PROTEOMIC AND LIPIDOMIC ANALYSIS OF MYCOBACTERIOPHAGES ZALKECKS AND POTATOSPLIT

by

Taylor Sorrell

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science in Agricultural and Biological Engineering

School of Agricultural and Biological Engineering

West Lafayette, Indiana

May 2022

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Kari Clase, Chair

School of Agricultural and Biological Engineering

Dr. Jackeline Marmolejo

Proteomics Senior Research Associate

Dr. Stephen Byrn

School of Pharmacy

Approved by:

Dr. Nathan S. Mosier

To my parents, for always encouraging me to believe in myself and pursue my dreams.

To my fiancé, for supporting me and standing by my side through it all.

To my friends, for making my time at Purdue unforgettable.

ACKNOWLEDGMENTS

I would first like to thank Dr. Kari Clase, my committee chair, for her unending support and guidance throughout my time as a master's student and as an undergraduate student. She continues to encourage me to pursue my passions for biotechnology and research. I am ever grateful for the opportunity to complete my master's under her guidance, to work for her, and to learn from her. I would like to thank Dr. Jackeline Marmolejo, my committee member, for her guidance, support, and patience throughout my research project. I would also like to thank the Purdue Proteomics Facility for their expertise and troubleshooting throughout the proteomic aspect of my project. I would like to thank Dr. Stephen Byrn, my committee member, for the support and guidance throughout my time as a master's student. Next, I would like to thank the Metabolite Profiling Facility, specifically Bruce Cooper, for their expertise when processing my lipidomics data.

I would like to thank the College of Engineering and the Department of Agricultural and Biological Engineering for providing me with funding, as a Graduate Teaching Assistant. I enjoyed the opportunity to pass on my passion for biotechnology to younger students and to continue to grow as a leader. I would also like to thank the Biotechnology Innovation and Regulatory Science (BIRS) Center for providing the funding for my research project.

Lastly, I would like to thank my family and friends for their unwavering support throughout my college career. I am extremely grateful for the amazing people and the amount of encouragement I have received throughout graduate school.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Ever since the invention of antibiotics nearly a century ago, the threat of antibiotic resistance has been gradually increasing. As antibiotics are continually prescribed, the rate at which bacteria are becoming resistant to antibiotics is increasing as well. It is projected that antibiotic resistance is one of the largest threats to overall world health, and bacteriophage therapy is one of the leading strategies to combat it. Bacteriophages are viruses that infect and kill specific host bacteria and can potentially be utilized to kill desired bacteria causing infections that are resistant to antibiotics.

The purpose of this research project is to learn more about the bacteriophage-host interaction through mass spectrometry and bioinformatic tools. This is done through the analysis of proteins and lipids that are produced when the bacteriophage infects the host bacteria. The growth curve of a Passage One From Frozen (P1FF) and a Passage Two From Frozen (P2FF) sample of *Mycobacterium smegmatis* was calculated to determine to optimum time for bacteriophage infection. Two bacteriophages were chosen, PotatoSplit and Zalkecks, the *Mycobacterium smegmatis* samples were infected, samples collected, and mass spectrometry performed. A large portion of this research project is based on the analysis of the proteins and lipids that are produced during each bacteriophage's infection. Proteomic and lipidomic strategies can be implemented to understand more about the bacteriophage-host interaction and discover any proteins and lipids that are produced at varying timepoints throughout the inoculation process. Bioinformatic tools can then be used to understand the potential functions of each protein or lipid and potential functions or applications of the bacteriophage in general, including the pathogenicity of each bacteriophage.

Determined from proteomic and lipidomic analysis, a list of all proteins and lipids found within each phage infected sample was made. An important trend discovered is that more phage proteins were expressed at later times during the phage infection – Hour 7 and Hour 10, whereas more bacterial proteins were expressed initially – Hour 0 and Hour 3. A case study to investigate the usage of different intensity types produced from mass spectrometry was completed. Overall, it was determined that both the number of phage proteins and bacterial proteins can differ depending on if LFQ or iBAQ intensity type data was used. Correlation between proteins and lipid ontology classes was performed and shows whether groups of lipids are upregulated or downregulated at

each time point. Understanding the function of lipid ontology groups and the type of regulation provides insight into how the phage or bacteria are potentially using the lipids produced. Some of the main findings include lipids that are involved in bacterial defense mechanisms/energy usage increase over time. Some correlation trends were not consistent across the different bacteriophages, which can be contributed to the different phage life cycles and therefore different phage-host interactions. Further investigation should be performed to determine the specific biological function of proteins and lipids to confidently make claims about potential applications for each phage. Also, further investigation should be performed to understand if the differences in results between bacteriophage PotatoSplit and Zalkecks are due to the varying life cycles.

1. INTRODUCTION

This chapter provides an overview of bacteriophages, protein bioinformatics, and the investigation of mycobacteriophage mass spectrometry research. Included in this chapter is a statement of purpose for the research performed, main research questions addressed, scope of the project, significance of the project, assumptions, limitations, and key terms.

1.1 Statement of Purpose

Bacteriophages are viruses that infect and kill bacteria and have many applications. One of the largest and most important applications of bacteriophages is to combat antibiotic resistance. As more bacteria are becoming antibiotic-resistant every year, a new treatment is becoming more urgent every year. Bacteriophage treatment has been investigated for decades and is deemed as one of the best solutions to treat antibiotic resistance, as they infect and kill bacteria. The purpose of this research is to use cutting-edge technology to determine and analyze the proteins and lipids that are products of the bacteriophage-host interaction. Figure 1-1 is a process flow diagram of the proteomic and lipidomic experiments that are performed in this research project.

Initially, the growth curves of multiple passages from frozen *Mycobacterium Smegmatis* were measured to determine the optimum phage infection time point. Bacteriophages PotatoSplit and Zalkecks were chosen for this research project. The growth curves of each bacteriophage were recorded to determine the exponential and stationary stages, and thus determine the time points to collect samples. Experiments were repeated but inoculating with the phage at the determined time and collecting samples at the determined times for both protein and lipid analysis. Mass spectrometry was run on all samples to determine the proteins and lipids present in the solution and produced by each phage during infection. Data analysis of the data collected on the proteins and lipids produced during bacteriophage infection is the basis for this study. This data gives insight into the bacteriophage-host interaction and potentially sheds light on phage applications (Hatfull et al., 2008).

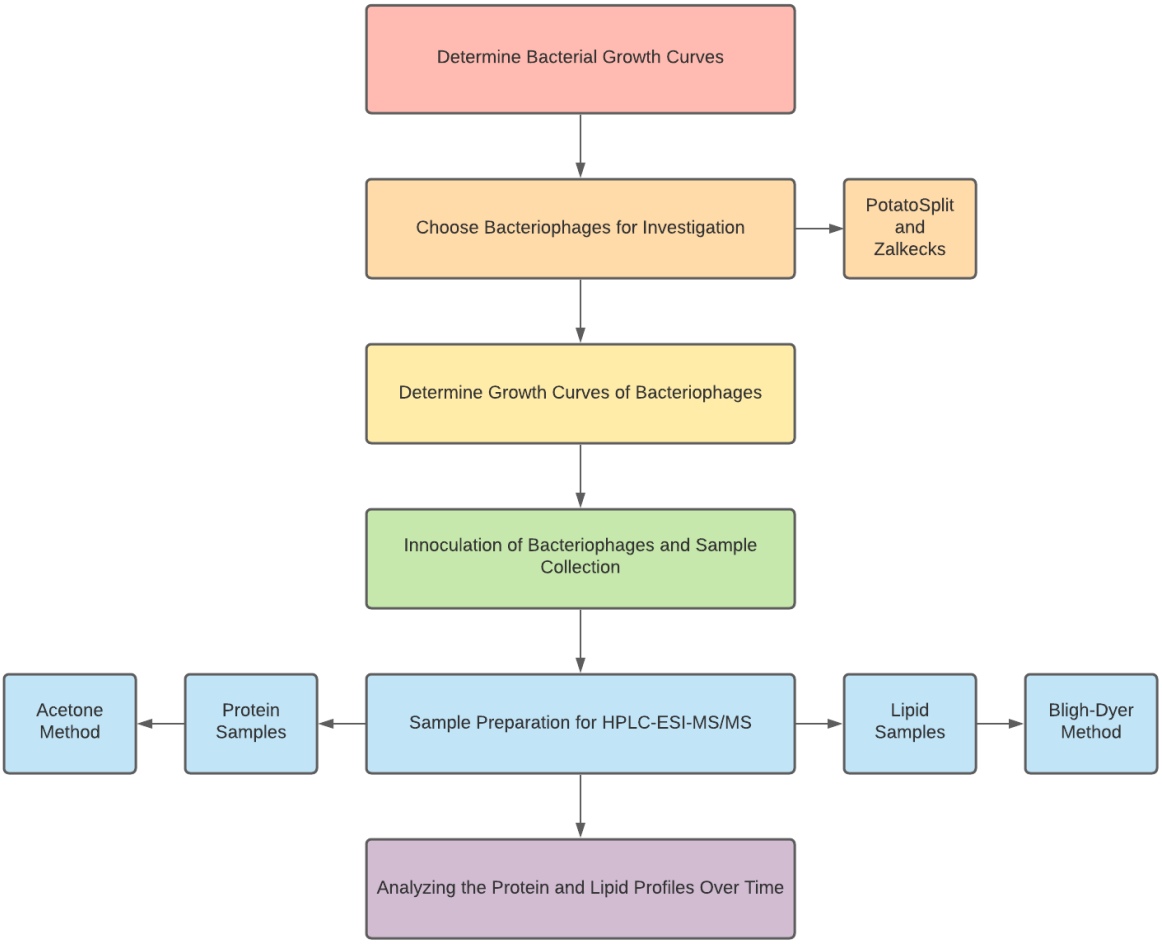


Figure 1-1. Process Flow Diagram of the Proteomic and Lipidomic Experiments in this Research Project

1.2 Research Questions

- Are significant proteins and lipids being produced at different timepoints during bacteriophage infection?
- Can proteins and lipids discovered through mass spectrometry be linked to biological function or applications of phages?
- Does the lifecycle of a phage impact the proteins and lipids produced during phage-host infection?

1.3 Scope

The scope of this project is to use high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry (HPLC-ESI-MS/MS) and various data analysis tools to analyze proteins and lipids produced by bacteriophages during bacterial-host infection. The goal was to determine characteristics of the bacteriophage-host interaction that can be linked to therapeutic or other applications by characterizing proteins and classifying lipids. MetaboAnalyst was used to determine statistical analysis of both the lipid and protein data collected. Proteins that were deemed to be significant were further studied using DAVID, a functional annotation bioinformatics microarray analysis program.

1.4 Significance

Antibiotic resistance continues to spread, and it is estimated by the Center for Disease Control that by the year 2050 more people will die from antibiotic resistance than all forms of cancer (CDC). Bacteriophages, viruses that infect and kill bacteria, are one of the leading solutions to antibiotic resistance. Although bacteriophages are a promising solution, there is still much unknown about them. Prior to acceptance as a safe and effective solution, the make-up and function of the entire genome of each phage need to be researched and understood. One way to gain extensive knowledge of the bacteriophage is through proteomic and lipidomic studies. Understanding the proteins and lipids that are produced during phage infection can provide clarification on each bacteriophage's function and the pathogenesis of its infection of bacteria. Research is being performed on the effectiveness and breadth of bacteriophages to determine their function and potential applications.

1.5 Assumptions and Limitations

Assumptions:

- The six protein samples that had low intensities are correct and reliable after they were reprocessed through mass spectrometry.
- Plastic contamination in protein and lipid results are due to using multiple brands/types of micropipette tips during sample prep.
- Mycobacterium smegmatis and mycobacteriophage samples are pure and do not contain biological contaminants.
- The lipid database reliably matches the MS2 spectra of the lipids in the bacterial host and bacteriophage-treated samples.

Limitations:

- The phage-host interaction was only recorded up to hour 10. There may be significant findings after hour 10 but these cannot be inferred.
- Limited functional analysis can be completed on the identified proteins and lipids due to the discovery aspect of this project and having limited resources available.
- MS-DIAL was used to identify the lipids in the mass spectrometry samples.

Delimitations:

- MetaboAnalyst was the only program used to calculate statistical significance in the protein and lipid mass spectrometry samples.
- Only the program DAVID was used for protein pathway analysis.

2. LITERATURE REVIEW

This chapter will review relevant literature to provide a background of information useful to the project and research goals.

2.1 Bacteriophages

2.1.1 Bacteriophage Structure

Bacteriophages are viruses that infect and kill a host bacterium (Kutter et al., 2010). Although there are many types and shapes of bacteriophages, they can be broken down into similar structures. All bacteriophages have a capsid head that contains genetic information, a tail that acts as a pipe to securely transport DNA into the host upon infection, and tail fibers that aid in attaching to the host, as shown in Figure 2-1 (White & Orlova, 2019). The tail fibers have host recognition mechanisms allowing bacteriophages to be host-specific (White & Orlova, 2019).

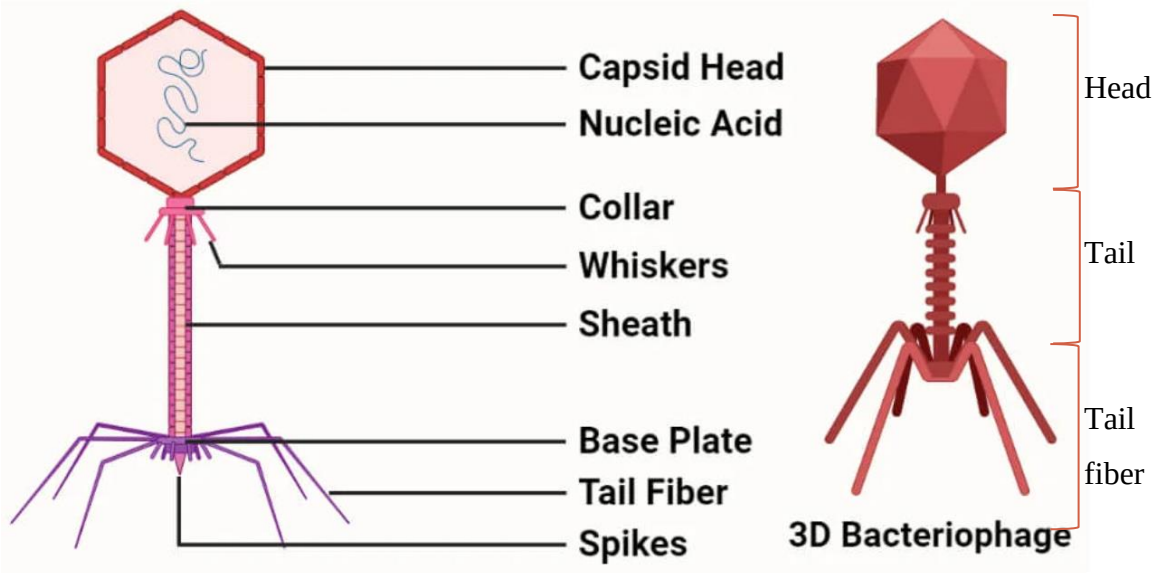


Figure 2-1. General structure of a bacteriophage (Sapkota, 2020)

Mycobacterial diseases, such as Mycobacterium tuberculosis (Tuberculosis) and Mycobacterium lepare (leprosy), have been devastating to the human population, which is what gives rise to the

importance in studying Mycobacteriophages, or bacteriophages that infect Mycobacterium (Hatfull, 2018a). There are three morphotypes of bacteriophages – *Siphoviridae*, *Myoviridae*, and *Podoviridae* – although only *Siphoviridae* and *Myoviridae* are present in mycobacteriophages due to physical blockage of the bacterial cell wall (Hatfull, 2018a). The defining characteristic between the three morphotypes is the length and contractibility of the tail.

2.1.2 Bacteriophage Life Cycles

There are two modes of replication for bacteriophages: the lytic life cycle and the lysogenic life cycle. Both of these life cycles are demonstrated in Figure 2-2 (Boundless, 2020).

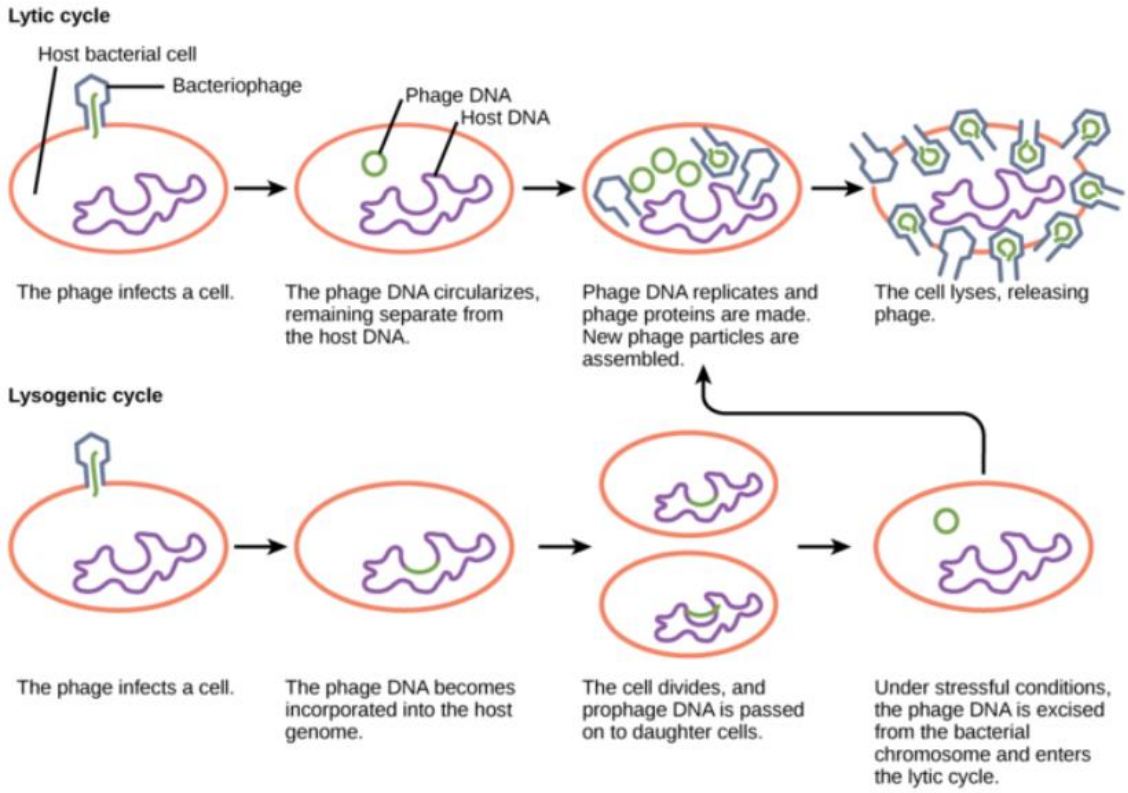


Figure 2-2. Diagram of Lytic and Lysogenic life cycles (Boundless, 2020)

The lytic cycle is named after the term *lysis*, or rupturing of a membrane, because that is the mechanism this type of phage uses. The phage infects the host but remains circular to prevent detection from the host. The phage then replicates inside the cell, reassembles into multiple functioning phages, and lyses the cell wall to release all the phage particles (White & Orlova,

2019). The lysogenic pathway is slightly different because it doesn't always lyse the bacteria membrane to release phage particles. The bacteriophage infects the host, and the phage DNA is incorporated into the host DNA rather than circularizing (Sapkota, 2020). When the cell divides there will now be phage DNA in each copy. Occasionally, often under stressful conditions, the phage DNA will be excised from the host chromosome and enter the lytic cycle where it will rupture the bacterium to produce many more lysogenic phages (Boundless, 2020).

2.1.3 Bacteriophage Clusters

Bacteriophages are grouped into "clusters" based on their overall nucleotide sequence similarity and other defining characteristics (Hatfull, 2018b). Bacteriophages that do not fit into any cluster are considered "singletons"(Hatfull, 2018a). As more bacteriophages are sequenced, the size and number of clusters continue to grow, and singletons can be placed into these new clusters. As a part of the HHMI Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program, Purdue University students find and archive mycobacteriophages.

Cluster A3:

Cluster A is mainly comprised of mycobacteriophages, including PotatoSplit found at Purdue University, and is the largest cluster of phages (Hatfull, 2018a). This cluster is then broken down into subclusters, such as cluster A3 which PotatoSplit is in. Phages found in cluster A are known to be temperate siphoviridae that form large turbid plaques (Phages DB, n.d.-a). A siphoviridae can be characterized by a flexible short to medium-length tail (Phages DB, n.d.-a). One defining characteristic of Cluster A phages, especially A2 and A3, is that they are known to host a variety of bacteria, allowing for different applications (Guerrero-Bustamante et al., 2021). The unique characteristic of having such a large host range is one of the main reasons that Cluster A phages are important to study.

Cluster C1:

All mycobacteriophages that are determined to be Myoviridae are classified within Cluster C (Hatfull, 2018a). Myoviridae, unlike all other mycobacteriophages, have a contractile tail(Hatfull, 2018a). Cluster C has been experimentally determined to be lytic, due to many failed attempts to isolate a lysogen (Phages DB, n.d.-b). Zalkecks is a mycobacteriophage that was isolated at Purdue University and is classified within Cluster C1, a subcluster of Cluster C. The genomes of Cluster C1 mycobacteriophages show a potential common dsDNA packaging mechanism through direct terminal repeats (Oliveira et al., 2019). C1 mycobacteriophages are said to have broad-host ranges and therefore have the potential to be therapeutics (Oliveira et al., 2019).

2.2 Applications and Uses of Bacteriophages

2.2.1 Agriculture and Food Industry

Phage are known to have many applications, such as in the food industry, agriculture, pharmaceuticals, and much more. Scientists all around the world are turning to phage research for potential ideas when it comes to many hot topics and pressure points in today's world. One example of this includes the agriculture and food industry (Mahony et al., 2020). Bacteriophages or phage by-products can potentially be applied as natural alternatives to current food preservation techniques (Mahony et al., 2020). There are currently commercial products, such as Phage Guard S, E and Listex^{TM}, which claim to eliminate Salmonella, E. Coli, and Listeria, respectfully, on food products (Mahony et al., 2020). As more phages are discovered, there is potential for an application to control microorganisms that cause food to spoil, eliminate food-borne pathogens, and aid in food safety and preservation overall.

Not only can phages decrease the bacterial load of food and vegetable products, but they can be used to decrease the bacterial load in animals before they reach slaughterhouses. There are several known bacteria that live in animals that can cause harm to humans if consumed. An example of this would be *Campylobacter,* which is known to live in the intestinal tract of poultry (Nagel et al., 2016). Numerous studies have shown that supplementing poultry feed with *Campylobacter* phages can be an effective biocontrol strategy, as it largely decreases the bacterial levels in the birds (Nagel et al., 2016). There needs to be continued research for large-scale phage

application, especially in agriculture, to ensure that there are no environmental impacts and to minimize any chances of phage resistance.

One of the major benefits of using phage in the agriculture and food industries is the decrease in waste. Not only does more product yield mean less food waste and less environmental impact but it also decreases financial loss. In India alone, $1 billion are lost annually due to decreased milk yield caused by Bovine Mastitis (Nagel et al., 2016). Bovine Mastitis is caused by bacteria, which can potentially be controlled by phage applications.

2.2.2 Antibiotic Resistance

According to the CDC, "Antibiotic resistance happens when germs like bacteria and fungi develop the ability to defeat the drugs designed to kill them. Infections caused by antibiotic-resistant germs are difficult, and sometimes impossible, to treat." According to the World Health Organization (WHO), resistance in common bacteria has reached alarming levels in some parts of the world. Many scientists consider antibiotic resistance to be the biggest current threat to global health (Gordillo Altamirano & Barr, 2019). Phage therapy was introduced nearly a century ago, but the discovery of antibiotics halted phage therapy research. As antibiotic resistance is growing, the use of bacteriophages is one of the leading strategies to combat it (Gordillo Altamirano & Barr, 2019).

An antibiotic is a chemical that is designed to selectively disrupt certain bacterial processes whereas a phage is a virus that uses a bacterium to reproduce (Gordillo Altamirano & Barr, 2019). Although they have very different modes of action, they can yield similar results. Phage therapy only impacts the pathogenic bacteria at hand while antibiotics can have side effects mostly seen as microbiome damage (Gordillo Altamirano & Barr, 2019). The narrow host range of bacteriophages can be viewed as an advantage and a disadvantage. The specificity for a target bacteria host will reduce damage to the patient, but the narrow host range also minimizes the applications of the bacteriophage (Haq et al., 2012). Bacteria can become resistant to phage infection, similar to antibiotics, but the bacteriophage virus will naturally adapt to overcome the resistant bacteria (Haq et al., 2012).

2.2.3 Medical/Therapeutics

The definition of phage therapy is the therapeutic use of bacteriophages to treat pathogenic bacterial infections. Phage therapy is restrained to lytic phages only because lysogenic phages have a lower killing capacity and there are possible harmful effects of lysogenic conversion (Gordillo Altamirano & Barr, 2019). Continued research is needed to determine any "No Known Function" genes in a phage before infecting the host. Unknown function genes could potentially be dangerous when phage DNA is injected and expressed in the host cells. Genetic research of potentially therapeutic phages is necessary prior to infection, to ensure the safety of phage therapy (Gordillo Altamirano & Barr, 2019).

Currently, Georgia, Poland, and Russia have approved the use of phage therapy to treat infectious diseases (Guo et al., 2020). Phage therapy has entered Phase III clinical trials in several other countries as well (Guo et al., 2020). One of the major downsides to phage therapy is that each phage has such a limited host range. As more phages are discovered, sequenced, and tested every year more applications and treatments will be discovered.

2.2.4 Previous Studies:

Phages have been used in other countries for over a century. A study took place in the 1960s, where phages were tested for the prevention of *Shigella* infections in Georgia (Nagel et al., 2016). There are roughly 163 million cases a year with over 1 million deaths a year, due to *Shigella* infections (Kutter et al., 2010). Over 30,000 children participated in the study, either receiving a dosage of phage or a placebo. There was a statistically significant reduction in the *Shigella* infection rate in the phage-treated group (Kutter et al., 2010). *Shigella,* like other infection-causing bacteria, show signs of antibiotic resistance which increases motivation to study phages like the one in this study (Nagel et al., 2016).

2.3 Manufacturing

2.3.1 Scale up

As bacteriophages are studied and undergo clinical trials, it is important to understand how to optimize the manufacturing procedures to scale the process to meet demand and ensure it is a cost-effective treatment. A therapeutic is rendered useless if it is unable to be produced it at a large scale. A majority of current bacteriophage research is performed on the 1-20mL laboratory scale in batch reactions (Ali et al., 2019). Research on potential therapeutic bacteriophage cocktails but at a large scale is currently underway. A majority of current bacteriophage research is performed on the 1-20mL laboratory scale in batch reactions (Ali et al., 2019). Research on potential therapeutic bacteriophage cocktails has been scaled-up to be produced in 50L batch bioreactors, but this is the largest volume used in bacteriophage research to date (Ali et al., 2019). There are proposed ideas to adopt existing bioprocess engineering methods, such as those used to produce other biotherapeutics, for phage production (Mancuso et al., 2018). If these established procedures are to be used, it is important to consider the major differences between bacteriophage and the previous therapeutic that may cause implications (Mancuso et al., 2018).

2.3.2 Kinetics and Process

Increasing the scale of bacteriophage production or infection is not a linear process, as there are different limiting variables at each stage of the process. Some of these limiting variables can be considered in two types: key process input variables (KPIV) and key process output variables (KPOV) (Ali et al., 2019). Both of these variables need to be identified to optimize the manufacturing process and have control over the final product (Ali et al., 2019). Another important factor in phage production is the adsorption rate and ability of each phage to the host cells. An overall understanding of the reaction kinetics will be required by regulatory agencies, such as the FDA, to ensure quality control before being approved as a pharmaceutical (Mancuso et al., 2018).

Not only is it essential to determine the dynamics of phage manufacturing as mentioned above, but it will eventually become necessary to consider the type of reaction process used. In order to produce large enough quantities at a low enough cost for bacteriophage to be used as a wide-spread therapeutic, the development of a continuous production procedure will be necessary (Mancuso et al., 2018).

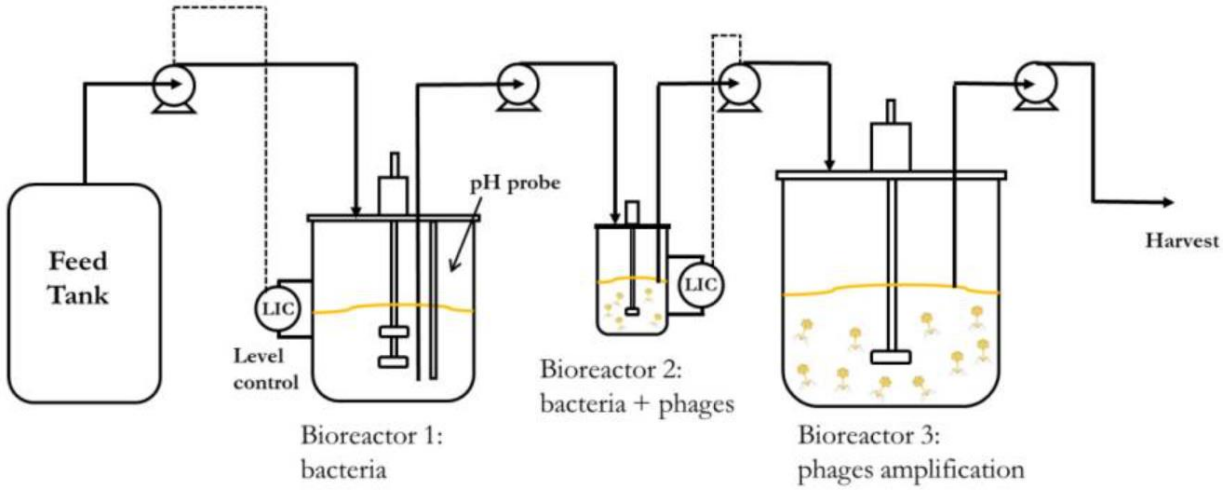


Figure 2-3. Proposed Continuous Bacteriophage Production Process Diagram (Mancuso et al., 2018)

Mancuso proposed a continuous production model, as shown in Figure 2-3, that implements a continuous feed tank and multiple bioreactor vessels to achieve continuous bacteriophage production. The first bioreactor has a constant working volume, the second has a variable volume to alter the dilution rates with a constant feed rate from the tank, and the third was operated in a semi-batch reactor (Mancuso et al., 2018).

2.3.3 Adsorption

Bacteriophage adsorption is the mechanism by which a phage particle attaches to the bacterial host and initiates the infection process. A bacteriophage can recognize a host through interactions between its' binding proteins and receptors on the bacterial cell surface (Bertozzi Silva et al., 2016). The adsorption process includes three main stages: initial contact, reversible binding, and irreversible attachment (Bertozzi Silva et al., 2016). After coming into contact with a host cell, the bacteriophage reversibly binds to the bacteria. This process is reversible for the scenario when the phage does not identify the receptor needed to determine a host match, then it will desorb and continue to search until it finds a correct host (Bertozzi Silva et al., 2016). Lastly, the irreversible attachment step occurs when the phage-binding domain recognizes the bacterial receptor and connects causing an irreversible enzymatic cleavage (Bertozzi Silva et al., 2016). This irreversible attachment triggers a series of conformational changes in the phage molecule, leading to the release of phage DNA into the host cell (Bertozzi Silva et al., 2016).

Phage adsorption is one of the most limiting variables of phage therapeutics or potential application because the ability of a phage to adsorb and the rate at which the phage adsorbs are critical to phage infection dynamics (Storms et al., 2010). Experiments can be performed to calculate the phage adsorption rate and efficiency of adsorption. “Adsorption efficiency” is a term used to quantify the adsorption capability of each individual phage and can be determined by the fraction of a phage population that irreversibly binds to a host cell compared to those that are free in the solution (Storms et al., 2010).

Adsorption rates can be impacted by factors other than phage characteristics, including several environmental factors. Adsorption rate tends to increase with a decrease in replication time in the bacterial host, which can be created using a richer media to increase cell growth (Hadas et al., 1997). Similarly, the adsorption rate increases with an increase in the total surface area of the infected cells in the media (Hadas et al., 1997). As there are more or larger bacterial cells in the media, there is an increased chance of a phage attaching to a bacterium and recognizing a receptor.

2.4 Omics

2.4.1 Proteomics of Mycobacteriophages

The goal of proteomics is the determine gene and cellular function on the protein level (Aebersold & Mann, 2003). Information on proteins produced by bacteriophages is found using bioinformatic tools such as DNA Master and PECAAN (Phages DB, n.d.-b). These programs draw conclusions to discern gene function by comparing it to the known coding regions of other phages. There are many databases that collect and publicly share phage information, such as the Phage Data Base, pFam, Conserved Domain, SCOPe, and more. Using these programs alone renders most of the gene functions as unknown.

Mass spectrometry is a powerful technology that can identify and precisely quantify thousands of proteins from samples, which aids in the identification of proteins (Aebersold & Mann, 2003). Using mass spectrometry can accurately identify proteins that can confirm the annotation or fill in the unknown areas in the annotation produced above. As mass spectrometry techniques improve and the field of proteomics advances, more proteins produced from phage infection will be discovered (Aebersold & Mann, 2003). The number and safety of bacteriophage

application continue to increase as more functions of proteins in bacteriophages are discovered and analyzed confidently.

2.4.2 Lipidomics of Mycobacteriophages

Lipidomics is the term used for the identification and quantification of lipids. Lipidomics is a drastically increasing research area due to its impact on metabolism, cancer, and disease research (DJ et al., 2017). A disruption of lipid metabolic enzymes and pathways are commonly linked to human diseases in many genetic studies (Wenk, 2010). In recent years, the fields of genomics and proteomics are exploded but lipidomics has not advanced as much, due to the lack of technology for analysis and lipid complexity (Wenk, 2005). As seen in Figure 2-4, the number of citations for the lipidomics field is drastically smaller than any other omics-related field, to date (Wenk, 2010). Lipids are not genetically encoded, like proteins are, so lipidomics requires much more extensive research and technology to gain similar information to other omics fields (Wenk, 2010).

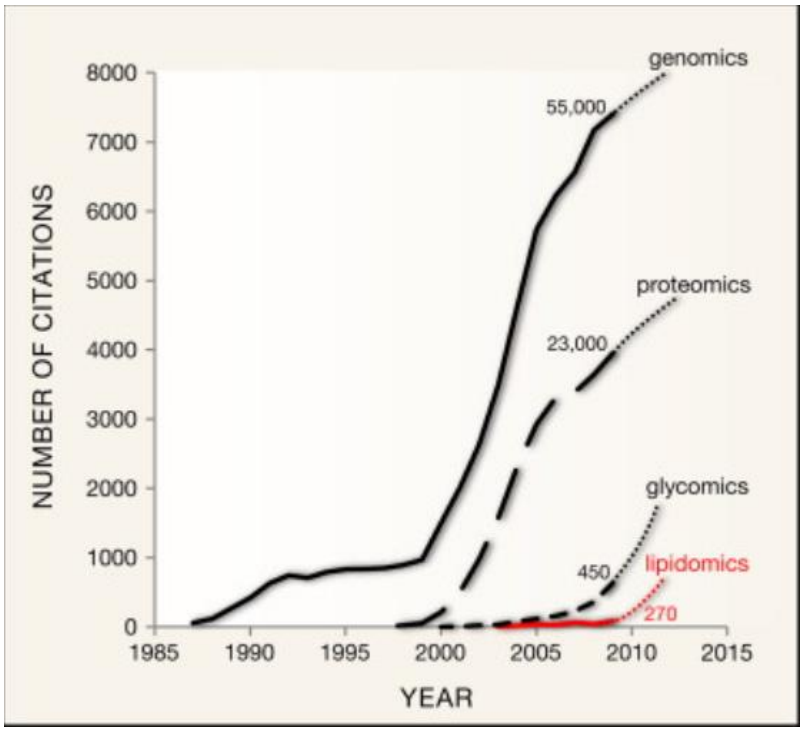


Figure 2-4. Plot showing the number of citations per omics-based research up to the year 2015 (Wenk, 2010)

Using the same mass spectrometry technology as proteomics, understanding of lipid components, such as lipid-based biomarkers, has significantly increased in recent years. The use of mass spectrometry to analyze lipid compounds was first reported in the 1990s, which is much later than similar fields (Wenk, 2010). Information on biomarkers and mechanisms of lipid pathways aids in the development of therapeutics because lipids are linked to many diseases (DJ et al., 2017). There is a large chemical diversity of lipids making it difficult or nearly impossible to analyze the entire lipidome in a single run (Wenk, 2010). Untargeted mass spectrometry is often used to return any potential hits because so little is known about lipids.

2.4.3 Mycobacterium Smegmatis

Mycobacterium smegmatis is a non-pathogenic and rapidly growing species, but specifically, strain mc^2 155 has been utilized as a molecular analysis tool (Fujiwara et al., 2012). This strain is a mutation of the original wild type but has many characteristics that make it ideal for genetic manipulations, such as a loss of the cell-clumping properties and a high rate of transformation (Etienne et al., 2005).

2.4.4 Bacteriophage-Host Interactions

Key characteristics of the relationship between a bacteriophage and its' host are determined by their interactions, including phage applications; phage life cycles; and potential omics analyses. As both bacteriophages and their hosts evolve over time, so does their relationship. Bacteria are constantly mutating and evolving new immune mechanisms due to the pressure from foreign invaders, including bacteriophages (Hampton et al., 2020). These immune mechanisms can defend against foreign invaders at different stages in their life cycle to increase the efficiency of immunity, as shown in Figure 2-5 (Hampton et al., 2020). Bacteriophages also evolve to counteract changes in the bacteria. Understanding this ever-changing relationship is key to understanding bacteriophage-based therapies, manufacturing, research, and biotechnology tools (Hampton et al., 2020).

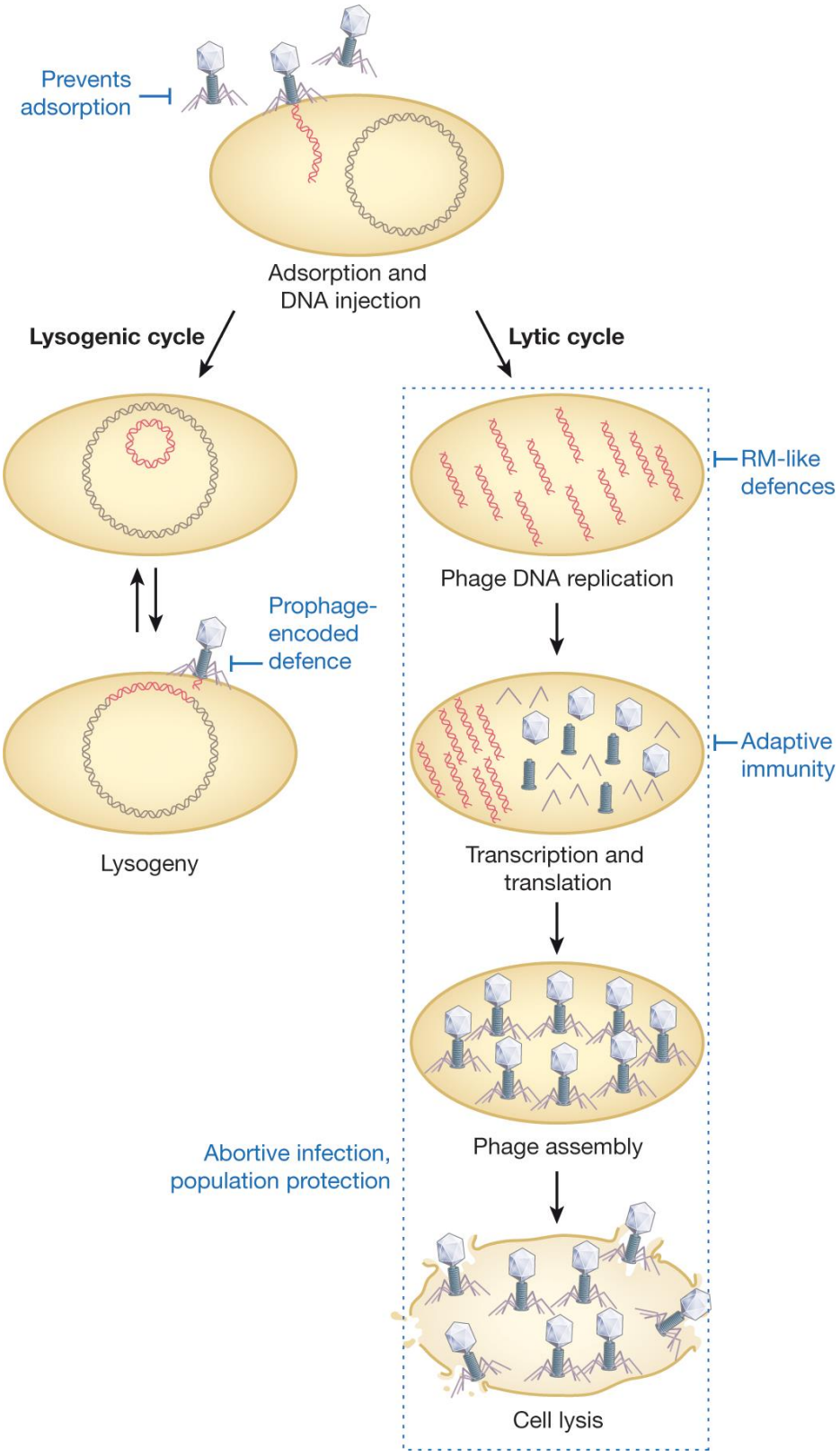


Figure 2-5. Diagram of anti-phage mechanisms of bacteria at different stages in the life cycle (Hampton et al., 2020)

Since the first step in infection is adsorption, bacteria have adsorption resistance techniques that prevent the bacteriophage from approaching and binding to the cell surface (Samaddar et al., 2016). There are many techniques that the bacteria use intracellularly to defend against bacteriophages. Some of these techniques include prophage-encoded defense, RM-like defense, receptor availability, abortive infection, and adaptive immunity (Hampton et al., 2020). RM-like defense, restriction-modification defense, allows the bacteria to detect restriction enzyme sequences to destroy any DNA that is inserted into the cell (Jacobs-Sera et al., 2012). This mechanism has a downside, if a large number of restriction enzyme recognition sequences are being targeted, the bacteriophage DNA will modify itself to become undetected by this system (Abedon, 2012). If all other mechanisms fail and the bacteriophage infects the host, the bacteria can use an abortive infection mechanism, the same mechanism that is used in apoptosis, to kill the phage but in the process kill itself (Abedon, 2012). With all of these mechanisms combined, the bacterial host has a wide range of strategies to prevent bacteriophage infection at all stages of the life cycle.

