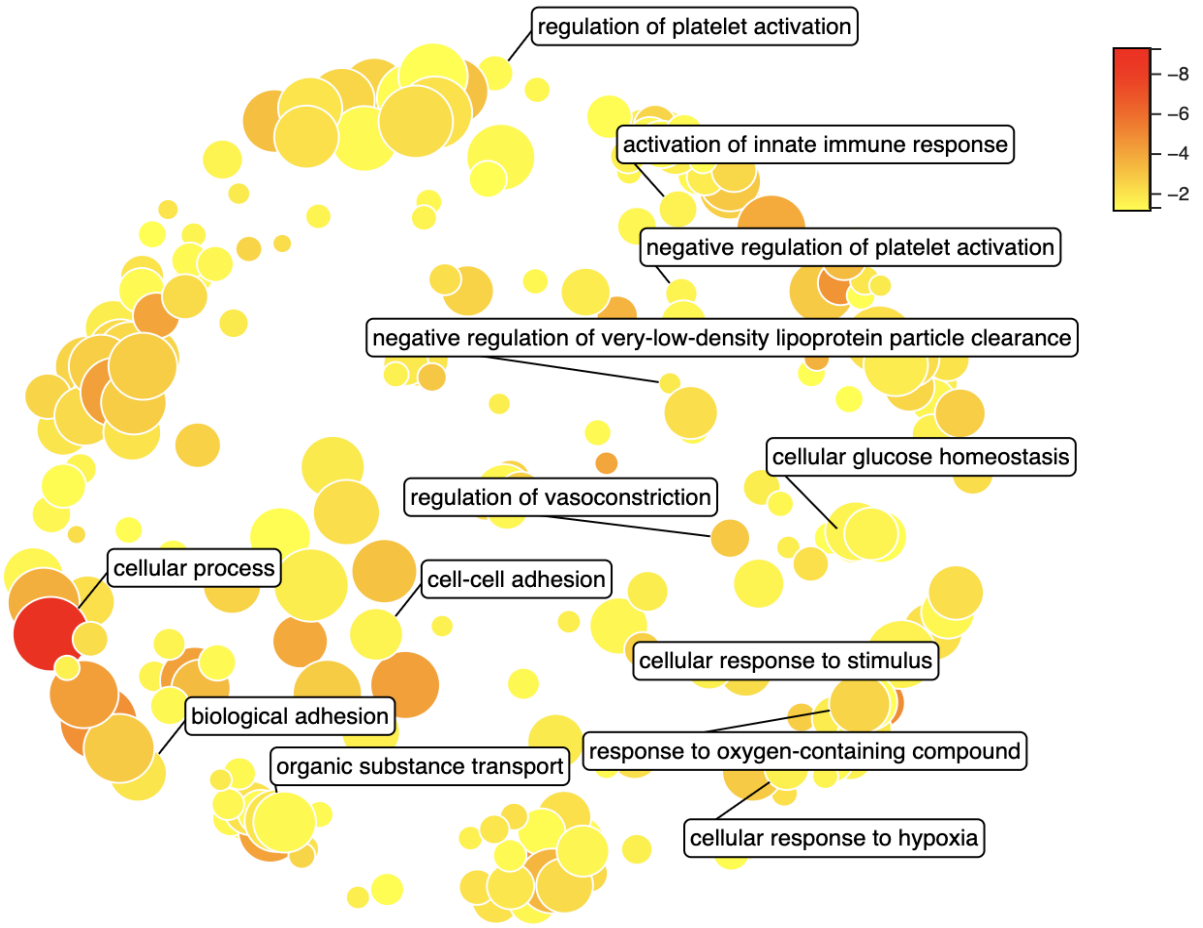


Figure 27. REVIGO semantic similarity of gene ontology (GO) biological processes related to plasma-specific proteins, which are proteins only detected in plasma. Results intentionally highlight plasma related biological processes found amongst the whole dataset. Proximity of circles indicates GO similarity (relatedness). Size of circles represent the frequency of the GO term in the EBI - GOA database; large circles indicate general gene ontology terms, while small circles indicate specific gene ontology terms. Colour of circles represent the user-provided log10 p-values of the protein ID, where red circles are more significant than yellow circles (legend). User-provided p-values calculated using Gene Ontology. Data only includes gene ontology IDs with $p < 0.05$.

Table 7. Top 20 proteins expressed in the plasma proteome.

HUMAN PROTEIN SYMBOL ORTHOLOG[1]	PROTEIN NAME[2]	FUNCTION[2]
COPB1	coatomer subunit beta	intracellular protein transport
MYH4	myosin 4	motor protein
HYI	putative hydroxypyruvate Isomerase	isomerase activity
ARMC8	armadillo repeat-containing protein 8	ubiquitin-protein ligase
CA5B	carbonic anhydrase 5B	reversible hydration of carbon dioxide
OSBPL1A	oxysterol binding protein like 1A	phospholipid binding
DNAH8	dynein axonemal heavy chain 8	microtubule motor activity
NXF1	nuclear RNA export factor 1	RNA export
LYL1	class A basic helix-loop-helix protein 18	DNA binding activity
DNAH7	dynein axonemal heavy chain 7	microtubule motor activity
AK9	adenylate kinase 9	nucleoside diphosphate kinase activity
ELF4	ETS-related transcription factor	DNA-binding transcription factor activity
PPME1	protein phosphatase methylesterase	protein phosphatase inhibitor activity
PLXNB1	plexin-B1	semaphorin receptor activity
OTOF	otoferlin	calcium ion binding activity
TRPS1	zinc finger transcription factor	transcriptional repressor
COG2	conserved oligomeric Golgi complex subunit 2	required for Golgi morphology and function
GDF5	growth/differentiation factor 5	growth factor involved in bone and cartilage formation
TCP1	T-complex protein 1 subunit alpha	molecular chaperone

[1]Human protein symbol ortholog: the list of identified rainbow trout protein accession numbers was blasted against the Human Uniprot Reference Proteome (Proteome ID#UP000005640) to obtain human protein orthologs based on sequence similarity.

[2]Protein name and function interpreted from Gene Cards and Uniprot and are specific to humans.

Table 8. Top 20 proteins expressed in the plasma-specific proteome, which are proteins only detected in plasma samples.

HUMAN PROTEIN SYMBOL ORTHOLOG [1]	PROTEIN NAME[2]	FUNCTION[2]
COPB1	coatomer subunit beta	intracellular protein transport
NXF1	nuclear RNA export factor 1	RNA export
LYL1	class A basic helix-loop-helix protein 18	DNA binding activity
AK9	adenylate kinase 9	nucleoside diphosphate kinase activity
GDF5	growth/differentiation factor 5	growth factor involved in bone and cartilage formation
MMAB	corrinoid adenosyltransferase	intracellular vitamin B12 metabolism
MRC1	macrophage mannose receptor 1	mediates the endocytosis of glycoproteins by macrophages
ZWINT	ZW10 interactor	kinetochore formation
CLPTM1L	cisplatin resistance-related protein 9	involved in apoptosis process
STAT3	signal transducer and activator of transcription 3	transcription activator activity
NEXN	nexilin	actin binding
RXFP1	relaxin receptor 1	G protein-coupled receptor activity
TMEM183A	transmembrane protein 183A	protein binding
CAP2	adenylyl cyclase-associated protein 2	actin binding
FUT9	fucosyltransferase 9	enzymatic activity
VRK1	serine/threonine-protein kinase	Golgi disassembly
ZNF277	zinc finger protein	transcription
LRRC18	leucine-rich repeat-containing protein	protein binding
SRF	serum response factor	regulates cytoskeleton, morphogenesis, and cell migration
MBL2	mannose-binding protein	innate immune response

[1]Human protein symbol ortholog: the list of identified rainbow trout protein accession numbers was blasted against the Human Uniprot Reference Proteome (Proteome ID#UP000005640) to obtain human protein orthologs based on sequence similarity.

[2]Protein name and function interpreted from Gene Cards and Uniprot and are specific to humans.

5. Discussion

The plasma proteome contains many different proteins from the whole body that constantly fluctuate to suit an organism's needs, rendering plasma an effective biofluid in non-lethal health monitoring (Califf, 2018; Lawrence et al., 2020; López-Maury et al., 2008). Currently, proteomic databases lack a substantial amount of annotation, such as tissue-specific information, related to environmentally relevant species (Consortium, 2020). This study aimed to discover tissue-specific information associated with the rainbow trout (*Oncorhynchus mykiss*) proteome, with a goal to understand the functions and mechanistic importance of blood plasma proteins. Considering literature surrounding multi-tissue proteomics comparisons are limited in environmental sciences, the proceeding discussion will mainly compare plasma to individual tissue proteomes, as well as highlight the function and origin of a subset of proteins found amongst various species. Overall, the proteome of plasma, brain, heart, gill, kidney, and liver were notably distinct from one another. Careful and thorough literature review could not find examples of other studies that have compared multiple tissue proteomes in fish. Sex-specific differences were not apparent, and vitellogenin domain-containing proteins were found amongst male and female samples in all tissues analyzed (Figure 5-7). Aside from the presence of vitellogenin, Li *et al*., (2016) also observed no obvious sex differences amongst male and female plasma proteins (C. Li et al., 2016). We did not expect to see vitellogenin in male and female fish, as well as in all tissue types analyzed. This is because vitellogenin, an egg-yolk precursor protein, is typically only synthesized in the liver of mature female fish, which then travels to the ovaries via the bloodstream

(Sugawara, 2011; Sullivan & Yilmaz, 2018). There were four rainbow trout proteins detected in our analysis that were identified as vitellogenin domain-containing proteins, which had a 29-43% sequence similarity to the Apolipoprotein B-100 human ortholog. An evolutionary relationship has been noted amongst vitellogenin and mammalian proteins, such as apolipoproteins, and phylogenetic studies indicate vitellogenin may share a common ancestor with large lipid transfer proteins (Avarre, Lubzens, & Babin, 2007). Baker (1988) also mentions that partially conserved sequences of vitellogenin likely evolved to include other functions aside from a food source for developing embryos, such as lipid transport and metabolism (Baker, 1988). Based on the preceding evidence in our study and in literature, we suspect that many of the proteins identified as vitellogenin in this study were actually lipoproteins where the peptides in our samples matched to conserved sequences and domains between the two protein families. Overall, 10,000 proteins were identified across all six sample groups, almost half of which were non-redundant. Among the tissues, the greatest number of proteins were identified in the gill, followed by heart, liver, kidney, and brain. The proceeding discussion will briefly compare the number of identified proteins in each tissue type to those found in literature, followed by an in-depth analysis of tissue-specific plasma proteins and their functions. As there is a lack of information surrounding tissue origin in fish proteomes, we decided to compare the tissue-specific plasma proteins found in this study to tissue-specific annotations found in human proteomic databases.

Gill tissue contributions to the plasma proteome

The gills made up the largest fraction of identified the total number of identified proteins in this study, with 2,132 (20%) proteins detected. Li *et al.*, (2020) identified 1,691 gill proteins in three-spined stickleback fish, while De Souza *et al.*, (2009) detected 5,716 proteins in zebrafish gill (De Souza, MacCormack, Wang, Li, & Goss, 2009; J. Li & Kültz, 2020), which demonstrates that our method was able to detect a similar number of proteins compared to other studies. Within the present study, the gill and plasma proteome shared 141 unique proteins that were not present in any of the other tissues analyzed, 72 of which contained significant differences in abundance (Figure 18 & Table 3). Considering tissue-specific information is missing amongst fish proteomes in functional databases, and gills and lungs both serve as respiratory organs (Haddad M & S., 2021 ; Roberts & Smith, 2011), we opted to compare the 72 significantly different gill proteins detected in plasma from this study to human protein orthologs expressed in lungs using online databases. 3 out of the 72 gill proteins detected in plasma with differential abundance were also known to be expressed in lungs: Bnipl, Kif23, and Ptger4 (Table 3). The preceding three proteins have not been widely studied thus we took a closer look at all gill-specific proteins detected in plasma and found that Alas1, Dspp, and Pou2f1 resulted in relevant information associated with gill function. Specifically, gills act as a first-line of defense from the external environment and hold a wide range of adaptive immune molecules (Maki & Dickerson, 2003; Xu et al., 2016; Yu, Wang, Huang, Ding, & Xu, 2020). Alas1, a gill-specific plasma protein detected in our study, is known to play an essential role in innate immunity, and is associated with neutrophil development in

zebrafish (Lian et al., 2018). Additionally, DSPP, a protein found in teeth, is suggested to be a homologue to *stm-l* and *starmaker* genes found in medaka and zebrafish, respectively. The medaka gene *stm-l* is expressed in otoliths, teeth, gills, and kidney (Bajoghli, Ramialison, Aghaallaei, Czerny, & Wittbrodt, 2009). Pou2f1, an octamer-binding protein, was upregulated in gill samples of zinc-deprived fish, and a similar protein, Pou3f3 was identified as an essential gene in gill cover formation (Barske et al., 2020; Zheng et al., 2010). Research surrounding gill-specific protein functions are limited, though general gill functions include gas exchange, immunity and xenobiotic transport (Bury, Schnell, & Hogstrand, 2014; Maki & Dickerson, 2003; Stott, Schnell, Hogstrand, Owen, & Bury, 2015; Xu et al., 2016; Yu et al., 2020). A closer look into the biological processes associated with gill-specific plasma proteins in this study included cytokinesis, receptor activity, response mechanisms, immunity, xenobiotic transport, and cartilage development (Figure 19). It is interesting to see that even though our gill-specific plasma proteins were searched against human-specific databases, biological processes typical to general gill functions were still emphasized.

Heart tissue contributions to the plasma proteome

1,860 proteins were identified in the rainbow trout heart, making up 18% of the total proteins identified. This is more than the 1,375 proteins detected in zebrafish heart and 736 proteins identified in juvenile rainbow trout ventricles from other studies (Dindia, Alderman, & Gillis, 2017; J. Zhang, Lanham, Peterson, Heideman, & Li, 2010). There were 130 heart-specific plasma proteins detected in this study, 97 of which contained

significantly different abundances (Figure 15). 18 out of the 97 heart-specific plasma proteins identified in our rainbow trout have also been expressed in human heart: Pde4b, Ntn4, Ctdspl2, Adgrf3, Myo10, Mybpc3, Shh, Frrs1, Znf423, Lrp4, Slca4, Pnkp, Trim27, Hes1, Nyap1, Setbp1, Znf280d, and Myo19. Ntn4 (Netrin-4) is crucial in blood vessel formation, has been detected in zebrafish vasculature systems, and is known to play a role in angiogenesis (Lambert, Coissieux, Laudet, & Mehlen, 2012; Lu et al., 2004). Mybpc3 (myosin-binding protein) mutations have been linked to cardiac-specific disorders (e.g., hypertrophic cardiomyopathy), and genetic knock-down of Mybpc3 in zebrafish resulted in ventricular hypertrophy, increased myocardial wall thickness, and diastolic heart failure (Carrier, Mearini, Stathopoulou, & Cuello, 2015; Chen et al., 2013). Shh (sonic hedgehog) is a conserved protein involved in embryonic development; studies suggest Shh plays a role in heart development and maintenance. For instance, SHH showed a protective role against reperfusion injuries after myocardial infarctions, and SHH null mice resulted in cardiac defects (Levin, Johnson, Sterna, Kuehn, & Tabin, 1995; Paulis et al., 2015; Washington Smoak et al., 2005). Slc2a4, a Glut4 receptor, is known to be highly expressed in insulin-responsive tissues, such as the heart (Bryant, Govers, & James, 2002; Capilla, Díaz, Hou, Planas, & Pessin, 2010; Watson, Kanzaki, & Pessin, 2004). Additionally, the heart and muscle are considered the major sites of Glut4 distribution in fish (Capilla et al., 2010; Hall, Short, & Driedzic, 2006; S. Li et al., 2018). Biological processes linked to the heart-specific plasma proteins in this study included enzymatic activity, receptor and transport, transcription, DNA damage repair, muscle activity, and heart and vasculature development (Figure 16). Zhang *et al.*, (2010) also

found that the major functions associated with zebrafish heart proteins were ion channel/transport, and enzymatic activity, such as protein kinase and ATPase activity (J. Zhang et al., 2010). In brief, the heart-specific plasma proteins detected in this study showed relevant biological processes and expression patterns pertaining to heart function found in literature, demonstrating the capability of plasma to hold a substantial level of information surrounding cardiac health.