2.5 Phage Genome and Bioinformatics

2.5.1 Phage Genome

Bacteriophages were of the first complete genomes to be sequenced due to their ease of isolation and relatively small size (Hatfull et al., 2008). A high degree of genetic diversity and novel genetic sequences within phage genomes leads to interest in early evolutionary origins and unexplored genes (Hatfull et al., 2008). Genomes can be annotated efficiently and effectively using bioinformatics software that can determine gene location and potential protein functions. As technology continues to improve, more information will be discovered about phage genes and their accompanying functions which will lead to more applications of bacteriophages.

Prior to bioinformatic improvements, a proteins' structure would be determined experimentally which would lead to an explanation of its' functional properties (Loewenstein et al., 2009). Because structure and sequence tend to lead to functional properties, much can be found about an unknown protein's function based on homology to known proteins. Many bioinformatics programs are based upon the idea that statistics can be used to determine if two proteins are similar

enough to be considered homologous. The protein structure predicted from the amino acid sequence may not always align with the experimentally observed structure (Matthews, 1975).

2.5.2 Determining Homology

Bacteriophage discovery and research is a relatively new field. Many databases and programs have little information that is directly sourced from bacteriophages, but they do have substantial information on bacterial proteins, which can often translate to baseline information for phage research.

As mentioned above, homology to known proteins can dictate information about an unknown protein so multiple programs are based on this theory. Programs such as PSI-BLAST use an iterative approach to detect distant relationships between proteins (Söding et al., 2005). The iteration of BLASTp in this program allows for a more accurate and more comprehensive result than just using BLASTp alone.

Another algorithm that determines homology between two proteins is called the Hidden Markov Model (HMM). This algorithm is the foundation for or has been implemented into multiple programs, such as HHPred, HMMER, and more. HHPred is the standard program that is used in the SEA-PHAGES coursework at Purdue University for bacteriophage annotation. HHPred was the first program to implement a pairwise comparison of HMMs and overall is a server to determine protein homology and structure prediction (Söding et al., 2005). HMMER also predicts homology by searching widely used sequence databases and implementing HMMs (Potter et al., 2018).

2.5.3 Predicting Structure

Many programs are used only to predict the structure of a protein, such as I-TASSER, PyMOL, Phyre2, and more. I-TASSER works by making 3-D models of the protein and comparing them against known proteins in the Protein Data Base (PDB). PyMOL is deemed as a cross-platform molecular graphic tool, but it also makes 3-D visuals of proteins and other biological entities. Similarly, Phyre2 is a suite of tools that can be used to predict and analyze protein structures and functions. Phyre2 can build 3-D models, predict binding sites, and analyze the effect of amino acid variants in the genome (Kelley et al., 2015). Some of these programs are much more

time-intensive and can take days or even weeks to get results for one protein. When analyzing an entire genome, it is unrealistic to use these intensive programs for every protein, but they can provide much greater detail for specific proteins of interest.

3. METHODOLOGY

3.1 Growing M. smegmatis Cell Cultures and Determining M. Smegmatis Growth Curves

A sample from the -80°C frozen stock of Mycobacterium smegmatis strain mc^2 155 was reconstituted by streaking on an LB agar plate. The streaked plate was incubated for 60 hours at 37°C, before a single colony was picked to form the Passage 1 From Forzen stock of *Mycobacterium smegmatis* strain mc^2 155 was reconstituted by streaking on an LB agar plate. The streaked plate was incubated for 60 hours at 37°C before a single colony was picked to form the Passage One From Frozen stock (P1FF). P1FF was made by inoculating 7H9 liquid medium with the picked plaque and was incubated for 72 hours in a shaking incubator set to 37°C and 250 RPM. The 7H9 medium used includes 50 mL of 7H9 Middlebrook broth with 0.05% Tween80, 1mM calcium chloride, 10% AD supplement, 0.02% glycerol, 50 ug/ml carbenicillin, and 10 ug/ml cycloheximide. Next, a 50mL Passage Two From Frozen (P2FF) stock was made in a 250mL flask and incubated in the same shaking incubator for 72 hours at 37°C with 250 RPM. P1FF medium was diluted in a 1:1000 ratio with 7H9 liquid medium, without Tween80, to form the P2FF stock.

The growth curve of the M smegmatis cultures was determined by measuring the OD600 value of the P2FF stock every couple hours a a total of 52.5 hours. The growth curve for M. Smegmatis can be shown by plotting the OD600 values over time. After inoculating with P1FF, the P2FF culture takes 35 hours before it will be used during the bacteriophage experiments, to gather protein and lipid samples. The P1FF and P2FF stocks will be created using the same recipe and process for the entire research project, including the inoculation of bacteriophages.

3.2 Preparation of Mycobacteriophage Lysates

Two bacteriophages, from diverse clusters, were chosen to infect *M. smegmatis* cell cultures. Both of these bacteriophages were isolated at Purdue University and are of interest because of potential applications due to proteins in their genome. A small sample of each phage's lysate is kept in the -80C. These frozen stocks were used to streak LB agar plates, which were then incubated at 37C for 48 hours. Serial dilutions were performed on each phage by picking a single plaque from each LB agar plate. To do this, 10 microliters of each dilution are added to a bacterial

culture tube that contains 250 microliters of the P2FF *M. smegmatis* sample. After 10 minutes, 3mL of top agar is added to the tube and then poured immediately onto a labeled LB agar plate. Once all of the dilution samples have been plated, they are incubated at 37C for 48 hours. If a webbed plate is formed from the dilution process, it is flooded with 5 mL of phage buffer and left in the 4C refrigerator overnight or for 12 hours. The next morning the lysate is collected off the plate using a syringe and filtered through a 0.22-micron filter, into a new 15mL conical tube. A serial dilution is performed on this lysate to calculate the titer of it. Using a more diluted plate, count the number of plaques present and calculate the lysate titer using the formula shown in Figure 3-1. If the titer value is above $5 * 10^9(\frac{pfu}{mL})$ then it is high enough and ready to use in the experiments, otherwise the lysate needs to be amplified to increase the titer.

$$lysate\ titer\ (\frac{pfu}{mL}) = \frac{number\ of\ plaques}{volume\ of\ lysate\ used} * 10^3 * dilution\ factor\ of\ plate$$

Figure 3-1. The equation used to determine titer of a lysate in (pfu/mL) (Poxleitner et al., 2018).

3.3 Inoculation of M. Smegmatis with Phage Lysate and Measuring Growth Curve

The growth curve of both P1FF and P2FF M. smegmatis samples were determined. An OD600 value of 1.5 was used to indicate the mid-exponential growth phase in the M. smegmatis samples. In the P1FF growth curve, it occurred at 25 hours after inoculation. A mid-exponential P1FF sample was used to make the P2FF, and once the P2FF reached the mid-exponential stage it can be used in the following experiments. 12mL of each P2FF sample is aliquoted into 15mL conical tubes. These tubes were centrifuged at 500 x g for 10 minutes before 10 mL of the supernatant was removed. The bacteria in the tubes were resuspended with 10mLof broth culture media before being infected with either phage buffer, bacteriophage Zalkecks, or bacteriophage PotatoSplit at an MOI of 10. A Multiplicity Of Infection (MOI) value of 10 was used to determine the amount of each phage lysate to add to the P2FF samples, as shown in Figure 3-2. After being infected, the contents of each conical tube were transferred into a sterile 50mL flask and incubated at 37C with an agitation of 250rpm.

$$MOI = \left(\frac{phage\ titer\ \left(\frac{pfu}{mL}\right) * volume\ of\ phage\ titer}{\left(\frac{M.smegmatis\ cells}{mL} = OD600\ value * 10^8\right) * volume\ of\ bacteria}\right.$$

$$10 = \frac{(3.0 * 10^{10}) * (x)}{(1.6 * 10^8) * (2)} * 1000 \rightarrow x = 106.67\ \mu L\ of\ lysate$$

Figure 3-2. MOI equation and an example calculation to find the volume of lysate needed to achieve an MOI value of 10.

The OD600 values were taken of each phage bacteria and control samples at time intervals of 0, 2, 4, 6, 8, 10, and 12 hours for PotatoSplit, 0, 1, 2, 3, 5, 7, and 10 hours for Zalkecks, and 0, 3, 7, and 10 hours for the final experiment which included both phages and a control. A nanophotometer NP80 was used to take the OD600 values and a 7H9 liquid medium without Tween80 was used as the blank standard solution. At each time point after infection in the final experiment 5mL of each subsample was collected. 1mL was used to determine the OD600 value and four microcentrifuge tubes were filled with 1mL each for further protein and lipid extraction, and a backup sample for each. This was repeated three times at each time point and for each phage and control, to obtain biological triplicates. This process is visually represented in Figure 3-3, which is a process flow diagram of all the steps taken during the M. smegmatis bacteriophage infection, OD600 growth curve determination, and protein/lipid sample collection.

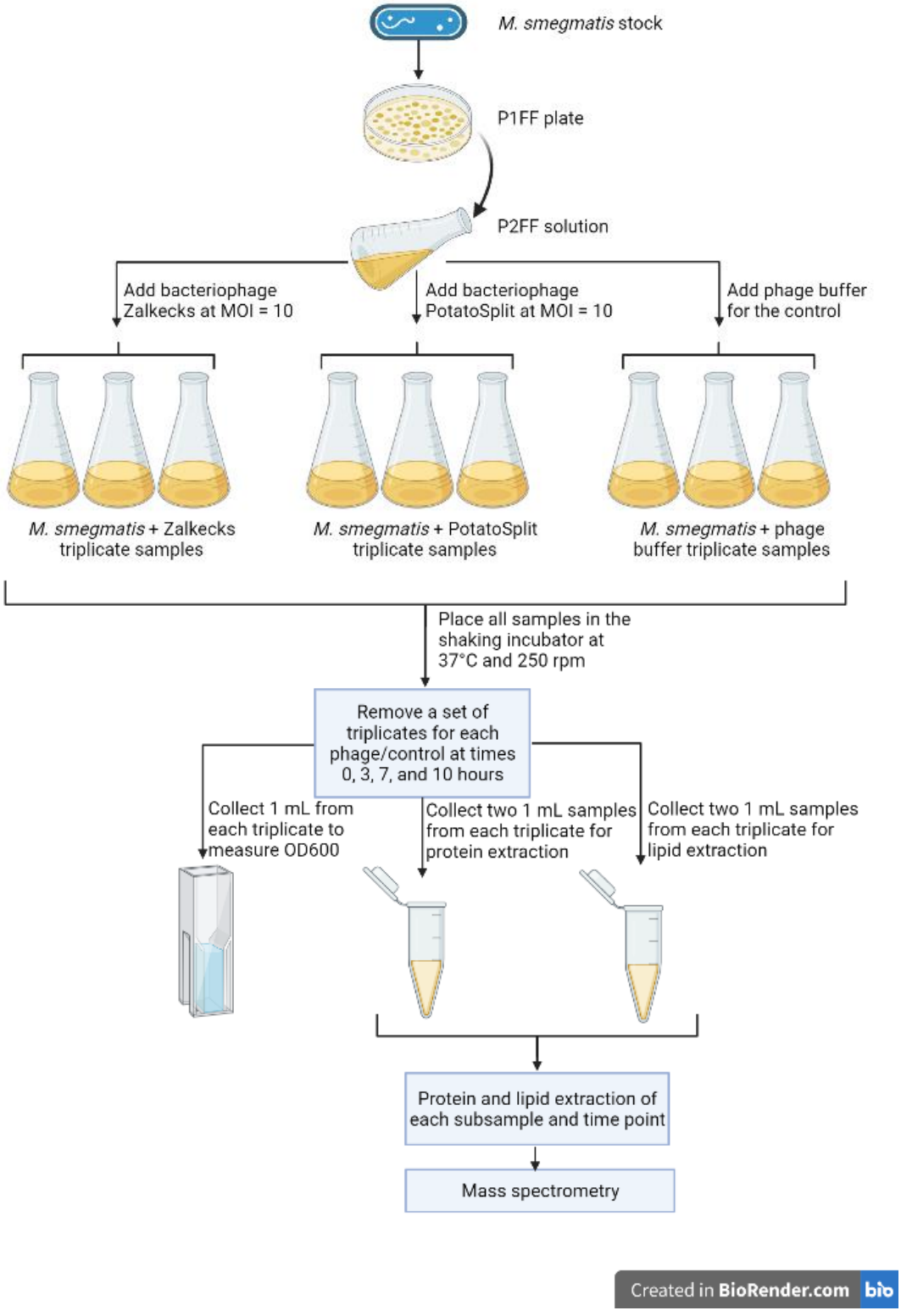


Figure 3-3. Process flow diagram of the inoculation of M. smegmatis with each bacteriophage lysate or phage buffer and determining growth curves.

3.4 Sample Preparation

3.4.1 Washing Samples

Samples need to be washed to remove any Tween80 detergent or residue to not damage the equipment used later on in the analysis process. For all sample preparations, 1 mL of the inoculated M. smegmatis solution was collected. The samples were processed in the 4C centrifuge at 14,000 rpm for 10 minutes, and the supernatant was removed to leave only a pellet. To ensure the sample is free from contaminants, it is washed three times with Phosphate Buffer Saline (PBS). 850 microliters of PBS were added to the pellet and resuspended, before being centrifuged under the same conditions. The supernatant was removed, and this process was repeated three times. After completing all three washes, the sample was transferred to a precellys tube, and 300 microliters of a HEPES and PMSF cocktail were added. To make this cocktail, the final ratio should be 1% PMSF, 1% protease inhibitors, and 98% HEPES. The precellys tubes were processed at 6200 rpm for three rounds of 20 seconds each, to lyse the cells, in a Bertin precellys evolution machine (Bertin Instruments, n.d.). The samples were removed from the precellys tube and placed into microcentrifuge tubes, using a needle tip to ensure no glass beads are in the sample.

3.4.2 BCA Protein Assay

The BCA protein assay can be used to determine the concentration of protein and therefore the amount of each sample needed to get 40 micrograms of protein in each mass spec sample. Each sample was diluted in a 1:10 ratio with double distilled water. 10 microliters of each diluted sample were placed in a 96 well plate, along with 10 microliters of the BCA standards. The BCA reaction mix is made of a 50:1 ratio of solutions A and B, and 200 microliters of it were added to each well. After 30 minutes of incubating at 37C in the Spectra Max PLUS384 microplate spectrophotometer (Trusted Laboratory Solutions, n.d.), the plate was shaken once before reading the protein concentrations on the SoftMax Pro software (V1.18) (Molecular Devices, n.d.). This reading produces the concentration of protein in the sample and indicates the volume of sample that is needed to achieve 40 micrograms of protein or lipids and will be used in the following methods. Some samples that had a lower concentration were constrained by the total volume of samples available.

3.4.3 Bligh-Dyer Extraction

The volume needed to achieve 300 micrograms of protein from the BCA assay was added to a 1.5ml microcentrifuge tube for lipid extraction. The total volume of each tube was made to be 200 microliters by adding the correct amount of ABC. Next, 500 microliters of chloroform and 400 microliters of meOH were added to the microcentrifuge tubes, and they were vortexed for 10 minutes. 200 microliters of double distilled water were added and the samples were centrifuged at 4000 rpm for 10 minutes. This causes the sample to separate into three phases: metabolites on the top, proteins in the middle, and lipids on the bottom. Each layer was collected and added to a centrifuge tube of its own. All of the samples were placed in the speedvac, with no heat, to dry for 1-2 hours. The dried lipids and metabolites were placed in the -80C freezer until they could be processed on the mass spectrometry, but the protein samples need further preparation.

3.4.4 Acetone Extraction

For protein extraction, the volume needed to achieve 50 micrograms of protein from the BCA assay was added to a 1.5ml microcentrifuge tube. Four times this volume of -20C acetone was added to each sample. The samples were left overnight in the -20C freezer. The following morning the samples were centrifuged at 14,000 rpm at 4C for 10 minutes. This will cause the protein in the sample to precipitate. The supernatant is removed and the protein pellets are dried on the speedvac with no heat for 1-2 hours. The centrifuge tubes with the dried lipids will be stored at -80C until mass spectrometry can be performed.

3.4.5 Reduction/ Alkylation

10 microliters of 10 mM dithiothreitol (DTT) and 10 microliters of 8M urea are added to each of the previously pelleted and dried protein centrifuge tubes. They are incubated with agitation for one hour at 800rpm in a 37C thermomixer. TEP mix is made from 97.5% Acetonitrile (ACN), 2% Iodoethanol, and 0.5% Triethyl phosphine (TEP). 10 microliters of this TEP mix are added to each protein sample and it is incubated again for one hour at 800 rpm in a 37C thermomixer. After the incubation period is complete, the samples are dried for 1-2 hours in the speedvac with no heat applied.

3.4.6 Digestion

Trypsin was dissolved in 25 mM ABC to make a 0.05 microgram/microliter solution. A 2-microgram vial of Trypsin makes 40 microliters of the solution when using the 0.05 microgram/microliter ratio. Samples were resuspended with 20 microliters o enzyme mixture each, to produce a ratio of 1:50 enzyme to protein. Due to the timing of sample prep, samples were not processed on the barocycler. After adding the trypsin, samples were left to digest overnight on the thermomixer set at 800 rpm and 37C. Samples were removed from the thermomixer after 16 hours and continued to be processed.

3.4.7 Sample Clean-Up

A nest tube with a 1.5 mL conical tube was made and the Pierce Peptide Desalting Spin Column (Pierce TM, n.d.) was conditioned with 300 microliters of ACN. The nested tubes were centrifuged at 5000 rpm for one minute. 300 microliters of MilliQ water with 0.1% TFA solution was added before being centrifuged for another minute at 5000 rpm. Next, 300 microliters of MilliQ water with 0.1% TFA are added again and centrifuged for another minute at 5000 rpm, before emptying the conical tube into waste. The samples are dissolved in 0.1% TFA, using up to 300 microliters of solution, loaded into the centrifuge, and ran for another minute at 5000 rpm. After ensuring that the sample ran through the column, the column can be washed by adding 100 microliters of 0.1% TFA in MilliQ water and centrifuging for another minute at 5000 rpm. This step is repeated two additional times by adding 300 microliters of 0.1% TFA in MilliQ water and centrifuging for one minute at 5000 rpm. After that, ensure that all liquid has passed through the column, then move the column into a new tube. The protein samples were eluted by adding 300 microliters of 50% ACN and 0.1% TFA then centrifuging for one minute at 5000 rpm. This was repeated two more times by adding 300 microliters of 50% ACN and 0.1% TFA then centrifuging for another minute at 5000 rpm. The eluted samples were dried on the speedvac with 45C heat for about three hours. After the samples have been dried, they are stored in the -80C freezer until mass spectrometry can be performed.

3.5 Liquid Chromatography and Tandem Mass Spectrometry (LC-MS/MS)

The samples were resuspended and transferred to HPLC vials prior to MS analysis. A 60 μL mixture of 30% ACN, 50% methanol, and 20% water was used to resuspend the lipid samples. Lipid samples were placed in a cold sonication bath for five minutes, centrifuged at 13,000 rpm for six minutes, before the supernatant was transferred to HPLC vials.

Protein MS Analysis

The proteins were resuspended in a 10 μL mixture of 3% ACN and 0.1% formic acid in water, vortexed for five minutes, centrifuged at 13,000 rpm for five minutes and transferred to HPLC vials. Both oter, vortexed for five minutes, centrifuged at 13,000 rpm for five minutes, and transferred to HPLC vials. Both of the phage-infected samples and the control samples were analyzed by a reverse-phase high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry (HPLC-ESI-MS/MS) using the Dionex Ultimate 3000 RSLC nano system (UltiMateTM 3000 RSLCnano System, n.d.) coupled to the Q-Exactive High-Field (HF) Hybrid Quadrupole Orbitrap MS (Q ExactiveTM Plus Hybrid Quadrupole-Orbitrap TM Mass Spectrometer, n.d.) and a Nano-electrospray Flex ion source (Nanospray FlexTM Ion Sources, n.d.). Reverse phase peptide separation was performed using a trap column (300 μm ID × 5 mm) packed with 5 μm 100 Å PepMap C18 medium, and then separated on a separated on an Aurora UHPLC C18 packed emitter column (25-cm long × 75 μm ID) with 1.6 μm 120 Å, while maintained at 40C.

Mobile phase solvent A was 0.1% formic acid (FA) in water and solvent B was 0.1% FA in 80% acetonitrile (ACN). Loading buffer was 98% water, 2% CAN, and 0.1% FA. Reverse phase separation of peptides was performed by loading into the trap column in a loading buffer for 5-min at 5 μL/min flow rate and eluted from the analytical column with a linear 75-min linear gradient of 27% of buffer B, then changing to 45% of B at 100 min, 100% of B at 105-min at which point the gradient was held for 7 min before reverting to 2% of B at 112.1-min. A flow rate of 150 nL/min was used to separate the peptides from the analytical column. The mass spectrometer was operated in the standard data-dependent acquisition and positive ion mode with a minimum intensity threshold of 5.0E4 and a minimum Automatic Gain Control target of 1.0E3.

Higher energy collision dissociation at a normalized collision energy setting of 27% was used to fragment the precursor ion.

The resolution of Orbitrap mass analyzer was set to 120,000 and 15,000 at 100 m/z for MS1 and MS2, respectively, with a maximum injection time of 100 ms for MS1 and 20 ms for MS2. The dynamic exclusion was set at 60s to avoid repeats by scanning identical peptides. The charge state was set at 2-7 with 2 as a default charge and a mass tolerance of 10 ppm for both high and low masses. A mass range of 350-1,600 m/z was used to collect the full scan MS1 spectra and MS2 had a fixed mass of 100 m/z. The AGC target of 34E6 for MS1 and 1e5 for MS2, as well as a spray voltage of 2.6kV was set. Each treatment had a triplicate of samples that were processed in LC-MS/MS to achieve sufficient statistical power. At the beginning of each run instrument optimization and recalibration were performed using the Pierce calibration solution.

Lipid MS Analysis

An Agilent 1290 Infinity II UPLC, coupled to an Agilent 6545 quadrupole time-of-flight tandem mass spectrometer (6545 Quadrupole Time-of-Flight LC/MS, n.d.) was used to analyze the lipid extracts. Using 10 µL of a mixture of ACN: methanol: water (3 :5 :2 v/v ratio) samples were resuspended and 8 µL were loaded to a Waters ACQUITY UPLC® BEH C18 1.7 µm columns (ACQUITY UPLC BEH C18 Columns, n.d.) with a controlled temperature of 45°C. Mobile phase A consists of 10mM ammonium acetate in water with 0.1% formic acid mobile phase B consists of 10mM ammonium acetate in a 50% isopropyl alcohol: 49.9% acetonitrile: 0.1% formic acid. The binary pump used these mobile phases at a flow rate of 0.4 ml/min. The liquid chromatography gradient was of 35% B at 0 minutes, 80% at 5 minutes, and 100% B at 10 minutes. After a 5-minute hold, the gradient then returned to 35% in 2 minutes and a 4-minute hold. An ESI capillary voltage cap of 35000V was used in the mass analyzer, along with a nebulizer gas pressure of 35psig, a skimmer of 35V, a fragmentor of 135V, and a sheath gas temperature of 320°C with a flow of 8L/min. Profile mode was used to collect mass spectrums with a range of 100-1200 m/z at a scan rate of 5 spectra/s with 200 min/spectrum for MS1 and a scan rate of 3 spectra/s with 333.3 min/spectrum for MS2. The raw data was analyzed using MS-DIAL (RIKEN Center for Sustainable Resource Science, 2020) with the MSP spectral kit of 13,303 unique compounds in

positive mode. The mass, retention time, and intensity of the compounds' positive ions [M+H]+ were obtained for the phage-treated samples.

3.6 Data Analysis

3.6.1 Lipid Data Analysis

The raw MS/MS data were processed using MS DIAL (v4) (RIKEN Center for Sustainable Resource Science, 2020), using positive ion mode. During data collection, the mass tolerance of MS1 was set to 0.001Da and 0.01Da for MS2. The MS1 mass range was set to show results between 0 and 2000Da. The retention time was set to begin at 0.5 minutes and end at 20 minutes. The MS1 mass range was set for 50 to 1200 Daltons. The maximum number of charged molecules was set to two and the number of threads was set to four. A minimum peak height was set to an amplitude of 300, during peak detection. A linear weight moving average smoothing method, using smoothing level 3, a minimum peak width of 5, and a mass slice width of 3 was selected. A sigma window of 0.1 and a cut-off of 5 for MS/MS abundance were implemented during MS2 detection. All default settings were used during lipid identification. [M+2H]2+, [M+H-H2O]+, [M+H]+, [M+Na]+, and [M+NH4]+ were the adducts selected when analyzing the phage-treated samples. The raw data was searched against the MS-DIAL lipid database and a minimum identification score of 80% was used as a cutoff.

The results produced from this search were filtered even further by removing any results without MS2 spectra data and any hits that come from the blanks. All of these results were compared to the reference matches and given a ranking of either confident, unsettled, or unknown. Any result that very closely matched the reference data was deemed as a confident match and any results that were mostly accurate were deemed as unsettled. All of the results that had little to no similarity to the reference data were deemed unknown and further studied using MS-Finder to predict a potential chemical formula of the unknown lipids.

After filtering all the results, the peak intensities of each result were used to determine the significance of that hit through MetaboAnalyst (v5.0) (Meinicke et al., 2008). The data was analyzed in three main group comparisons: bacteriophage PotatoSplit vs the control, bacteriophage Zalkecks vs the control, and bacteriophage Zalkecks vs bacteriophage PotatoSplit. Both univariate and multivariate investigations were performed to gain a broader understanding of the proteins

found within each comparison of samples. A univariate analysis of a linear model with covariate adjustment was completed in MetaboAnalyst. This produced a list of significant lipids and their associated fold change value (FC) and p-value. Fold change can be calculated as the ratio between two group means before normalization of the data occurred, and relays how much the sample has changed over time. A LOG(FC) threshold of 1.5 along with a p-value cut-off of 0.05 was used to filter the results to be significant during the univariate analysis. A multivariate analysis, called ANOVA Simultaneous Component Analysis (ASCA) was also completed in MetaboAnalyst. This analysis is used to identify major patterns in regard to two factors (phenotype and time), as well as their interaction. This analysis allows for a direct comparison to determine if a lipid is significant based on time, phenotype, or the interaction of time and phenotype. A cut-off value of 0.9 for leverage and 0.05 for p-value was implemented during the ASCA model, to only display significant findings. All of the significant lipids were grouped based on their ontology group, which was reported by MS-DIAL nomenclature. These groups could potentially give a deeper understanding of the lipids produced and why they are utilized by the phage.

3.6.2 Protein Data Analysis

The raw MS/MS data were processed using MaxQuant (v1.6.0.16) (Max Planck Institute of Biochemistry, n.d.) where the accompanying Andromeda search engine searched against the Uniprot *M. smegmatis* FASTA file (Mycolicibacterium Smegmatis (Strain Mc(2)155), n.d.) along with a reverse-decoy database and a common contaminant database (Max Planck Institute of Biochemistry, 2008). Data were searched using trypsin enzyme digestion with a cutoff of up to 2 missed cleavages. A 1% False Discovery Rate (FDR) for both protein and peptide levels, and a minimum peptide length of seven amino acids were used in MaxQuant. MS/MS fragment ions tolerance of ± 20 ppm, precursor mass tolerance of ± 10 ppm, and alkylation of cysteine and oxidation of methionine was set as fixed and variable modifications. Non-redundant and non-unique peptides assigned to the protein group with most other peptides are considered razor peptides. The “unique plus razor peptides” were used for peptide quantitation. LFQ intensities were used as the relative protein abundance measurement when comparing samples. To be included in the final analysis, proteins needed at least one unique peptide and at least two MS/MS counts detected.

Initially the MaxQuant data was filtered by removing proteins with contamination or reverse identification. If results only appeared in one of the biological triplicates, those data were also removed. Proteins that had no MS/MS counts were also removed from the dataset. If a protein had multiple protein IDs, it was filtered down to one ID to limit redundancy. After all the filtering steps, the data were normalized and MetaboAnalyst (Meinicke et al., 2008) was utilized to perform statistical analysis. The data was analyzed in three main group comparisons: bacteriophage PotatoSplit vs the control, bacteriophage Zalkecks vs the control, and bacteriophage Zalkecks vs bacteriophage PotatoSplit. Both univariate and multivariate investigations were performed to gain a broader understanding of the proteins found within each comparison of samples. A univariate analysis of a linear model with covariate adjustment was completed in MetaboAnalyst. This produced a list of significant proteins and their associated fold change value (FC) and p-value. Fold change can be calculated as the ratio between two group means before normalization of the data occurred, and relays how much the sample has changed over time. A LOG(FC) threshold of 1.5 along with a p-value cut-off of 0.05 was used to filter the results to be significant during the univariate analysis. A multivariate analysis, called ANOVA Simultaneous Component Analysis (ASCA) was also completed in MetaboAnalyst. This analysis is used to identify major patterns in regard to two factors (phenotype and time), as well as their interaction. This analysis allows for a direct comparison to determine if a protein is significance based on time, phenotype, or the interaction of time and phenotype. A cut-off value of 0.9 for leverage and 0.05 for p-value was implemented during the ASCA model, to only display significant findings.

DAVID, a functional annotation bioinformatics microarray program (DAVID, n.d.), was used to analyze Gene Ontology and Kyoto Encyclopedia of Gene and Genomes (KEGG) pathway enrichment of significantly related proteins. A fold change analysis was performed, and the results were exported from MetaboAnalyst for each time point of each sample comparison. The results from MetaboAnalyst were loaded into DAVID in two groups per time point. First all negative fold change values are loaded and processed, then repeated for positive fold change. This produces a p-value that determines if the fold change of a protein is significant or not. Using a p-value cut-off of 0.05 or a -log10(p-value) or 1.3, the significant proteins were grouped and made into bar graphs using JMP. The fold change analysis determines if a protein is either up or down regulated by the phage and provide insight into how and why the virus uses each protein.

3.7 Correlation of Protein and Lipid Results

The correlation between protein and lipids is determined after a list of all significant proteins and all significant lipids was made. To determine correlation, the results from both the protein and lipid analysis are input into one JMP sheet. Response screening is performed using the proteins/lipids as they Y-variable and time as the X-variable. This returns log worth of the p-value, log worth of the False Discovery Rate, effect size, and the standard deviation. The p-value, FDR, and means are downloaded from this analysis and transposed into a new table and filtered based on the time point. Next, the data is analyzed using the multivariate method in JMP selecting the data, the combined list of proteins and lipids, as the Y-variable and analyzing by time. This results in a table of each timepoint, with a p-value that shows if the relationship between a protein and lipid are significant or not. A p-value of 0.05 is used to determine significance of the correlation. All of this data is used to make heatmaps, which visually demonstrates when a protein and lipid are either positively or negatively correlated. It is important to note that this analysis determines correlation between proteins and lipids found in the samples and does not determine causation between them.

3.8 MS/MS Data Intensity Investigation

All of the analysis done during this research project was with the LFQ data produced from MS/MS results. LFQ data is normalized to exclude outliers and compare samples to represent the ratio changes in different samples. An investigation into iBAQ data was performed, to determine if further analysis should be completed using this data to provide a different scope of results. iBAQ is represented as the total intensity divided by the identified peptides for one protein. Although they are very similar types of data, they may produce different results, so it is important to investigate the difference. To determine if the number of results when using iBAQ data is different that when using LFQ data, the raw MS/MS data of each was loaded into MetaboAnalyst. The results were downloaded for each of the analyses performed throughout this research project: univariate, multivariate time, multivariate phenotype, and multivariate interaction. This data was then filtered into the number of hits per phage proteins and the number of hits per bacteria proteins. All of this data is overlayed and made into figures to determine the difference between using LFQ and iBAQ data from the MS/MS results.

3.9 Clean-up and Archival of Previously Discovered Mycobacteriophages

Purdue University is required to send archived samples of all the phages discovered to the University of Pittsburgh, as a part of the SEA-PHAGES program. To be archived, lysates need to have a titer of at least 5.0E9, free of any contaminants, and frozen in the proper tubes at -80C. The SEA-PHAGES Discovery Guide contains all of the protocols used during this process. The main protocols implemented during phage discovery and archival include phage purification (protocol 6.1), serial dilutions (protocol 6.2), collecting plate lysates (protocol 6.3), making webbed plates (protocol 7.1), and archiving a phage lysate (protocol 7.3).

4. RESULTS

4.1 Bacterial Growth Curves

Growth curves of both the P1FF and P2FF *M. smegmatis* samples are made to understand the growing patterns of each bacteria solution, by plotting OD600 values over time. These charts, as seen below in Figure 4-1 and Figure 4-2, are used to determine the time to collect P1FF to make P2FF and the time to infect each bacteriophage at. The time chosen needs to be during the exponential growth phase of each sample to allow for optimum growing and infecting conditions.

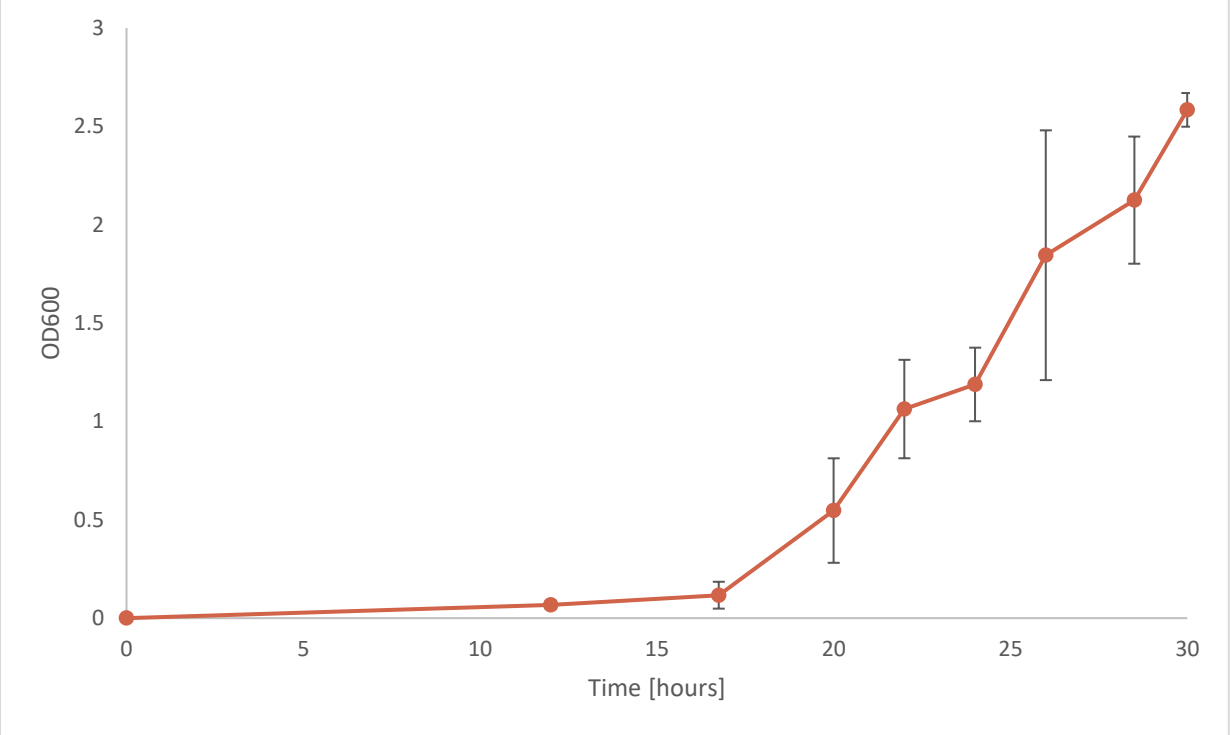


Figure 4-1. The average OD600 value of the M. smegmatis P1FF solution recorded over a 30-hour period. Error bars represent one standard deviation from the average value.

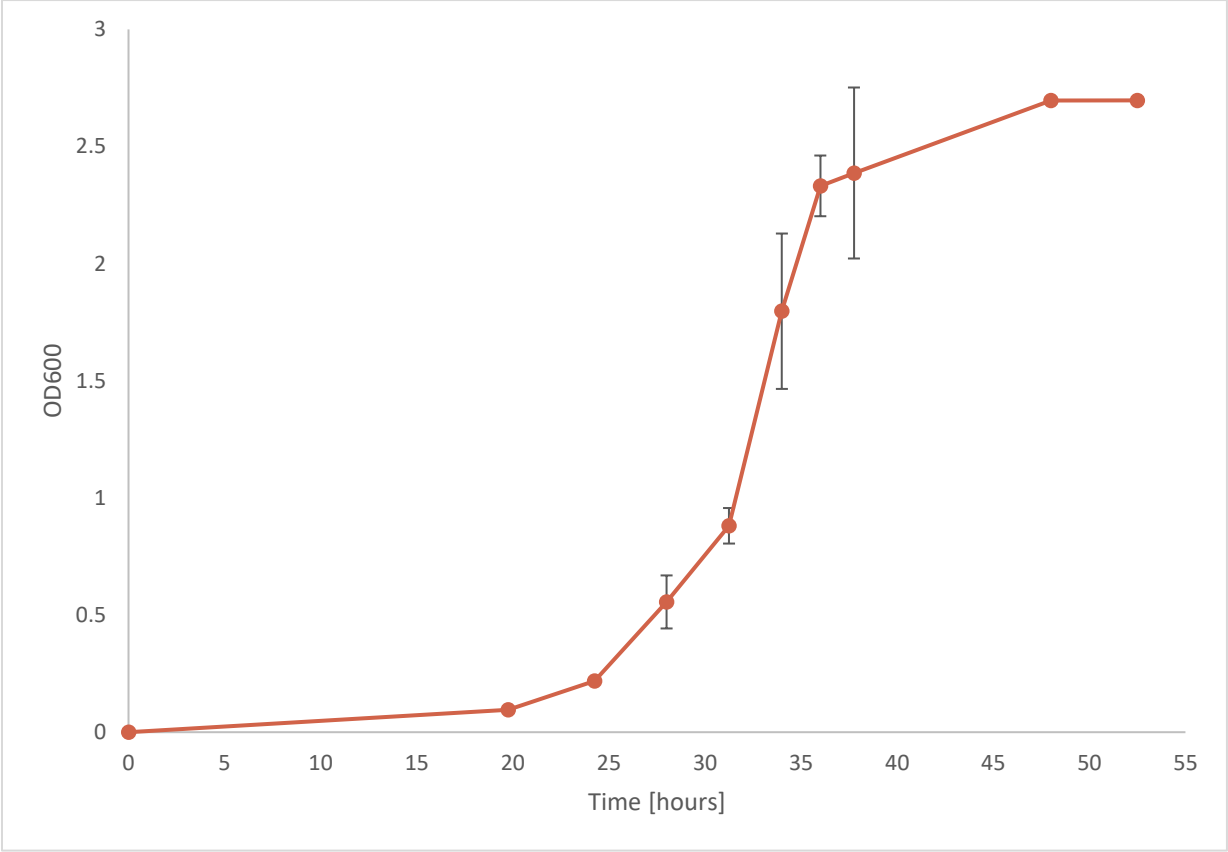


Figure 4-2. The average OD600 value of the M. smegmatis P2FF solution recorded over a 52.5-hour period. Error bars represent one standard deviation from the average value.

4.2 Phage Infection Growth Curves

Growth curves were made for two Mycobacterium smegmatis samples, one infected with a bacteriophage and the other infect with phage buffer as the control. The OD600 levels were recorded for each sample over time, and repeated for the different bacteriophages, to get Figure 4-3 and Figure 4-4 below. These phage-infection growth curves were used to plan and create a procedure to run the final experiment in which protein and lipids samples were collected. The main goal of these curves is to demonstrate where infection plateaus and therefore the timepoints to collect protein and lipid samples at. Figure 4-5 shows the growth curves from the final experiment, and each OD600 timepoint coincides with the times that samples were collected for further analysis.

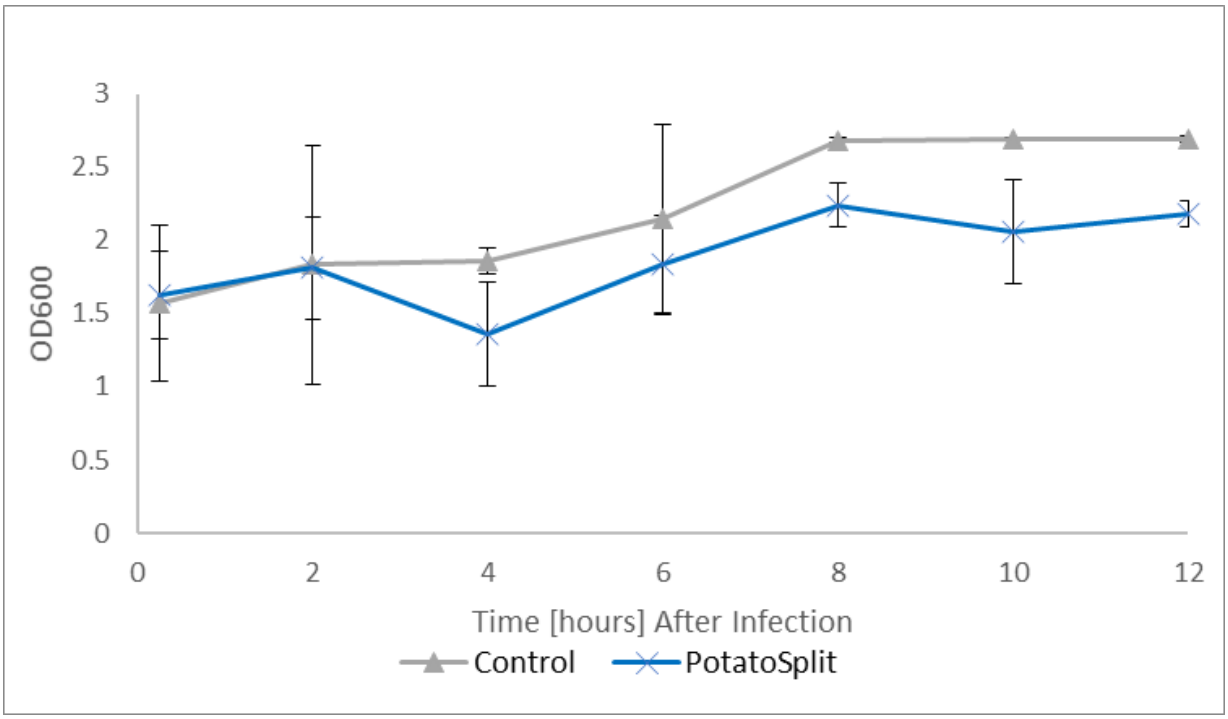


Figure 4-3. The average OD600 value of each sample, recorded over a 12-hour period. The grey represents the control, which is M. Smegmatis infected with phage buffer only. Blue represents M. smegmatis infected with mycobacteriophage PotatoSplit. Each sample has three biological replicates, and the error bars represent one standard deviation from the average value.

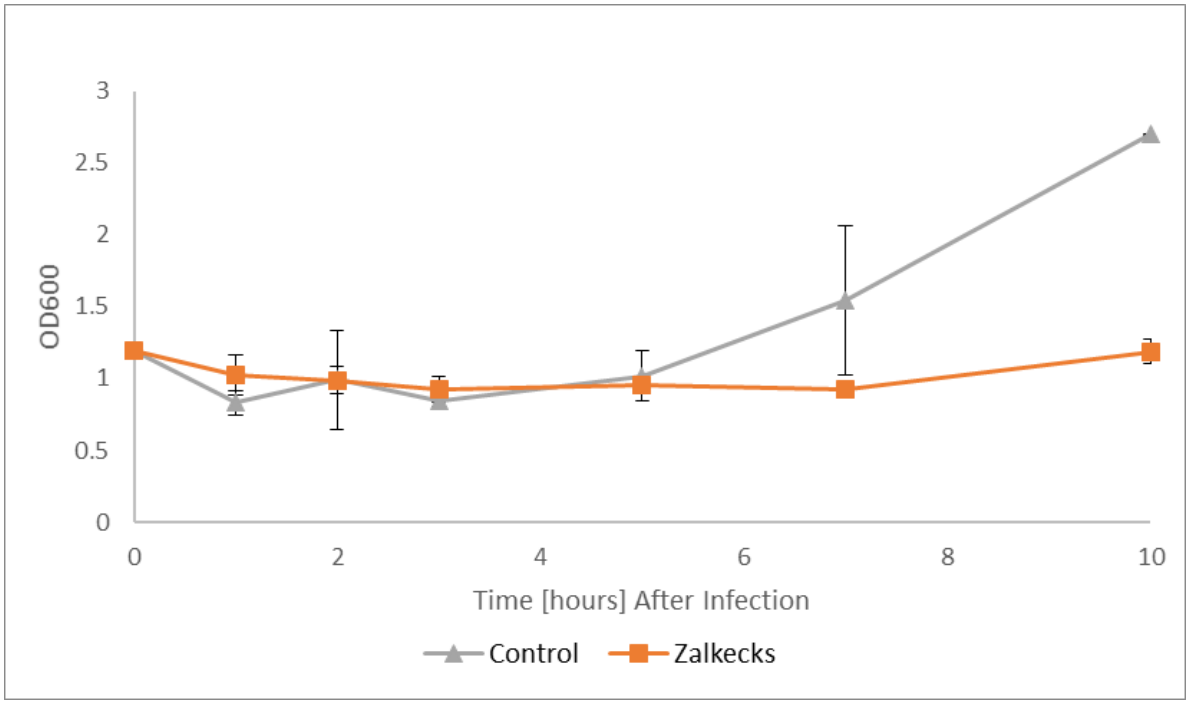


Figure 4-4. The average OD600 value of each sample, recorded 14-hour period. The grey data represents the control which is M. Smegmatis infected with phage buffer only. Orange represents M. smegmatis infected with mycobacteriophage Zalkecks. Each sample has three biological replicates, and the error bars represent one standard deviation from the average value.

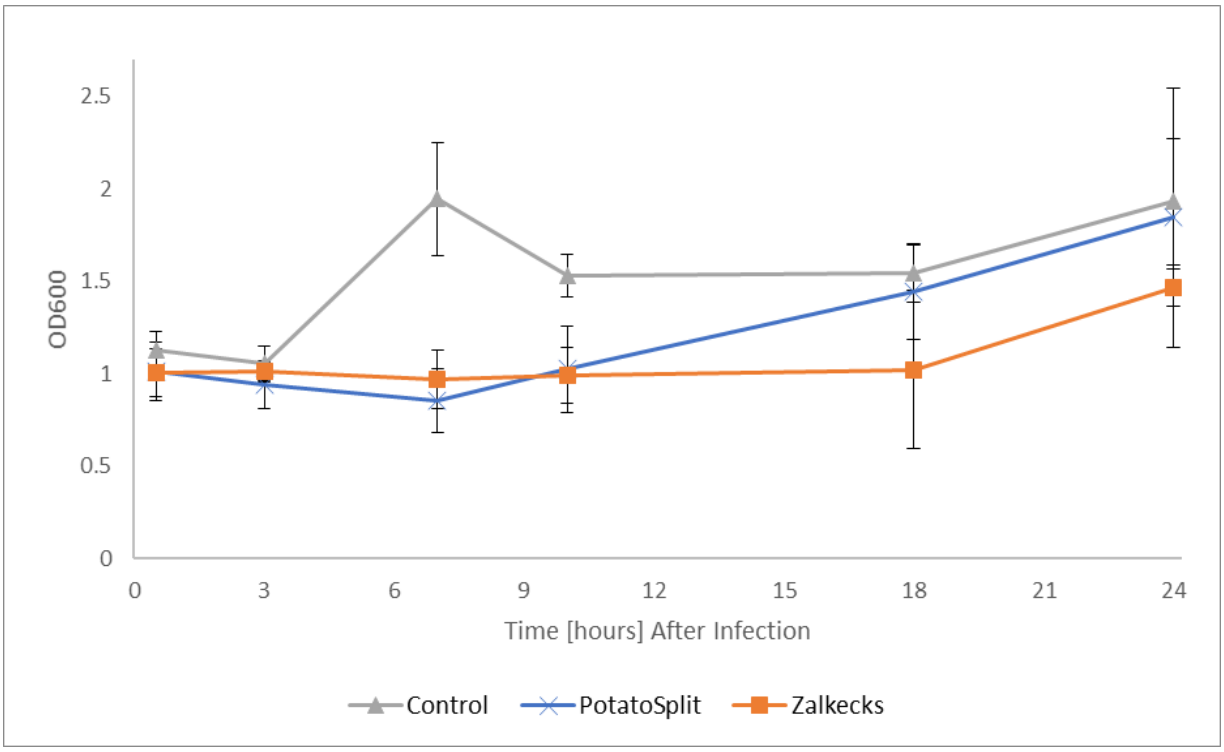


Figure 4-5. The average OD600 value of each sample recorded over a 24-hour period. The grey data represents the control which is M. Smegmatis infected with phage buffer only. Blue represents M. smegmatis infected with mycobacteriophage PotatoSplit, and orange is M. smegmatis infected with mycobacteriophage Zalkecks. Each sample has three biological replicates, and the error bars represent one standard deviation from the average value.

4.3 Examining Lipids from Phage-Treated Samples

This section displays the results of all lipids found when each phage-infected sample is compared with the phage buffed infected control samples or another phage-infected sample. MetaboAnalyst was used to perform univariate and multivariate analysis. Each comparison below has a heatmap distribution of each lipid, at the different time points, in each sample. The first table displays significant lipids based on a univariate analysis of a linear model with covariate adjustment, using a p-value cut-off of 0.05 to determine significance of lipids found in the phage-infected samples as compared to the control samples or the other phage-infected samples. The second table displays significant lipids based on a multivariate analysis, to determine the significance of each lipid based on the phenotype, time, or interaction of the samples, using a cut-off value of 0.9 for leverage and 0.05 for p-value. A visual representation of the multivariate analysis, and the relationship between lipids that are significant based on time, phenotype, or interaction, can be seen in the Venn diagram in Figure 4-7, Figure 4-9, and Figure 4-11 for each comparison.

4.3.1 PotatoSplit Compared Against the Control

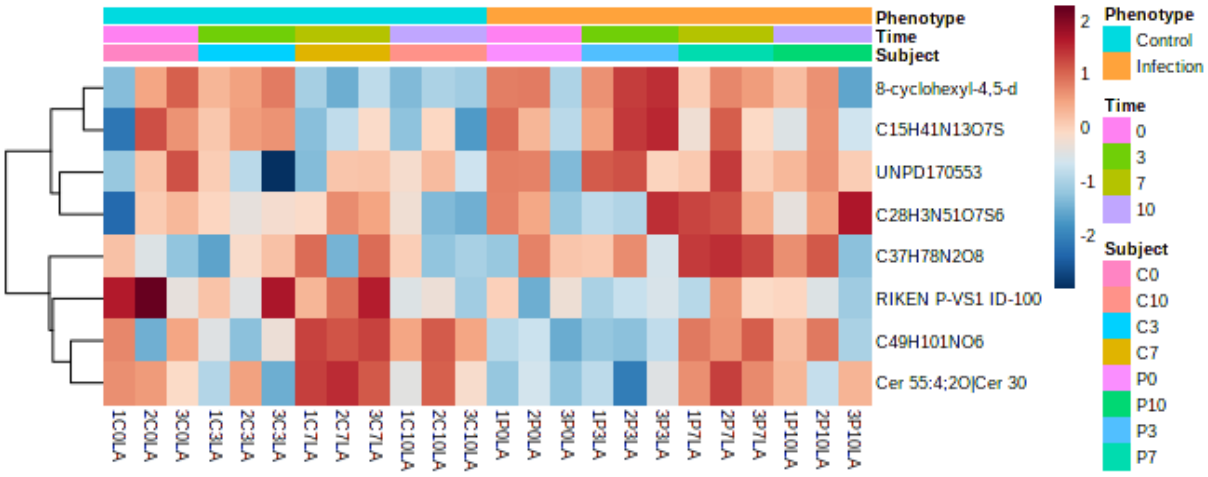


Figure 4-6. Heatmap of all significant ANOVA lipids found in mycobacteriophage PotatoSplit treated samples when compared against the control samples, at each time point taken.

Table 4-1. PotatoSplit MS2 acquired lipid results from a MetaboAnalyst linear model with covariate adjustments, when compared against the control.

	POTATOSPLIT LIPIDS	ONTOLOGY	LOG(FC)	P.VALUE
COMPARISON TO CONTROL	RIKEN P-VS1 ID-1009 from Mouse_Muscle _WT_CTX0_Ctr		-0.99817	0.004984
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl) -3,8-dimethyl-4-[2- (5-oxo-2,5-dihydrofuran-3-yl)ethyl]-decahydronaphthalen-1-yl acetate	Ether MGDG	0.81397	0.012404
	C49H101NO6	Cer_NS	-0.606	0.025974
	C37H78N2O8		0.79003	0.026626
	UNPD170553		0.855	0.028244
	Cer 55:4;2O|Cer 30:3;2O/25:1	Cer_NS	-0.6255	0.029785
	C28H3N51O7S6		0.74822	0.042504
	C15H41N13O7S		0.69209	0.046576

Table 4-2. PotatoSplit ANOVA Simultaneous Component Analysis (ASCA) significant lipid results when compared against the Control to model phenotype and time effects and their interaction.

	POTATOSPLIT VS CONTROL LIPIDS	ONTOLOGY	LEVERAGE	SPE
PHENOTYPE	RIKEN P-VS1 ID-1009 from Mouse_Muscle_WT_CTX0_Ctr		0.037672	3.70E-32
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl)-3,8-dimethyl-4-[2-(5-oxo-2,5-dihydrofuran-3-yl)ethyl]-decahydronaphthalen-1-yl acetate	Ether MGDG	0.025051	3.70E-32
	C28H3N51O7S6		0.021167	0
	C15H41N13O7S		0.01811	0
	MGDG O-19:2|MGDG O-17:2_2:0	Ether MGDG	0.017116	7.40E-32
	PS 30:8	PS	0.016838	3.70E-32
	UNPD200843		0.01679	3.70E-32
	2-amino-N1,N9-bis[11-hydroxy-2,5,9-trimethyl-1,4,7,14-tetraoxo-6,13-bis(propan-2-yl)-1H,2H,3H,4H,5H,6H,7H,9H,10H,13H,14H,16H,17H,18H,18aH-pyrrolo[2,1-i]1-oxa-4,7,10,13-tetraazacyclohexadecan-10-yl]-4-methyl-3-oxo-3H-phenoxazine-1,9-dicarboximidic acid		0.016107	3.70E-32
	C32H84N10O3S9		0.016051	3.70E-32

Table 4-2. continued

PHENOTYPE	CER(D18:0/23:0)		0.013819	7.40E-32
	Cer 55:4;2O|Cer 30:3;2O/25:1	Cer_NS	0.014793	7.40E-32
	C29H71N6O2PS	DG	0.013169	7.40E-32
	PE 32:0	PE	0.013093	7.40E-32
	C29H55N11S3		0.012783	3.70E-32
	C27H4N5O45PS4		0.01278	3.70E-32
	C33H83N15S3		0.012661	3.70E-32
	Cer 57:4;2O|Cer 30:3;2O/27:1	Cer_NS	0.012006	3.70E-32
INTERACTION	PE 35:0	PE	0.024896	0.081461
	ethyl-XTP		0.036426	0.29822
	2-furancarboxylicacid[4-[[1-(4-ethoxycarbonyl phenyl)-3,5-dioxo-4-pyrazolidinylidene]methyl]-2-methoxyphenyl] ester		0.035926	0.000985
	C20H40NO3P	Cer_HDS	0.032439	0.11429
	C25N5O2PS5		0.028843	0.35681
	Katsumadain B;(+)-Katsumadain B		0.028725	0.003002
	UNPD180606	VAE	0.028344	0.098602
	C28H60O19S	PI	0.028114	0.33174
	Butoctamide hydrogen succinate	Fatty Amide	0.027963	0.011444
	C13H7NO2S12		0.02708	0.12492
	EI 1511-3 (KnapSack)		0.026525	0.000673
	C26H34N20O4P2		0.025691	0.01459
	DG 44:11	DG	0.024479	0.007043
	C19H22N8O5P2S		0.024255	0.02042
	UNPD13907		0.023061	0.003667
	NAE 19:4	NAE	0.0229	0.14375
	C22H60N16O		0.021996	0.01711
	C33H71N5O4P2S	SL	0.020994	0.0064
	C17H67N33O38P2S		0.020867	1.3361
	DG 64:17	DG	0.020663	0.001625
	TG 38:5|TG 8:0_15:1_15:4	TG	0.020423	0.012577
	MINEs-479764		0.020578	0.096703
	N-acetyl-3,5,11,18-tetrahydroxyoctadecyl-2-amine		0.020363	0.2348
	TG 67:17|TG 15:4_16:4_36:9	TG	0.020236	0.005366
	DG 32:0	DG	0.020009	0.045883
	MINEs-322604		0.020002	0.027003
	Cer 21:2;2O|Cer 12:2;2O/9:0	Cer_NS	0.019479	0.012891

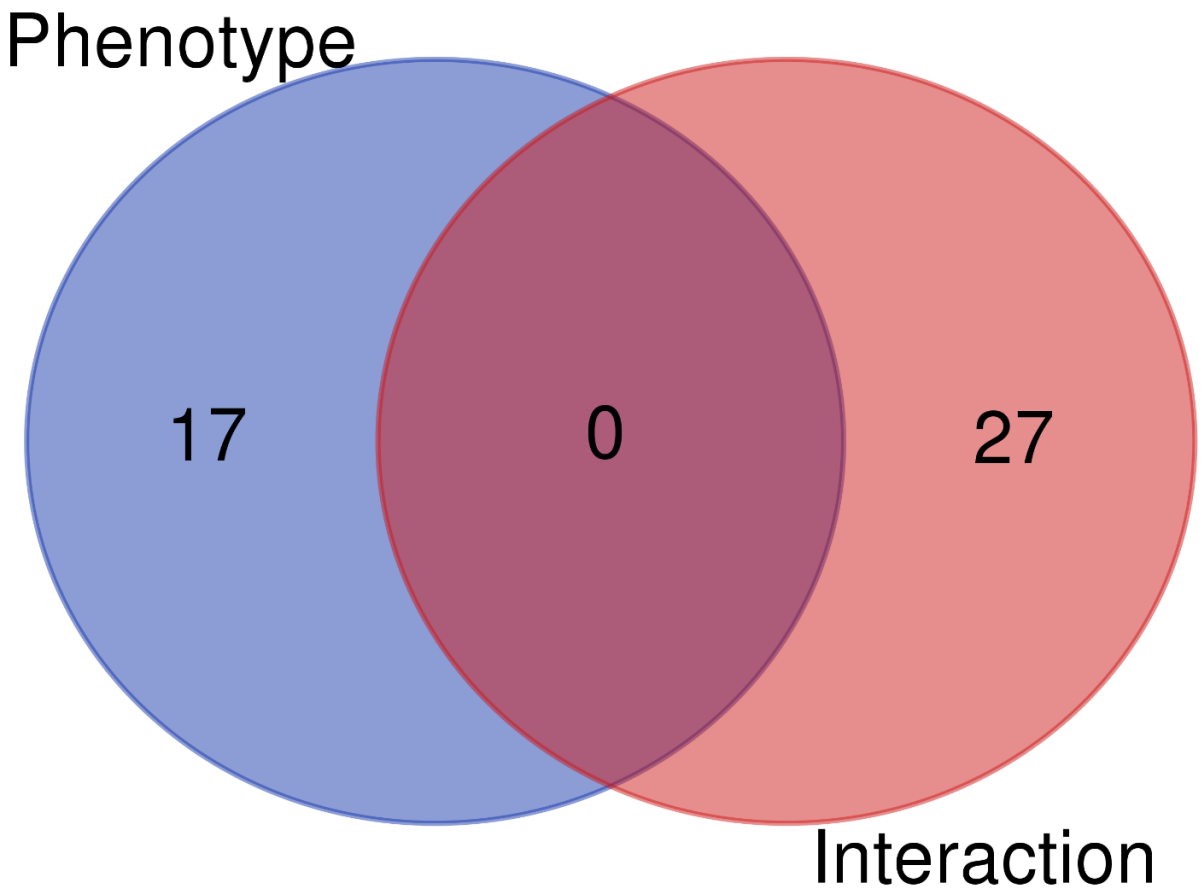


Figure 4-7. Multivariate analysis results, when PotatoSplit is compared against the control, displayed in a Venn Diagram to show the relationship between lipids that are significant based on phenotype or the interaction between time and phenotype.

4.3.2 Zalkecks Compared Against the Control

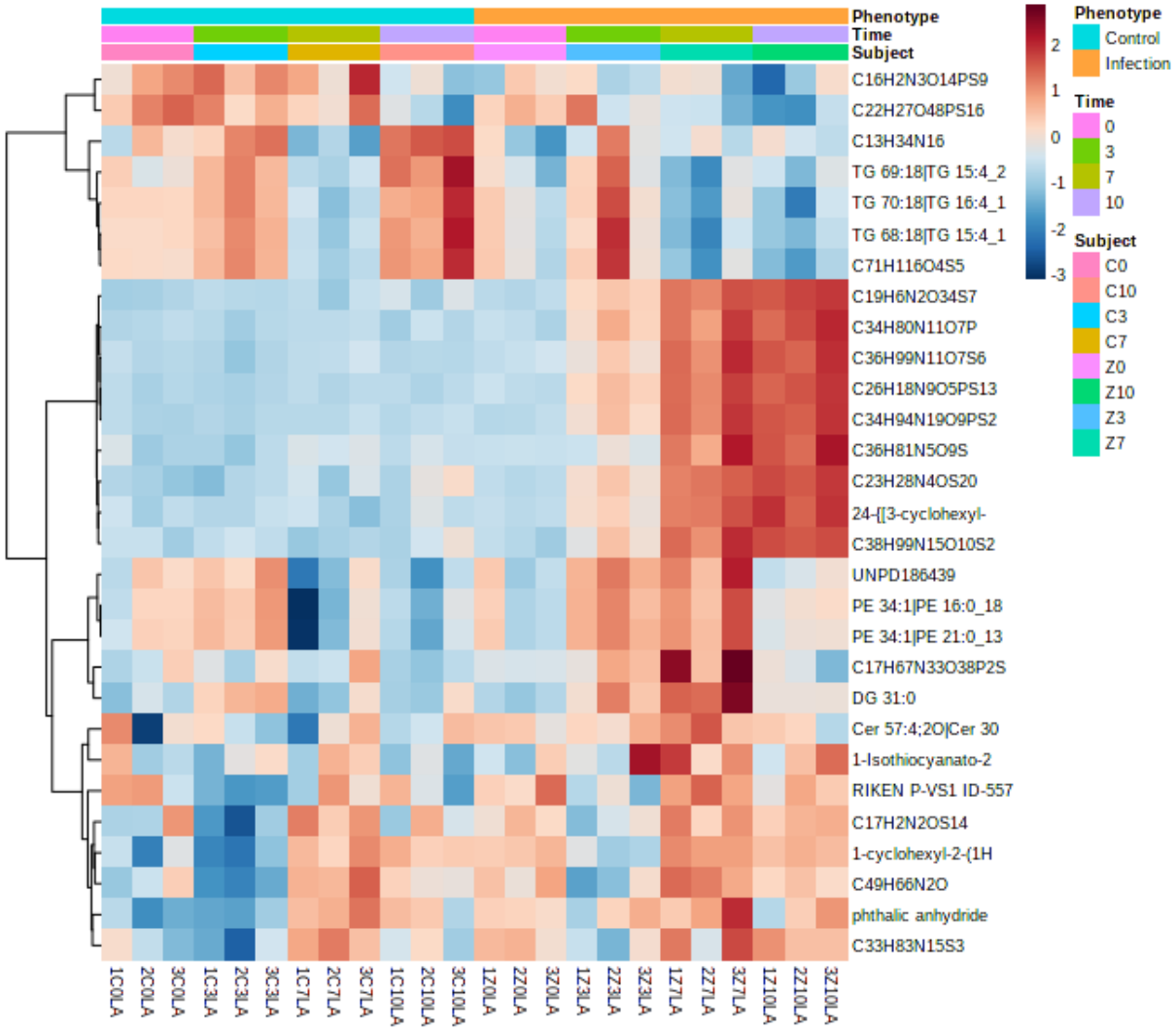


Figure 4-8. Heatmap of all significant ANOVA lipids found in mycobacteriophage Zalkecks treated samples when compared against the control samples, at each time point taken.

Table 4-3. Zalkecks MS2 acquired lipid results from a Metaboanalyst linear model with covariate adjustments, when compared against the control.

	ZALKECKS LIPIDS	ONTOLOGY	LOG(FC)	P.VALUE
COMPARISON TO CONTROL	C26H18N9O5PS13		1.4497	2.51E-06
	C19H6N2O34S7		1.3885	4.50E-06
	C36H99N11O7S6		1.4024	5.64E-06
	C34H80N11O7P		1.405	6.68E-06
	C34H94N19O9PS2		1.379	6.84E-06
	C23H28N4OS20		1.3075	8.26E-06
	24-{[3-cyclohexyl-5-(4-hydroxyoxan-4-yl)phenyl]methyl}-17-[2-(3,3-dimethyloxiran-2-yl)-2-hydroxyethyl]-2-hydroxy-6-(3-hydroxy propyl)-4,8,12,13,17-pentamethyl-8-[2-(methyl amino) ethyl] -22,26-diazaheptacyclo [14.11.1.1??,??.0?,??. 0?,?.0??,??.0??,??] non acosa -16(28), 20, 23(29), 24-tetraene-7,15-dione		1.351	1.64E-05
	C38H99N15O10S2		1.2531	7.35E-05
	C36H81N5O9S		1.1916	8.53E-05
	C16H2N3O14PS9		-1.0674	0.001463
	TG 69:18|TG 15:4_22:6_32:8		-0.86599	0.006117
	1-cyclohexyl-2-(1H-imidazol-2-ylmethyl)-1,3,4,9-tetrahydropyrido[3,4-b]indole		0.73775	0.006594
	phthalic anhydride		0.81889	0.008283
	TG 68:18|TG 15:4_15:4_38:10	TG	-0.79465	0.012635
	C33H83N15S3		0.74563	0.014265
	C71H116O4S5		-0.79417	0.014699
	C17H67N33O38P2S		0.77746	0.015338
	TG 70:18|TG 16:4_16:4_38:10	TG	-0.78605	0.016414
	C22H27O48PS16		-0.69829	0.017426
	DG 31:0	DG	0.76221	0.019543
	UNPD186439	DG	0.76439	0.024769
	RIKENP-VS1 ID-5576from Mouse_Muscle_ fads2KO_N_ Ctr		0.71181	0.019706
	Cer 57:4;2O|Cer 30:3;2O/27:1	Cer_NS	0.84342	0.025114
	1-Isothiocyanato-2-phenylethane		0.80883	0.025993
	PE 34:1|PE 16:0_18:1	PE	0.77278	0.028354
	PE 34:1|PE 21:0_13:1	PE	0.75749	0.030241
	C17H2N2OS14		0.61673	0.038365
	C49H66N2O		0.53181	0.040442
	C13H34N16	NAOrn	-0.65042	0.04557

Table 4-4. Zalkecks ANOVA Simultaneous Component Analysis (ASCA) significant lipid results when compared against the Control to model phenotype and time effects and their interaction.

	ZALKECKS VS CONTROL LIPIDS	ONTOLOGY	LEVERAGE	SPE
PHENOTYPE	C26H18N9O5PS13		0.04199	1.48E-31
	C19H6N2O34S7		0.038517	0
	C16H2N3O14PS9		0.022761	0
	TG 69:18|TG 15:4_22:6_32:8	TG	0.014983	3.70E-32
	Cer 57:4;2O|Cer 30:3;2O/27:1	Cer_NS	0.014212	3.70E-32
	phthalic anhydride		0.013397	1.85E-31
INTERACTION	DG 24:3	DG	0.03293	0.26198
	C38H99N15O10S2		0.032323	0.01693
	PE 34:1|PE 21:0_13:1	PE	0.032223	0.000899
	PE 34:1|PE 16:0_18:1	PE	0.032194	0.014396
	DG 31:0	DG	0.03123	0.205
	C33HNO38S8		0.031697	0.44466
	ethyl-XTP		0.02588	0.002294
	TG 70:18|TG 16:4_16:4_38:10		0.025469	0.10508
	C34H94N19O9PS2		0.025386	0.15609
	C36H99N11O7S6		0.025206	0.11709
	C34H80N11O7P		0.024794	0.34295
	C36H81N5O9S		0.024207	0.000458
	C17H12N2O34S8		0.024029	0.13539
	C26H18N9O5PS13		0.024008	0.2633
	C25N5O2PS5		0.023304	0.019054
	C13H34N16	NAOrn	0.023096	0.01496
	C19H6N2O34S7		0.022801	0.11074
	TG 68:18|TG 15:4_15:4_38:10	TG	0.02171	0.1536
	PE 32:1|PE 11:0_21:1	PE	0.030486	0.12284
	DG 49:6	DG	0.030313	0.40161
	24-{[3-cyclohexyl-5-(4-hydroxyoxan-4-yl)phenyl]methyl}-17-[2-(3,3-dimethyloxiran-2-yl)-2-hydroxyethyl]-2-hydroxy-6-(3-hydroxypropyl)-4,8,12,13,17-pentamethyl-8-[2-(methyl amino)ethyl]-22,26-diazaheptacyclo[14.11.1.1??,??.0?, ??.0?,?.0??, ??.0??,??]nonacosa-16(28),20,23(29),24-tetraene-7,15-dione		0.030286	0.03547
	C71H116O4S5		0.028359	0.14582
	Cer 51:2;2O|Cer 30:0;2O/21:2	Cer_NDS	0.021501	0.80445
	UNPD186439	DG	0.027795	0.006888

Table 4-4. continued

INTERACTION	C20H40NO3P	CER_HDS	0.028127	0.18051
	N"-{[1-(5-cyclohexylidene-3-hydroxy-6-methylheptan-2-yl)-3a,7,8-trihydroxy-9a,11a-dimethyl-5-oxo-1H,2H,3H,3aH,5H,5aH,6H,7H,8H,9H,9aH,9bH,10H,11H,11aH-cyclopenta[a]phenanthren-10-yl]methyl}guanidine		0.020362	0.41996

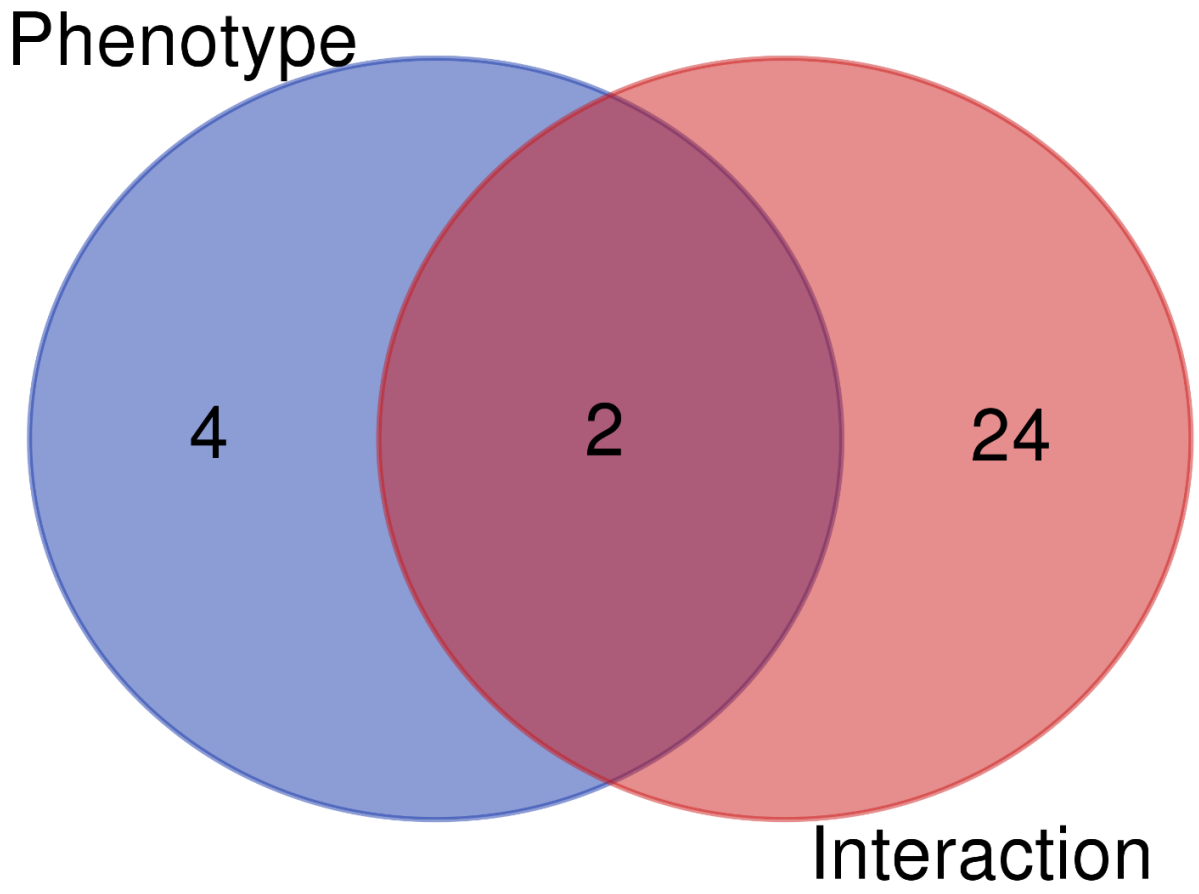


Figure 4-9. Multivariate analysis results, when Zalkecks is compared against the control, displayed in a Venn Diagram to show the relationship between proteins that are significant based on phenotype or the interaction between time and phenotype.

4.3.3 Zalkecks Compared Against PotatoSplit

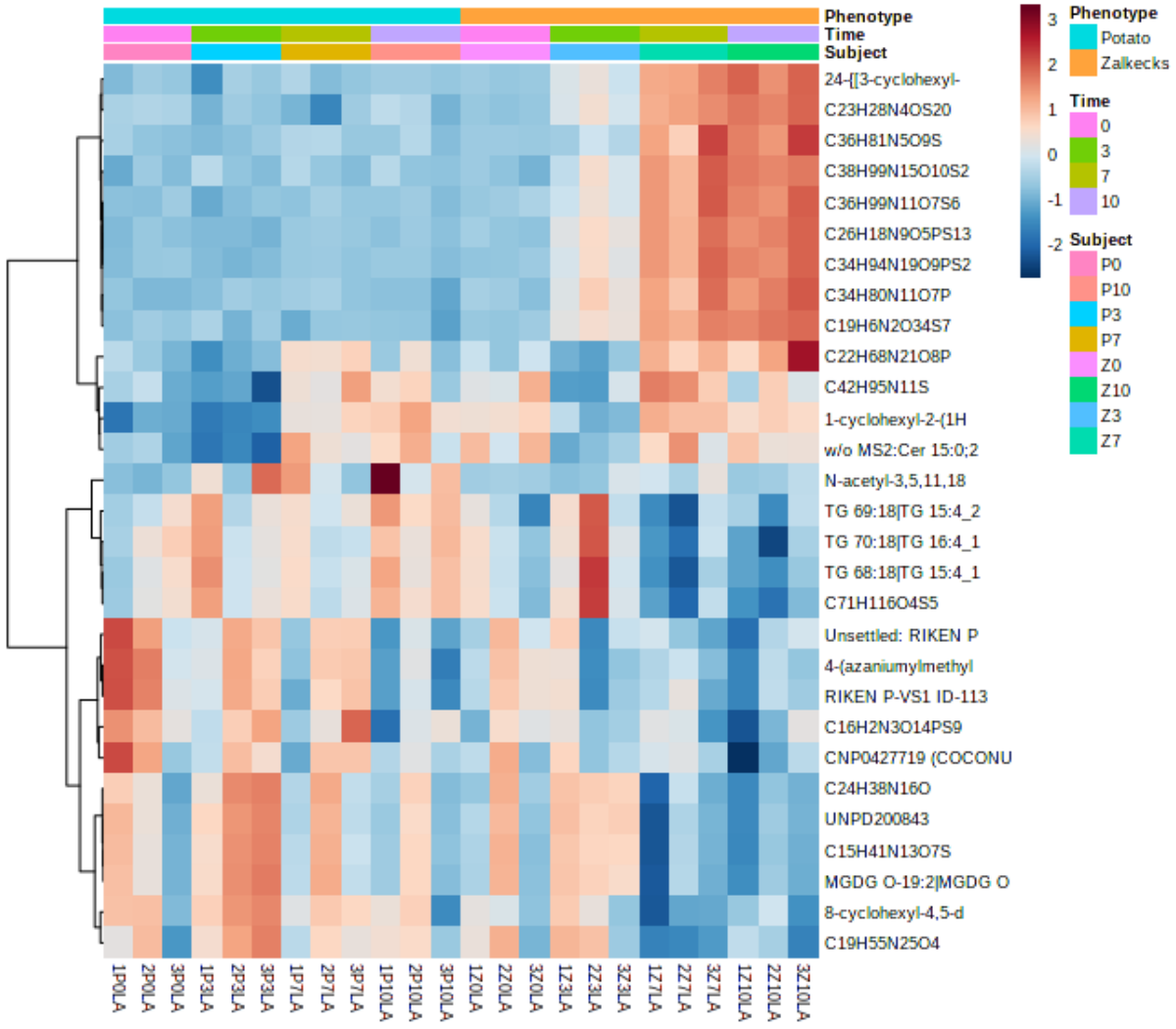


Figure 4-10. Heatmap of all significant ANOVA lipids found in mycobacteriophage PotatoSplit and mycobacteriophage Zalkecks treated samples, when compared against each other, at each time point taken.

Table 4-5. Zalkecks MS2 acquired lipid results from a Metaboanalyst linear model with covariate adjustments, when compared against PotatoSplit.

	ZALKECKS LIPIDS	ONTOLOGY	LOG(FC)	P.VALUE
COMPARISON TO POTATOSPLIT	C26H18N9O5PS13		1.4275	2.86E-06
	C34H80N11O7P		1.4279	4.67E-06
	C36H99N11O7S6		1.3959	5.98E-06
	C34H94N19O9PS2		1.3916	6.03E-06
	C19H6N2O34S7		1.4269	7.55E-06
	24-{[3-cyclohexyl-5-(4-hydroxyoxan-4-yl)phenyl]methyl}-17-[2-(3,3-dimethyloxiran-2-yl)-2-hydroxyethyl]-2-hydroxy-6-(3-hydroxypropyl)-4,8,12,13,17-pentamethyl-8-[2-(methylamino)ethyl]-22,26-diazaheptacyclo [14.11.1.1??,??.0?,??.0?,?.0??,??.0??,??] nonacosa-16(28),20,23(29),24-tetraene-7,15-dione		1.351	1.12E-05
	C38H99N15O10S2		1.2737	3.45E-05
	C23H28N4OS20		1.3116	5.41E-05
	C36H81N5O9S		1.1633	0.000187
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl)-3,8-dimethyl-4-[2-(5-oxo-2,5-dihydrofuran-3-yl)ethyl]-decahydronaphthalen-1-yl acetate	Ether MGDG	-1.0405	0.002139
	1-cyclohexyl-2-(1H-imidazol-2-ylmethyl)-1,3,4,9-tetrahydropyrido[3,4-b]indole		0.76695	0.003441
	TG 69:18|TG 15:4_22:6_32:8	TG	-0.86306	0.014937
	C22H68N21O8P		0.69455	0.015237
	C16H2N3O14PS9		-0.82642	0.020667
	C71H116O4S5		-0.80404	0.022377
	TG 70:18|TG 16:4_16:4_38:10	TG	-0.8044	0.022641
	w/o MS2:Cer 15:0;2O/33:0;O	Cer_HDS	0.62087	0.023433
	C19H55N25O4		-0.76299	0.024116
	C15H41N13O7S		-0.71839	0.024743
	MGDG O-19:2|MGDG O-17:2_2:0	Ether MGDG	-0.70788	0.026881
	TG 68:18|TG 15:4_15:4_38:10	TG	-0.82641	0.018394
	4-(azaniumylmethyl)-5-(7-cyclohexyl-2-{2-[3-(hydroxymethyl)phenyl]ethyl}hept-4-en-1-yl)-8-{3-[(3-hydroxyphenyl)methyl]-1-phenylcyclohexyl}-1,4,4a,5,8,8a-hexahydronaphthalene-1-carboxylate		-0.71131	0.030514
	C42H95N11S	Cer_NDS	0.59451	0.033289
	UNPD200843		-0.65524	0.037122

Table 4-5. continued

POTATOSPLIT	UNSETTLED: RIKEN P-VS1 ID-476 FROM CELL_HEK293_WT_N_CTR		-0.70587	0.040422
	CNP0427719 (COCONUT)		-0.72777	0.039224
	N-acetyl-3,5,11,18-tetrahydroxyoctadecyl-2-amine		-0.72059	0.047445
	RIKEN P-VS1 ID-11326 from Mouse_Aorta_ApoEKO_N_F1EPA		-0.65562	0.049198
	C24H38N16O	DG	-0.63674	0.049348

Table 4-6. Zalkecks ANOVA Simultaneous Component Analysis (ASCA) significant lipid results when compared against PotatoSplit to model phenotype and time effects and their interaction.

	ZALKECKS VS POTATOSPLIT LIPIDS	ONTOLOGY	LEVERAGE	SPE
PHENOTYPE	C26H18N9O5PS13		0.042133	2.96E-31
	C19H6N2O34S7		0.042095	2.96E-31
	C36H99N11O7S6		0.040284	2.96E-31
	C23H28N4OS20		0.035568	1.48E-31
	C38H99N15O10S2		0.033541	2.96E-31
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl)-3,8-dimethyl-4-[2-(5-oxo-2,5-dihydrofuran-3-yl)ethyl]-decahydronaphthalen-1-yl acetate	Ether MGDG	0.022385	2.96E-31
	TG 69:18|TG 15:4_22:6_32:8	TG	0.015401	1.85E-31
INTERACTION	C19H22N8O5P2S		0.02257	0.43143
	C17H67N33O38P2S		0.043858	1.2332
	Cer 57:4;2O|Cer 30:3;2O/27:1	Cer_NS	0.029455	0.004042
	DG 49:6	DG	0.02871	0.85594
	C40H95N15O	Cer_NS	0.02784	1.3219
	C23H28N4OS20		0.027704	0.020511
	DG 31:0	DG	0.027014	1.525
	C71H116O4S5		0.026065	0.80598
	C19H6N2O34S7		0.024782	0.000344
	UNPD186439		0.024008	0.59086
	C30H56N2O2		0.023761	1.4204
	C36H81N5O9S		0.023445	0.19625
	w/o MS2:RIKEN P-VS1 ID-6673 from Mouse_Adrenal Glands_WT_N_Ctr		0.022447	0.96523
	PE 34:1|PE 21:0_13:1	PE	0.022386	0.38372

Table 4-6. continued.