Liver tissue contributions to the plasma proteome

1,692 proteins were detected in the liver, accounting for 16% of the total number of proteins identified in our study. Other studies have identified 780 proteins in rainbow trout liver using 2D-gel electrophoresis, 6770 proteins using data-dependent and data-independent MS/MS analysis, and 2,433 proteins identified in bacteria-challenged rainbow trout liver (Causey et al., 2018; Martin et al., 2001; Quan et al., 2021). The liver and plasma proteome shared 120 unique proteins that were not present in any of the other tissues analyzed, 86 of which contained significant differences in abundance (Figure 24). Upon comparing the liver-specific plasma protein in this study to protein expression in human databases, 9 out of the 86 proteins were predominantly expressed in liver: Cish, Col20a1, Fam200A, Cyp24a1, Mylk3, Mgat5b, Erbb2, Ces2, and Cyp2c8. Of note, CYP24A1, has been associated with liver diseases, such as non-alcoholic fatty liver disease in humans (Wang et al., 2021). The overexpression of the tyrosine kinase ERBB2 has been associated with hepatocellular damage and hepatocellular carcinoma in humans (Döring, Calvisi, & Dombrowski, 2021; Shi et al., 2019). CES2, a

carboxylesterase, has been documented in liver, intestine, and kidney, and is known to play a role in detoxification (Chalhoub et al., 2021; Satoh & Hosokawa, 2006). Additionally, Li *et al.*, (2016) found that CES2 knock-down resulted in fatty liver disease in mice, and CES2 overexpression improved glucose tolerance in mice fed a high-fat diet (Y. Li et al., 2016). Major biological processes associated with the proteins detected amongst the plasma and liver proteome in the present study included cytokine signaling, metabolism, enzymatic and receptor activity, lipid transport, detoxification, and fatty acid oxidation (Figure 25). Limited proteomics studies in fish include liver function unless associated with a chemical exposure, but generally, fish liver functions include detoxification, vitellogenin production, and metabolism (Bruslé & Anadon, 1996), all of which were highlighted as major functions of liver-specific plasma proteins in this study.

Kidney tissue contributions to the plasma proteome

A total of 1,454 kidney proteins were detected in our study, making up 13% of the total proteins identified. Saxena *et al*., (2011) identified 313 proteins in zebrafish kidney, which includes a collective count of three different gel-based proteome mapping methods (Saxena et al., 2011), which suggests that liquid chromatography is likely a better method for capturing more proteins from kidney. The kidney and plasma proteome shared 98 unique proteins that were not present in any of the other tissues analyzed, 52 of which contained significant differences in abundance. After data-mining kidney-specific plasma proteins in this study to tissue-specific information found in human databases, we identified 1 protein among our data that is also known to be expressed in the kidney,

Cyp27b1, which plays a role in vitamin D metabolism. CYP27B1 expression is found at extremely high levels in mammalian kidney, which is expected considering vitamin D is hydroxylated in the liver and kidney (Chun et al., 2014; Jones, Prosser, & Kaufmann, 2018; Lawson, Fraser, Kodicek, Morris, & Williams, 1971; Zehnder et al., 1999). Aside from vitamin metabolism, some biological processes associated with proteins unique to plasma and kidney include regulation of vitamin D, and amino acid metabolism (Figure 22). Li *et al.*, (2020) describes the vital role kidneys play in amino acid metabolism using a wide range of mechanisms extending from reabsorption of free amino acids, to the metabolism and synthesis of a wide range of amino acids (X. Li, Zheng, & Wu, 2020). Other studies indicate that the major functions associated with kidney proteins include binding and catalytic activity, as well as regulation, amino acid metabolism, and inflammatory response (Saxena et al., 2011). Additionally, our study found that kidney-specific plasma proteins contained binding and cation ATPase activity (Table 4), which coincides with ion binding and transport playing a major role in freshwater fish physiology (Thibaut et al., 2019). Although the number of kidney-specific plasma proteins found in this study resulted in minimal kidney expressed proteins found in human databases, the overall biological processes and functions associated with kidney-specific plasma proteins detected in this study did overlap with general kidney functions associated with fish.

Brain tissue contributions to the plasma proteome

Specifically, 1,436 proteins were identified in the rainbow trout brain, making up 14% of the total proteins identified. Literature surrounding the brain proteome in fish demonstrates varying numbers of identified brain proteins; Singh *et al.*, (2010) identified 8,475 proteins in zebrafish brain using LC-ESI-MS/MS, Gebriel *et al.*, (2014) identified 95 labelled proteins in zebrafish using 2D-gel electrophoresis and SDS-PAGE, and Purushothaman *et al.*, (2015) identified at least 78 differentially regulated proteins in the brain during light disturbance (Gebriel et al., 2014; Purushothaman et al., 2015; Singh, Rakesh, Ramamoorthy, Saradhi, & Idris, 2010). The brain and plasma proteome shared 77 unique proteins that were not present in any of the other tissues analyzed, 50 of which contained significant differences in abundance (Figure 12). 8 out of the 50 brain-specific plasma proteins were found to be expressed in human brain: Gdi1, Abcd7, Rimbp2, Esrrb, Sox11, Kcng1, Rasgrf2, and Bbox1. Of note, GDI1, a GDP dissociation inhibitor involved with GTPases of the Rab family, has been associated with X-linked mental retardation, where the gene is upregulated during the early stages of brain differentiation, and GDI1 knock-out mice showed impairments in hippocampus-dependent mechanisms (P. D'Adamo et al., 1998; Patrizia D'Adamo et al., 2002). ACBD7, an acyl-CoA binding protein, has been detected in mouse hypothalamus, as well as in neurons (Lanfray et al., 2016; Lanfray & Richard, 2017). RIMBP2, which are Rab3-interacting molecule binding proteins, are associated with tethering calcium channels to presynaptic active zones (Coppola et al., 2020; Kaeser et al., 2011). Sox11 proteins have been linked to a wide range of functions in fish development, one of which includes nervous system

development (Hu, Wang, & Du, 2021). Specifically, Sox11 has been detected in the brain and eye of the large yellow croaker, and is associated with nervous system development in zebrafish and orange-spotted grouper (De Martino et al., 2000; Hu et al., 2021; L. Zhang, Zhu, Lin, Zhang, & Zhang, 2010). RASGRF2, a calcium/calmodulin-regulated guanine-nucleotide exchange factor that activates Ras GTPases, has been shown to influence alcohol-induced reinforcement via the mesolimbic dopamine pathway (Stacey et al., 2012). Furthermore, the full list of 77 proteins unique to plasma and brain have been associated with binding and transport, enzymatic activity, synaptic transmission, exchange factors, regulatory functions, signaling, and behaviour. Singh *et al.*, (2010) found that binding, catalytic activity, hydrolase activity and ATPase activity constituted the top functions in proteins detected in zebrafish brain. The preceding study also noted that zebrafish proteins detected in the brain were linked to synapse formation, transport, signal transduction, signaling, and regulation (Singh et al., 2010). As well, Gebriel *et al.*, (2014) noted that zebrafish brain proteins were predominantly involved in binding and catalytic activity (Gebriel et al., 2014). Overall, the brain-specific plasma proteins detected in this study contain a wide-range of biological processes associated with general brain function, demonstrating the ability of plasma to hold a substantial level of information on neurological health.