INTERACTION	TG 70:18|TG 16:4_16:4_38:10	TG	0.02176	0.78524
	Katsumadain B;(+)-Katsumadain B		0.022007	1.4709
	PE 34:1|PE 16:0_18:1	PE	0.02285	0.23157
	EI 1511-3 (KnapSack)		0.022555	0.8306
	UNPD13907		0.022517	0.30237
	C26H34N20O4P2		0.022471	0.25708
	phthalic anhydride		0.020742	0.93564
	1,2,3,4-Tetrahydro-b-carboline-1,3-dicarboxylic acid		0.020724	0.14169
	C16H2N3O14PS9		0.014121	1.85E-31
	TG 68:18|TG 15:4_15:4_38:10	TG	0.01412	3.70E-32
	C34H80N11O7P		0.020407	0.00992
	C34H94N19O9PS2		0.020318	0.005838
	C38H99N15O10S2		0.020293	0.11576
	24-{[3-cyclohexyl-5-(4-hydroxyoxan-4-yl)phenyl]methyl}-17 -[2-(3,3-dimethyloxiran-2-yl)-2-hydroxyethyl]-2-hydroxy-6-(3-hydroxypropyl)-4,8,12,13,17-pentamethyl-8-[2-(methylamino) ethyl]-22,26-diazaheptacyclo[14.11.1.1??,??.0?,??.0?,?.0??, ??.0??,??]nonacosa-16(28),20,23 (29),24-tetraene-7,15-dione		0.02021	0.000352
	C36H99N11O7S6		0.01939	0.017525
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr		0.019372	0.038238

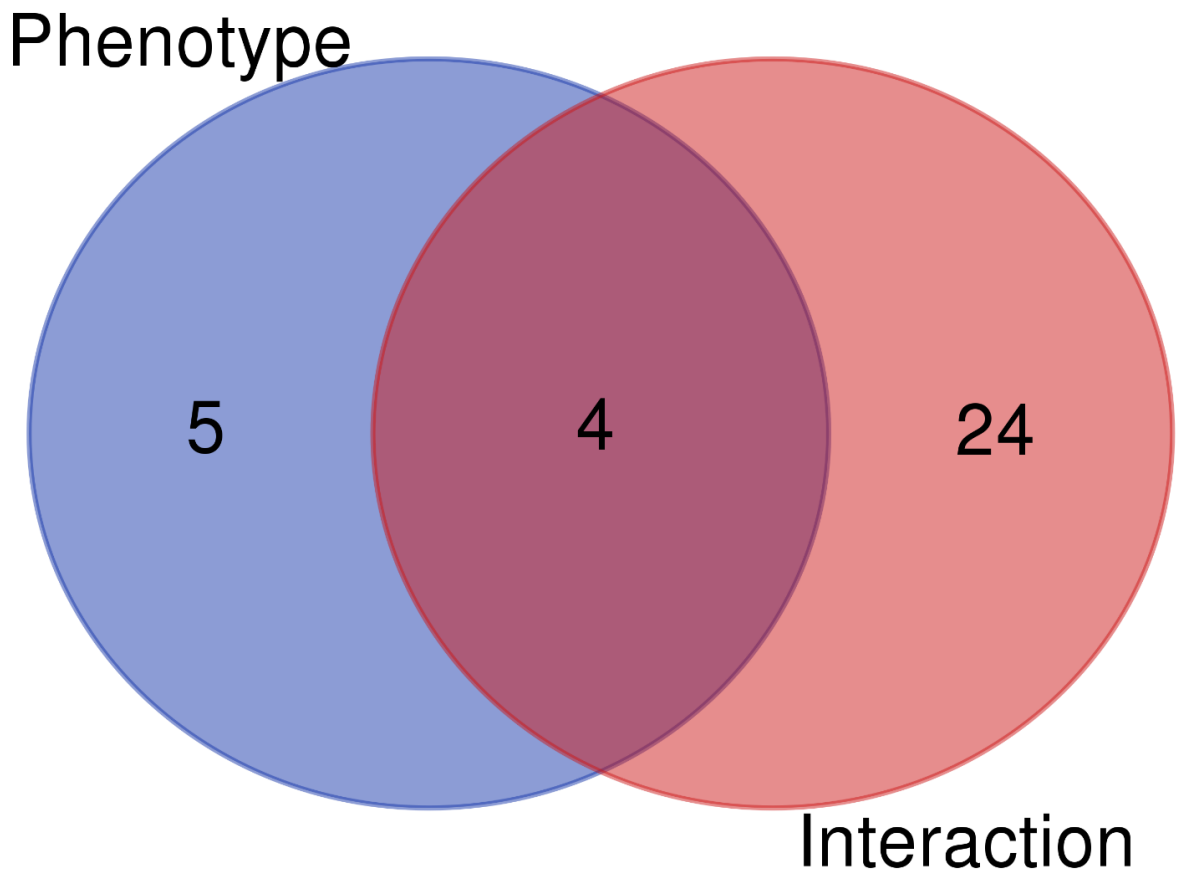


Figure 4-11. Multivariate analysis results, when Zalkecks is compared against PotatoSplit, displayed in a Venn Diagram to show the relationship between proteins that are significant based on phenotype or the interaction between time and phenotype.

4.4 Examining Proteins from Phage-Treated Samples

This section displays the results of all proteins found when each phage-infected sample is compared with the phage buffed infected control samples or another phage-infected sample. MetaboAnalyst was used to perform univariate and multivariate analysis. Each comparison below has a heatmap distribution of each protein, at the different time points, in each sample. Appendix Table 1, Appendix Table 3, and Appendix Table 5 displays significant proteins based on a univariate analysis of a linear model with covariate adjustment, using a p-value cut-off of 0.05 to determine significance of proteins found in the phage-infected samples as compared to the control samples or the other phage-infected samples. Appendix Table 2, Appendix Table 4, and Appendix Table 6 display significant proteins based on a multivariate analysis, to determine the significance of each protein based on the phenotype, time, or interaction of the samples, using a cut-off value of 0.9 for leverage and 0.05 for p-value. A visual representation of the multivariate analysis, and the relationship between proteins that are significant based on time, phenotype, or interaction, can be seen in the Venn diagram in Figure 4-13, Figure 4-19, and Figure 4-25 for each comparison. The fold change and regulation type of each protein was also analyzed through DAVID, a functional annotation bioinformatics microarray analysis program.

4.4.1 PotatoSplit Compared Against the Control

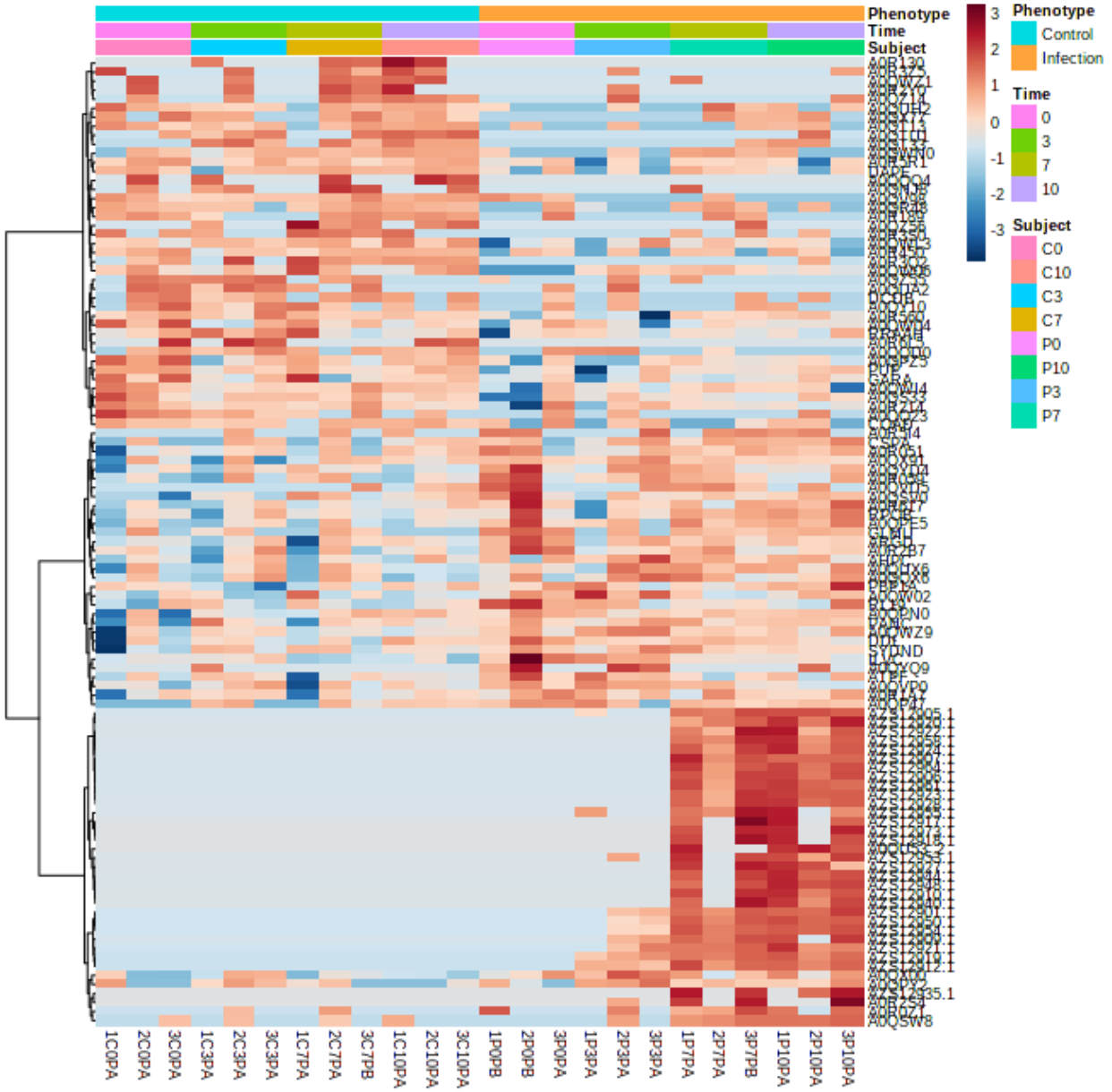


Figure 4-12. Heatmap of all significant ANOVA proteins found in mycobacteriophage PotatoSplit treated samples when compared against the control samples, at each time point taken.

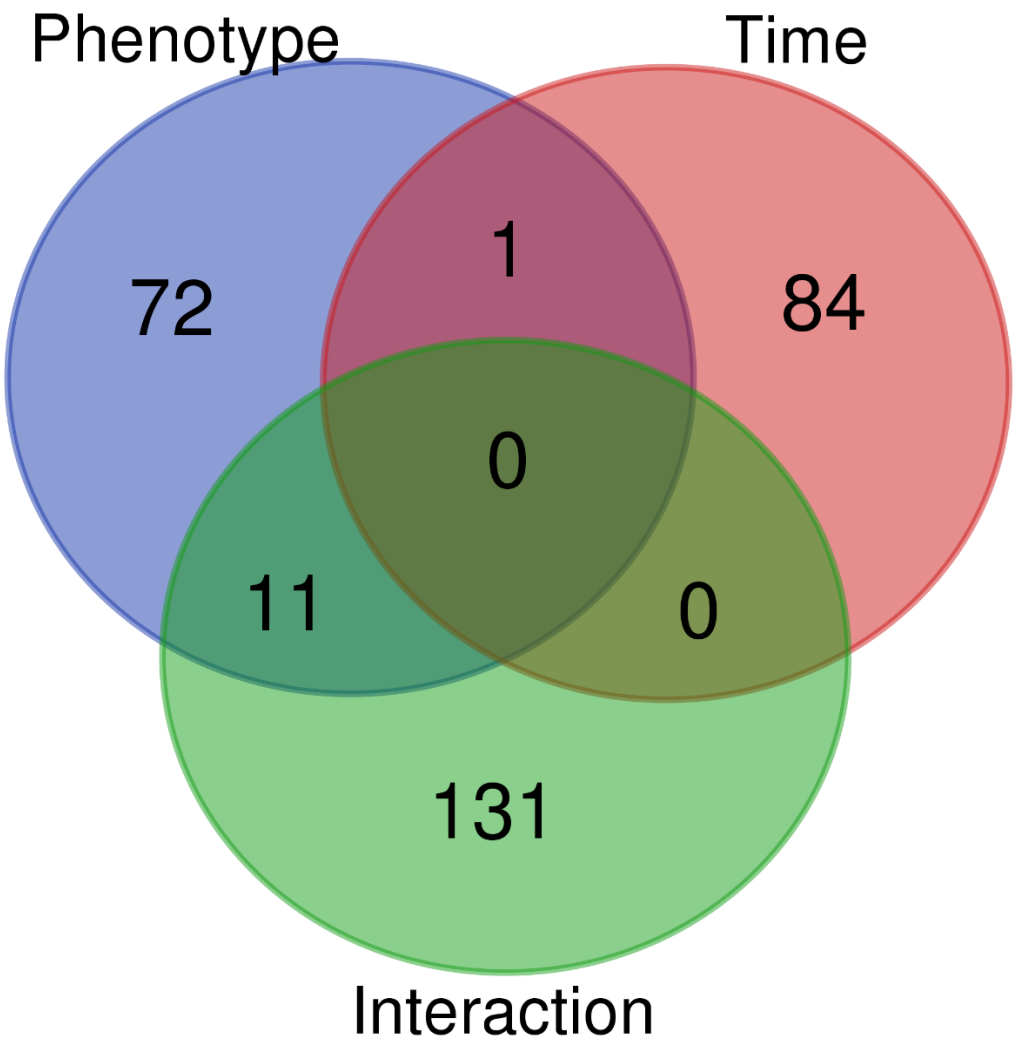


Figure 4-13. Multivariate analysis results, when PotatoSplit is compared against the control, displayed in a Venn Diagram to show the relationship between proteins that are significant based on time, phenotype, or interaction.

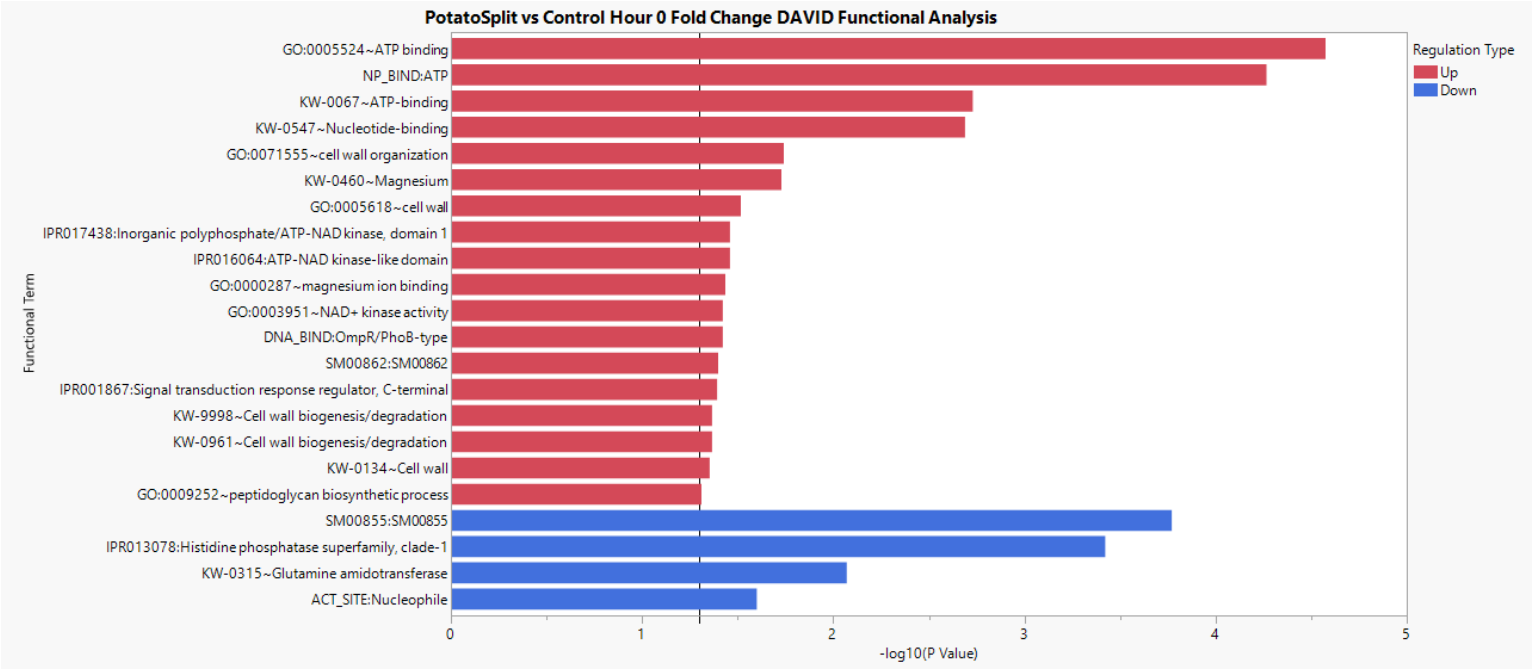


Figure 4-14. Bar graph showing the significance and regulation type of proteins found in hour 0 samples, when PotatoSplit is compared to the control. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

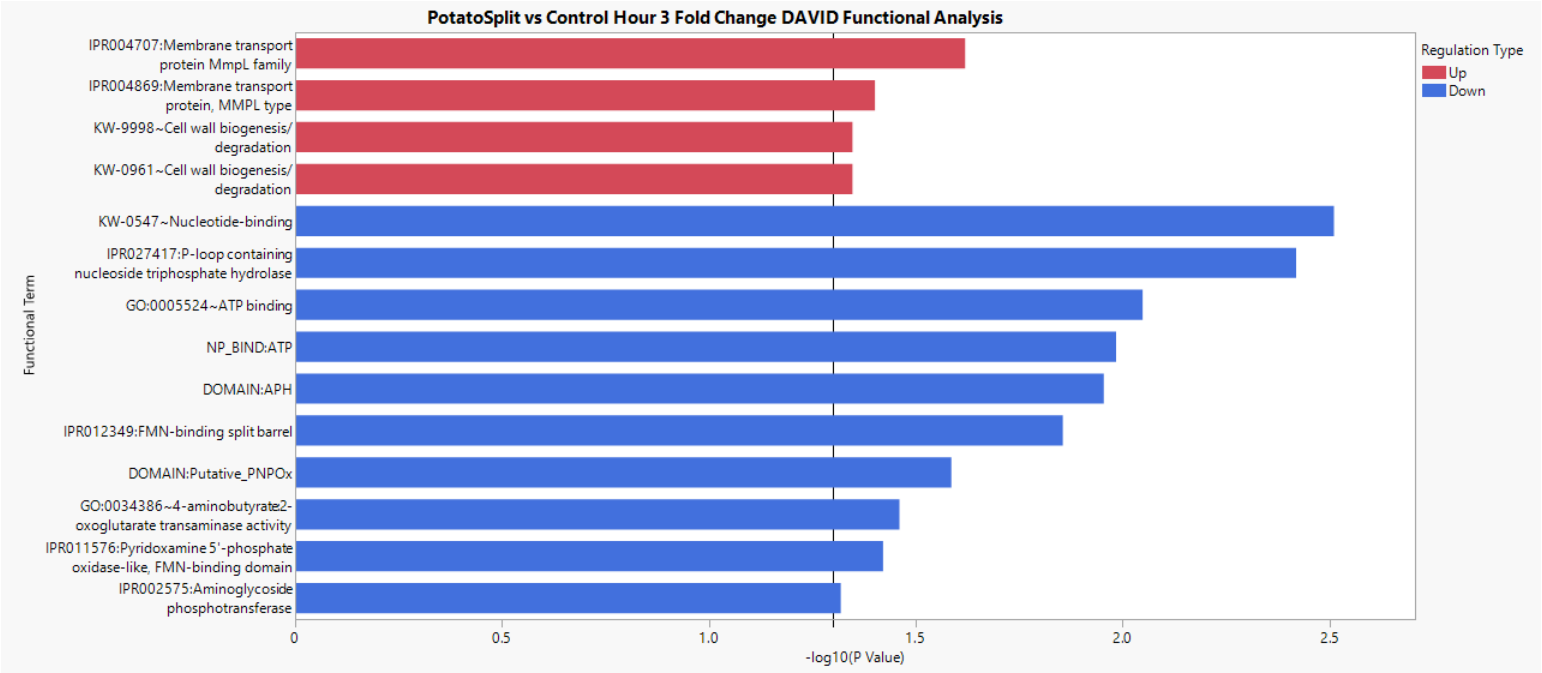


Figure 4-15. Bar graph showing the distribution and regulation type of proteins found in hour 3 samples, when PotatoSplit is compared to the control. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

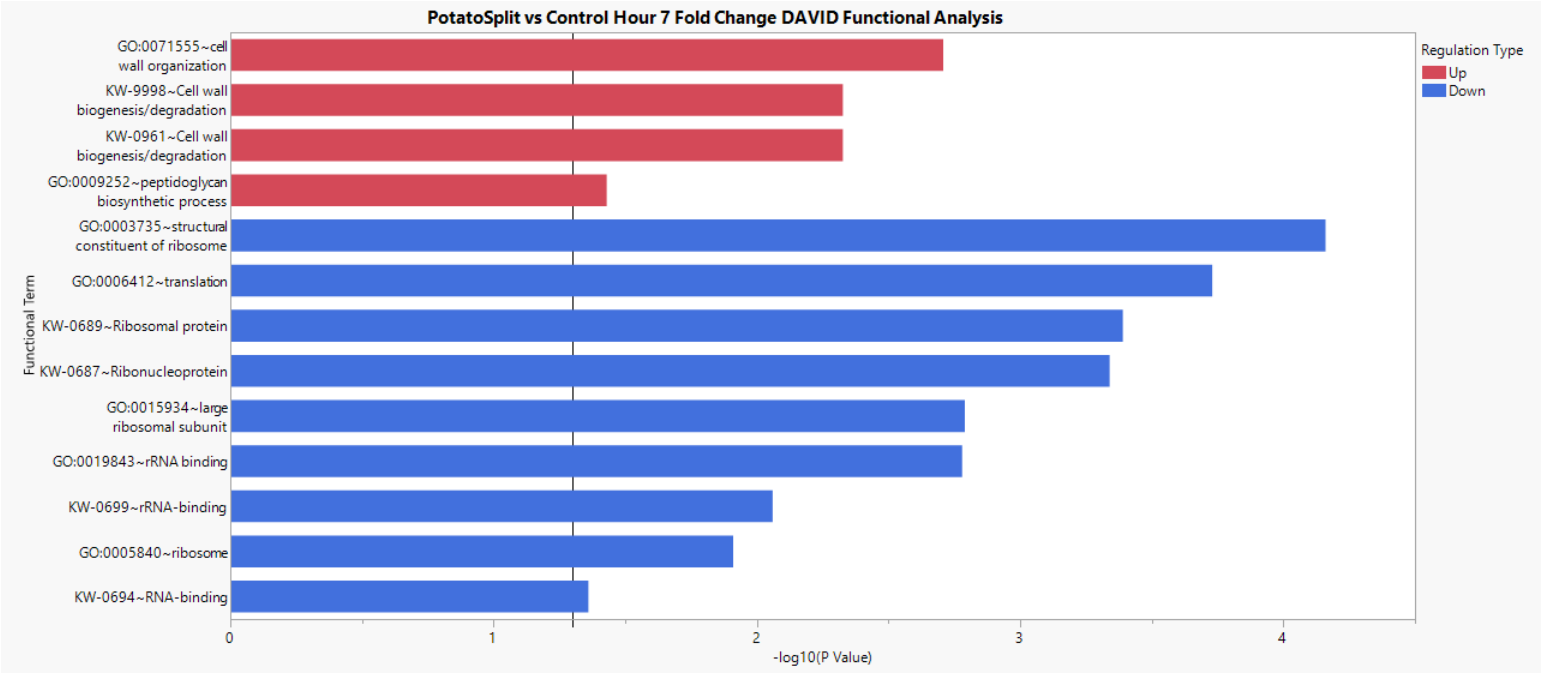


Figure 4-16. Bar graph showing the distribution and regulation type of proteins found in hour 7 samples, when PotatoSplit is compared to the control. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

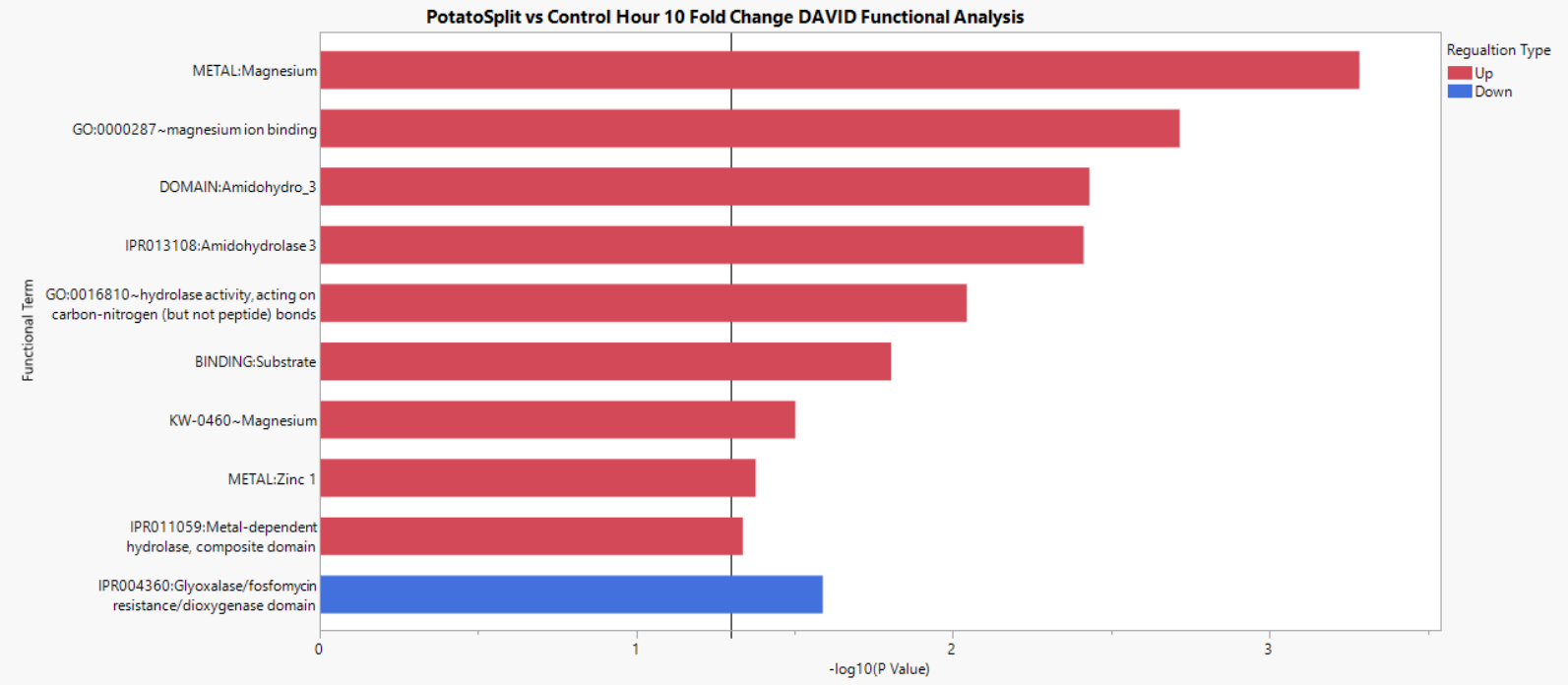


Figure 4-17. Bar graph showing the distribution and regulation type of proteins found in hour 10 samples, when PotatoSplit is compared to the control. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

4.4.2 Zalkecks Compared Against the Control

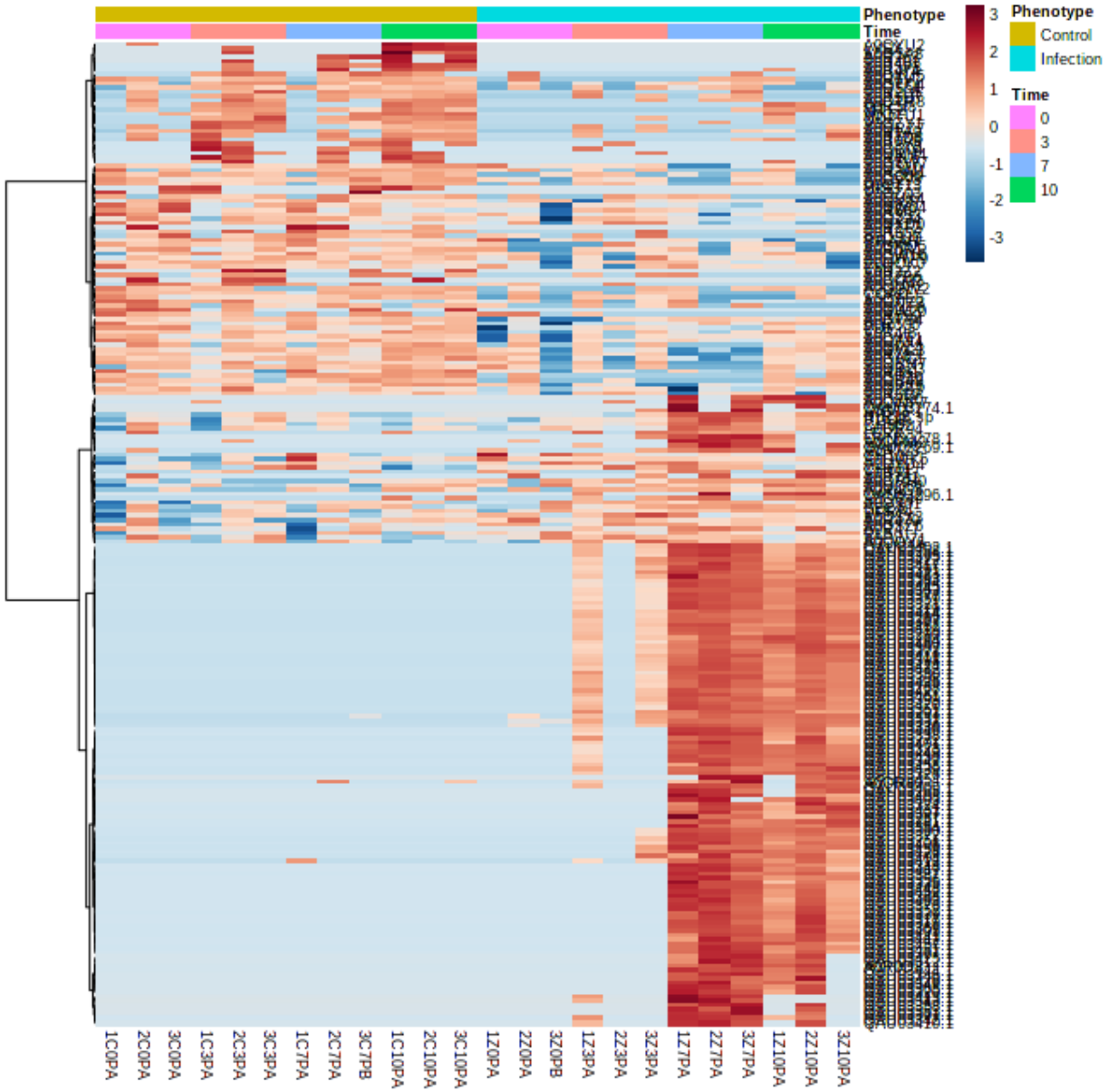


Figure 4-18. Heatmap of all significant ANOVA proteins found in mycobacteriophage Zalkecks treated samples when compared against the control samples, at each time point taken.

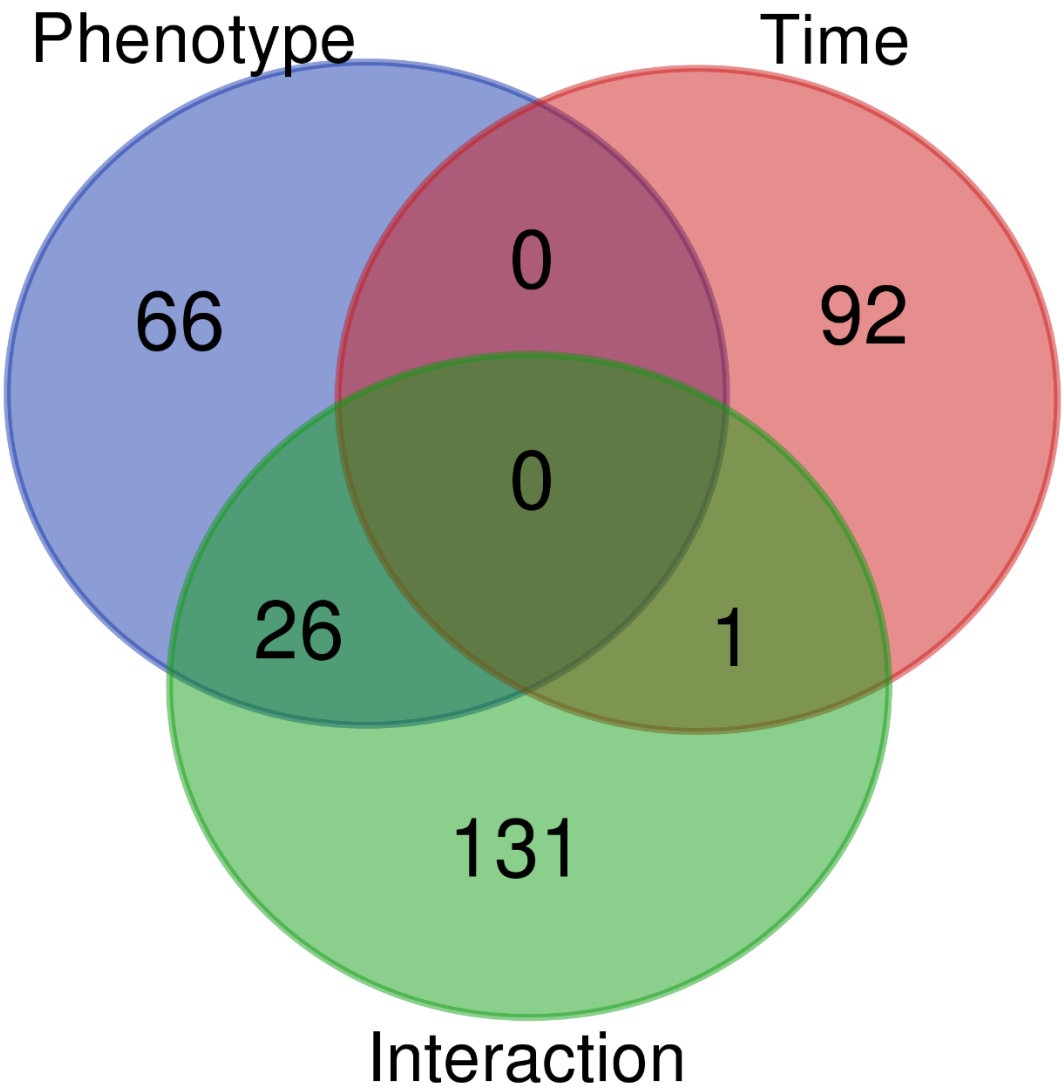


Figure 4-19. Multivariate analysis results, when Zalkecks is compared against the control, displayed in a Venn Diagram to show the relationship between proteins that are significant based on time, phenotype, or interaction.

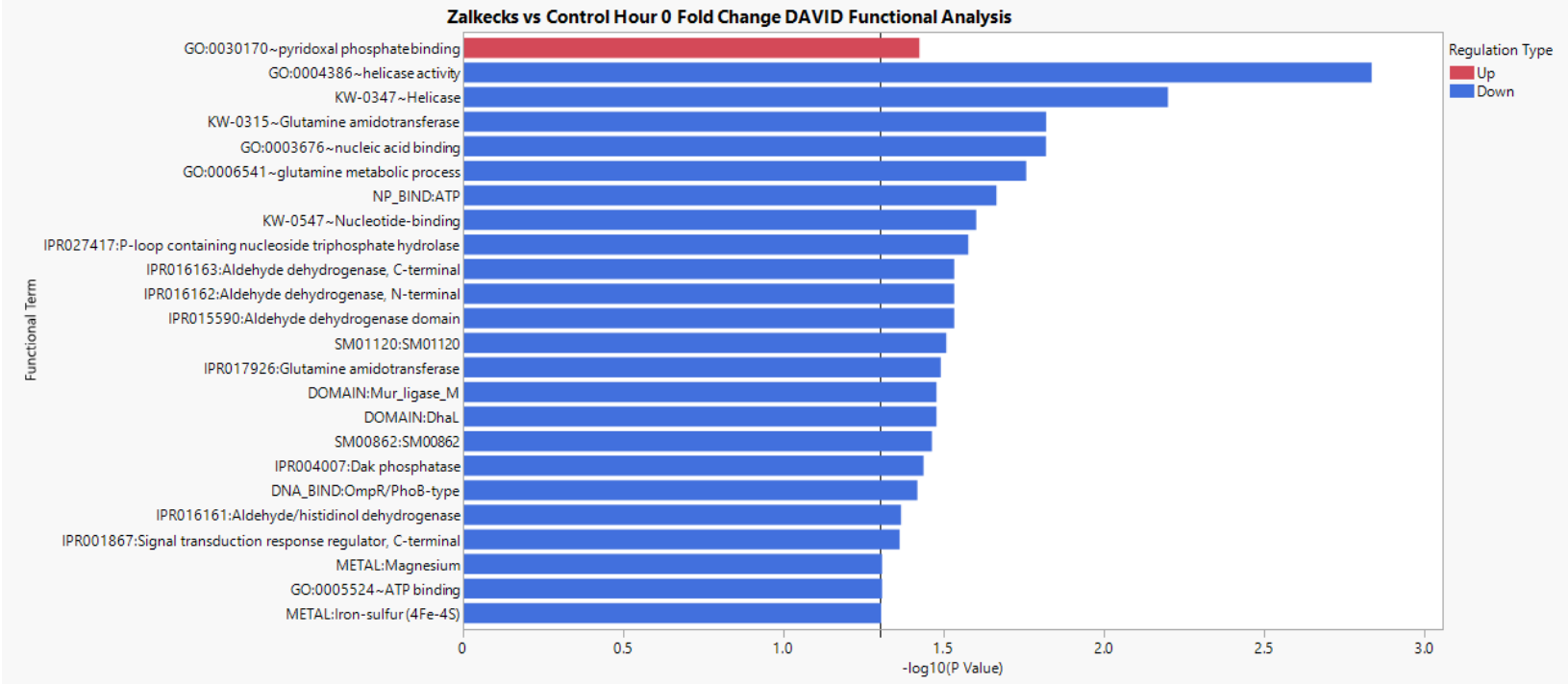


Figure 4-20. Bar graph showing the distribution and regulation type of proteins found in hour 0 samples, when Zalkecks is compared to the control. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

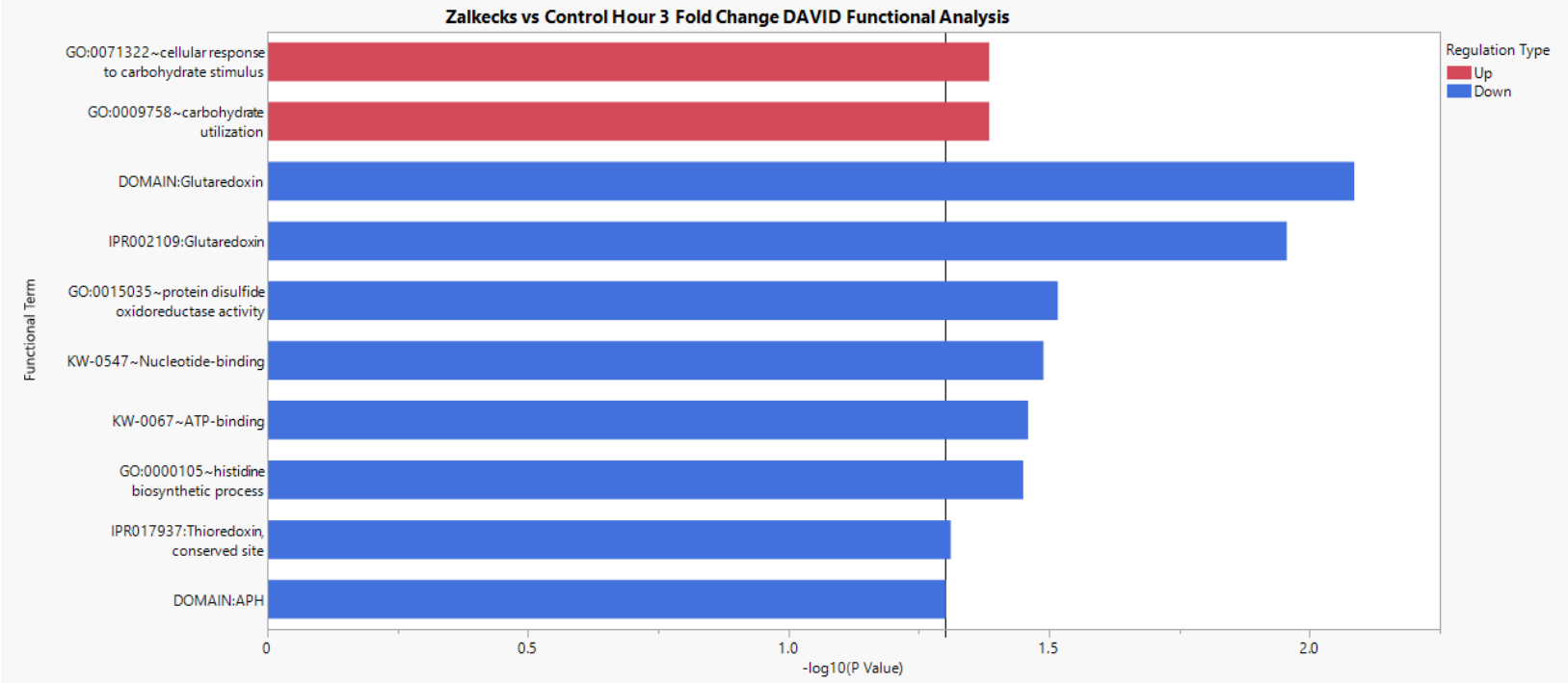


Figure 4-21. Bar graph showing the distribution and regulation type of proteins found in hour 3 samples, when Zalkecks is compared to the control. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