Plasma-specific proteins

1,990 proteins were identified in plasma, constituting 18% of the total number of proteins identified. In other studies, LC-MS/MS analysis of enriched rainbow trout plasma resulted

in the detection of 1,495 proteins, proteome analysis of zebrafish plasma resulted in a total of 132 proteins, while another study on zebrafish plasma identified 959 plasma proteins which suggests that our plasma proteome method is very comprehensive (Babaei et al., 2013; C. Li et al., 2016; Morro et al., 2020). However, the differences amongst the total number of proteins identified in this study and in literature are likely due to species differences and method of protein identification. For instance, Babaei *et al.*, (2013) detected 15x less plasma proteins, though their study examined plasma proteins in zebrafish, a much smaller fish species with limited sample volume (Babaei et al., 2013). Of note, the preceding study used 50 µg of protein, while our study used 1000 µg of protein per sample. Additionally, De Souza *et al.*, (2009) detected more than 2x the number of gill proteins than our study, though their methods included an additional step of sequential solubilization prior to protein identification, and almost half of their proteins were linked back to single-peptide matches (De Souza et al., 2009).

When comparing our full list of plasma proteins to those found in the Plasma Proteome Database (PPD), a total of 1,337 of our proteins (67%) were matched to PPD proteins (Nanjappa et al., 2014). Based upon the results from the present study, the origin of plasma proteins can be grouped into three major categories, (1) plasma-specific proteins, (2) ubiquitous non-specific tissue proteins, and (3) tissue-specific proteins (Figure 10). The former part of this discussion predominately focused on tissue-specific proteins, which contained evidence surrounding plasma and its capability to hold a substantial amount of information pertaining to multiple organs in the body. Almost half the plasma proteins (43%) originated from multiple organs, which supports the findings of Li *et al.*,

(2016) who compared zebrafish plasma to tissue-specific information found in online databases (C. Li et al., 2016). In that study, they found plasma proteins to be predominantly expressed in the whole body, and that the liver made up the largest proportion of the proteins detected in plasma. In the present study, the gill made up the largest proportion of plasma proteins, followed by heart, liver, brain, and kidney. It is important to note that the preceding study by Li *et al.*, (2016) compared their plasma proteome to online databases that contained tissue-specific information relevant to humans, meaning the gill proteome could not be considered.

Next, we will discuss the full plasma proteome, with a focus on plasma-specific proteins. Comparison of the top 20 most abundant proteins in the full plasma proteome to the plasma-specific proteome showed five common proteins between the two groups: Copb1, Nxf1, Lyl1, Ak9, and Gdf5. Aside from Gdf5, none of the preceding proteins were specifically linked to plasma in literature. GDF5, a growth factor, is a secretion protein with expression in blood cells and has been associated with cartilage and bone development (Abd Elazeem, Abdelaleem, & Mohamed, 2017; Uhlén et al., 2015). The major functions found amongst plasma proteins in this study included immunity, transport, cell communication and adhesion, regulation and response, secretion, and signaling (Figure 26 & 27). Binding, lipid transport, enzymatic activity, and oxygen binding were also linked to protein in literature (C. Li et al., 2016). Many of the functions that are specific to plasma make sense when considering that plasma is a component of blood, and functions like oxygen binding and lipid transport would be expected in blood-plasma.

Overall, the plasma proteome contained proteins and functions specific to plasma, as well as multiple organs, demonstrating its power to be a non-lethal diagnostic option in environmentally relevant species. This is especially significant considering the current thinking is that proteins found in plasma are a result of “tissue-leakage”, which is generally looked at as serving no purpose in circulation (Geyer et al., 2017). Recently, tissue-specific proteins as biomarkers for diagnostics continues to gain momentum in medicinal biology, though the same cannot be said about environmental sciences, possibly due to the lack of information surrounding the proteome of environmentally relevant species (Geyer et al., 2017; Pernemalm et al., 2019; H. Zhang et al., 2007). In our study we see that tissue-specific proteins, or “tissue leakage” proteins, detected in plasma displayed defined physiological functions associated with each given tissue type, such as xenobiotic transport in gills, lipid metabolism in liver, and even artery development in heart, further exemplifying the ability of plasma to serve as a practical biofluid in biomarker discovery in fish. There were also plasma specific proteins that clearly point to a big role in immune function, lipid transport, and cell-to-cell signalling and communication.

Ubiquitous non-specific tissue proteins, which were proteins detected in plasma and multiple tissues, were not analyzed or discussed in this current study, though could be considered as “tissue leakage” proteins which could be eliminated from biomarker studies.

Conclusion

Analysis of the rainbow trout plasma, brain, heart, gill, kidney, and liver proteomes resulted in over 10,000 protein identifications with general functions like binding and catalytic activity to tissue-specific functions like synaptic transmission, and heart development. The plasma proteome contained a wide-range of proteins that could be classified as (1) plasma-specific proteins, (2) ubiquitous non-specific tissue proteins, and (3) tissue-specific proteins. Plasma proteins were compared on a tissue-specific basis, relating both expression and functional patterns found in this study and in the literature, respectively. The information gained in this study can be applied to update proteomic database annotations linked to tissue-specificity in rainbow trout, which is otherwise lacking. This could then help to improve research surrounding biomarker discovery in easily accessible biofluids from fish and wildlife, which in turn can be used to ameliorate environmental health monitoring. Specifically, using plasma to monitor environmental health supports the three Rs of animal research, refinement, reduction, and replacement. Recent evidence from our lab suggests that plasma can safely be collected non-lethally in the field environment (Pollard, 2021). Of note, tissue-specific plasma proteins held a substantial amount of functional information specific to each tissue type examined, signifying the capability of plasma to serve as a useful biofluid in biomarker discovery. Considering most environmentally relevant species lack annotated proteomes, future avenues of research could include repeating this study on other sentinel species, as well as evaluating how the plasma proteome and tissue representation changes under various

conditions, or at different time points, to determine how dynamic and/or static the plasma proteome is, and if certain vertebrate species have different levels of dynamicity.

6. Bibliography

Abd Elazeem, M. I., Abdelaleem, E. A., & Mohamed, R. A. (2017). Genetic influence of growth and differentiation factor 5 gene polymorphism (+104T/C) on the development of knee osteoarthritis and its association with disease severity. *European journal of rheumatology, 4*(2), 98-103. doi:10.5152/eurjrheum.2017.160093

Aedo, J. E., Fuentes-Valenzuela, M., Molina, A., & Valdés, J. A. (2019). Quantitative proteomics analysis of membrane glucocorticoid receptor activation in rainbow trout skeletal muscle. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 32*, 100627. doi:https://doi.org/10.1016/j.cbd.2019.100627

Avarre, J.-C., Lubzens, E., & Babin, P. J. (2007). Apolipocrustacein, formerly vitellogenin, is the major egg yolk precursor protein in decapod crustaceans and is homologous to insect apolipophorin II/I and vertebrate apolipoprotein B. *BMC Evolutionary Biology, 7*(1), 3. doi:10.1186/1471-2148-7-3

Babaei, F., Ramalingam, R., Tavendale, A., Liang, Y., Yan, L. S. K., Ajuh, P., . . . Lam, Y. W. (2013). Novel Blood Collection Method Allows Plasma Proteome Analysis from Single Zebrafish. *Journal of Proteome Research, 12*(4), 1580-1590. doi:10.1021/pr3009226

Babaheydari, S. B., Keyvanshokooh, S., Dorafshan, S., & Johari, S. A. (2016). Proteome changes in rainbow trout (Oncorhynchus mykiss) fertilized eggs as an effect of triploidization heat-shock treatment. *Animal Reproduction Science, 166*, 116-121. doi:10.1016/j.anireprosci.2016.01.010

Bajoghli, B., Ramialison, M., Aghaallaei, N., Czerny, T., & Wittbrodt, J. (2009). Identification of starmaker-like in medaka as a putative target gene of Pax2 in the otic vesicle. *Developmental Dynamics, 238*(11), 2860-2866. doi:https://doi.org/10.1002/dvdy.22093

Baker, M. E. (1988). Is vitellogenin an ancestor of apolipoprotein B-100 of human low-density lipoprotein and human lipoprotein lipase? *Biochem J, 255*(3), 1057-1060. doi:10.1042/bj2551057

Barske, L., Fabian, P., Hirschberger, C., Jandzik, D., Square, T., Xu, P., . . . Crump, J. G. (2020). Evolution of vertebrate gill covers via shifts in an ancient Pou3f3 enhancer. *bioRxiv*, 2020.2001.2027.918193. doi:10.1101/2020.01.27.918193

Baumgarner, B. L., Bharadwaj, A. S., Inerowicz, D., Goodman, A. S., & Brown, P. B. (2013). Proteomic analysis of rainbow trout (Oncorhynchus mykiss) intestinal epithelia: Physiological acclimation to short-term starvation. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 8*(1), 58-64. doi:https://doi.org/10.1016/j.cbd.2012.11.001

Bruslé, J., & Anadon, G. G. (1996). *The structure and function of fish liver*: Routledge

Bryant, N. J., Govers, R., & James, D. E. (2002). Regulated transport of the glucose transporter GLUT4. *Nat Rev Mol Cell Biol, 3*(4), 267-277. doi:10.1038/nrm782

Bury, N. R., Schnell, S., & Hogstrand, C. (2014). Gill cell culture systems as models for aquatic environmental monitoring. *Journal of Experimental Biology, 217*(5), 639-650. doi:10.1242/jeb.095430

Califf, R. M. (2018). Biomarker definitions and their applications. *Experimental biology and medicine (Maywood, N.J.), 243*(3), 213-221. doi:10.1177/1535370217750088

Capilla, E., Díaz, M., Hou, J. C., Planas, J. V., & Pessin, J. E. (2010). High basal cell surface levels of fish GLUT4 are related to reduced sensitivity of insulin-induced translocation

toward GGA and AS160 inhibition in adipocytes. *American journal of physiology. Endocrinology and metabolism, 298*(2), E329-E336. doi:10.1152/ajpendo.00547.2009
Carrier, L., Mearini, G., Stathopoulou, K., & Cuello, F. (2015). Cardiac myosin-binding protein C (MYBPC3) in cardiac pathophysiology. *Gene, 573*(2), 188-197. doi:10.1016/j.gene.2015.09.008
Causey, D. R., Pohl, M. A. N., Stead, D. A., Martin, S. A. M., Secombes, C. J., & Macqueen, D. J. (2018). High-throughput proteomic profiling of the fish liver following bacterial infection. *BMC Genomics, 19*(1), 719. doi:10.1186/s12864-018-5092-0
Chalhoub, G., Kolleritsch, S., Maresch, L. K., Taschler, U., Pajed, L., Tilp, A., . . . Haemmerle, G. (2021). Carboxylesterase 2 proteins are efficient diglyceride and monoglyceride lipases possibly implicated in metabolic disease. *Journal of Lipid Research, 62*, 100075. doi:https://doi.org/10.1016/j.jlr.2021.100075
Chen, Y. H., Pai, C. W., Huang, S. W., Chang, S. N., Lin, L. Y., Chiang, F. T., . . . Tsai, C. T. (2013). Inactivation of Myosin binding protein C homolog in zebrafish as a model for human cardiac hypertrophy and diastolic dysfunction. *J Am Heart Assoc, 2*(5), e000231. doi:10.1161/jaha.113.000231
Chun, R. F., Blatter, E., Elliott, S., Fitz-Gibbon, S., Rieger, S., Sagasti, A., . . . Hewison, M. (2014). Cloning of a functional 25-hydroxyvitamin D-1α-hydroxylase in zebrafish (Danio rerio). *Cell biochemistry and function, 32*(8), 675-682. doi:10.1002/cbf.3071
Consortium, T. U. (2020). UniProt: the universal protein knowledgebase in 2021. *Nucleic acids research, 49*(D1), D480-D489. doi:10.1093/nar/gkaa1100
Coppola, U., Olivo, P., D'Aniello, E., Johnson, C. J., Stolfi, A., & Ristoratore, F. (2020). Rimbp, a New Marker for the Nervous System of the Tunicate Ciona robusta. *Genes, 11*(9), 1006. doi:10.3390/genes11091006
D'Adamo, P., Menegon, A., Lo Nigro, C., Grasso, M., Gulisano, M., Tamanini, F., . . . Toniolo, D. (1998). Mutations in GDI1 are responsible for X-linked non-specific mental retardation. *Nat Genet, 19*(2), 134-139. doi:10.1038/487
D'Adamo, P., Welzl, H., Papadimitriou, S., Raffaele di Barletta, M., Tiveron, C., Tatangelo, L., . . . Toniolo, D. (2002). Deletion of the mental retardation gene Gdi1 impairs associative memory and alters social behavior in mice. *Human Molecular Genetics, 11*(21), 2567-2580. doi:10.1093/hmg/11.21.2567
De Martino, S., Yan, Y.-L., Jowett, T., Postlethwait, J. H., Varga, Z. M., Ashworth, A., & Austin, C. A. (2000). Expression of sox11 gene duplicates in zebrafish suggests the reciprocal loss of ancestral gene expression patterns in development. *Developmental Dynamics, 217*(3), 279-292. doi:https://doi.org/10.1002/(SICI)1097-0177(200003)217:3<279::AID-DVDY6>3.0.CO;2-S
De Souza, A. G., MacCormack, T. J., Wang, N., Li, L., & Goss, G. G. (2009). Large-scale proteome profile of the zebrafish (Danio rerio) gill for physiological and biomarker discovery studies. *Zebrafish, 6*, 229+. Retrieved from https://link-gale-com.uproxy.library.dc-uoit.ca/apps/doc/A209408619/AONE?u=ko_acd_uoo&sid=bookmark-AONE&xid=e1f12ece
Dindia, L. A., Alderman, S. L., & Gillis, T. E. (2017). Data for iTRAQ-based quantification of the cardiac proteome of rainbow trout (Oncorhynchus mykiss) at rest and with exercise training. *Data in Brief, 13*, 32-36. doi:10.1016/j.dib.2017.05.016

Döring, P., Calvisi, D. F., & Dombrowski, F. (2021). Nuclear ErbB2 expression in hepatocytes in liver disease. *Virchows Archiv, 478*(2), 309-318. doi:10.1007/s00428-020-02871-z

Gebriel, M., Prabhudesai, S., Uleberg, K.-E., Larssen, E., Piston, D., Bjørnstad, A. H., & Møller, S. G. (2014). Zebrafish brain proteomics reveals central proteins involved in neurodegeneration. *Journal of Neuroscience Research, 92*(1), 104-115. doi:https://doi.org/10.1002/jnr.23297

Geyer, P. E., Holdt, L. M., Teupser, D., & Mann, M. (2017). Revisiting biomarker discovery by plasma proteomics. *Molecular systems biology, 13*(9), 942-942. doi:10.15252/msb.20156297

Haddad M, & S., S. (2021). Physiology, Lung. . In. Retrieved from StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK545177/

Hall, J. R., Short, C. E., & Driedzic, W. R. (2006). Sequence of Atlantic cod (Gadus morhua) GLUT4, GLUT2 and GPDH: Developmental stage expression, tissue expression and relationship to starvation-induced changes in blood glucose. *J Exp Biol, 209*(Pt 22), 4490-4502. doi:10.1242/jeb.02532

Hu, Y., Wang, B., & Du, H. (2021). A review on sox genes in fish. *Reviews in Aquaculture, 13*(4), 1986-2003. doi:https://doi.org/10.1111/raq.12554

Huang, Z., Gallot, A., Lao, N. T., Puechmaille, S. J., Foley, N. M., Jebb, D., . . . Teeling, E. C. (2016). A nonlethal sampling method to obtain, generate and assemble whole blood transcriptomes from small, wild mammals. *Mol Ecol Resour, 16*(1), 150-162. doi:10.1111/1755-0998.12447

Jones, G., Prosser, D. E., & Kaufmann, M. (2018). The activating enzymes of vitamin D metabolism (25-and 1α-hydroxylases). In *Vitamin D* (pp. 57-79): Elsevier.