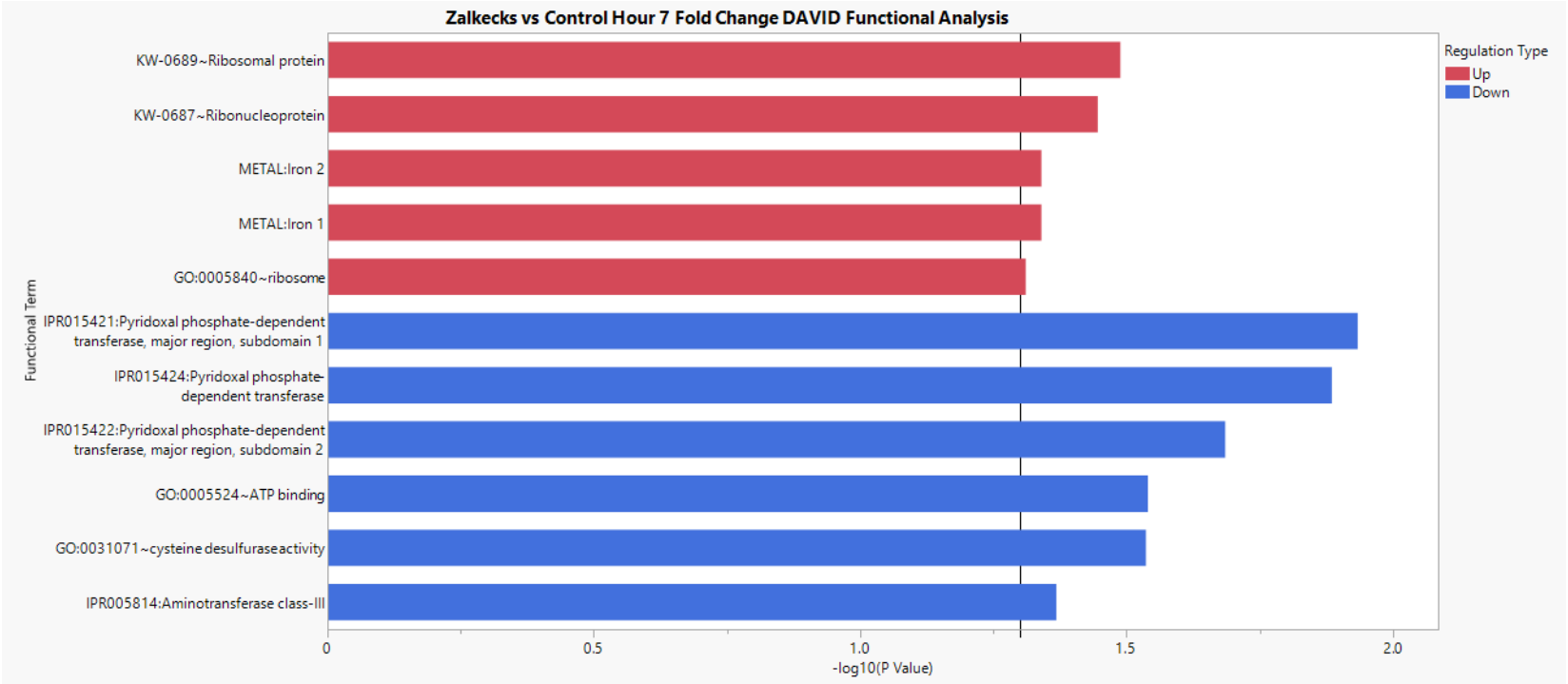


Figure 4-22. Bar graph showing the distribution and regulation type of proteins found in hour 0 samples, when Zalkecks is compared to the control. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

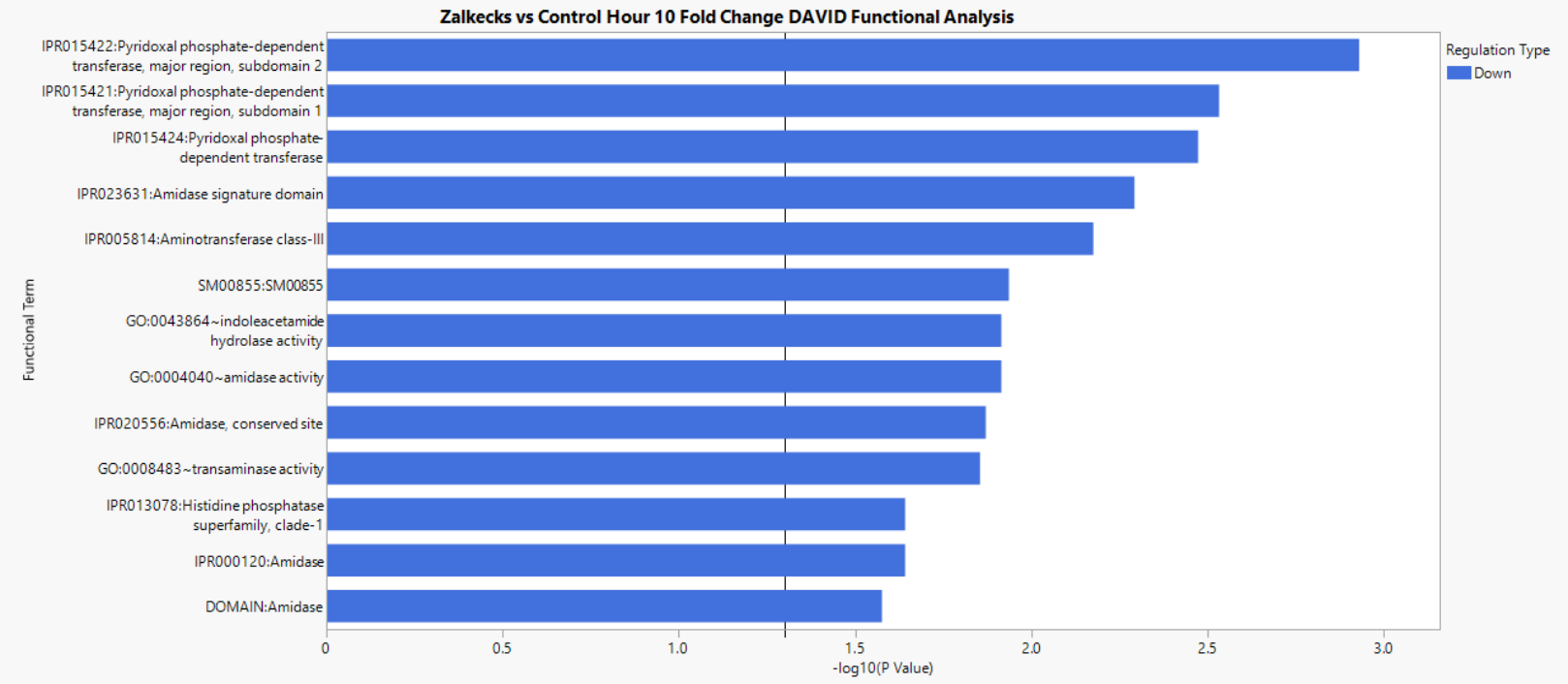


Figure 4-23. Bar graph showing the distribution and regulation type of proteins found in hour 0 samples, when Zalkecks is compared to the control. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

4.4.3 Zalkecks Compared Against PotatoSplit

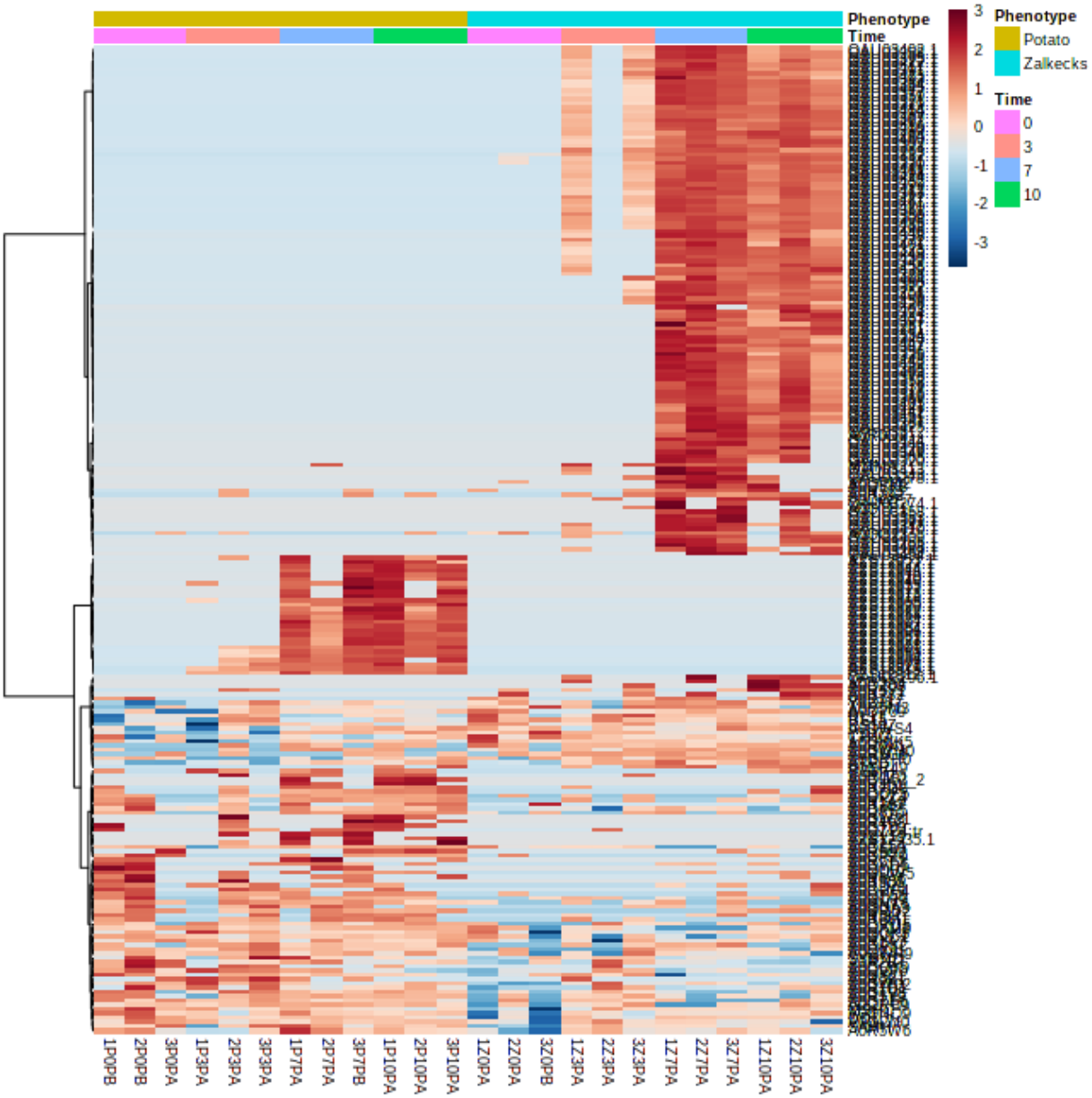


Figure 4-24. Heatmap of all significant ANOVA proteins found in mycobacteriophage PotatoSplit and mycobacteriophage Zalkecks treated samples, when compared against each other, at each time point taken.

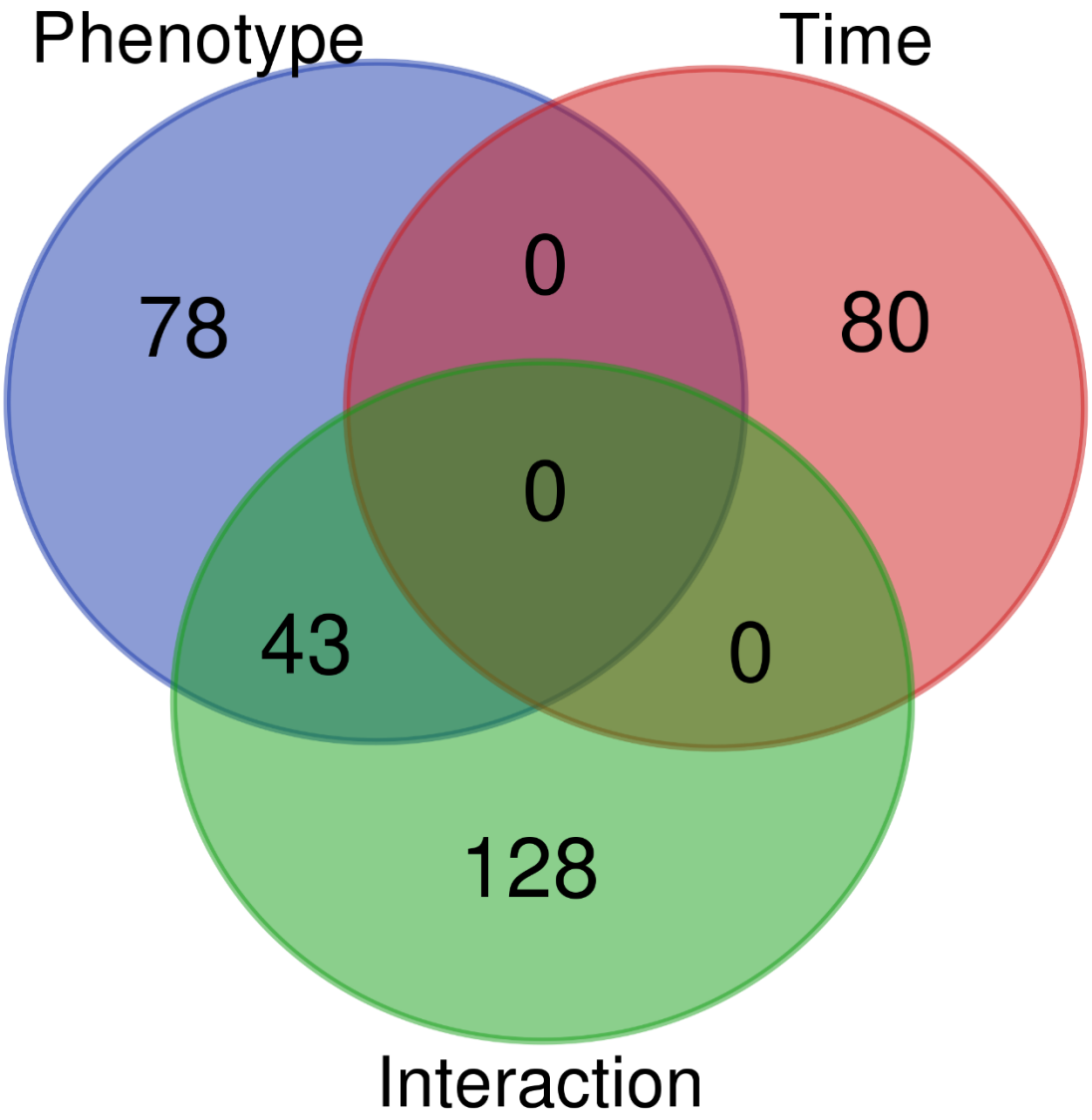


Figure 4-25. Multivariate analysis results, when Zalkecks is compared against PotatoSplit, displayed in a Venn Diagram to show the relationship between proteins that are significant based on time, phenotype, or interaction.

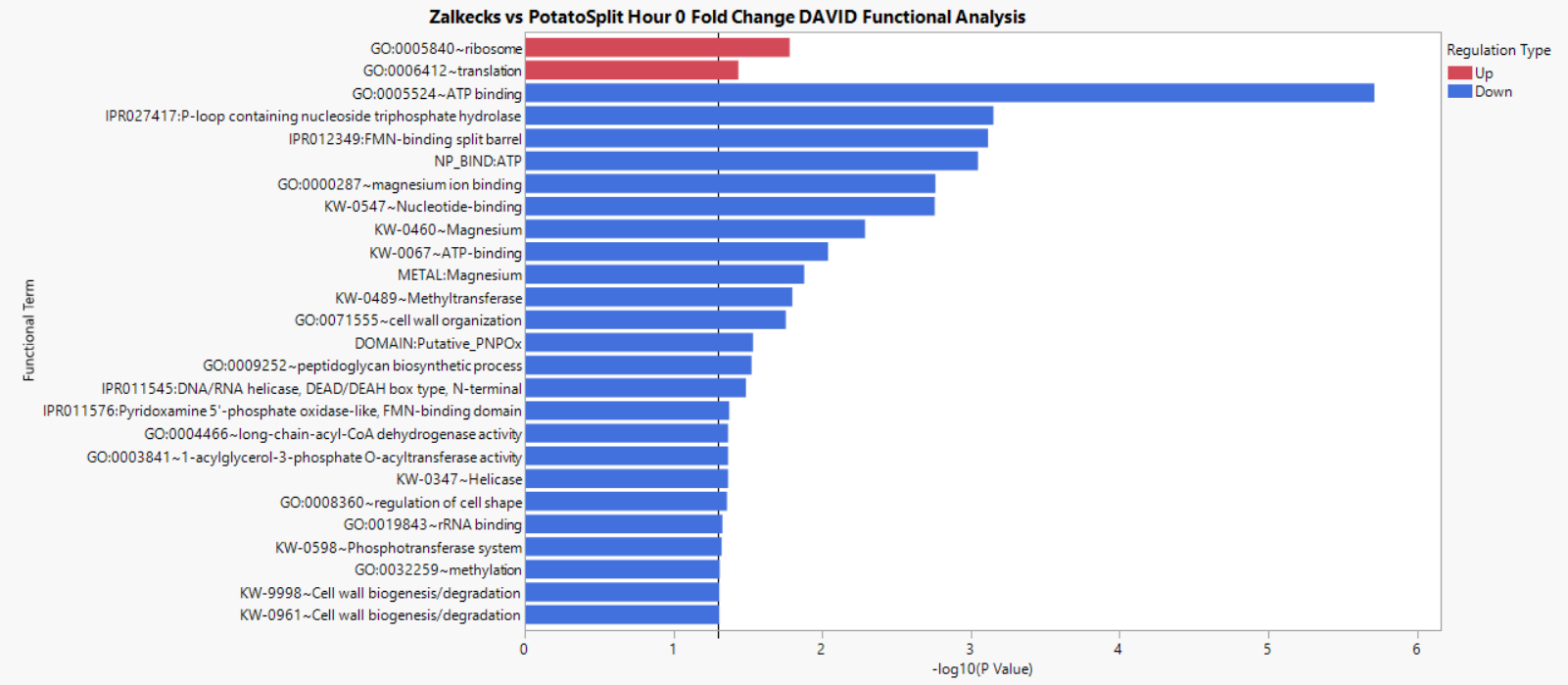


Figure 4-26. Bar graph showing the distribution and regulation type of proteins found in hour 0 samples, when Zalkecks is compared to PotatoSplit. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

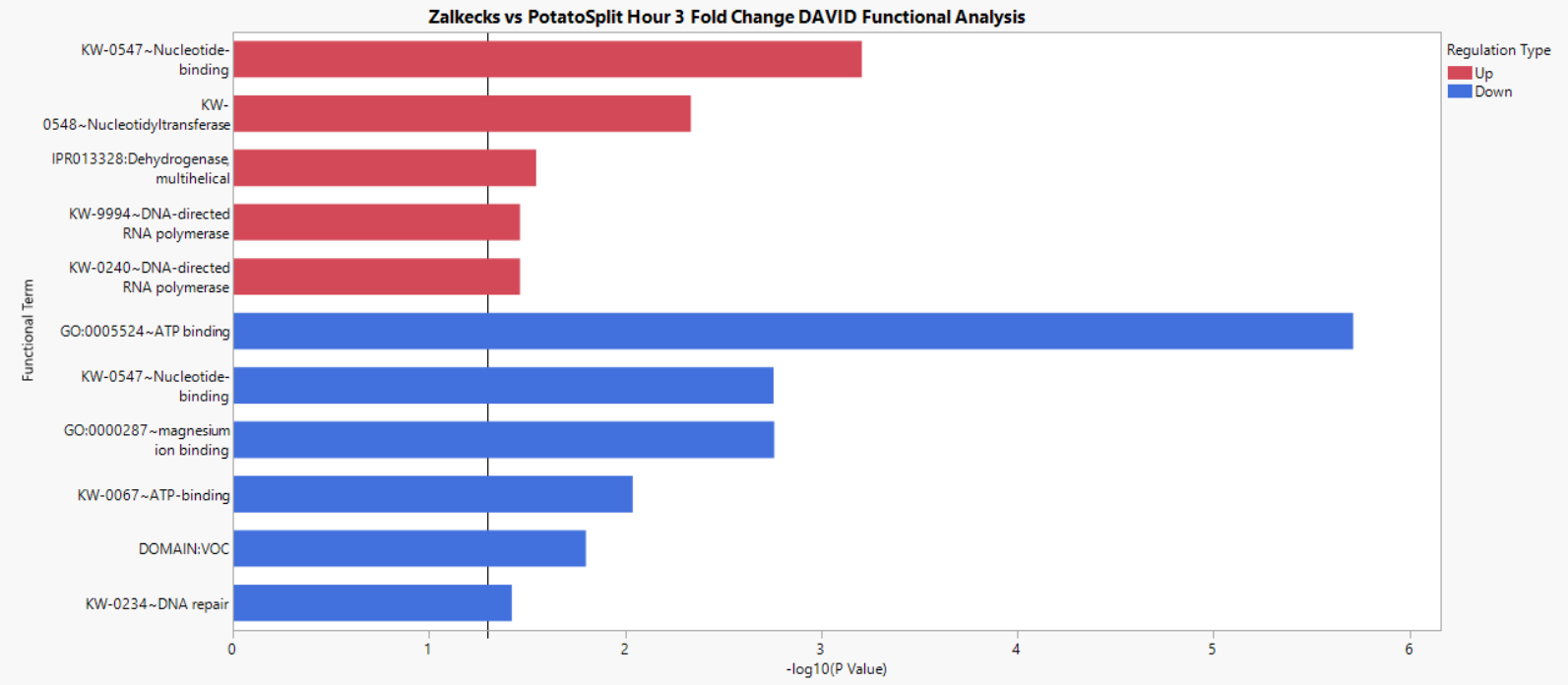


Figure 4-27. Bar graph showing the distribution and regulation type of proteins found in hour 3 samples, when Zalkecks is compared to PotatoSplit. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

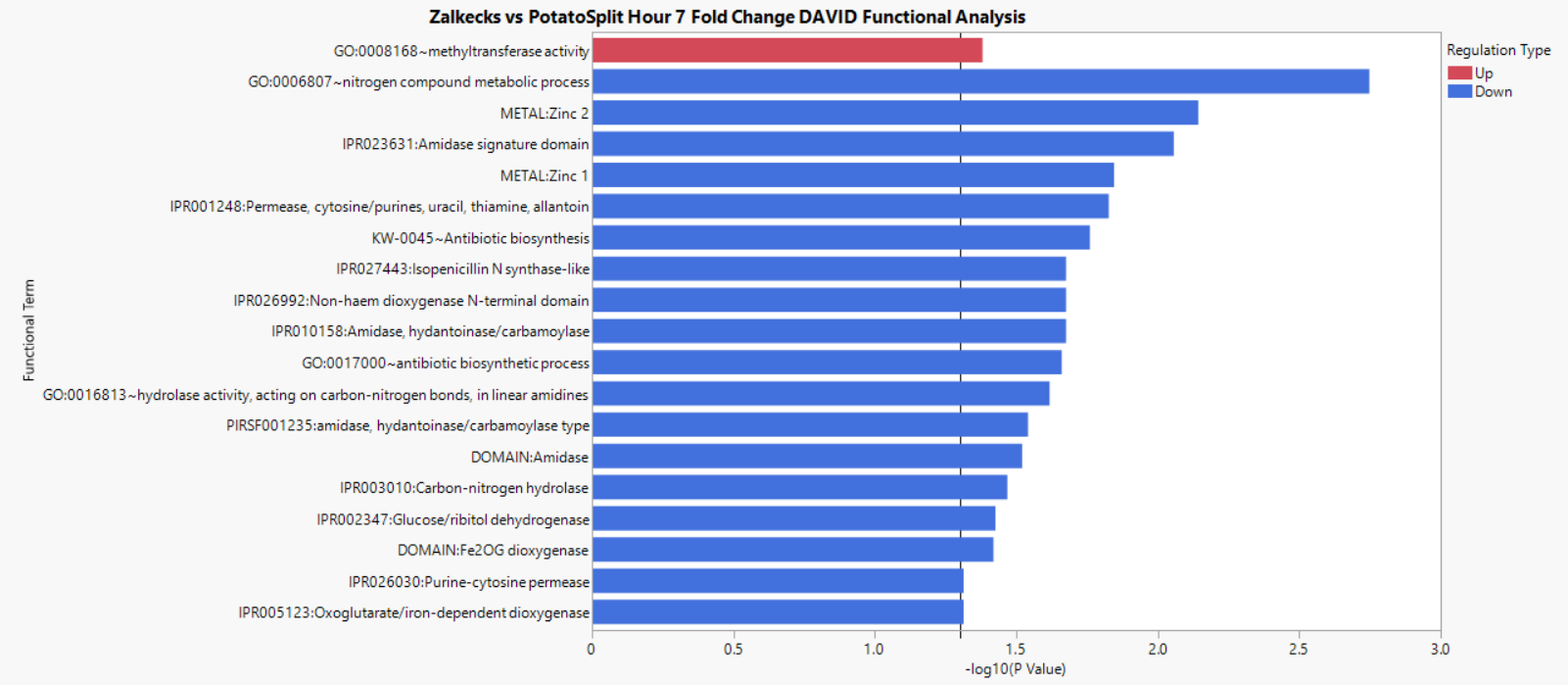


Figure 4-28. Bar graph showing the distribution and regulation type of proteins found in hour 7 samples, when Zalkecks is compared to PotatoSplit. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

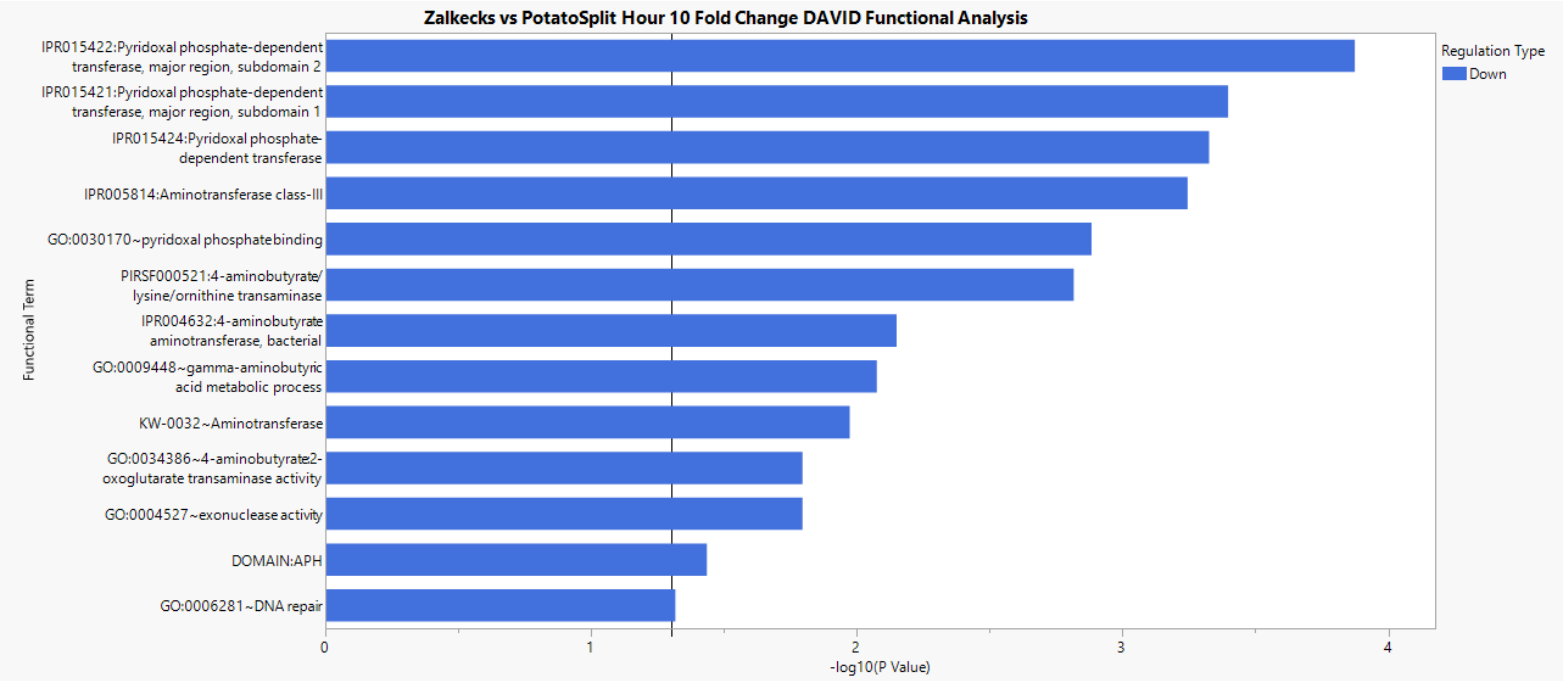


Figure 4-29. Bar graph showing the distribution and regulation type of proteins found in hour 10 samples, when Zalkecks is compared to PotatoSplit. Red demonstrates up-regulated proteins and blue represents down-regulated proteins.

4.5 Examining the Correlation of Proteins to Lipids in Phage-Treated Samples

This section aims to find any correlation between the proteins and lipids found in the phage-treated samples when compared with the phage buffer infected control samples or other phage-infected samples. MetaboAnalyst was used to collect the significant protein and lipid data. JMP was implemented to compare the two sets of data and determine correlation between protein and lipids at each time point. Below are heatmaps that represent either a positive or negative correlation between proteins and lipids, at each timepoint, when focusing on phage proteins only. Plots that contain all of the proteins, both phage and bacteria, are located in the appendix, as Appendix Figure 1 through Appendix Figure 44.

4.5.1 PotatoSplit Compared Against the Control

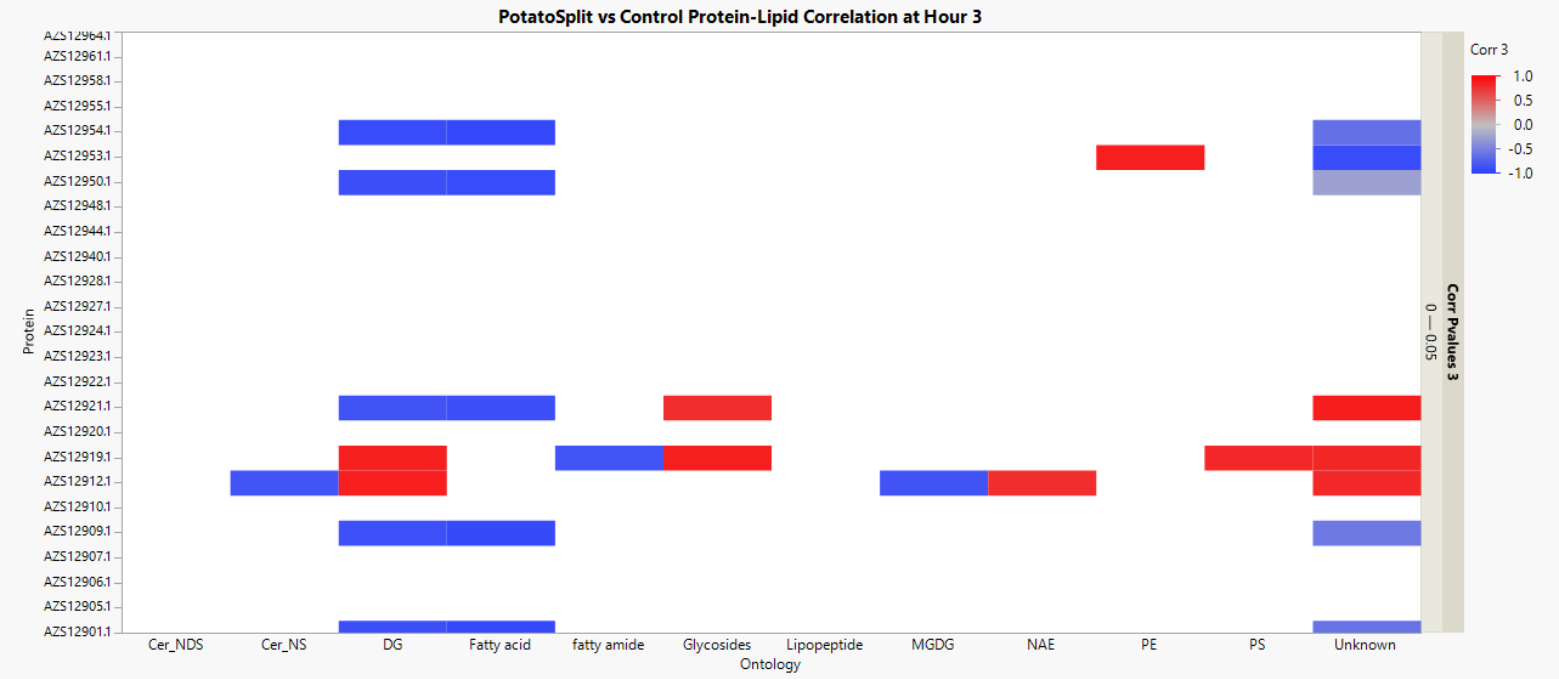


Figure 4-30. Heatmap of the correlation of phage proteins and all lipids, that have a p-value below 0.05 at Hour 3, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

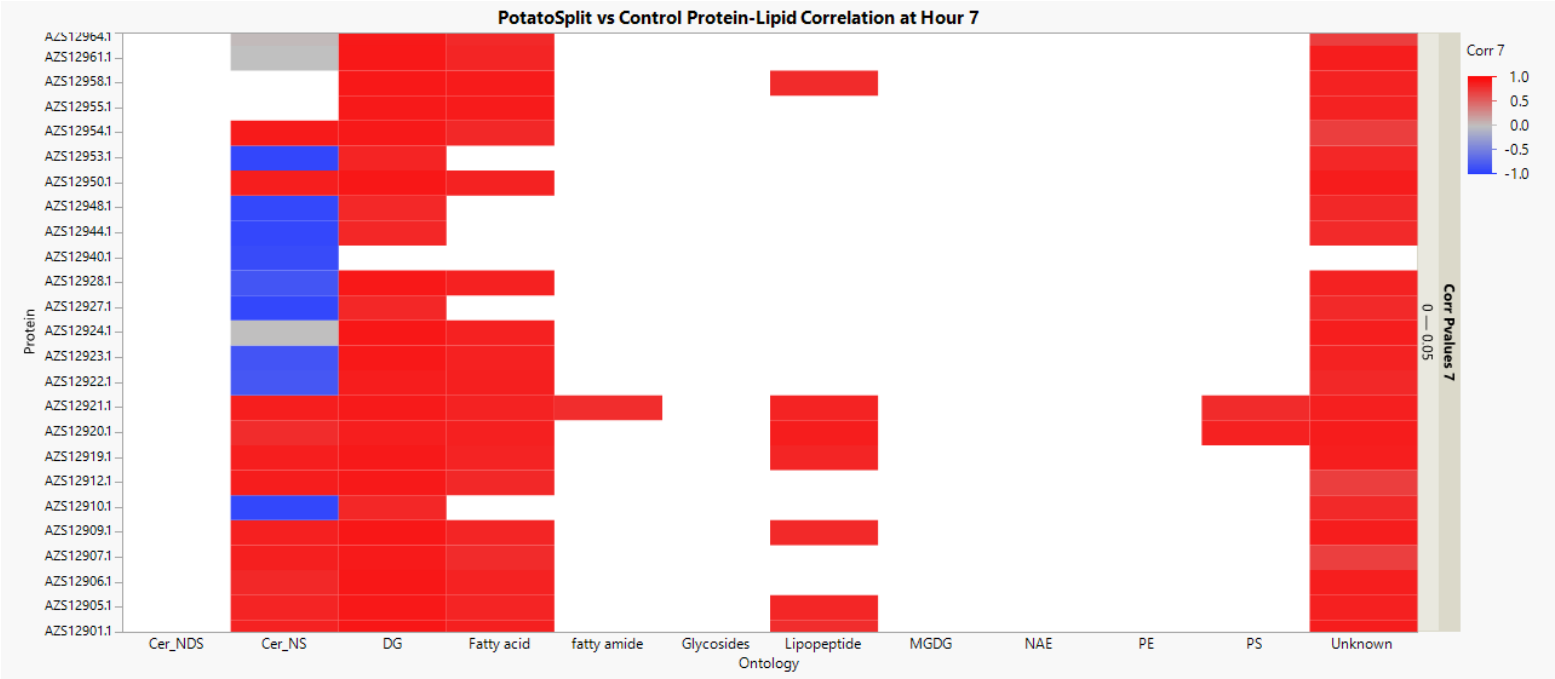


Figure 4-31. Heatmap of the correlation of phage proteins and all lipids, that have a p-value below 0.05 at Hour 7, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

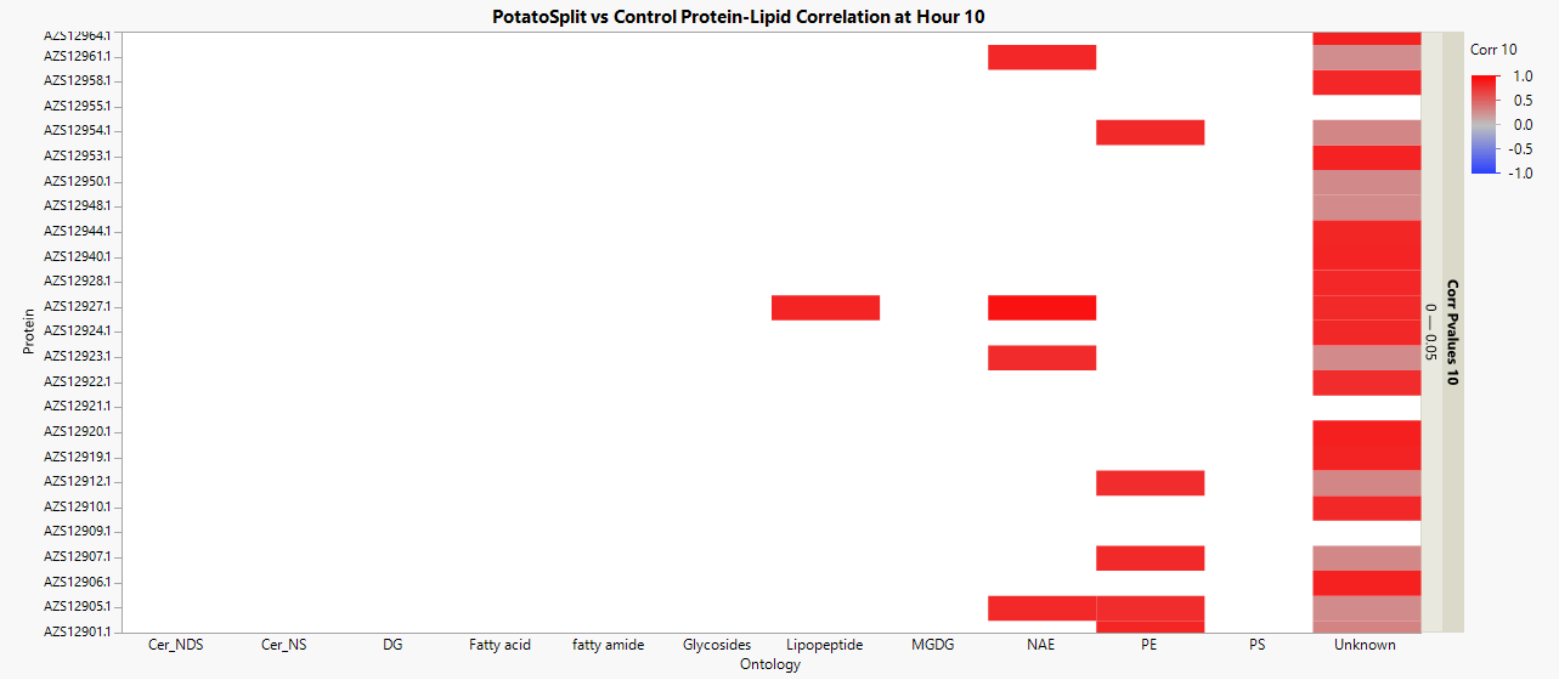


Figure 4-32. Heatmap of the correlation of phage proteins and all lipids, that have a p-value below 0.05 at Hour 10 for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

4.5.2 Zalkecks Compared Against the Control

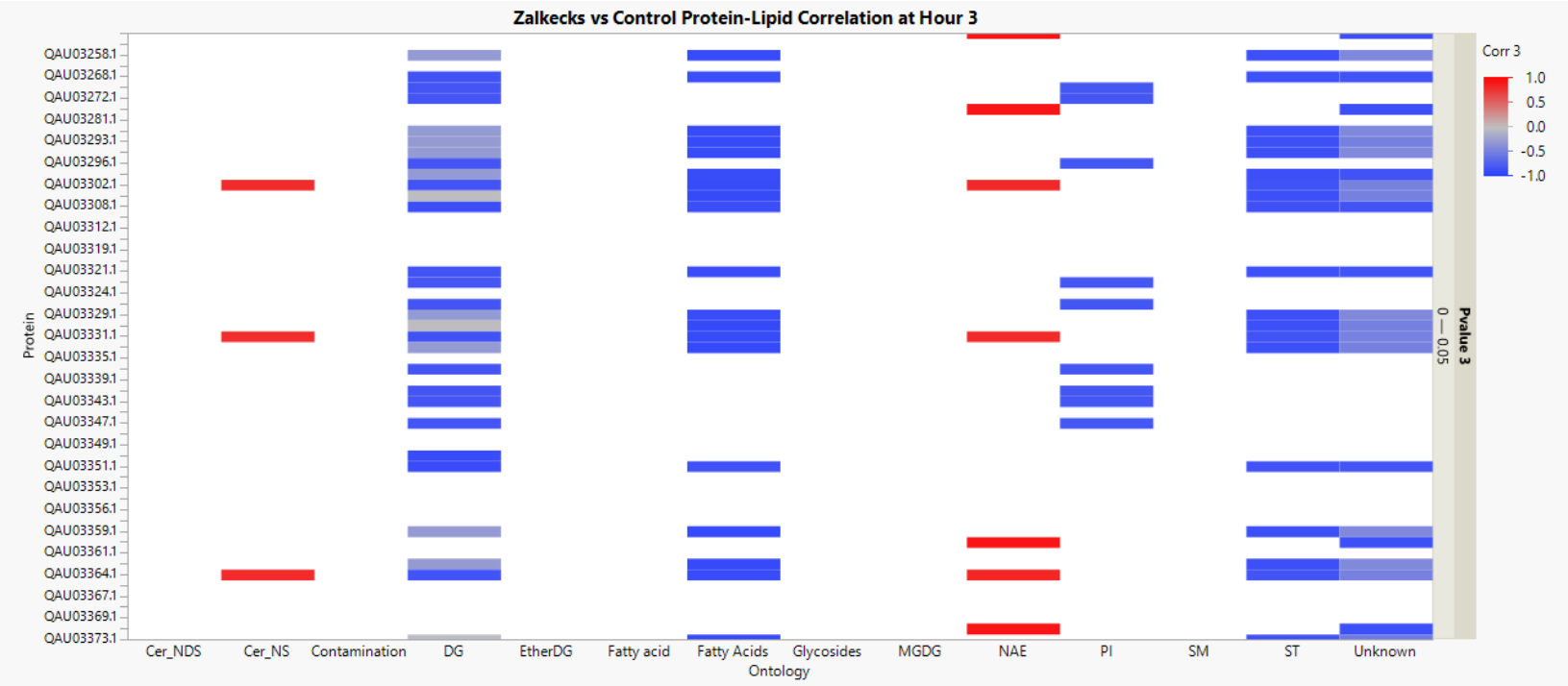


Figure 4-33. Heatmap of the correlation of phage proteins and all lipids, that have a p-value below 0.05 at Hour 3, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