Kaeser, P. S., Deng, L., Wang, Y., Dulubova, I., Liu, X., Rizo, J., & Südhof, T. C. (2011). RIM proteins tether Ca2+ channels to presynaptic active zones via a direct PDZ-domain interaction. *Cell, 144*(2), 282-295. doi:10.1016/j.cell.2010.12.029

Kumar, G., Hummel, K., Noebauer, K., Welch, T. J., Razzazi-Fazeli, E., & El-Matbouli, M. (2018). Proteome analysis reveals a role of rainbow trout lymphoid organs during Yersinia ruckeri infection process. *Scientific reports, 8*(1), 13998. doi:10.1038/s41598-018-31982-6

Lambert, E., Coissieux, M.-M., Laudet, V., & Mehlen, P. (2012). Netrin-4 acts as a pro-angiogenic factor during zebrafish development. *The Journal of biological chemistry, 287*(6), 3987-3999. doi:10.1074/jbc.M111.289371

Lanfray, D., Caron, A., Roy, M.-C., Laplante, M., Morin, F., Leprince, J., . . . Richard, D. (2016). Involvement of the Acyl-CoA binding domain containing 7 in the control of food intake and energy expenditure in mice. *eLife, 5*, e11742. doi:10.7554/eLife.11742

Lanfray, D., & Richard, D. (2017). Emerging Signaling Pathway in Arcuate Feeding-Related Neurons: Role of the Acbd7. *Frontiers in neuroscience, 11*, 328-328. doi:10.3389/fnins.2017.00328

Lawrence, M. J., Raby, G. D., Teffer, A. K., Jeffries, K. M., Danylchuk, A. J., Eliason, E. J., . . . Cooke, S. J. (2020). Best practices for non-lethal blood sampling of fish via the caudal vasculature. *Journal of Fish Biology, 97*(1), 4-15. doi:https://doi.org/10.1111/jfb.14339

Lawson, D. E., Fraser, D. R., Kodicek, E., Morris, H. R., & Williams, D. H. (1971). Identification of 1,25-dihydroxycholecalciferol, a new kidney hormone controlling calcium metabolism. *Nature, 230*(5291), 228-230. doi:10.1038/230228a0

Levin, M., Johnson, R. L., Sterna, C. D., Kuehn, M., & Tabin, C. (1995). A molecular pathway determining left-right asymmetry in chick embryogenesis. *Cell, 82*(5), 803-814.
Li, C., Tan, X. F., Lim, T. K., Lin, Q., & Gong, Z. (2016). Comprehensive and quantitative proteomic analyses of zebrafish plasma reveals conserved protein profiles between genders and between zebrafish and human. *Scientific reports, 6*, 24329-24329. doi:10.1038/srep24329
Li, J., & Kültz, D. (2020). Proteomics of Osmoregulatory Responses in Threespine Stickleback Gills. *Integrative and Comparative Biology, 60*(2), 304-317. doi:10.1093/icb/icaa042
Li, S., Li, Z., Sang, C., Zhang, J., Chen, N., & Huang, X. (2018). Glucose transporters in pearl gentian grouper (Epinephelus fuscoguttatus ♀ × E. lanceolatus ♂): Molecular cloning, characterization, tissue distribution and their expressions in response to dietary carbohydrate level. *Aquaculture Research, 49*(1), 253-264. doi:https://doi.org/10.1111/are.13455
Li, X., Zheng, S., & Wu, G. (2020). Amino Acid Metabolism in the Kidneys: Nutritional and Physiological Significance. In (Vol. 1265, pp. 71-95).
Li, Y., Zalzala, M., Jadhav, K., Xu, Y., Kasumov, T., Yin, L., & Zhang, Y. (2016). Carboxylesterase 2 prevents liver steatosis by modulating lipolysis, endoplasmic reticulum stress, and lipogenesis and is regulated by hepatocyte nuclear factor 4 alpha in mice. *Hepatology, 63*(6), 1860-1874. doi:https://doi.org/10.1002/hep.28472
Li, Z. H., Li, P., Sulc, M., Hulak, M., & Randak, T. (2012). Hepatic Proteome Sensitivity in Rainbow Trout after Chronically Exposed to a Human Pharmaceutical Verapamil. *Molecular & Cellular Proteomics, 11*(1). doi:10.1074/mcp.M111.008409
Lian, J., Chen, J., Wang, K., Zhao, L., Meng, P., Yang, L., . . . Zhang, Y. (2018). Alas1 is essential for neutrophil maturation in zebrafish. *Haematologica, 103*(11), 1785-1795. doi:10.3324/haematol.2018.194316
López-Maury, L., Marguerat, S., & Bähler, J. (2008). Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics, 9*(8), 583-593. doi:10.1038/nrg2398
Lu, X., Le Noble, F., Yuan, L., Jiang, Q., De Lafarge, B., Sugiyama, D., . . . Thomas, J.-L. (2004). The netrin receptor UNC5B mediates guidance events controlling morphogenesis of the vascular system. *Nature, 432*(7014), 179-186.
Maki, J. L., & Dickerson, H. W. (2003). Systemic and cutaneous mucus antibody responses of channel catfish immunized against the protozoan parasite Ichthyophthirius multifiliis. *Clin Diagn Lab Immunol, 10*(5), 876-881. doi:10.1128/cdli.10.5.876-881.2003
Martin, S. A. M., Cash, P., Blaney, S., & Houlihan, D. F. (2001). Proteome analysis of rainbow trout (Oncorhynchus mykiss) liver proteins during short term starvation. *Fish Physiology and Biochemistry, 24*(3), 259-270. doi:10.1023/A:1014015530045
Martyniuk, C. J., & Simmons, D. B. (2016). Spotlight on environmental omics and toxicology: a long way in a short time. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 19*, 97-101. doi:https://doi.org/10.1016/j.cbd.2016.06.010
Mathew, J., & Varacallo, M. (2020). Physiology, Blood Plasma. In *StatPearls*. Treasure Island (FL): StatPearls Publishing
StatPearls Publishing LLC.
Morro, B., Doherty, M. K., Balseiro, P., Handeland, S. O., MacKenzie, S., Sveier, H., & Albalat, A. (2020). Plasma proteome profiling of freshwater and seawater life stages of rainbow

trout (Oncorhynchus mykiss). *PloS one, 15*(1), e0227003-e0227003. doi:10.1371/journal.pone.0227003

Nanjappa, V., Thomas, J. K., Marimuthu, A., Muthusamy, B., Radhakrishnan, A., Sharma, R., . . . Prasad, T. S. K. (2014). Plasma Proteome Database as a resource for proteomics research: 2014 update. *Nucleic acids research, 42*(Database issue), D959-D965. doi:10.1093/nar/gkt1251

Nynca, J., Arnold, G., Fröhlich, T., & Ciereszko, A. (2017). Proteomic identification of rainbow trout blood plasma proteins and their relationship to seminal plasma proteins. *PROTEOMICS, 17*(11), 1600460. doi:https://doi.org/10.1002/pmic.201600460

Nynca, J., Arnold, G. J., Fröhlich, T., Otte, K., & Ciereszko, A. (2014). Proteomic identification of rainbow trout sperm proteins. *PROTEOMICS, 14*(12), 1569-1573. doi:10.1002/pmic.201300521

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., . . . Xia, J. (2021). MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic acids research, 49*(W1), W388-W396. doi:10.1093/nar/gkab382

Paulis, L., Fauconnier, J., Cazorla, O., Thireau, J., Soleti, R., Vidal, B., . . . Lacampagne, A. (2015). Activation of Sonic hedgehog signaling in ventricular cardiomyocytes exerts cardioprotection against ischemia reperfusion injuries. *Scientific reports, 5*(1), 7983. doi:10.1038/srep07983