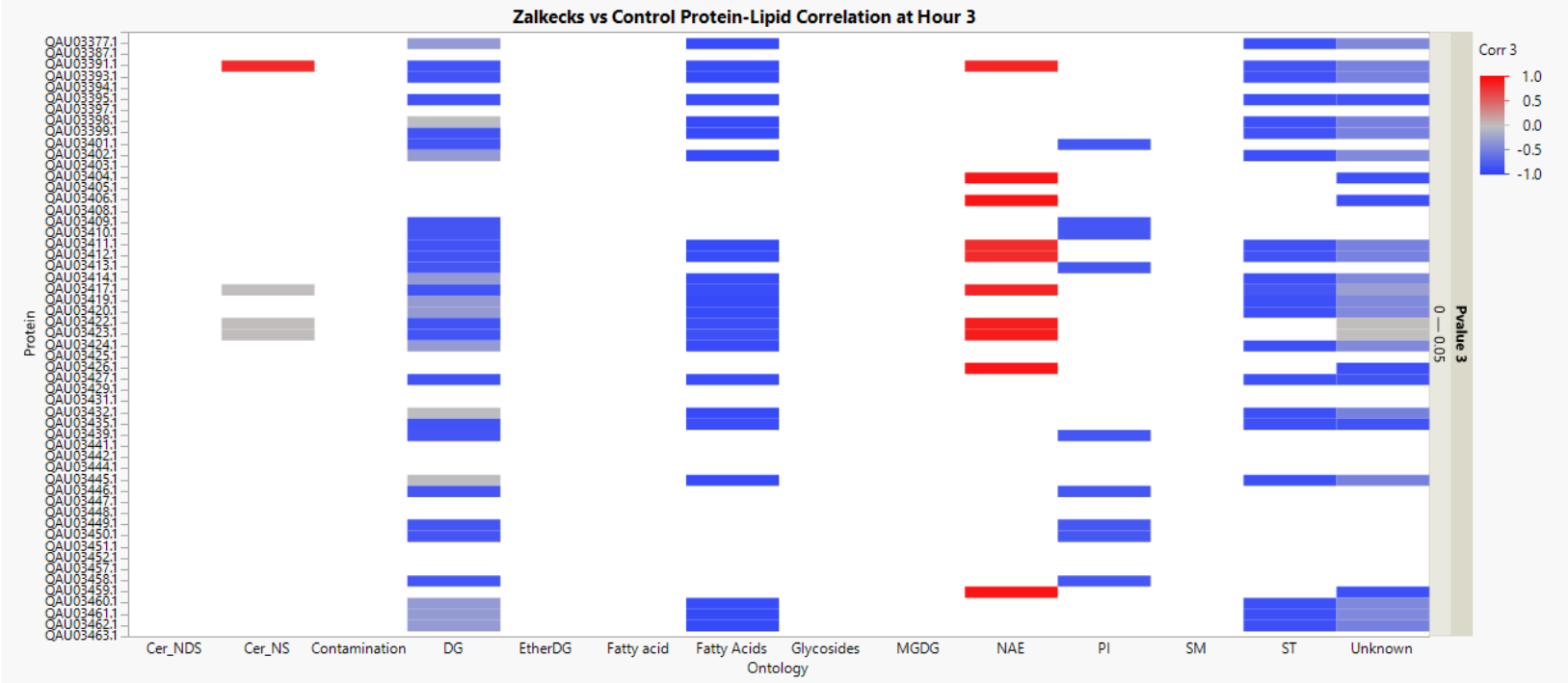


Figure 4-34. Heatmap of the correlation of phage proteins and all lipids, that have a p-value below 0.05 at Hour 3, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

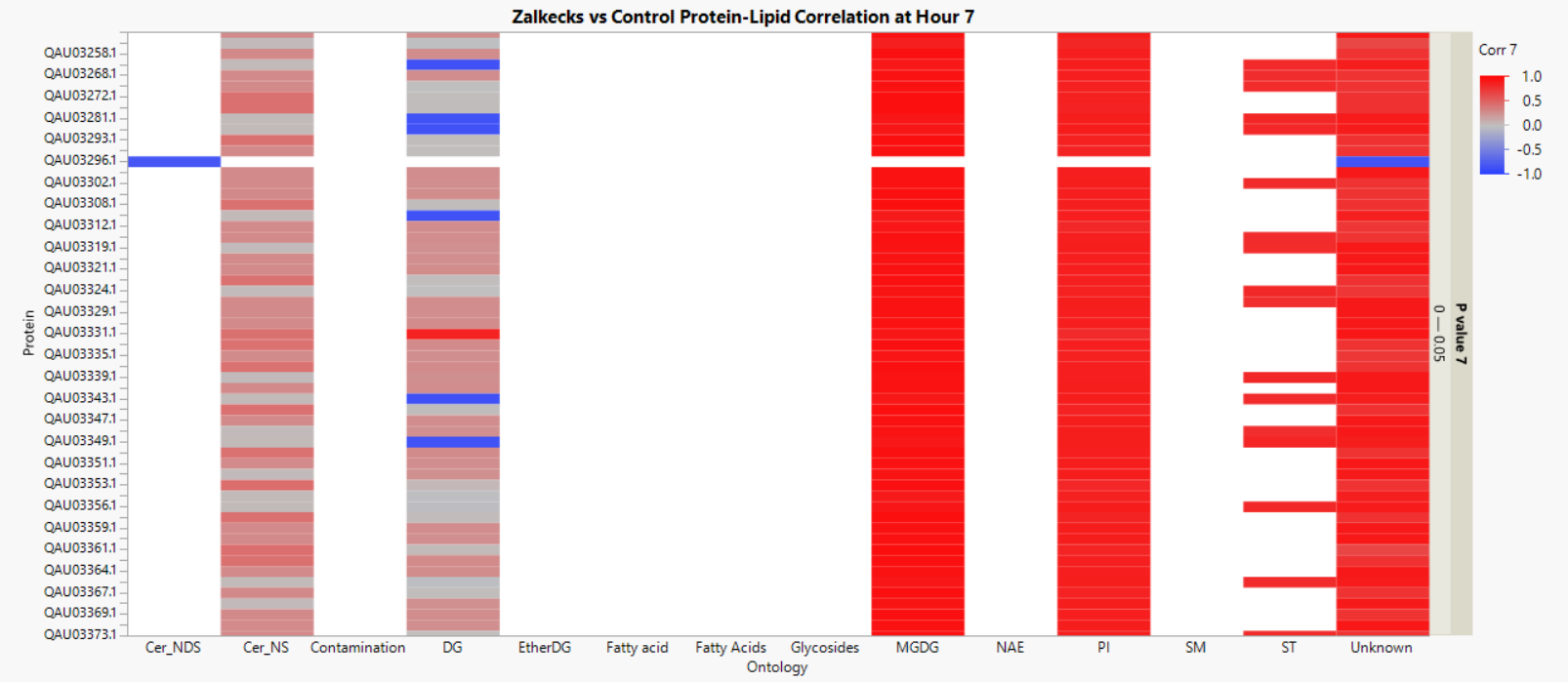


Figure 4-35. Heatmap of the correlation of phage proteins and all lipids, that have a p-value below 0.05 at Hour 7, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

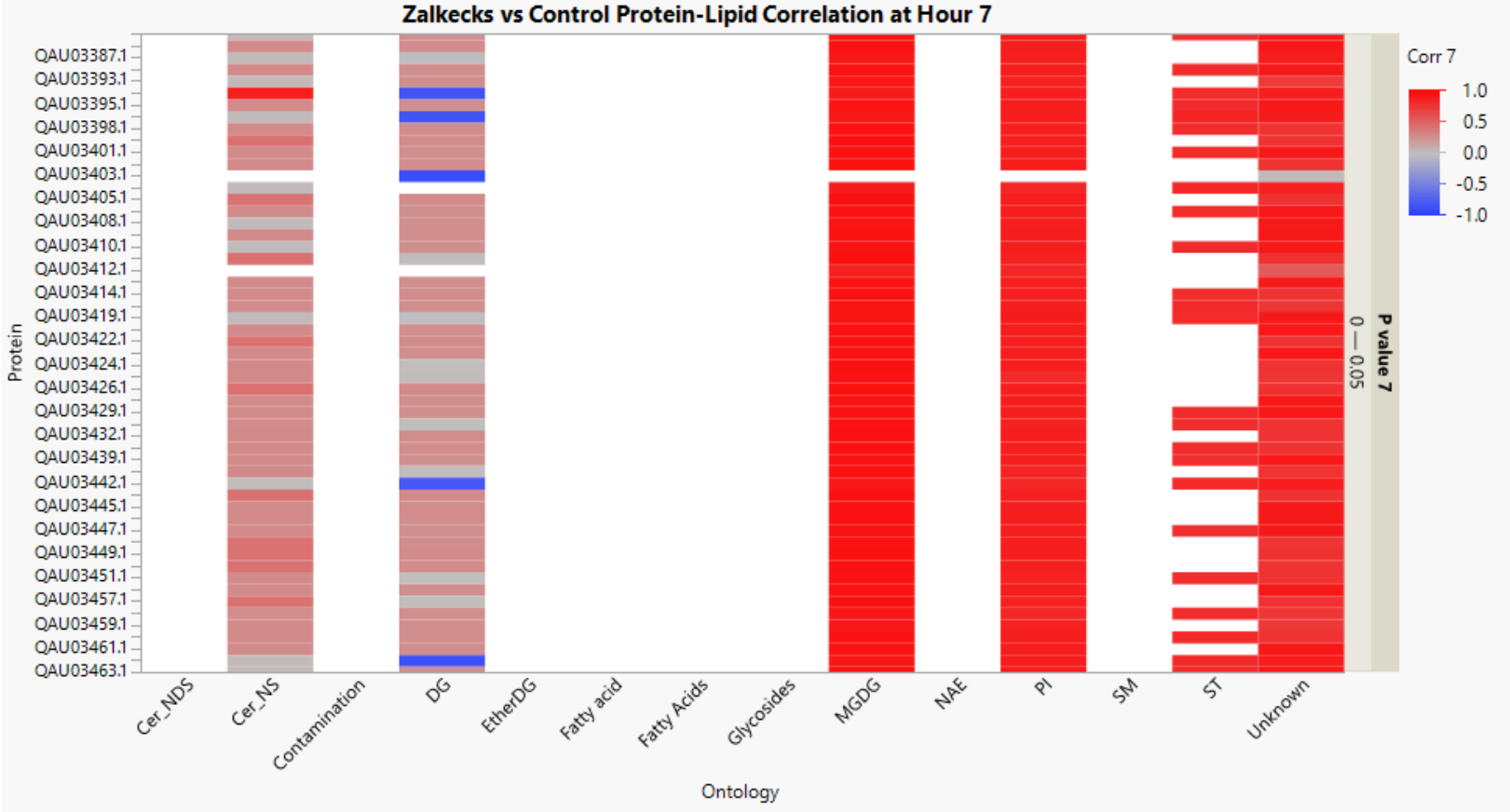


Figure 4-36. Heatmap of the correlation of phage proteins and all lipids, that have a p-value below 0.05 at Hour 7, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

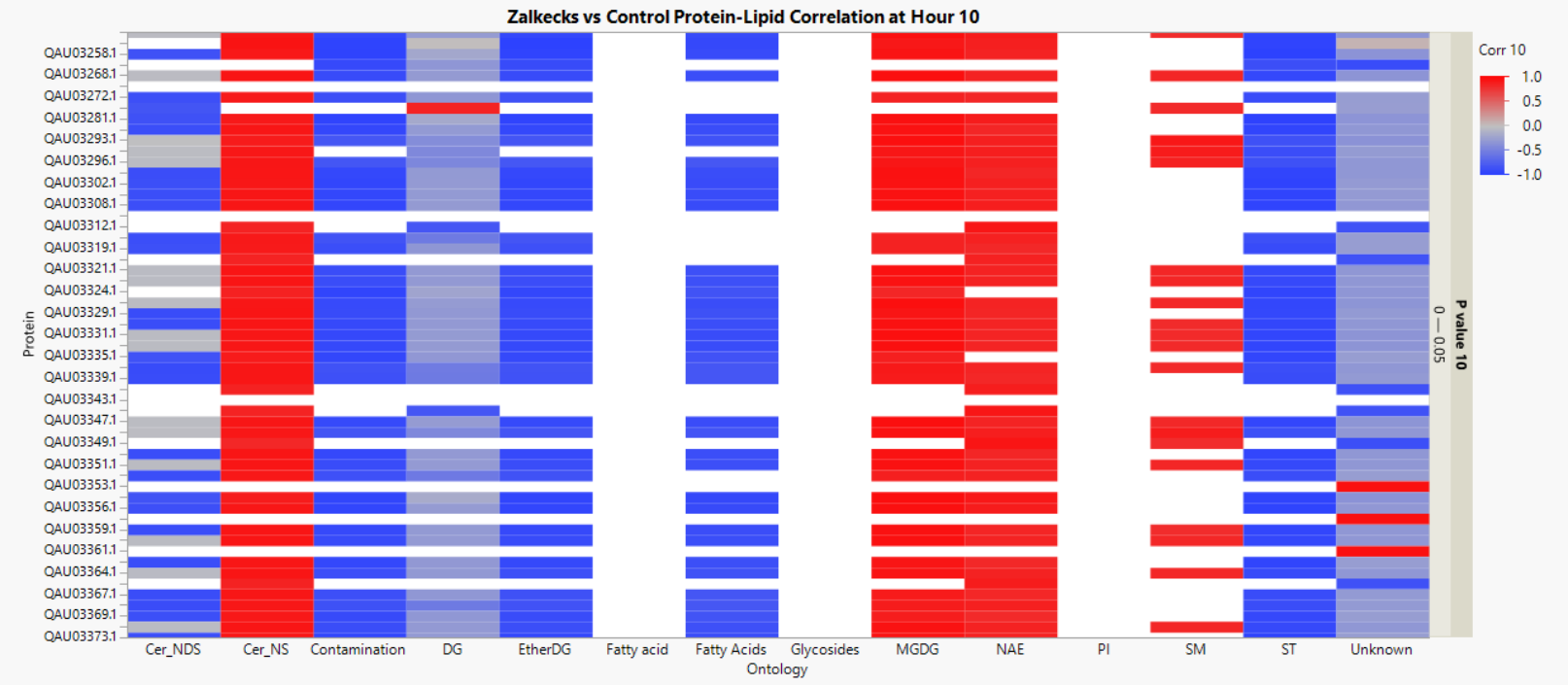


Figure 4-37. Heatmap of the correlation of phage proteins and all lipids, that have a p-value below 0.05 at Hour 10, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

Figure 4-38. Heatmap of the correlation of phage proteins and all lipids, that have a p-value below 0.05 at Hour 10, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

4.6 Comparison of MS Intensity Data Type

Bar graphs were made to determine the difference in results between using LFQ and iBAQ data that is produced from the MS/MS analysis. Both sets of data were loaded into MetaboAnalyst and the results from each of the analyses performed throughout this research project: univariate, multivariate time, multivariate phenotype, and multivariate interaction. The data was then split into bacteria proteins and lipid proteins, before being counted, to show a change in each type of protein. Figure 4-39, Figure 4-40, and Figure 4-41 below represent this data for each of the sample comparisons.

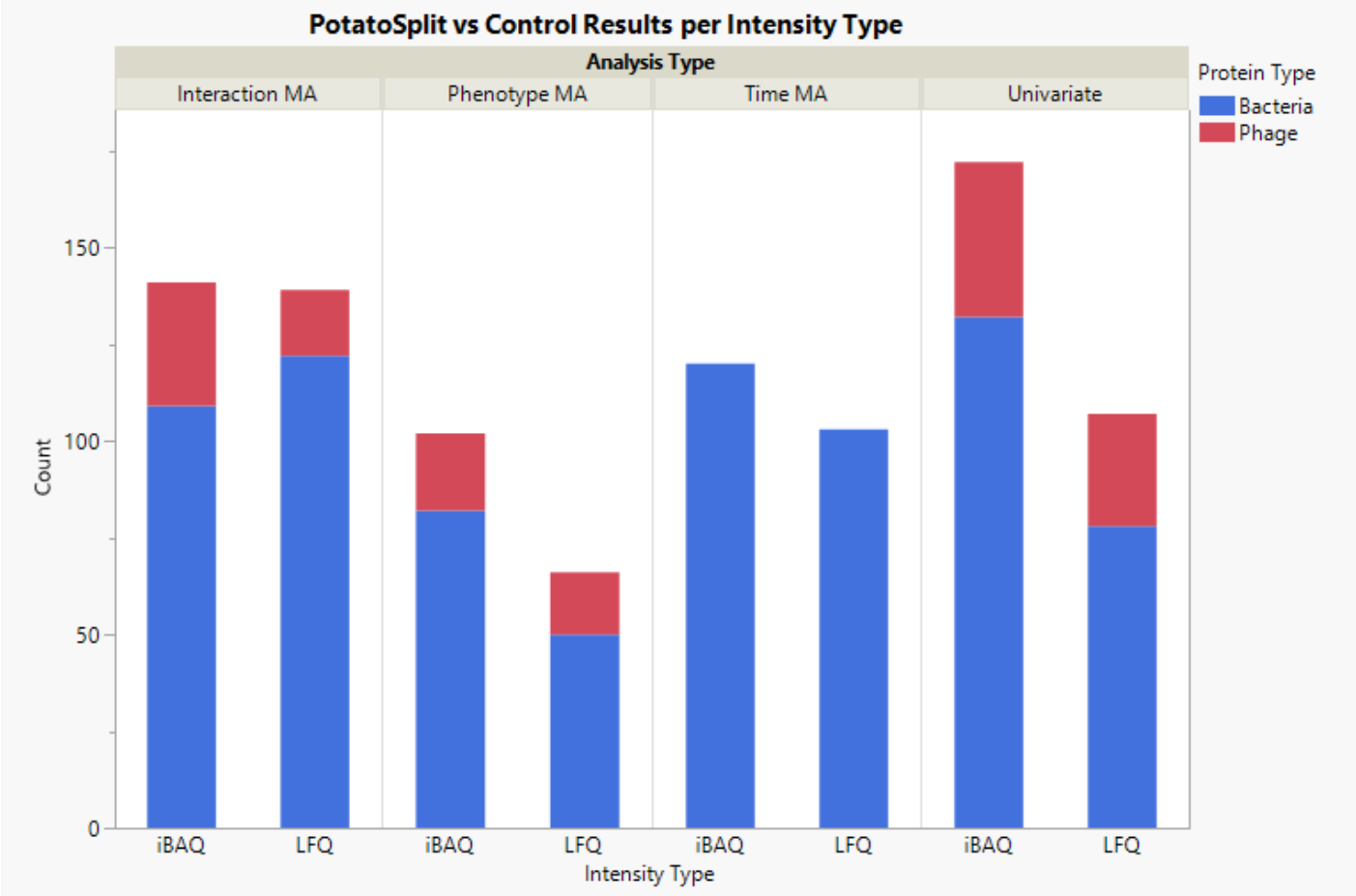


Figure 4-39. Comparison of LFQ vs iBAQ MS/MS data type for the PotatoSplit vs the control experiments. Each section is a different analysis from MetaboAnalyst. Blue represents bacteria proteins and red represents phage proteins.

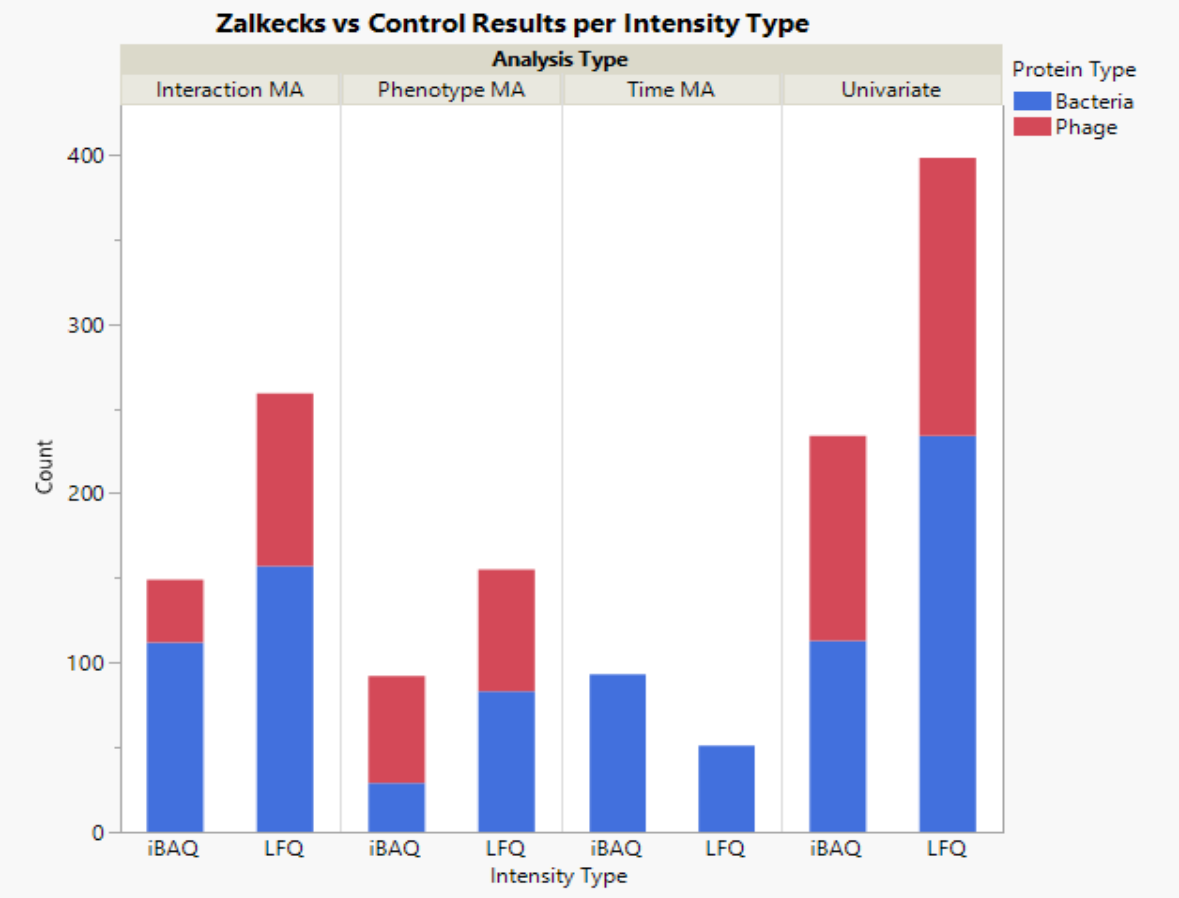


Figure 4-40. Comparison of LFQ vs iBAQ MS/MS data type for the Zalkecks vs the control experiments. Each section is a different analysis from MetaboAnalyst. Blue represents bacteria proteins and red represents phage proteins.

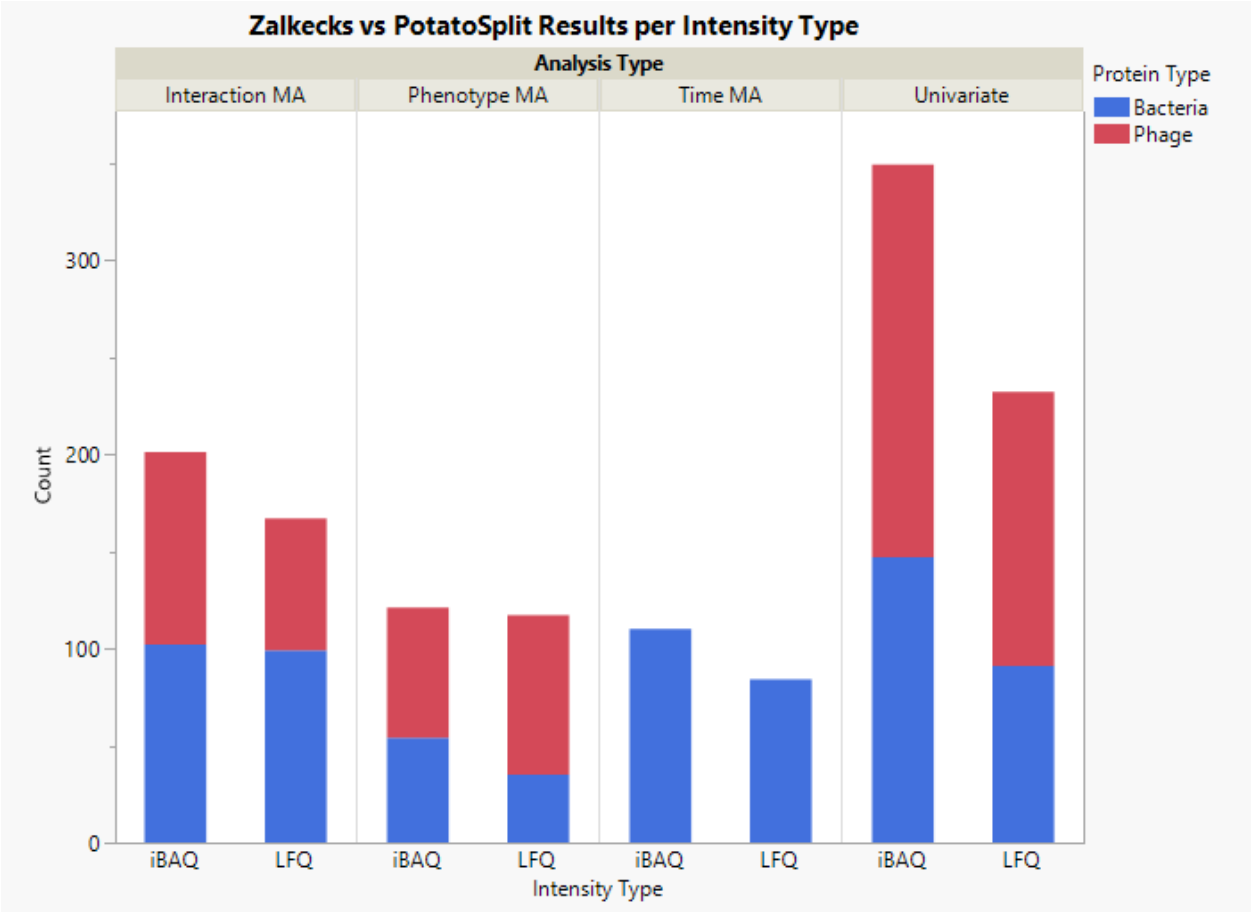


Figure 4-41. Comparison of LFQ vs iBAQ MS/MS data type for the Zalkecks vs PotatoSplit experiments. Each section is a different analysis from MetaboAnalyst. Blue represents bacteria proteins and red represents phage proteins.

5. DISCUSSION

5.1 M. smegmatis Growth Curves

The growth curve of Mycobacterium smegmatis was constructed by plotting the recorded OD600 values over time, as shown in Figure 4-1 and Figure 4-2. Figure 4-1 represents the P1FF samples plotted over 30 hours, which was used to determine the exponential phase of the bacteria. The exponential phase of the P1FF bacteria was determined to be 25 hours with an OD600 value of roughly 1.5, which will be used as the guide to make the P2FF solution. Figure 4-2 represents the P2FF samples plotted over 55 hours, which was used to determine the optimum time to inoculate the samples with a bacteriophage. At the mid-exponential phase, a time of 32 hours and an OD600 value of 1.3 was determined to be the best time to inoculate the samples with a phage. Inoculating during an exponential phase is necessary to promote a strong population of bacteria for the phages to interact with and allow adequate time to collect samples.

After determining inoculation timepoints, growth curves with each bacteriophage were constructed. Using two P2FF samples, one was infected with a phage and the other with phage buffer, and the OD600 values were recorded over time.

Figure 4-3 represents the PotatoSplit infected experiment which was inoculated at an average OD600 value of 2.00 after 31.5 hours of growth.

Figure 4-4 represents the Zalkecks infected experiment which was inoculated at an average OD600 value of 1.19 after 32.5 hours of growth. The phage-treated samples stay within a similar range of OD600 values as the control, which shows that the amount of phage infected in the medium was accurate. If the amount of phage was too high, it would’ve killed all of the host bacteria quickly or if too low there would not be sufficient interaction and therefore limited proteins and lipids produced. From these plots, it can be seen that the OD600 values for the bacteria begin to plateau at about 8-10 hours for each experiment. This shows that the phage population has grown and is beginning to kill bacteria at the same or similar rate to which bacteria are reproducing. This means that the phages are having a large impact on the bacteria population by hour 10, suggesting that this time point should contain more proteins and lipids than at earlier times. Another reason for the plateau in bacteria around hour 10 is that the bacteria may switch

energy utilization from reproduction to defense mechanisms to ward off phage infection. These defense mechanisms produce proteins and lipids which will potentially be collected during the experiment.

These experiments were repeated but with three P2FF samples, with an average OD600 value of 1.34: one infected with Zalkecks, one infected with PotatoSplit, and one infected with phage buffer as the control. The OD600 growth curves of this overall experiment is shown in Figure 4-5. When comparing the growth curves of Zalkecks and PotatoSplit, one can notice similar overall trends, but the OD600 values and patterns are slightly different. Mycobacteriophage Zalkecks had a lower OD600 value than mycobacteriophage PotatoSplit. This could potentially be due to the fact that bacteriophage PotatoSplit and Zalkecks are from different clusters and have different life cycles, therefore they interact with the host in different ways. PotatoSplit is a temperate phage, meaning that it can switch between the lytic and lysogenic pathways depending on the environment and stress factors, whereas Zalkecks is always in the lytic pathway. This may explain why it's OD600 values of PotatoSplit were not as low as that of Zalkecks, either it wasn't as effective at killing or interacting with the host. It was during this overall experiment that the protein and lipid samples were collected to perform protein and lipid extraction and be analyzed in Mass Spectrometry. Time points 0, 3, 7, and 10 hours were used to collect samples to display the entire phage-host interaction, ending at the bacterial growth plateau seen at hour 10.

5.2 Investigation of Lipids

Data retrieved from mass spectrometry was filtered and analyzed through multiple programs to determine the significance of lipids when comparing different testing groups. To begin with, the raw data was input into MS-DIAL to be filtered. Initially, results that were found in the blank solution are removed, as they are contaminants and not from the phage experiment samples. The MS2 spectra data was then directly compared against reference data to determine confidence in the match in three categories: confident, unsettled, and unknown. All unknown results were analyzed even further using MS-FINDER. This allowed for comparison to hits that were potentially similar, to determine if they matched or not based on the same confidence scale. After analyzing all the unknown lipid hits, those that are considered either confident or unsettled are extracted and made into a table. These filtered lipids and the associated intensities from mass spectrometry are statistically analyzed in MetaboAnalyst. The data is compared in the following

groups for lipid analysis: PotatoSplit compared to the control, Zalkecks compared to the control, and Zalkecks compared to PotatoSplit. These groups are compared to show how each phage deviates from a standard baseline but also to show the difference between each phage.

In MetaboAnalyst, a univariate analysis of a linear model with covariate adjustment, using a p-value cut-off of 0.05 to determine the significance of lipids in each comparison is used to make heatmaps. Heatmaps plot the distribution of lipids over time and allow for comparison of each sample at the same time point but also allow one to see how lipid concentration changes in the same sample of time. Figure 4-6 displays the heatmap that shows the distribution of all lipids found in both PotatoSplit infected samples and the control, over time. Figure 4-8 displays the heatmap that shows the distribution of all lipids found in both Zalkecks infected samples and the control, over time. Figure 4-10 displays the heatmap that shows the distribution of all lipids found in both Zalkecks infected samples and PotatoSplit infected samples, over time. The data that is used to make these plots are represented in Table 4-1, Table 4-3, and Table 4-5, respectively. Some of the main trends that can be seen in each of the heatmaps: lipids that increase overall timepoints of a sample, lipids that decrease overall time points of a sample, lipids that are mostly prevalent in bacteria-heavy timepoints of phage-treated samples (0 and 3 hours), lipids that are mostly prevalent in phage-heavy timepoints of phage-treated samples (7 and 10 hours), and lipids that are mostly found in either the control or a phage-treated sample but not both.

A multivariate analysis, called ANOVA Simultaneous Component Analysis (ASCA), is used to identify major patterns regarding two factors, phenotype, and time, as well as their interaction to understand when and why a lipid is significant. Using a cut-off value of 0.9 for leverage and 0.05 for p-value, the ASCA model was performed in MetaboAnalyst. This analysis is represented in Table 4-2 for PotatoSplit vs control, Table 4-4 for Zalkecks vs control, and Table 4-6 for Zalkecks vs PotatoSplit. In all three of the group comparisons, there were no lipids that were deemed significant due to changes in time. Each of these tables is represented by a Venn diagram, displayed in Figure 4-7, Figure 4-9, and Figure 4-11 respectively. The Venn diagrams show counts of lipids from each category, and if there is any overlap between categories in the ASCA model results. When comparing PotatoSplit and the control samples, there are no lipids significant due to both the phenotype and the interaction. When comparing Zalkecks and the control, there are two lipids significant due to both the phenotype and the interaction. When

comparing Zalkecks and PotatoSplit, there are four lipids significant due to both the phenotype and the interaction.

Other than trends, specific data can be extrapolated from this data. MS-DIAL groups lipids based on the ontology classification of the lipids but does not provide a definitive function of each lipid. These lipid groups are represented and discussed in the protein-lipid correlation section of the discussion below. The results from MS-DIAL were ran against several databases, including the RIKEN MetaDatabase, COCONUT, and LOTUS. When looking at lipids that matched to products in these databases, potential functions can be investigated as long as the model matches the scope of the research.

5.3 Investigation of Proteins

Data retrieved from mass spectrometry was filtered and analyzed through multiple programs to determine the significance of proteins when comparing different testing groups. To begin with, the raw data was processed through MaxQuant, a qualitative proteomics software, for peptide identification. The filtered protein data and associated intensities from mass spectrometry are statistically analyzed in MetaboAnalyst. The data is compared in the following groups for protein analysis: PotatoSplit compared to the control, Zalkecks compared to the control, and Zalkecks compared to PotatoSplit. These groups are compared to show how each phage deviates from a standard baseline but also to show the difference between each phage.

In MetaboAnalyst, a univariate analysis of a linear model with covariate adjustment, using a p-value cut-off of 0.05 to determine the significance of proteins in each comparison is used to make heatmaps. Heatmaps plot the distribution of proteins over time and allow for comparison of each sample at the same time point but also allow one to see how protein concentration changes in the same sample of time. It is important to note that protein names that end in ".1" are virus-related proteins and all other proteins are of bacterial descent. Some of the main trends that can be seen in each of the heatmaps: proteins that increase overall timepoints of a sample, proteins that decrease overall time points of a sample, proteins that are mostly prevalent in bacteria-heavy timepoints (0 and 3 hours) in a phage-treated sample, proteins that are mostly prevalent in phage-heavy timepoints (7 and 10 hours) in a phage-treated sample, or proteins that are only prevalent in either a phage-infected sample or the control sample. Figure 4-12 displays the heatmap that shows the distribution of all proteins found in both PotatoSplit infected samples and the control, over time.

Figure 4-10 displays the heatmap that shows the distribution of all proteins found in both Zalkecks infected samples and the control, over time. Figure 4-24 displays the heatmap that shows the distribution of all proteins found in both Zalkecks infected samples and PotatoSplit infected samples, over time. Virus-related proteins are mostly found in the phage-infected samples and are distinctively expressed in the later hours of the phage samples (7 and 10 hours). There are a few results for virus-proteins that appear in the control side of the plot, as seen in Figure 4-18. Since these results are not present in every triplicate of the sample timepoint and appear to be sparsely and randomly distributed, this phage contamination can most likely be associated with the normalization of the data. ta. When comparing Zalkecks to PotatoSplit, it can be seen in Figure 4-24, that there are two main sets of virus proteins: those mostly found in Zalkecks and those mostly found in PotatoSplit. Both sets of proteins are prevalent in the later time periods (hours 7 and 10) still. The data that is used to make these plots are represented in Appendix Table 1, Appendix Table 3, and Appendix Table 5, respectively.

A multivariate analysis, called ANOVA Simultaneous Component Analysis (ASCA), is used to identify major patterns in regard to two factors, phenotype, and time, as well as their interaction to understand when and why a protein is significant. Using a cut-off value of 0.9 for leverage and 0.05 for p-value, the ASCA model was performed in MetaboAnalyst. This analysis is represented in Appendix Table 2 for PotatoSplit vs control, Appendix Table 4 for Zalkecks vs control, and Appendix Table 6 for Zalkecks vs PotatoSplit. Each of these tables is represented by a Venn diagram, displayed in Figure 4-13, Figure 4-19, and Figure 4-25 respectively. The Venn diagrams show counts of lipids from each category, and if there is any overlap between categories in the ASCA model results. When comparing PotatoSplit and the control samples, there are 11 proteins significant due to both the phenotype and the interaction, while only one protein is significant due to both phenotype and time. There are no overlapping proteins between time and interaction, nor all three categories. When comparing Zalkecks and the control, there are 26 proteins significant due to both the phenotype and the interaction, while only one protein is shared between time and interaction. There are no overlapping proteins between time and phenotype, nor all three categories. When comparing Zalkecks and PotatoSplit, there are 43 proteins significant due to both the phenotype and the interaction. There are no overlapping proteins between time and interaction, time and phenotype, nor all three categories.

A fold change analysis was performed on the protein data through MetaboAnalyst, at each time point for each comparison. The data was then processed through DAVID to analyze the Gene Ontology of significantly related proteins in two groups: positive fold change and negative fold change. All results from DAVID that were significant, with a p-value of 0.05 or below, were made into bar graphs shown in Figure 4-14 through Figure 4-17 for PotatoSplit vs the control; Figure 4-20 through Figure 4-23 for Zalkecks vs the control; and lastly Figure 4-26 through Figure 4-29 for Zalkecks vs PotatoSplit samples. These figures show the significance of each protein as well as if it is up or downregulated by the phage. PotatoSplit vs the control samples tend to have many more up-regulated proteins than down-regulated proteins, which is different from each other comparison. For example, hour 10 had nine up-regulated proteins and only one downregulated protein for PotatoSplit vs the control but both Zalkecks vs the control and Zalkecks vs PotatoSplit had no up-regulated proteins in hour 10. Zalkecks vs the control samples had very little upregulation except during hour 7, although this time point still had a majority of down-regulated proteins. The same is true for Zalkecks vs PotatoSplit samples, except it was hour 3 that had the most upregulation. Further investigation of the functional terms that were deemed significant in the DAVID functional analysis is performed below.

Free metals within a media drastically increase the phage-dependent killing of bacteria. Metals that are known to trigger this response include magnesium, iron, and calcium (Ma et al., 2018). An upregulation of magnesium production or magnesium ion binding is found within PotatoSplit Hour 0 and PotatoSplit Hour 10 while up-regulation of iron production is found within Zalkecks Hour 7. A downregulation of magnesium binding is found within Zalkecks Hour 0, Zalkecks vs PotatoSplit Hour 0, and Zalkecks vs PotatoSplit Hour 3.

Phages can release proteins that inhibit bacterial ATPase which prevents the bacterial from being able to utilize energy to defend itself against phage infection (Chung et al., 2014). ATP binding proteins are upregulated and found within PotatoSplit Hour 0. ATP binding proteins are downregulated in PotatoSplit Hour 3, Zalkecks Hour 0, Zalkecks Hour 3, Zalkecks Hour 7, Zalkecks vs PotatoSplit Hour 0, and Zalkecks vs PotatoSplit Hour 3.

Phages can recognize molecules on the surface of bacterial cells to target them for infection. Bacteria with a cell wall deficiency have a higher chance of escaping phage infection because the phage can’t recognize them as well (Ongenae et al., 2021). Potentially, a bacteria would be able to decrease cell wall organization as a phage defense mechanism. Cell wall organization and related

proteins are upregulated in PotatoSplit Hour 0, PotatoSplit Hour 3, and PotatoSplit Hour 7, while they are downregulated in Zalkecks vs PotatoSplit Hour 0.

Often times bacteria have membrane transport systems that are used to intake nutrients from the environment for energy gradients within the cell (Jeckelmann & Erni, 2020). Membrane transport systems are necessary to provide all the nutrients for the cell to properly function. Membrane transport proteins are up regulated in PotatoSplit Hour 3.

As a crucial step in the replication process, a phage will recruit cellular ribosomes to translate viral mRNAs. This allows the phage to take control of cellular functions and signaling pathways to regulate their activity. This disables the bacterial defense mechanisms and ensures that the viral proteins will be produced in the infected cells (Walsh & Mohr, 2011). Ribosomal proteins and those related to them are upregulated in Zalkecks Hour 7, Zalkecks vs PotatoSplit Hour 0. While they are downregulated in PotatoSplit Hour 7.

RNA-binding proteins are able to bind to target RNA on another cell. They have impacts with regulation of gene expression and can form mRNA-protein complexes that have signaling capabilities (Oliveira et al., 2017). RNA-binding proteins can potentially impact cellular function, transport, and localization (Zhang & Wu, 2020). RNA-binding proteins are down-regulated in PotatoSplit Hour 7, Zalkecks vs PotatoSplit Hour 0. RNA-binding proteins are not up regulated in any of the three sample comparisons.