Pernemalm, M., Sandberg, A., Zhu, Y., Boekel, J., Tamburro, D., Schwenk, J. M., . . . Lehtiö, J. (2019). In-depth human plasma proteome analysis captures tissue proteins and transfer of protein variants across the placenta. *eLife, 8*, e41608. doi:10.7554/eLife.41608

Pino, L. K., Searle, B. C., Bollinger, J. G., Nunn, B., MacLean, B., & MacCoss, M. J. (2020). The Skyline ecosystem: Informatics for quantitative mass spectrometry proteomics. *Mass Spectrometry Reviews, 39*(3), 229-244. doi:https://doi.org/10.1002/mas.21540

Purushothaman, S., Saxena, S., Meghah, V., Meena Lakshmi, M. G., Singh, S. K., Brahmendra Swamy, C. V., & Idris, M. M. (2015). Proteomic and gene expression analysis of zebrafish brain undergoing continuous light/dark stress. *J Sleep Res, 24*(4), 458-465. doi:10.1111/jsr.12287

Quan, J., Kang, Y., Li, L., Zhao, G., Sun, J., & Liu, Z. (2021). Proteome analysis of rainbow trout (Oncorhynchus mykiss) liver responses to chronic heat stress using DIA/SWATH. *Journal of Proteomics, 233*, 104079. doi:https://doi.org/10.1016/j.jprot.2020.104079

Roberts, H. E., & Smith, S. A. (2011). Disorders of the respiratory system in pet and ornamental fish. *Vet Clin North Am Exot Anim Pract, 14*(2), 179-206, v. doi:10.1016/j.cvex.2011.03.004

Satoh, T., & Hosokawa, M. (2006). Structure, function and regulation of carboxylesterases. *Chemico-Biological Interactions, 162*(3), 195-211. doi:https://doi.org/10.1016/j.cbi.2006.07.001

Saxena, S., Singh, S. K., Lakshmi, M. G., Meghah, V., Sundaram, C. S., Swamy, C. V., & Idris, M. M. (2011). Proteome profile of zebrafish kidney. *J Proteomics, 74*(12), 2937-2947. doi:10.1016/j.jprot.2011.07.021

Shi, J.-H., Guo, W.-Z., Jin, Y., Zhang, H.-P., Pang, C., Li, J., . . . Zhang, S.-J. (2019). Recognition of HER2 expression in hepatocellular carcinoma and its significance in postoperative tumor recurrence. *Cancer Medicine, 8*(3), 1269-1278. doi:https://doi.org/10.1002/cam4.2006

Singh, S., Rakesh, K., Ramamoorthy, K., Saradhi, A., & Idris, M. (2010). Proteome profile of zebrafish brain based on gel LC-ESI MS/MS analysis. *Journal of Proteomics and Bioinformatics, 3*, 135-142. doi:10.4172/jpb.1000132

Smith, R. W., Wang, J., Bucking, C. P., Mothersill, C. E., & Seymour, C. B. (2007). Evidence for a protective response by the gill proteome of rainbow trout exposed to X-ray induced bystander signals. *PROTEOMICS, 7*(22), 4171-4180. doi:10.1002/pmic.200700573

Stacey, D., Bilbao, A., Maroteaux, M., Jia, T., Easton, A. C., Longueville, S., . . . Consortium, t. I. (2012). RASGRF2 regulates alcohol-induced reinforcement by influencing mesolimbic dopamine neuron activity and dopamine release. *Proceedings of the National Academy of Sciences, 109*(51), 21128-21133. doi:10.1073/pnas.1211844110

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., . . . Lancet, D. (2016). The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinformatics, 54*, 1.30.31-31.30.33. doi:10.1002/cpbi.5

Stott, L. C., Schnell, S., Hogstrand, C., Owen, S. F., & Bury, N. R. (2015). A primary fish gill cell culture model to assess pharmaceutical uptake and efflux: Evidence for passive and facilitated transport. *Aquatic Toxicology, 159*, 127-137. doi:https://doi.org/10.1016/j.aquatox.2014.12.007

Sugawara, T. (2011). Chapter 68 - Screening systems for endocrine disruptors. In R. C. Gupta (Ed.), *Reproductive and Developmental Toxicology* (pp. 893-902). San Diego: Academic Press.

Sullivan, C. V., & Yilmaz, O. (2018). Vitellogenesis and Yolk Proteins, Fish. In M. K. Skinner (Ed.), *Encyclopedia of Reproduction (Second Edition)* (pp. 266-277). Oxford: Academic Press.

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PloS one, 6*(7), e21800. doi:10.1371/journal.pone.0021800

Thibaut, L., apos, Honore, Emilie, F., Beatrice, C., Romain, G., . . . Eva, B.-B. (2019). Are European sea bass as euryhaline as expected? Intraspecific variation in freshwater tolerance. *Marine Biology, 166*. Retrieved from https://link-gale-com.uproxy.library.dc-uoit.ca/apps/doc/A592695586/AONE?u=ko_acd_uoo&sid=bookmark-AONE&xid=c2b2b38a

Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., . . . Narechania, A. (2003). PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res, 13*(9), 2129-2141. doi:10.1101/gr.772403

Thomas, P. D., Kejariwal, A., Guo, N., Mi, H., Campbell, M. J., Muruganujan, A., & Lazareva-Ulitsky, B. (2006). Applications for protein sequence–function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic acids research, 34*(suppl_2), W645-W650. doi:10.1093/nar/gkl229

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., . . . Pontén, F. (2015). Tissue-based map of the human proteome. *Science, 347*(6220), 1260419. doi:doi:10.1126/science.1260419

Vizcaíno, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., . . . Hermjakob, H. (2015). 2016 update of the PRIDE database and its related tools. *Nucleic acids research, 44*(D1), D447-D456. doi:10.1093/nar/gkv1145

Wang, M., Zhang, R., Wang, M., Zhang, L., Ding, Y., Tang, Z., . . . Wang, J. (2021). Genetic Polymorphism of Vitamin D Family Genes CYP2R1, CYP24A1, and CYP27B1 Are

Associated With a High Risk of Non-alcoholic Fatty Liver Disease: A Case-Control Study. *Front Genet, 12*, 717533. doi:10.3389/fgene.2021.717533

Washington Smoak, I., Byrd, N. A., Abu-Issa, R., Goddeeris, M. M., Anderson, R., Morris, J., . . . Meyers, E. N. (2005). Sonic hedgehog is required for cardiac outflow tract and neural crest cell development. *Dev Biol, 283*(2), 357-372. doi:10.1016/j.ydbio.2005.04.029

Watson, R. T., Kanzaki, M., & Pessin, J. E. (2004). Regulated membrane trafficking of the insulin-responsive glucose transporter 4 in adipocytes. *Endocr Rev, 25*(2), 177-204. doi:10.1210/er.2003-0011

Wulff, T., Jokumsen, A., Hojrup, P., & Jessen, F. (2012). Time-dependent changes in protein expression in rainbow trout muscle following hypoxia. *Journal of Proteomics, 75*(8), 2342-2351. doi:10.1016/j.jprot.2012.02.010

Xu, Z., Takizawa, F., Parra, D., Gómez, D., von Gersdorff Jørgensen, L., LaPatra, S. E., & Sunyer, J. O. (2016). Mucosal immunoglobulins at respiratory surfaces mark an ancient association that predates the emergence of tetrapods. *Nat Commun, 7*, 10728. doi:10.1038/ncomms10728

Yu, Y., Wang, Q., Huang, Z., Ding, L., & Xu, Z. (2020). Immunoglobulins, Mucosal Immunity and Vaccination in Teleost Fish. *Front Immunol, 11*, 567941. doi:10.3389/fimmu.2020.567941

Zahn-Zabal, M., Michel, P.-A., Gateau, A., Nikitin, F., Schaeffer, M., Audot, E., . . . Lane, L. (2019). The neXtProt knowledgebase in 2020: data, tools and usability improvements. *Nucleic acids research, 48*(D1), D328-D334. doi:10.1093/nar/gkz995