5.4 Correlation of Proteins and Lipids

The results produced from protein analysis and lipid analysis were used to determine the correlation between proteins and lipids within each sample. The data is compared in the following groups for protein analysis: PotatoSplit compared to the control and Zalkecks compared to the control. These groups are compared to show how each phage deviates from a standard baseline and to understand more about each virus. Using a combination of MetaboAnalyst to gather the data and JMP to manipulate and analyze the data, Figure 4-30 through Figure 4-38 are made and only represent the phage proteins in the samples. Figure 4-30, Figure 4-31, and Figure 4-32 represent the PotatoSplit vs the control samples at hours 3, 7, and 10, respectively. Figure 4-33, Figure 4-34, Figure 4-35, Figure 4-36, Figure 4-37, and Figure 4-38 represent the Zalkecks vs the control samples in time order. For Zalkecks samples, each timepoint has two images due to the number of results. Neither PotatoSplit nor Zalkecks samples have any results at Hour 0. Each

figure is a heatmap of a time point and sample comparison, that shows either a positive or negative correlation between a protein and lipid. Similar heatmaps showing every protein in the sample can be found in the appendix as Appendix Figure 1 through Appendix Figure 20 for PotatoSplit vs the control and Appendix Figure 21Appendix Figure 44 through Appendix Figure 44 for Zalkecks vs the control. It is important to note that this analysis is only aimed to discover a correlation between the proteins and lipids discovered in the samples, it does not lead to any consensus on causation between them.

Each correlation plot compares the lipids based on their ontology groups, identified by the MS-DIAL program. The lipid ontology groups found within PotatoSplit samples when compared to the control are Cer_NDS, Cer_NS, DG, Fatty acid, Fatty amide, Glycosides, Lipopeptides, MGDG, NAE, PE, PS, and Unknown. Whereas the lipid ontology groups found within Zalkecks samples when compared to the control are Cer_NDS, Cer_NS, Contamination, DG, EtherDG, Fatty acids, Glycosides, MGDG, NAE, PI, SM, ST, and Unknown. It is important to note that there is no contamination group in the PotatoSplit data, at any timepoints, when there is contamination within the Zalkecks data. Other notable differences are that Zalkecks samples have EtherDG, PI, SM, and ST ontology groups whereas PotatoSplit has Fatty amide, Lipopeptides, PE, and PS ontology groups. Having different lipid ontology groups in each phage could potentially relate to the different life cycles and mechanisms that each phage has.

Ceramides: Cer_NDS and Cer_NS

Ceramides are a lipid class, defined by sphingosine linked to a fatty acid. There are many different types of ceramides, two of which are Cer_NDS and Cer_NS. Cer_NDS is a ceramide class with non-hydroxy fatty acids and sphinganines. Cer_NS is a ceramide class with non-hydroxy fatty acids and 4-sphingenines (Masukawa et al., 2009). Ceramides are lipid messengers that have key roles in regulating membranes and can directly impact viral lifecycle and therefore the pathogenicity of a phage. Some of their roles include serving as receptors for viral entry, forming microdomains that cluster entry receptors, enabling receptors to acquire the optimum conformation, regulating cell surface expression, and forming viral replication sites (Eckmann & Becker, 2021). Sphingolipid metabolism is also controlled by ceramides. Sphingolipids can regulate viral uptake and intracellular trafficking. Not to mention that sphingolipids mediate the

release of new virion from cells that have been infected with a virus (Beckmann & Becker, 2021). As ceramides as an entire class are directly linked to many components of the viral life cycle, it can be attributed that both Cer_NDS and Cer_NS will have a large impact on the pathogenicity of each phage.

When analyzing the correlation plots, it can be seen that a majority of the results for PotatoSplit samples have Cer_NS and not Cer_NDS hits. Almost all of the Cer_NS results can be found during the Hour 7 period, and not during the other times. This timeline could potentially relate to the phage-host interaction. The ceramides play a role in membrane signaling, forming viral replication sites, and releasing virion from infected cells. At Hour 7, the phage has infected the bacteria and is increasing in reproduction. Potentially the ceramides are at their highest concentration here because they are being used for membrane signaling and to release more phages from the infected cells. For Zalkecks samples, the Cer_NS is also found in Hour 7 whereas Cer_NDS is not. Different from PotatoSplit samples, both ceramides are found during Hour 10. At Hour 10, the Cer_NDS are all negatively correlated whereas the Cer_NS are all positively correlated. This timeline also coincides with that of PotatoSplit: at later hours the phage is focusing on replicating and using ceramides for signaling and releasing more phages.

DG and EtherDG

A diglyceride (DG) is a glyceride with two fatty acid chains that are attached to a glycerol molecule covalently, through an ester linkage. Diglycerides are known to function as second messenger lipids. It has been shown through multiple studies that DG can provide a specific signal that is required for the rupturing of cells (Shahnazari et al., 2011). Another lipid ontology group found within the correlation data is EtherDG, which is short-hand for Ether-linked Diacylglycerols. EtherDG is a subset of DG, so the two lipid ontology groups should have very similar biological functions.

When analyzing the correlation plots, it can be seen that DG is mostly present throughout each phage infection but changes between regulation types depending on time. For the PotatoSplit samples, DG is partially present at Hour 3 and is mostly downregulated. By Hour 7, DG is correlated to almost all of the proteins and it is entirely upregulated but by Hour 10 there is no correlation of DG to any of the proteins. For Zalkecks samples, at Hour 3 a large portion of proteins

are correlated to DG, and they are all down regulated. By Hour 7 almost all proteins are correlated to the DG group with a majority being up regulated but some still being down regulated. By Hour 10 almost all proteins are correlated to DG and are down regulated or neutral (not up or down regulated).

The EtherDG lipid ontology group is only found within the Zalkecks correlation plots because there are no lipids relating to this group within the PotatoSplit data. When looking at the Zalkecks correlation plots, there are no proteins correlated to the EtherDG lipid ontology group at either Hour 3 or Hour 7. During Hour 10, however, almost all of the proteins are correlated to the EtherDG ontology group, and they are all downregulated.

PotatoSplit only has results for the DG ontology group. Within PotatoSplit data, the DG ontology group gets progressively more correlated over time, but after Hour 7 it completely drops off and has no correlated proteins at Hour 10. Within Zalkecks data, both DG and EtherDG follow the same trend by becoming increasingly correlated to proteins over time. Since diglycerides and ether-linked diglycerides are used during signaling and the rupturing of cells, it would align with phage infection that these lipids would be found increasingly in the later time periods.

Fatty Acid, Fatty Amide, and NAE

In bacteria, the primary role of fatty acids is to be a hydrophobic portion of the membrane and act as the building blocks of cell membranes (Cronan & Thomas, 2009). Fatty acids have antibacterial properties and are known to inhibit the growth of bacteria (Desbois & Smith, 2010). Many organisms take advantage of this function and use fatty acids to defend against parasitic bacteria. Fatty acids are antibacterial because they disrupt the electron transport chain and ultimately interfere with energy production. Fatty acids may also inhibit enzyme activity, decrease nutrient uptake, or cause direct lysis of bacterial cells (Desbois & Smith, 2010). Fatty amides, also called fatty acid amides, are similar to fatty acids but consist of aliphatic acids with varying amines connected by an amide linkage (*Fatty Acid Amides | Cyberlipid*, n.d.). Fatty acid amides have similar functions to fatty acids, including antibacterial properties. N-acylethanolamine (NAE) is a type of fatty acid amide, therefore will have similar if not the same biological functions as fatty acid amides.

Upon further investigation of the correlation plots, it can be seen that during Hour 3 of the PotatoSplit samples, there were some proteins that were correlated to the fatty acid lipid group, and they were all down regulated. During Hour 7 a majority of proteins were correlated to the fatty acid group, and they were all up regulated, whereas at Hour 10 there were no proteins that correlated to the fatty acid group. In the Zalkecks samples, roughly half of the proteins were correlated to the fatty acid group, and they were all down regulated. During Hour 7, however, there was no protein-lipid correlation present for fatty acids. Whereas at Hour 10, almost all proteins were correlated to the fatty acid group and were down regulated. It seems as though PotatoSplit and Zalkecks had opposite results when it comes to protein-fatty acid correlation, at Hours 7 and 10. The difference in fatty acid concentration and implementation at different time periods could be due to the phage life cycle, but further investigation would need to be completed to determine the correlation between fatty acid production and the phage life cycle.

Fatty amides are only prevalent in the PotatoSplit samples and not found within Zalkecks samples. At Hour 3 there is one protein correlated to fatty amides in PotatoSplit and it is downregulated. There is also one protein correlated to fatty amides at Hour 7, but it is up regulated. PotatoSplit has no proteins correlated to fatty amides during Hour 10.

Upon further investigation of the correlation plots, PotatoSplit has one protein that is positively correlated with the NAE lipid ontology group during Hour 3 of the experiment, meaning that the NAE group is up regulated at this time. During Hour 7, however, there are no proteins correlated to the NAE lipid ontology group. There are four proteins that are correlated to the NAE group at Hour 10 of the PotatoSplit data, and they are all up regulated. The Zalkecks data has a similar trend with more results. During Hour 3 of the Zalkecks data, nearly one-fourth of the proteins are correlated to the NAE lipid group and all of them are up regulated. Similar to PotatoSplit, there are no correlated proteins to this lipid group at Hour 7. Almost all of the proteins at Hour 10 are correlated to the NAE group, and again they are all up regulated.

Glycosides

Glycosides are known to have antimicrobial activities, that can be strengthened or weakened depending on the strain of bacteria in question (Shimamura, 2012). Several well-known antibiotics are derived from glycosides (*Glycoside | Biochemistry | Britannica*, n.d.). These main

functions of glycosides align with the principles of phage-host infection. The phages are wanting to infect and eventually kill the bacteria, and glycosides are used as antibiotics to kill bacteria.

Both PotatoSplit and Zalkecks samples have the glycosides lipid ontology group present in the correlation charts, but only PotatoSplit has any proteins that are correlated to that group. PotatoSplit has two proteins that are correlated to glycosides, and they are both up regulated. Considering that glycosides can act as antibiotics and kill bacteria, it makes sense that they are prevalent in PotatoSplit over Zalkecks. Since Zalkecks is a lytic phage, it is in its life cycle to rupture the bacterial cells. PotatoSplit, on the other hand, is a lysogenic phage that only a small portion of the time will rupture the bacterial cells. Potentially PotatoSplit would need the aid of a glycoside to kill the cells. This would also align with the glycosides being present earlier on in the relationship, as the phage is overcoming bacterial mechanisms to block phage infection and may need more help rupturing cells at that point in time.

Lipopeptides

Lipopeptides are mostly found in bacteria and are made of a fatty acid linked to a peptide chain (Vecino et al., 2021). Lipopeptides often act as surfactants because they interact with the surface of cells. They are also known to have biological functions against pathogenic microorganisms and perform as an anti-bacterial agent (Vecino et al., 2021). Due to these functions, lipopeptides are often investigated for use in the pharmaceutical industry and align well with the ideology of the phage-host interaction.

The lipopeptide lipid ontology group is only found in the PotatoSplit correlation data. Not only are there no protein-lipid correlations within Zalkecks, but lipopeptides are not found in the samples. There are no correlations found in Hour 3 of the PotatoSplit data. A large portion of the proteins at Hour 7 are correlated with lipopeptides and all are up regulated. During Hour 10 of PotatoSplit samples, there is one protein correlated to the lipopeptide group and it is up regulated as well. Considering that lipopeptides can act as anti-bacterial agents, it makes sense that the phage would utilize and produce them, especially in the lysogenic phage PotatoSplit. Being produced mostly at Hour 7 shows that the phage has started infection and is working to kill the bacteria, which is why the lipopeptides have increased correlation at that point.

MGDG

MGDG is the shorthand notation for monogalactosyldiacylglycerol, which is the most abundant membrane lipid in the biosphere. MGDG has a similar structure to most membrane lipids with a polar head group with a hydrophobic tail (Kobayashi et al., 2007). Since MGDG is so abundant, it is found in most membranes including bacterial cell membranes. A loss of MGDG can have a significant impact on other lipids, especially those in the membrane (Masuda et al., 2011). Membrane lipids often play a crucial role in signaling and informing the cell about external environmental factors (Barák & Muchová, 2013). Potentially a bacterial cell could use a membrane lipid, such as MGDG, to sense a phage in the environment and signal to the cell to start using energy for phage defense mechanisms to prevent infection.

PotatoSplit has one protein that is correlated with the MGDG lipid ontology group during Hour 3, and it is down regulated. Neither Hour 7 nor Hour 10 have any correlated proteins to the MGDG group for PotatoSplit samples. Zalkecks, on the other hand, had no correlation at Hour 3 but did have correlation at Hours 7 and 10. At Hour 7, there were two proteins correlated to the MGDG lipid ontology group and they were both up regulated. Nearly all of the proteins were correlated to MGDG during Hour 10 of the Zalkecks samples, and they were all up regulated as well. Since MGDG is a membrane lipid and works as a signaling molecule for the cell, the difference in the timeline of MGDG correlation could be due to the different lifecycles of PotatoSplit and Zalkecks.

Plasma Membrane Phospholipids: PE, PS, and PI

As per the name, plasma membrane phospholipids play a large role in the membrane. They are composed of a polar head group attached to two nonpolar hydrophobic fatty acid tails. These phospholipids directly impact the permeability barrier by controlling the movement of molecules and ions into or out of the cells (Lin & Weibel, 2016). Phospholipids also regulate the spatial and temporal position of membrane proteins, which again directly impact many cellular functions (Lin & Weibel, 2016). Changes in the number of phospholipids produced by a bacteria can alter cellular envelope formation, bacterial fitness, and the ability of the bacteria to adapt to environmental stressors (Rowlett et al., 2017). Since phospholipids are so prevalent in the membrane, they have a large impact on cellular signaling and metabolic pathways as well.

The PE lipid ontology group is only prevalent in the PotatoSplit data. There are no results related to the PE group within the Zalkecks data, so this lipid ontology group doesn't appear in the Zalkecks correlation plots. When looking at the correlation plots for PotatoSplit at Hour 3, there is one protein correlated to the PE lipid ontology group and it is up regulated. At hour 7 there are no proteins correlated to the PE lipid ontology group within the PotatoSplit data. About one-fourth of the proteins become positively correlated to the PE group during Hour 10 which means that they are all up regulated.

Similar to the PE ontology group, the PS lipid ontology group is only prevalent in the PotatoSplit data. The PS group does not appear in the Zalkecks correlation plots because there are no lipid results related to this group within the Zalkecks data. When considering the PotatoSplit data, there is one protein correlated to the PS ontology group at Hour 3 and it is up regulated. During Hour 7, there are two proteins that are both up regulating and correlated to the PS ontology group. However, at Hour 10 there are no proteins correlated to the PS lipid ontology group.

The PI lipid ontology group is not found within the PotatoSplit correlation plots because there are no lipids relating to this group within the PotatoSplit data. When looking at the Zalkecks correlation plots, about one-third of the proteins are correlated to the PI lipid ontology group and they are all down regulated. Almost all proteins are correlated to the PI group at Hour 7 and they are all up regulated at this point. During Hour 10 of the Zalkecks data, there are no proteins correlated to the PI lipid group.

Out of the three plasma membrane phospholipids found in the correlation data, only the PE and PS groups are found within PotatoSplit and only PI is found within Zalkecks. Based on the main functions of phospholipids, it aligns with the phage-host interaction. Potentially the bacteria are producing more membrane phospholipids to combat the stressors in the environment or to use different metabolic pathways. An increase in phospholipids could open a new metabolic pathway that redirects energy toward a phage defense mechanism, rather than unnecessary mechanisms.

SM

Sphingomyelin (SM) is a phospholipid that is derived from sphingosine. A sphingosine can also be considered a ceramide, which was covered in the discussion above. Sphingolipids can regulate viral uptake and intracellular trafficking. Not to mention that sphingolipids mediate the

release of new virion from cells that have been infected with a virus (Beckmann & Becker, 2021). Sphingomyelins are common in the cell membrane and are known to have a function in signal transduction and signaling pathways (Merill & Sweeley, 1996).

The SM lipid ontology group is only found within the Zalkecks correlation plots because there are no lipids relating to this group within the PotatoSplit data. When looking at the Zalkecks correlation plots, there are no proteins correlated to the SM lipid ontology group during Hour 3 and Hour 7 of the correlation plots. During Hour 10 of the Zalkecks data, nearly half of the proteins are correlated to the SM ontology group, and they are all up regulated. Since the correlation of proteins to the SM lipid ontology group increase over time, the function of releasing new virion cells from those infected aligns well with the phage infection timeline. As time continues, they will be increasing the infection rate and more cells will be rupturing to release new virion cells which means that more sphingomyelin may be needed.

ST

Sterols (ST) are essential when it comes to cell structure and function. They are found in all membranes and interact with phospholipids to regulate membrane permeability and function (Yu et al., 2021). Sterols are also a precursor to many hormones that directly impact many biological processes and signaling pathways. of the time, bacteria only produce sterols when under extreme conditions (Dufourc, 2008). Phages are known to be responsive to changes in intracellular sterol levels (Yu et al., 2004).

The ST lipid ontology group is only found within the Zalkecks correlation plots because there are no lipids relating to this group within the PotatoSplit data. When looking at the Zalkecks correlation plots, nearly half of the proteins at Hour 3 are correlated to the ST lipid ontology group, and they are all down regulated. Conversely, at Hour 7 nearly half of the proteins are correlated to the ST ontology group, but they are all up regulated at this point. At Hour 10 of the Zalkecks data, almost all of the proteins are correlated to the ST lipid ontology group and are down regulated. Knowing that sterols are mostly produced under extreme conditions and that phages can be responsive to sterols, it coincides with the phage infection timeline. With increasing environmental stressors due to phage infection, the bacteria would potentially be producing more sterols over time.

Contamination

The contamination group of lipids was only represented in the Zalkecks compared to the control samples. There are only hits for contamination during the Hour 10 time period, found in Figure 4-37 and Figure 4-38. There could be many potential sources for contamination, but when looking at Figure 4-8, the heatmap of all lipid results in Zalkecks compared to the control samples, phthalic anhydride is a result. Upon further investigation, phthalic anhydride has a wide range of commercial applications including its use in the plastics industry. Phthalic anhydride was not found as a result in either of the other datasets, which aligns with there only being contamination in the Zalkecks data. The contamination ontology group found in the Zalkecks correlation data can be tentatively attributed to plastic contamination during sample prep, as shown by the phthalic anhydride hits in the lipid results. Plastic contamination could potentially have been caused by changing brands or types of micropipette tips, or other plastic material used.

5.5 Investigation of Similarities Between Phages Zalkecks and PotatoSplit

Based on the heatmaps represented in Figure 4-6, Figure 4-8, and Figure 4-10 there are little to no similarities of significant lipids found in both bacteriophages Zalkecks and PotatoSplit. Further investigation was performed to determine if there are any lipids found within both Zalkecks and PotatoSplit that may not be considered significant and therefore do not appear in the heatmaps or other analyses throughout this research project. There are a total of 263 lipids found in both Zalkecks and PotatoSplit and can be found in the appendix as Appendix Table 7. Out of these lipids, the most commonly shared lipid ontology group is the Unknown group. Other than that, they both share lipids found in the Cer_NS, DG, EtherMGDG, Fatty Amide, Glycosides, NAE, PC, PE, PE-Cer, Spirostans, and ST lipid ontology groups. For the most part, there are very few results in each ontology group aside from DG, PE, and Unknown groups.

A diglyceride (DG) is a glyceride with two fatty acid chains that are attached to a glycerol molecule covalently, through an ester linkage. Diglycerides are known to function as second messenger lipids and can provide a specific signal that is required for the rupturing of cells (Shahnazari et al., 2011). Since diglycerides are used during signaling and the rupturing of cells, they would align with both the lytic and lysogenic lifecycles and therefore coincides with being found in both phage samples.

Plasma membrane phospholipids (PE) play a large role in the membrane and are composed of a polar head group attached to two nonpolar hydrophobic fatty acid tails. These phospholipids directly impact the permeability barrier by controlling the movement of molecules and ions into or out of the cells (Lin & Weibel, 2016). Phospholipids also regulate the spatial and temporal position of membrane proteins, which again directly impact many cellular functions (Lin & Weibel, 2016). Changes in the number of phospholipids produced by a bacteria can alter cellular envelope formation, bacterial fitness, and the ability of the bacteria to adapt to environmental stressors (Rowlett et al., 2017). Considering that plasma membranes have such an impact on cellular functions and membrane permeability, it aligns with being found in both Zalkecks and PotatoSplit samples.

Based on the heatmaps represented Figure 4-12, Figure 4-18, and Figure 4-24 there are little to no similarities of significant proteins found in both bacteriophages Zalkecks and PotatoSplit, especially phage related proteins. Further investigation was performed to determine if there are any proteins found within both Zalkecks and PotatoSplit that may not be considered significant and therefore do not appear in the heatmaps or other analyses throughout this research project. There are a total of 78 proteins are shared between bacteriophages Zalkecks and PotatoSplit and can be found in the appendix as Appendix Table 8. Looking back at the heatmaps of the protein results, there are bacterial proteins shared between the two phages but no phage proteins. Out of all the proteins, both significant and insignificant, there were no phage-related proteins shared between the two phages. Most of the 78 total proteins discovered can be seen in the heatmaps of significant proteins although there are a few new and insignificant ones.

When looking at each timepoint of the protein-lipid correlation plots, more direct comparisons can be drawn between phages Zalkecks and PotatoSplit. At Hour 3 both phages have the same groups correlated except Zalkecks has the addition of NAE, PI, and ST groups. At Hour 7 each phage shares the same groups except PotatoSplit has Fatty Acids and lipopeptides groups present and Zalkecks has MGDG, and PI groups present. Lastly, at Hour 10 each phage shares the same groups except PotatoSplit has the addition of the PE ontology group and Zalkecks has Cer_NDS/Cer_NS, contamination, DG, EtherDG, FA, MGDG, SM, and ST lipid ontology groups.

Referencing the ontology group functions described above, the following inferences relating to phage life cycle can be drawn. The NAE group is related to fatty acids which are known to inhibit enzyme activity, decrease nutrient uptake, or cause direct lysis of bacterial cells. The

direct lysis of bacteria coincides with being present in Zalkecks, a lytic phage, but not PotatoSplit, a lysogenic phage during Hour 3.

Lipopeptides are only found within PotatoSplit samples and are only significantly correlated to proteins at Hour 7. Lipopeptides act as surfactants by interacting with the surface of bacterial cells and they also have antibacterial properties. A lysogenic phage, such as PotatoSplit, would potentially utilize lipopeptides and their antibacterial properties to aid in killing of bacteria which aligns with lipopeptides being significant at later times in the phage-infection process. At Hour 7 Zalkecks has proteins correlated to the PI lipid ontology group and PotatoSplit does not. The PI group can alter cellular envelope formation, bacterial fitness, and the ability of the bacteria to adapt to environmental stressors. It is more likely that a lytic phage would have plasma membrane lipids rather than a lysogenic phage because this lipid group decreases bacterial fitness and in turn makes it easier for a phage to overtake the bacteria.

During Hour 10 the Zalkecks samples had drastically more lipid ontology groups correlated to proteins than the PotatoSplit samples. Lytic phages are generally more aggressive and always aim to kill the bacteria whereas lysogenic phages typically live within bacteria and replicate with it before killing it. The drastic increase of lipid ontology groups within Zalkecks aligns with the lytic lifecycle. Zalkecks produces more lipids because it is continually lysing bacteria and needs as many antibacterial lipids as it can produce. PotatoSplit on the other hand has infected bacteria by Hour 7, and since it lives within the bacteria a large portion of the time, so it doesn't need as many antibacterial agents to lyse to bacterial cells.

5.6 Investigation of MS Intensity Data Type

For the entirety of this research project LFQ intensity data produced from mass spectrometry was used during analysis. An investigation into the use of iBAQ intensity data was performed to determine if the mass spectrometry data type would alter the results produced during analysis. The data is compared in the following groups for protein analysis: PotatoSplit compared to the control, Zalkecks compared to the control, and Zalkecks compared to PotatoSplit. Using a combination of MetaboAnalyst to gather the data and JMP to manipulate and analyze the data, Figure 4-39, Figure 4-40, and were produced to demonstrate discrepancies between the two types of data. Figure 4-39 represents the PotatoSplit data, Figure 4-40 represents the Zalkecks data, and Figure 4-41 represents the Zalkecks compared to PotatoSplit data. The two analyses performed on

both sets of data are the univariate and multivariate analysis through MetaboAnalyst. The multivariate analysis is broken down into proteins that are impacted by time, phenotype, or the interaction between time and phenotype. The three components of the multivariate analysis and the univariate analysis represent the four sections in each of the intensity data type plots.

For the PotatoSplit compared to the control samples, the following trends can be extrapolated from the plots described above. When comparing iBAQ and LFQ data, there are a similar number of overall results but iBAQ produced more phage proteins during the interaction portion of the multivariate analysis. During the phenotype analysis, iBAQ produced many more results overall but there was a similar number of phage proteins. The time analysis didn't produce any phage proteins for either set of data but iBAQ had more bacterial proteins. Lastly, during the univariate analysis, both types of data produced a similar number of phage proteins but iBAQ had drastically more protein results overall.

For the Zalkecks compared to the control samples, the following trends can be understood from the MS data intensity plots. When looking at the interaction portion of the multivariate analysis the LFQ data produced nearly triple the number of phage proteins and had more bacterial proteins than the iBAQ data. After the phenotype portion of the multivariate analysis, the LFQ data produced more bacterial proteins although both data types have a similar number of phage proteins. The time analysis didn't produce any phage proteins for either set of data but iBAQ had more bacterial proteins. The LFQ data produced drastically more bacterial proteins and slightly more phage proteins than the iBAQ data during the univariate analysis.

When comparing the MS data intensity types for the Zalkecks compared to PotatoSplit data, the following trends can be understood. The iBAQ data produced more phage proteins but each data type had a similar number of bacterial proteins for the interaction portion of the multivariate analysis. LFQ and iBAQ data have similar overall numbers of results but the LFQ data produced more phage proteins and the iBAQ data produced more bacterial proteins during the phenotype multivariate analysis. Again, the time multivariate analysis did not produce any phage proteins but the iBAQ data produced more bacterial proteins. Lastly, the iBAQ data produced some more phage proteins but drastically more bacterial proteins than the LFQ data during the univariate analysis.

As described above, there are many discrepancies between the number of results produced, both with bacterial and phage proteins, when using the different mass spectrometry data types.

There is not a general trend between these two data types. Occasionally LFQ will have more phage protein results and occasionally iBAQ will have more phage protein results. This investigation shows that there are differences between the two data types but does not draw conclusions on which would produce more well-rounded or more accurate results during the research project. Further investigation of the different data types would be necessary to draw any conclusions about which data type is better for this type of research project.

6. CONCLUSION

This research project provided insight into the interaction between bacteriophages and their hosts. The use of mass spectrometry and multiple bioinformatic tools identified and analyzed proteins and lipids produced during the phage-host interaction. Due to this being a discovery process, untargeted proteomics and lipidomics were the best method to identify any and all proteins and lipids in the samples. Correlation between each protein and each lipid was calculated to understand more about how the phage uses each and if they interact with each other, which can aid in the understanding of how phages use proteins and lipids during the infection process. Fold change analysis can be performed to understand if each protein being produced is significant and if it is up or down regulated by the phage. Also, a case study to learn more about the results produced from different mass spectrometry data types and conducted during this research project. As proteins and lipids are analyzed, the potential functions of each phage can be investigated, and a deeper understanding of the infection and pathogenicity of each phage can be determined.

During this study, the exponential phase of the P1FF bacteria was determined to be at 25 hours with an OD600 value of roughly 1.5, which was used to make the P2FF samples. The P2FF bacteria was determined to be at the mid-exponential phase after 32 hours with an OD600 value of roughly 1.3, and this time will be used as the optimum time to inoculate the samples with a bacteriophage. Multiple types of analysis were performed to get a list of all proteins and all lipids found during the untargeted proteomics and lipidomics studies were made. The analysis types for both protein and lipids include a univariate analysis and a multivariate analysis that compared time, phenotype, and their interaction.

Both a univariate and a multivariate method were used to identify proteins and lipids and determine their significance. Proteins and lipids were determined in three main sample comparisons: PotatoSplit vs the control, Zalkecks vs the control, and Zalkecks vs PotatoSplit. Proteins were further analyzed through a fold change analysis to determine if a protein is significant and if it is either up or down regulated by a phage, at each time point. Lipids were categorized into lipid classes or ontology groups and through the correlation research, showed how these groups change over time. Some specific lipids were investigated and mostly functions related to the cell membrane and to energy utilization were found.

The correlation between each protein and each lipid was determined and made into heatmaps. This study only shows the correlation of each and does not provide any insight into the causation between them. Each of the lipid ontology groups used during the correlational analysis was investigated to determine the biological function and convey understanding as to why the phage or bacteria may have produced those lipids. Comparison of the correlation between proteins and lipid ontology groups found in each phage were compared to understand how their function relates to the lifecycle of each phage.

There were little to no significant lipids shared between both phage Zalkecks and PotatoSplit, so further investigation into all lipid results was performed to determine if there is any overlap between the two phages. There are a total of 263 lipids found in both Zalkecks and PotatoSplit although most of these are not considered significant and do not appear in the heatmaps or other analyses. Out of the lipids in common, most of them fall into the DG, PE, and unknown lipid ontology groups. Since diglycerides (DG) and plasma membrane phospholipids (PE) are used during signaling and the rupturing of cells as well as bacterial fitness and the ability of a bacterium to adapt to environmental stressors, it would align with both the lytic and lysogenic lifecycles and therefore coincides with being found in both phage samples. The same investigation was performed for all of the protein results. Out of all the proteins, both significant and insignificant, there were no phage-related proteins shared between both Zalkecks and PotatoSplit. Most of the 78 total proteins discovered can be seen in the heatmaps of significant proteins although there are a few new and insignificant ones.

A case study was performed that investigated the difference between different mass spectrometry protein intensity types. Overall, it was determined that the number of results, for both bacteria related and phage related proteins, was different depending on if the LFQ data or the iBAQ data was used. This case study did not determine the optimal data type to use for further phage proteomic and lipidomic research but rather showed that there are differing results based on the mass spectrometry data type chosen.

Investigation and understanding of the phage-host interaction is necessary to determine applications, know any potential risks, and to ensure safety in their application. Before phage therapy because widespread, more knowledge about phages needs to be discovered and studied. This research has provided insight into the phage-host interaction of both Zalkecks and PotatoSplit, and the same needs to be done for any phage that is being considered for medicinal applications.

Now that proteins and lipids have been discovered for these two phages, specific biological functions of each should be determined to confidently make claims about the overall potential application of each phage. Furthermore, the protein correlation methods should be implemented on the Zalkecks vs PotatoSplit samples to understand more about how each of these phages compares to the other and provide insight into the impact of the phage life cycle on the correlation of proteins and lipids. Since a different number of results were determined during the MS/MS intensity data type investigation, this case study should be continued to understand the full impact the intensity type has on the results. Potentially, this research project could be repeated using iBAQ data instead of LFQ data, and different results may be produced which could provide a different perspective on each of the phages investigated.

APPENDIX

PotatoSplit Compared Against the Control Data

Appendix Table 1. PotatoSplit MS2 acquired protein results from a MetaboAnalyst linear model with covariate adjustments, when compared against the control.

	POTATOSPLIT PROTEINS	LOG(FC)	P.VALUE
COMPARISON TO CONTROL	AZS12919.1	1.466	0.000194
	AZS12912.1	1.4463	0.000236
	AZS12921.1	1.3321	0.000708
	AZS12901.1	1.3202	0.00079
	A0R3Q2	-1.3168	0.000815
	AZS12950.1	1.2821	0.001116
	AZS12954.1	1.2744	0.001195
	A0QV98	-1.248	0.001509
	AZS12909.1	1.2093	0.002109
	A0QPE5	1.1945	0.00239
	AZS12905.1	1.1911	0.002461
	A0QX00	1.1837	0.002619
	ILVC	1.1207	0.004384
	AZS12907.1	1.1104	0.00476
	AZS12964.1	1.1094	0.004797
	AZS12906.1	1.1091	0.004807
	AZS12961.1	1.1044	0.004989
	A0QTU1	-1.1007	0.005136
	AZS12923.1	1.0981	0.005244
	AZS12928.1	1.097	0.005288
	AZS12958.1	1.0969	0.005291
	AZS12924.1	1.0966	0.005304
	AZS12920.1	1.0879	0.005681
	AZS12953.1	1.0875	0.005698
	A0QT13	-1.0694	0.006553
	AZS12955.1	1.0694	0.006555
	A0QWZ9	1.0653	0.006765
	AZS12922.1	1.0389	0.008263
	A0QWL3	-1.0019	0.010862
	A0QQQ4	-0.9986	0.011125
	A0R6L5	-0.99769	0.011198
	AZS12944.1	0.99722	0.011237
	AZS12948.1	0.99615	0.011324

Appendix Table 1. Continued.

COMPARISON TO CONTROL	AZS12910.1	0.99167	0.011697
	AHPC	0.98909	0.011917
	A0R5R1	-0.98712	0.012087
	A0QYD4	0.98367	0.012391
	A0R450	-0.98252	0.012494
	AZS12940.1	0.97789	0.012914
	A0R130	-0.97534	0.013152
	GLMU	0.966	0.014054
	A0QW02	0.96384	0.01427
	A0R059	0.95933	0.014731
	A0R0Z1	0.95892	0.014773
	AZS12927.1	0.95421	0.01527
	A0QUX6	0.94432	0.01636
	A0QWI4	-0.93911	0.016961
	A0QS33	-0.93546	0.017395
	ARGD	0.93147	0.017879
	A0QQD0	-0.92367	0.01886
	A0QWZ1	-0.92158	0.019131
	A0R560	-0.91328	0.02024
	A0QPZ5	-0.91053	0.020619
	A0QQ23	-0.90662	0.02117
	RL19	0.90641	0.021201
	A0QR48	-0.90067	0.022032
	PBP1A	0.89702	0.022576
	PUP	-0.89505	0.022875
	DCDB	-0.88648	0.024212
	A0R1A7	0.8759	0.025958
	AZS12973.1	0.87206	0.026618
	A0QU53	0.87007	0.026965
	A0R189	-0.86884	0.027183
	AZS12918.1	0.86865	0.027216
	COAD	-0.86813	0.027308
	A0QP47	0.86435	0.027986
	DAPE	-0.86412	0.028028
	AZS12935.1	0.86364	0.028116
	A0QUH2	-0.85007	0.030681
	A0QT33	-0.84723	0.031244
	A0R2B7	0.84671	0.031348
	AZS12917.1	0.84665	0.03136
	DDL	0.84571	0.031548
	A0QNJ8	-0.84135	0.032436
	A0QX91	0.84092	0.032524

Appendix Table 1. Continued.

COMPARISON TO CONTROL	A0QQX6	0.83117	0.034589
	A0QZ56	-0.828	0.035285
	CSPA	0.82793	0.035301
	A0QPN0	0.82685	0.03554
	A0R2S4	0.82587	0.03576
	A0QW06	-0.82445	0.036079
	A0QWN0	-0.81945	0.037221
	A0R2Y0	-0.81847	0.037449
	A0R617	0.81721	0.037743
	PANC	0.8161	0.038003
	A0QZ55	-0.81409	0.03848
	RPOB	0.81308	0.038722
	SYDND	0.81035	0.03938
	A0R214	-0.8074	0.040102
	A0QUA2	-0.80662	0.040296
	A0QPY2	0.80471	0.04077
	A0QW04	-0.80167	0.041537
	A0QX77	-0.79819	0.04243
	A0R5I4	0.79658	0.042848
	A0QVD5	0.79239	0.043954
	A0R3S0	-0.79208	0.044035
	ATPF	0.79169	0.044139
	A0QZ14	-0.78872	0.044942
	A0QY10	-0.78855	0.044987
	A0R051	0.78763	0.045237
	A0QSV0	0.78634	0.045591
	RRAAH	-0.78473	0.046036
	A0QSW8	0.78422	0.046177
	A0R3Z5	-0.78153	0.04693
	GARA	-0.78037	0.047259
	A0QYQ9	0.77833	0.04784
	A0QVP0	0.77654	0.048353

Appendix Table 2. PotatoSplit ANOVA Simultaneous Component Analysis (ASCA) significant lipid results when compared against the control to model phenotype and time effects and their interaction.

	PROTEIN ID	LEVERAGE	SPE	PROTEIN ID	LEVERAGE	SPE
PHENOTYPE	AZS12921.1	0.004406	0	A0QWL3	0.002492	0
	AZS12950.1	0.004082	0	AZS12944.1	0.002469	3.70E-32
	AZS12954.1	0.004033	0	AZS12948.1	0.002464	1.85E-31
	AZS12909.1	0.003631	0	AZS12910.1	0.002442	3.70E-32
	A0QPE5	0.003543	0	AZS12940.1	0.002374	1.85E-31
	AZS12905.1	0.003523	0	A0R130	0.002362	1.85E-31
	A0QX00	0.003479	1.48E-31	A0R059	0.002285	7.40E-32
	ILVC	0.003118	0	AZS12927.1	0.002261	7.40E-32
	AZS12961.1	0.003029	0	A0QWI4	0.00219	7.40E-32
	A0QTU1	0.003008	1.48E-31	A0QS33	0.002173	7.40E-32
	AZS12953.1	0.002936	1.48E-31	A0QQD0	0.002118	1.85E-31
	A0QWL3	0.002492	0	A0R560	0.002071	7.40E-32
	AZS12944.1	0.002469	3.70E-32	A0QPZ5	0.002059	7.40E-32
	AZS12948.1	0.002464	1.85E-31	A0QQ23	0.002041	7.40E-32
	AZS12910.1	0.002442	3.70E-32	RL19	0.00204	1.85E-31
	AZS12940.1	0.002374	1.85E-31	PUP	0.001989	7.40E-32
	A0R130	0.002362	1.85E-31	AZS12973.1	0.001888	1.85E-31
	A0R059	0.002285	7.40E-32	A0QU53_2	0.00188	1.85E-31
	AZS12927.1	0.002261	7.40E-32	AZS12918.1	0.001874	7.40E-32
	A0QWI4	0.00219	7.40E-32	COAD	0.001871	3.70E-32
	A0QS33	0.002173	7.40E-32	A0QP47	0.001855	1.85E-31
	A0QQD0	0.002118	1.85E-31	AZS12935.1	0.001852	1.85E-31
	AZS12921.1	0.004406	0	A0QUH2	0.001794	1.85E-31
	AZS12950.1	0.004082	0	A0QT33	0.001782	0
	AZS12954.1	0.004033	0	A0R2B7	0.00178	1.85E-31
	AZS12909.1	0.003631	0	AZS12917.1	0.00178	0
	A0QPE5	0.003543	0	CSPA	0.001702	1.85E-31
	AZS12905.1	0.003523	0	A0QW06	0.001688	0
	A0QX00	0.003479	1.48E-31	A0QWN0	0.001667	7.40E-32
	ILVC	0.003118	0	A0R617	0.001658	7.40E-32
	AZS12961.1	0.003029	0	PANC	0.001654	7.40E-32
	A0QTU1	0.003008	1.48E-31	A0QZ55	0.001646	7.40E-32
	AZS12953.1	0.002936	1.48E-31	RPOB	0.001641	1.85E-31
	A0R214	0.001619	7.40E-32	A0QR04	0.001419	1.48E-31
	A0QW04	0.001596	0	A0QXI7	0.001417	7.40E-32
	A0R5I4	0.001576	3.70E-32	A0QX01	0.001416	3.70E-32
	A0QVD5	0.001559	1.85E-31	A0R003	0.001415	1.85E-31
	A0R3S0	0.001558	7.40E-32	A0QU45	0.001398	3.70E-32

Appendix Table 2. Continued.

PHENOTYPE	ATPF	0.001556	1.85E-31	A0QVM9	0.001395	0
	A0QY10	0.001544	1.48E-31	A0QRF5	0.001388	7.40E-32
	A0R051	0.00154	0	DOP	0.001379	1.85E-31
	A0QSV0	0.001535	7.40E-32	A0QVT1	0.00137	1.85E-31
	RRAAH	0.001529	7.40E-32	A0QZY1	0.001364	1.85E-31
	A0QSW8	0.001527	7.40E-32	A0R0I7	0.001356	3.70E-32
	A0R3Z5	0.001517	1.48E-31	NADD	0.001355	3.70E-32
	A0QYQ9	0.001504	7.40E-32	AZS12949.1	0.001354	3.70E-32
	A0QVP0	0.001497	7.40E-32	A0QQW8	0.001351	3.70E-32
	A0R0Z3	0.00147	1.85E-31	A0R594	0.001347	3.70E-32
	A0QV88	0.001449	1.85E-31	A0R2L3	0.001346	1.85E-31
	A0QWY8	0.001448	7.40E-32	A0QX64	0.001345	7.40E-32
	A0QRR4	0.00144	1.85E-31	A0R231	0.001343	7.40E-32
	A0QWD2	0.001438	7.40E-32	A0R0W4	0.001341	3.70E-32
	A0R212	0.00142	7.40E-32	Q3L887	0.001336	3.70E-32
TIME	A0R5P0	0.002473	3.0094	A0QQG1	0.001691	3.5342
	A0R3H8	0.00241	0.76452	A0QPD6	0.001688	0.60843
	A0QPW0	0.002387	0.21622	A0QX90	0.001687	2.7213
	DPPRS	0.002367	1.7713	GREA	0.001681	0.40899
	A0QSA9	0.002257	1.8091	A0R3T9	0.001659	0.13768
	CON__P17690	0.002182	0.30519	A0R0M4	0.001622	1.636
	A0R6P9	0.002177	0.46265	A0QY11	0.001607	1.3033
	A0QP45	0.002132	4.303	MFS55	0.001598	0.38283
	A0R4C5	0.002125	1.603	A0QQA1	0.001574	1.5641
	A0QR03	0.002119	1.2148	A0QPZ2	0.001572	4.1595
	A0QTT5	0.002115	0.91909	MUTT1	0.001566	5.2782
	A0QRB1	0.002072	0.1224	A0QVL4	0.001564	0.50014
	A0QPG7	0.002072	0.5882	A0R364	0.00156	1.5881
	A0QU91	0.002067	0.82705	A0QX50	0.001558	0.46025
	A0QYX6	0.002045	0.61797	A0R1I1	0.001551	3.3456
	A0QPM9	0.00202	2.1617	INO1	0.001543	4.5887
	A0QTI1	0.002006	0.43458	A0QR49	0.001543	4.0824
	A0QW24	0.001998	1.8257	A0QNR4	0.001537	0.74485
	RS13	0.001958	2.9005	A0R0W4	0.001536	0.32416
	A0QTT6	0.001951	1.0678	A0QUW4	0.001533	0.17903
	A0R379	0.001899	0.38203	A0QWR8	0.001529	0.068224
	A0QXV9	0.001872	1.2615	A0R3I2	0.001529	1.2545
	A0QWU5	0.001848	3.878	A0R049	0.001517	4.0129
	A0QZZ9	0.001827	0.71125	A0QXB1	0.001512	0.83145
	A0QUN5	0.001812	3.0094	A0QZA2	0.001512	0.87874
	A0QTP1	0.001803	0.76452	Y3950	0.001511	2.0564
	SEPF	0.001753	0.21622	A0QPY5	0.001502	1.8013

Appendix Table 2. Continued.