Zehnder, D., Bland, R., Walker, E. A., Bradwell, A. R., Howie, A. J., Hewison, M., & Stewart, P. M. (1999). Expression of 25-hydroxyvitamin D3-1alpha-hydroxylase in the human kidney. *J Am Soc Nephrol, 10*(12), 2465-2473. doi:10.1681/asn.V10122465

Zhang, H., Liu, A., Loriaux, P., Wollscheid, B., Zhou, Y., Watts, J., & Aebersold, R. (2007). Mass Spectrometric Detection of Tissue Proteins in Plasma. *Molecular & cellular proteomics : MCP, 6*, 64-71. doi:10.1074/mcp.M600160-MCP200

Zhang, J., Lanham, K. A., Peterson, R. E., Heideman, W., & Li, L. (2010). Characterization of the adult zebrafish cardiac proteome using online pH gradient strong cation exchange-RP 2D LC coupled with ESI MS/MS. *Journal of Separation Science, 33*(10), 1462-1471. doi:https://doi.org/10.1002/jssc.200900780

Zhang, L., Zhu, T., Lin, D., Zhang, Y., & Zhang, W. (2010). A second form of Sox11 homologue identified in the orange-spotted grouper Epinephelus coioides: analysis of sequence and mRNA expression patterns. *Comp Biochem Physiol B Biochem Mol Biol, 157*(4), 415-422. doi:10.1016/j.cbpb.2010.09.004

Zheng, D., Kille, P., Feeney, G. P., Cunningham, P., Handy, R. D., & Hogstrand, C. (2010). Dynamic transcriptomic profiles of zebrafish gills in response to zinc depletion. *BMC Genomics, 11*, 548-548. doi:10.1186/1471-2164-11-548

APPENDICES

Appendix A.

Proteomics Acquisition Methods

Acquisition Method Info

Method Name	Agilent_training_peptides_slope4.m
Method Path	D:\MassHunter\Methods\peptides\Agilent_training_peptides_slope4.m
Method Description	

Device List

Multisampler
Binary Pump
Column Comp.
Q-TOF

TOF/Q-TOF Mass Spectrometer

Component Name	MS Q-TOF	**Component Model**	G6545A
Ion Source	Dual AJS ESI	**Stop Time (min)**	No Limit/As Pump
Can wait for temp.	Enable	**Fast Polarity**	False
MS Abs. threshold	500	**MS Rel. threshold(%)**	0.010
MS/MS Abs. threshold	5	**MS/MS Rel. threshold(%)**	0.010

Time Segments

Time Segment #	Start Time (min)	Diverter Valve State	Storage Mode	Ion Mode
1	0	MS	Both	Dual AJS ESI

Time Segment 1

Acquisition Mode AutoMS2

MS Min Range (m/z)	200
MS Max Range (m/z)	3000
MS Scan Rate (spectra/sec)	3.00
MS/MS Min Range (m/z)	50
MS/MS Max Range (m/z)	3000
MS/MS Scan Rate (spectra/sec)	2.00
Isolation Width MS/MS	Medium (~4 amu)
Decision Engine	Native

Ramped Collision Energy

Charge	Slope	Offset
All	4	2

Auto MS/MS Preferred/Exclude Table

Mass	Delta Mass (ppm)	Charge	Type	Retention Time (min)	Delta Ret. Time (min)	Isolation Width	Collision Energy
921.9686	100	1	Exclude	0		Narrow (~1.3 amu)	

Precursor Selection

Max Precursors Per Cycle	10
Threshold (Abs)	500
Threshold (Rel)(%)	0.010
Precursor abundance based scan speed	Yes
Target (counts/spectrum)	25000.000
Use MS/MS accumulation time limit	Yes
Use dynamic precursor rejection	No
Purity Stringency (%)	100.000
Purity Cutoff (%)	30.000
Isotope Model	Peptides
Active exclusion enabled	Yes
Active exclusion excluded after (spectra)	2
Active exclusion released after (min)	0.20
Sort precursors	By abundance only

Static Exclusion Ranges

StartMZ	EndMZ
25	300

Charge State Preference

Selected Charges
2
3
>3

Instrument Parameters

Parameter	Value
Gas Temp (°C)	325
Gas Flow (l/min)	8
Nebulizer (psig)	35
SheathGasTemp	350
SheathGasFlow	11

Scan Segments

Scan Seg #	Ion Polarity
1	Positive

Scan Segment 1

Scan Source Parameters

Parameter	Value
VCap	4500
Nozzle Voltage (V)	1000
Fragmentor	180
Skimmer1	65
OctopoleRFPeak	750

ReferenceMasses

Ref Mass Enabled	Disabled
Ref Nebulizer (psig)	

Chromatograms

Chrom Type	Label	Offset	Y-Range
TIC	TIC	15	10000000
TIC	TIC	15	10000000

Name: Multisampler **Module:** G7167A

Sampling Speed	
Draw Speed	100.0 µL/min
Eject Speed	400.0 µL/min
Wait Time After Drawing	1.2 s
Injection	
Needle Wash Mode	Standard Wash
Injection Volume	2.00 µL
Standard Needle Wash	
Needle Wash Mode	Flush Port
Duration	10 s
High Throughput	
Injection Valve to Bypass for Delay Volume Reduction	No
Sample Flush-Out Factor	5.0
Overlapped Injection	
Overlap Injection Enabled	No
Needle Height Position	
Draw Position Offset	-1.0 mm
Use Vial/Well Bottom Sensing	No
Stop Time	
Stoptime Mode	No Limit
Post Time	
Posttime Mode	Off

Name: Binary Pump **Module:** G7112B

Flow	0.000 mL/min
Use Solvent Types	Yes
Low Pressure Limit	0.00 bar
High Pressure Limit	400.00 bar
Maximum Flow Gradient	100.000 mL/min²
Stroke A	
Automatic Stroke Calculation A	Yes
Stroke B	
Automatic Stroke Calculation B	Yes
Stop Time	
Stoptime Mode	Time set
Stoptime	50.00 min
Post Time	
Posttime Mode	Off

Solvent Composition

	Channel	Solvent 1	Name 1	Solvent 2	Name 2	Selected	Used	Percent (%)
1	A	H20		H20		Ch. 1	Yes	98.0 %
2	B	premixed ACN(95%) - H2O(5%)		ACN		Ch. 1	Yes	2.0 %

Timetable

	Time (min)	A (%)	B (%)	Flow (mL/min)
1	0.00 min	98.0 %	2.0 %	0.100 mL/min
2	2.00 min	98.0 %	2.0 %	0.100 mL/min

	Pressure (bar)
1	400.00 bar
2	400.00 bar

	Time (min)	A (%)	B (%)	Flow (mL/min)
3	27.00 min	60.0 %	40.0 %	0.100 mL/min
4	32.00 min	40.0 %	60.0 %	0.100 mL/min
5	32.01 min	15.0 %	85.0 %	0.100 mL/min
6	37.00 min	15.0 %	85.0 %	0.100 mL/min
7	37.01 min	98.0 %	2.0 %	0.100 mL/min

Name: Column Comp. **Module:** G7116A

Left Temperature Control	
Temperature Control Mode	Temperature Set
Temperature	40.0 °C
Enable Analysis Left Temperature	
Enable Analysis Left Temperature On	Yes
Enable Analysis Left Temperature Value	1.0 °C
Left Temp. Equilibration Time	0.0 min
Right Temperature Control	
Right temperature Control Mode	Temperature Set
Right temperature	40.0 °C
Enable Analysis Right Temperature	
Enable Analysis Right Temperature On	Yes
Enable Analysis Right Temperature Value	0.8 °C
Right Temp. Equilibration Time	0.0 min
Enforce column for run	
Enforce column for run enabled	No
Stop Time	
Stoptime Mode	As pump/injector
Post Time	
Posttime Mode	Off
Timetable	
Valve Position	Position 2 (Port 1 -> 2)
Position Switch After Run	Do not switch

	Pressure (bar)
3	400.00 bar
4	400.00 bar
5	400.00 bar
6	400.00 bar
7	400.00 bar