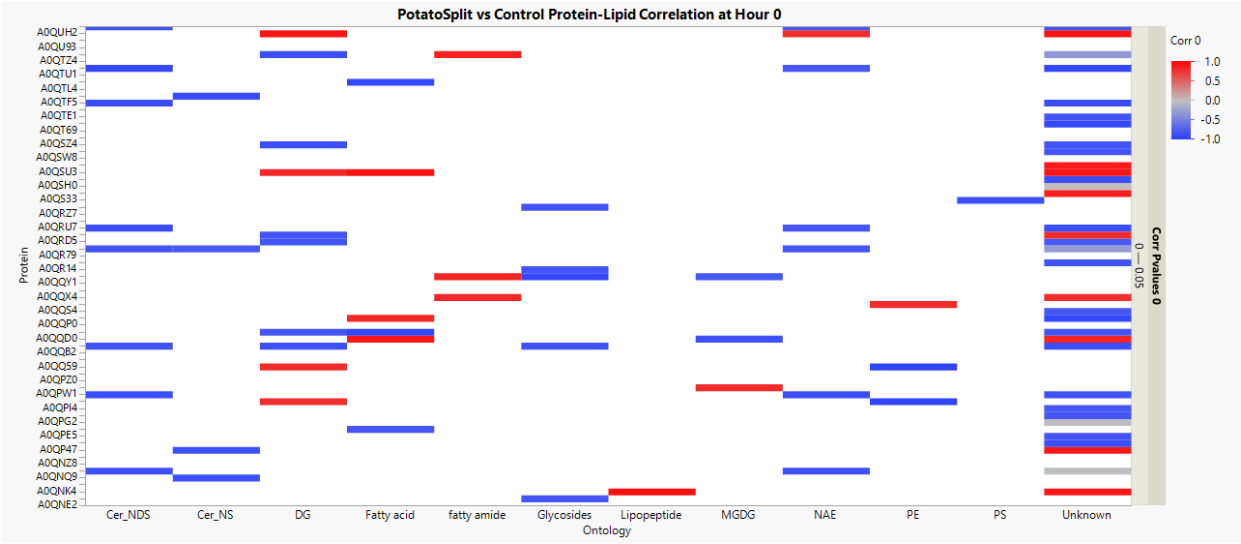
TIME	A0QW12	0.001734	1.7713	A0QNR6	0.001498	0.22021
	A0QUM7	0.001709	1.8091	A0QZ58	0.001489	0.049559
	A0R4K5	0.001706	0.30519	A0QW22	0.001479	0.72798
	A0R221	0.001703	0.46265	A0QY79	0.001474	1.0735
	A0QR14	0.001696	4.303	A0QZW4	0.00147	0.046646
	A0R4N7	0.001694	1.603	A0QYJ2	0.001463	0.39738
	EGTC	0.00146	0.83634	A0R101	0.001402	3.111
	A0R5K0	0.001459	1.7472	A0QR90	0.001399	0.68476
	Y1603	0.001455	0.11968	A0R238	0.001393	0.76287
	A0R3H7	0.001453	2.1679	A0QQF5	0.001389	0.64913
	A0QVL5	0.001452	0.94857	A0R5T7	0.001386	0.95174
	G3P	0.001445	2.2156	A0QRZ4	0.001372	0.79129
	A0R430	0.00143	0.66398	A0R2V5	0.001368	2.7092
	A0R0D4	0.001428	0.75113	PDXS	0.001364	0.63669
	A0QWT4	0.001427	2.1769	A0R2J0	0.001357	0.17253
	A0QSX6	0.001412	1.187			
INTERACTION	A0R2M2	0.003827	0.01301	A0QZ83	0.002339	0.79884
	A0R1P4	0.003734	0.28829	A0R717	0.002334	1.7533
	A0R2V8	0.003635	0.1416	A0R161	0.002332	0.08476
	A0QNQ9	0.003562	0.32607	CON__P02533	0.002331	0.052929
	A0R1A9	0.003472	0.025977	A0R2Z9	0.002331	0.32661
	A0QPG2	0.003389	0.42109	A0QR82	0.002324	0.013085
	A0R3Y4	0.00329	0.10348	MTF2	0.00232	0.81776
	LPRG	0.003221	0.19408	A0R069	0.00231	0.20274
	A0QT96	0.003213	0.017584	A0R2F0	0.002309	0.23419
	A0QYH7	0.003097	0.12311	A0QPQ5	0.002294	0.75109
	A0QXY7	0.003047	0.13774	A0QQX4	0.002294	0.12144
	A0QW09	0.00304	0.76333	A0QTF5	0.002287	0.36483
	RS16	0.003035	0.053464	A0QSY1	0.002278	0.059073
	A0R648	0.003007	0.86279	A0QNZ3	0.00227	0.10625
	A0R524	0.002989	0.009308	A0R0C8	0.00227	1.5777
	AZS12907.1	0.00297	0.25724	A0QS01	0.002264	0.094651
	A0R6A8	0.002961	0.92834	A0R3Q0	0.002258	0.12647
	A0QQX0	0.002951	0.035619	GCSH	0.002258	1.0795
	A0QY23	0.002945	0.18815	A0R3W0	0.002257	0.016474
	BFRB	0.002937	0.13427	A0R3C9	0.00225	0.023915
	RL36	0.002839	0.50435	METK	0.002249	0.37717
	AZS12964.1	0.002806	0.64075	A0R2B5	0.002248	0.12588
	A0QSV0	0.002805	0.46127	A0QY04	0.002247	0.4162
	AZS12906.1	0.002803	0.64493	A0QSW8	0.002244	0.2495
	THIE	0.0028	0.047746	A0QRJ6	0.002236	0.12524
	AZS12958.1	0.002795	0.47993	A0R6Z0	0.002221	0.12441

Appendix Table 2. Continued.

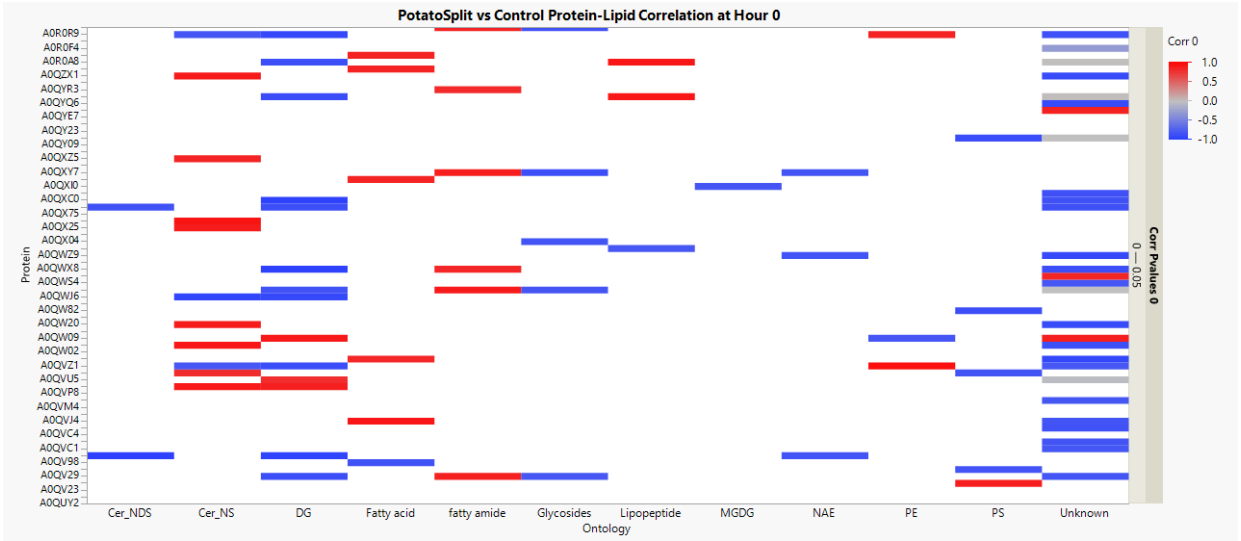
INTERACTION	AZS12961.1	0.00279	0.60705	A0R3P3	0.002216	0.48714
	A0QV52	0.002777	0.32995	KHSE	0.002216	0.99245
	A0R597	0.002766	2.5024	A0QR28	0.002213	0.38747
	A0QU56	0.002765	1.196	A0R4S0	0.002212	0.12387
	AZS12924.1	0.002733	0.65334	A0QY84	0.002209	0.12373
	A0QY08	0.002724	0.038241	A0QTZ4	0.002209	1.6089
	A0QTD8	0.002722	0.10754	A0QW16	0.002207	0.68027
	A0QRB4	0.00272	0.21176	DISA	0.002203	0.12339
	A0R0F4	0.002714	0.71623	A0QQV7	0.00219	0.12263
	AZS12928.1	0.002708	0.73994	DPDS	0.002181	0.29954
	A0QW36	0.0027	0.011373	A0R160	0.002178	0.12198
	A0QPW1	0.0027	0.57123	A0R3S3	0.002172	0.00672
	AZS12923.1	0.002693	0.81308	AZS12901.1	0.002159	0.28958
	A0QTQ5	0.002687	0.2959	A0R097	0.00215	0.21503
	A0QRZ5	0.002678	0.003491	RS182	0.002147	0.74098
	A0R293	0.002657	0.019903	CHDC	0.002146	0.12743
	AZS12905.1	0.002634	0.60606	A0QP13	0.002142	0.11999
	A0R5Q7	0.002627	0.44929	WHIA	0.002126	0.3041
	A0R461	0.002624	0.32758	A0QPW2	0.002126	0.10337
	TATA	0.002615	0.21274	A0QS33	0.002123	0.001286
	A0QT14	0.002594	0.087268	A0QSG5	0.002123	0.19664
	TOPON	0.002561	0.049899	A0QXH9	0.00211	0.15598
	A0R0G9	0.002556	0.34904	A0QV26	0.002109	1.004
	RRF	0.002553	4.28E-06	A0QXM6	0.002104	0.44228
	A0QVQ2	0.002523	0.24903	A0QZY7	0.002101	1.2179
	A0QW02	0.002516	0.22776	A0QX81	0.002099	0.034579
	A0QQP0	0.002515	1.2088	AZS12910.1	0.002093	1.1103
	AZS12920.1	0.002495	1.463	A0QS44	0.002092	1.6508
	AZS12954.1	0.002489	0.3099	A0QZG2	0.002088	0.056971
	A0R407	0.002488	0.40707	A0R711	0.002087	0.52435
	A0QWJ3	0.002485	0.044202	A0R5R5	0.002076	0.46743
	AZS12922.1	0.002474	0.52092	A0QV23	0.002074	1.5515
	MSHD	0.00247	0.28447	A0QUY6	0.002069	0.31317
	MSPA;sp	0.002467	0.16953	A0QR51	0.002067	0.014265
	A0R2C3	0.002461	0.03185	AZS12927.1	0.002065	0.50818
	A0R408	0.002459	0.019561	A0QPR6	0.002061	0.11545
	A0QR53	0.002443	0.16163	A0R5I8	0.002058	0.30285
	RL29	0.00244	1.1286	AZS12940.1	0.002051	1.0023
	A0R2T3	0.002425	0.69729	A0R593	0.002049	0.01177
	AZS12950.1	0.002418	0.30962	MSHA	0.002042	0.038034
	A0R612	0.002417	0.012056	A0QNL3	0.002041	0.19921
	A0R6M5	0.002391	0.01612	A0R1D1	0.002035	0.036393

Appendix Table 2. Continued.

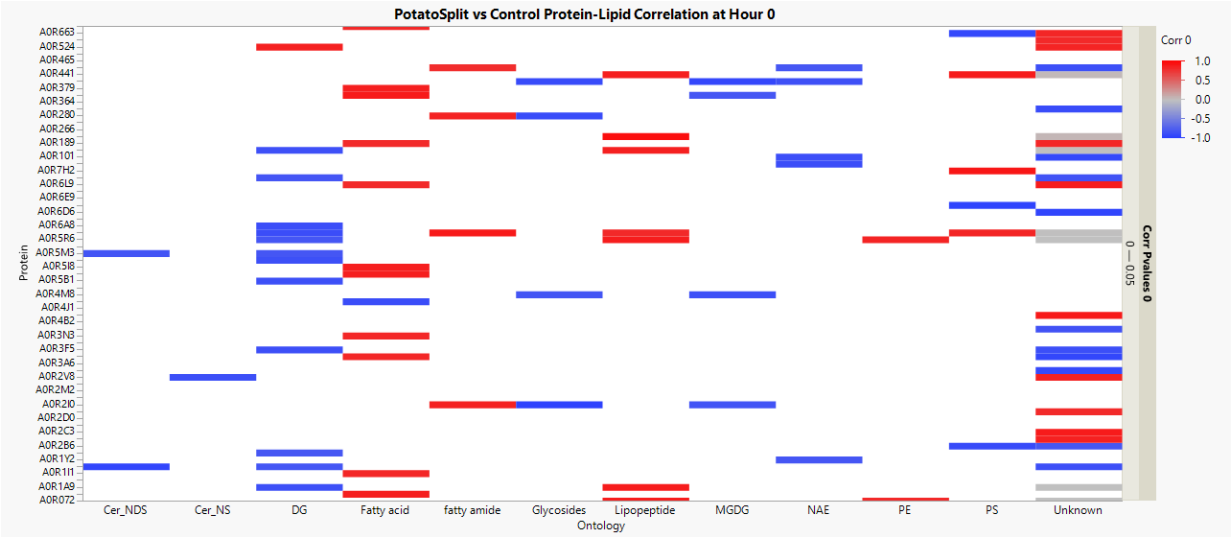
INTERAC	A0QVB6	0.002388	1.0439	AZS12944.1	0.002034	1.6007
	A0QQY1	0.002385	0.63961	A0QVG8	0.002029	0.000348
	A0R5R3	0.002346	0.6819	A0R3H5	0.00202	0.11314



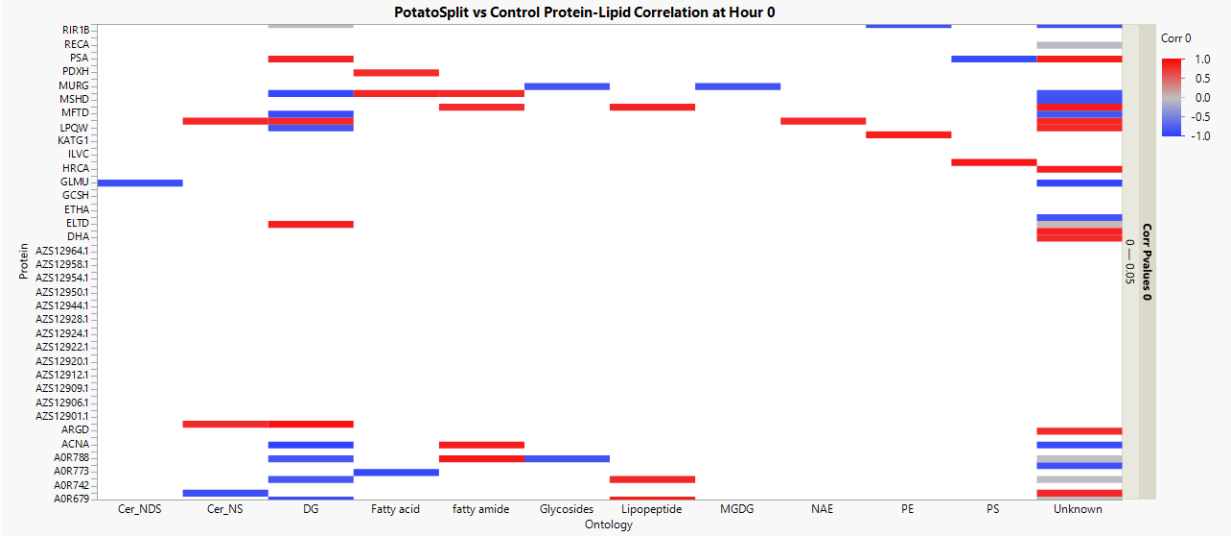
Appendix Figure 1. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



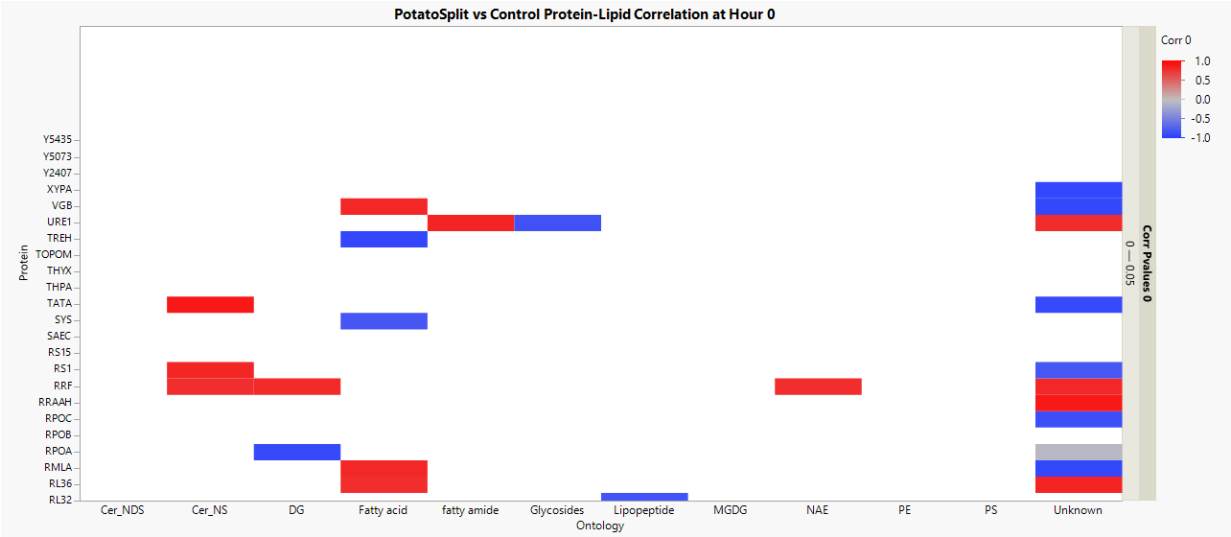
Appendix Figure 2. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



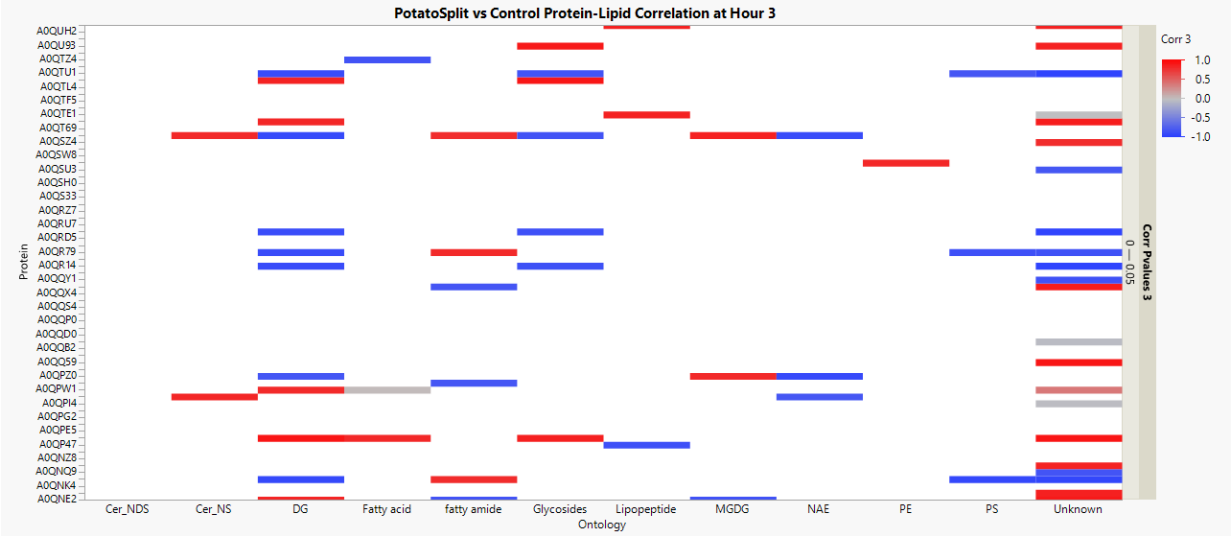
Appendix Figure 3. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



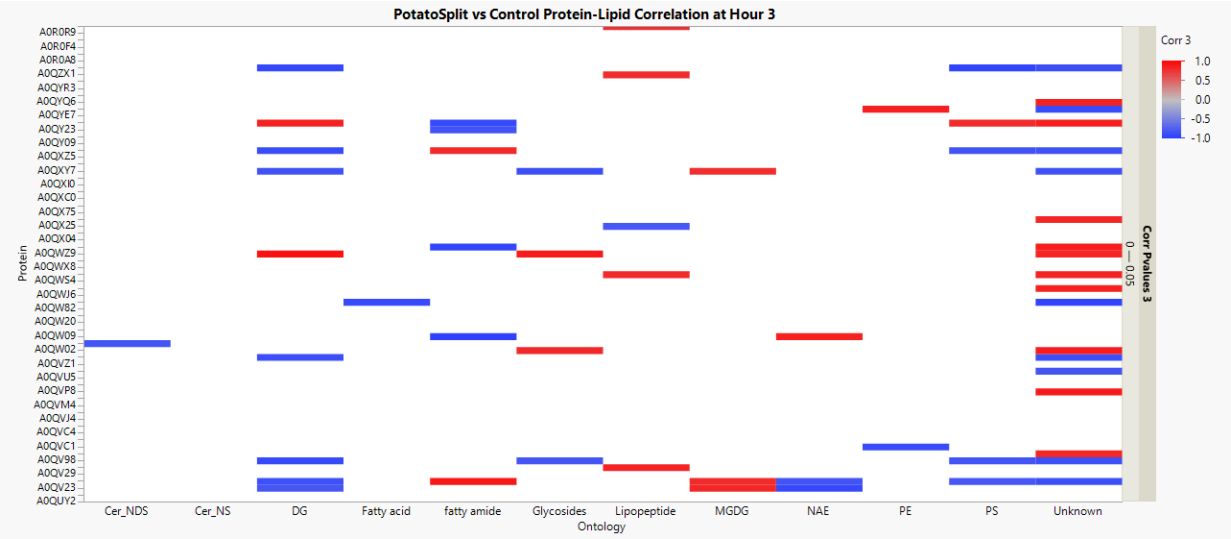
Appendix Figure 4. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



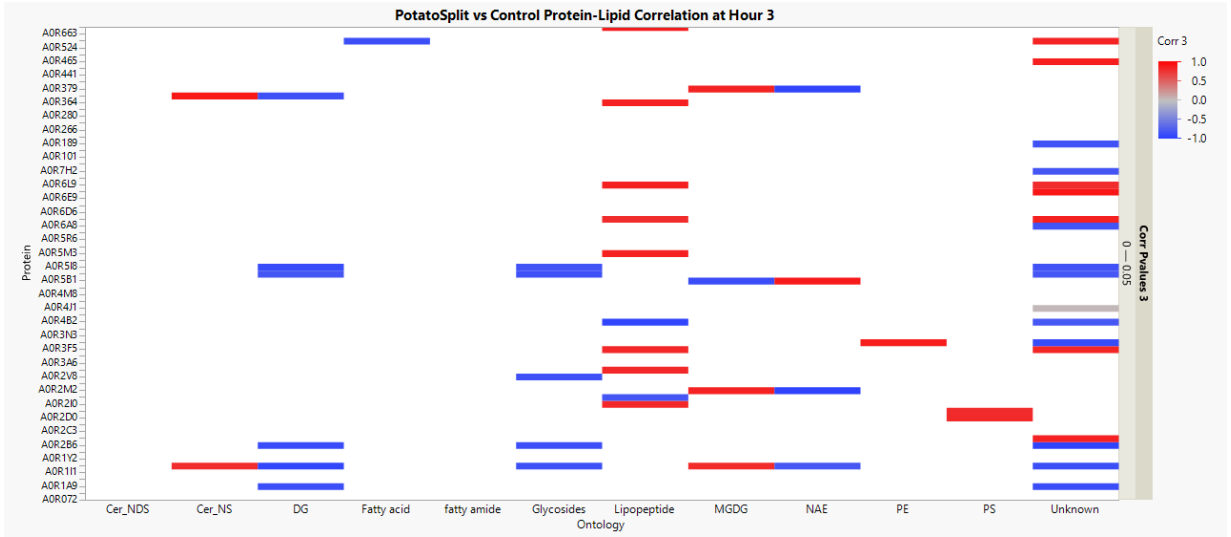
Appendix Figure 5. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



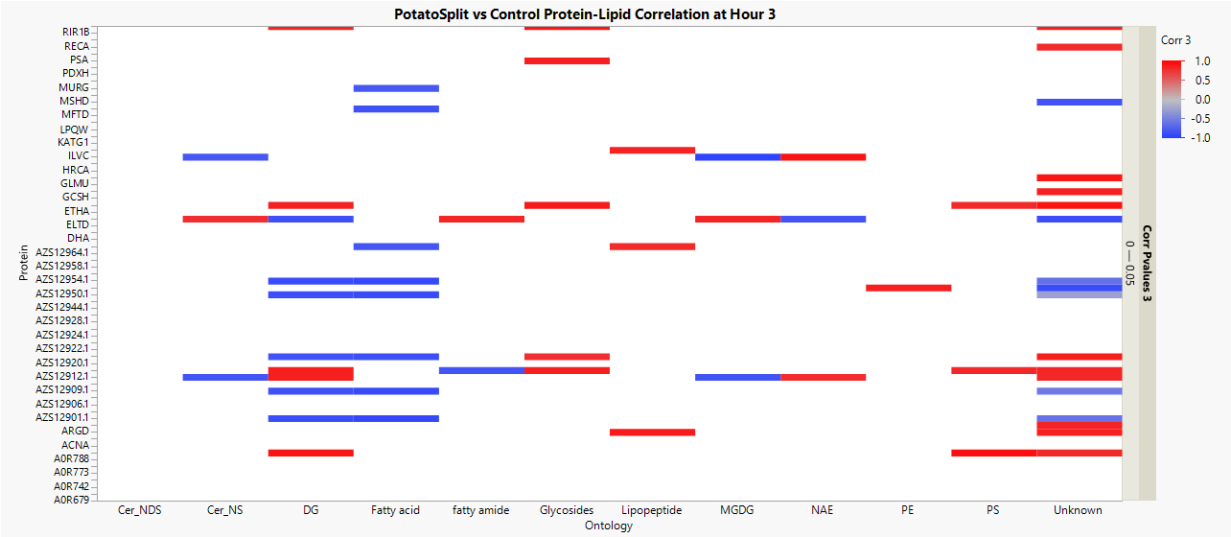
Appendix Figure 6. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



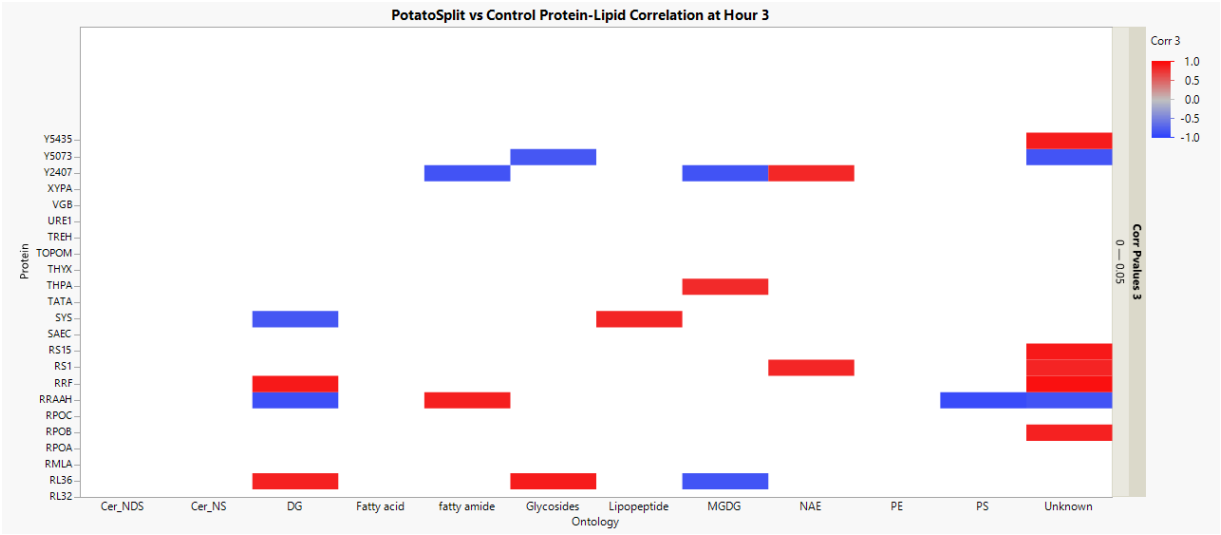
Appendix Figure 7. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



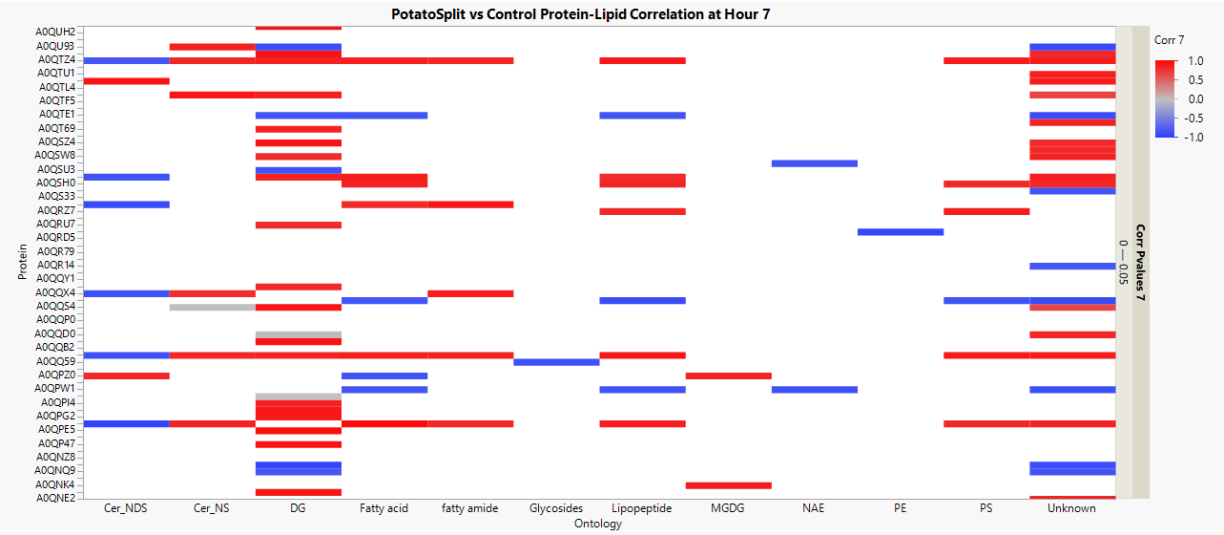
Appendix Figure 8. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



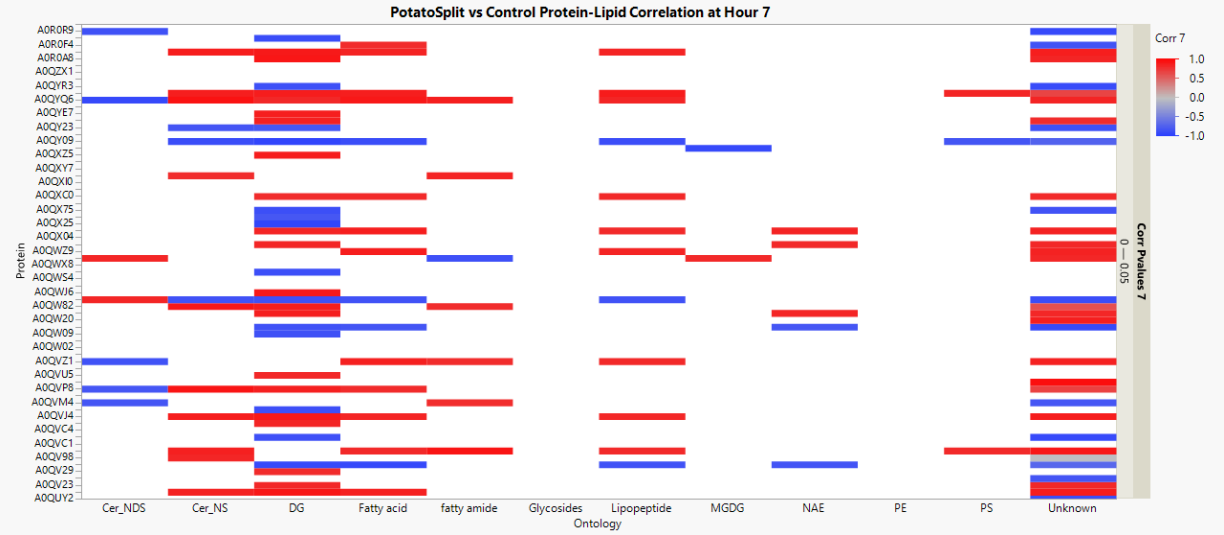
Appendix Figure 9. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



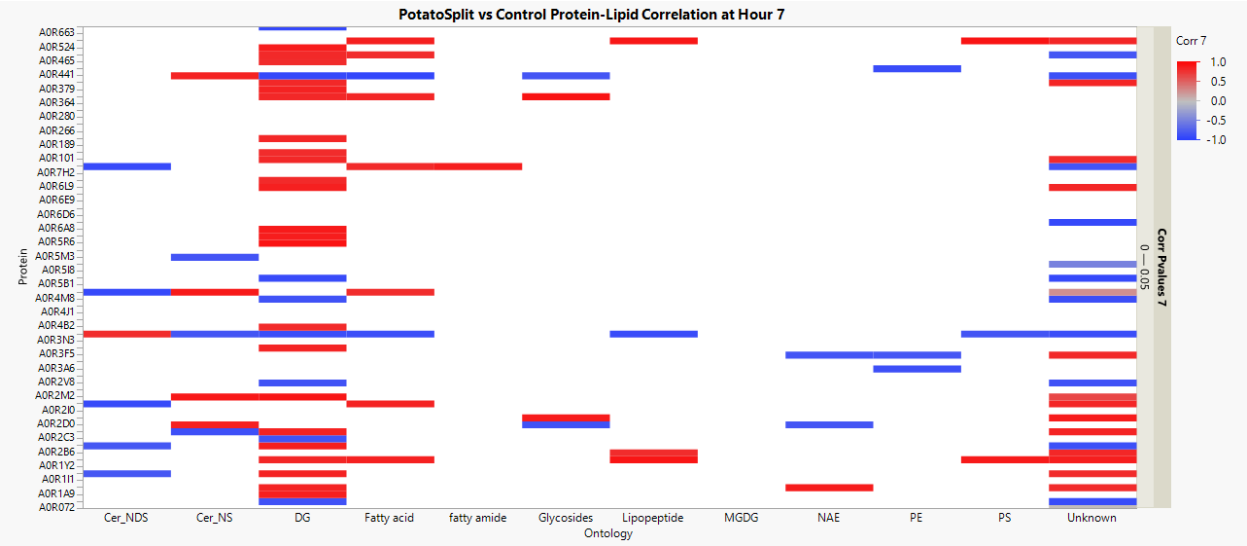
Appendix Figure 10. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



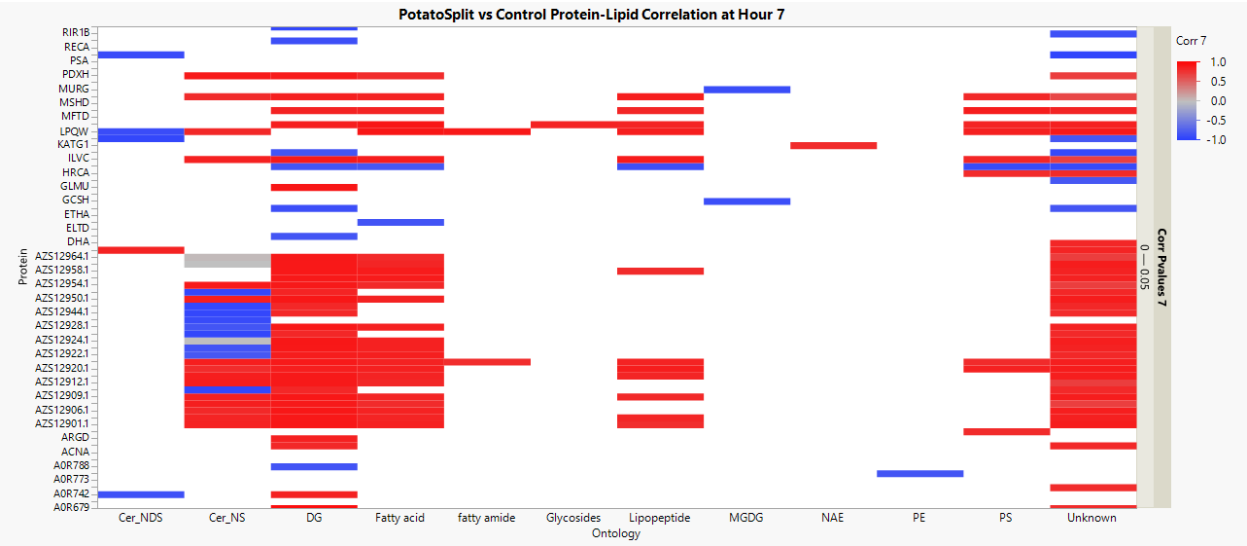
Appendix Figure 11. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



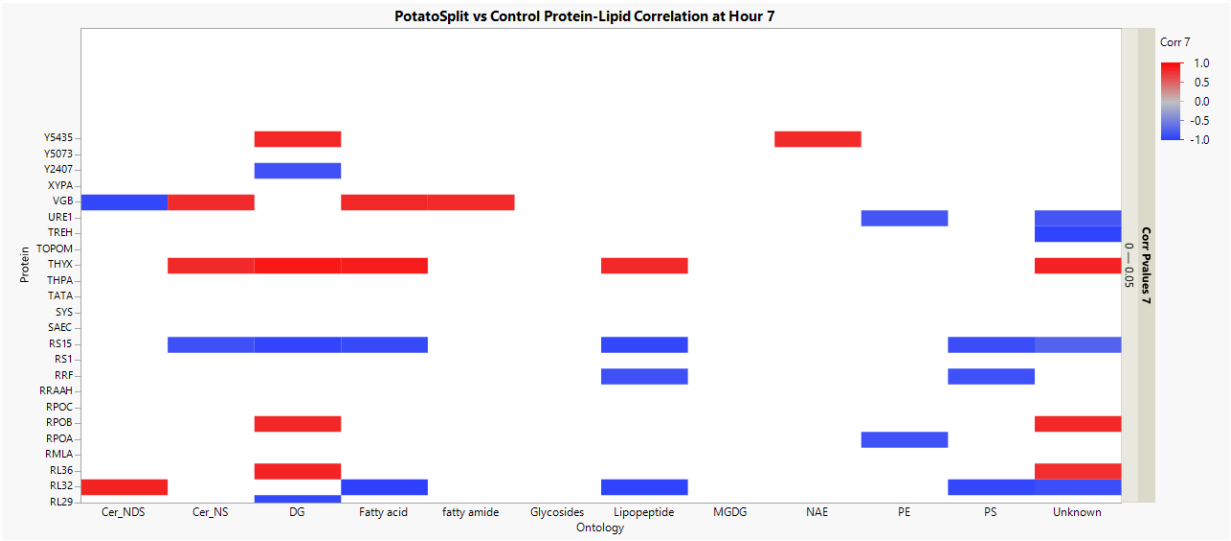
Appendix Figure 12. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



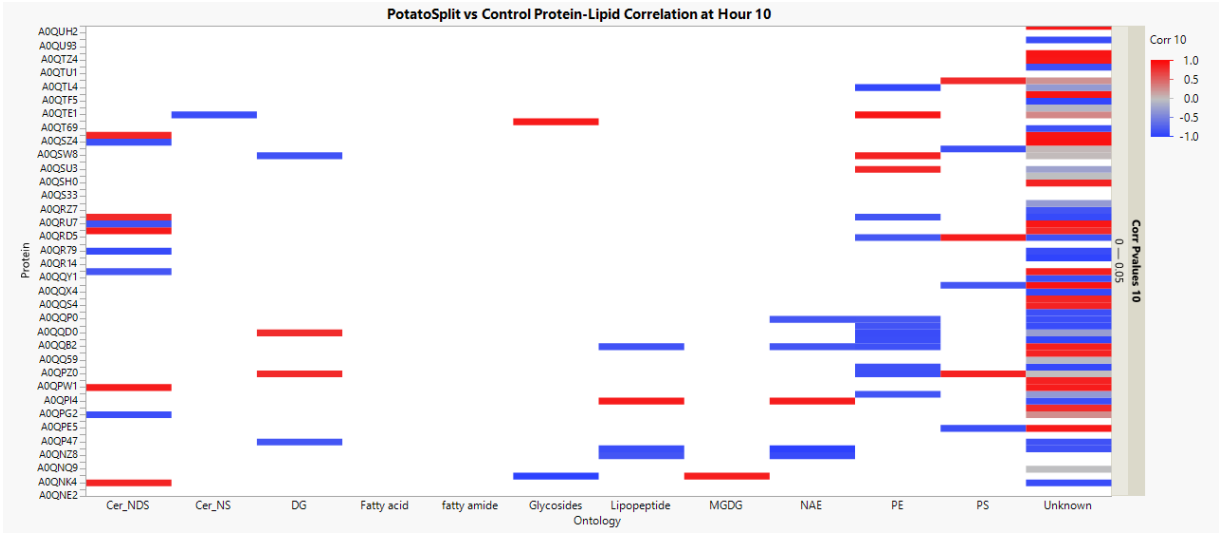
Appendix Figure 13. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



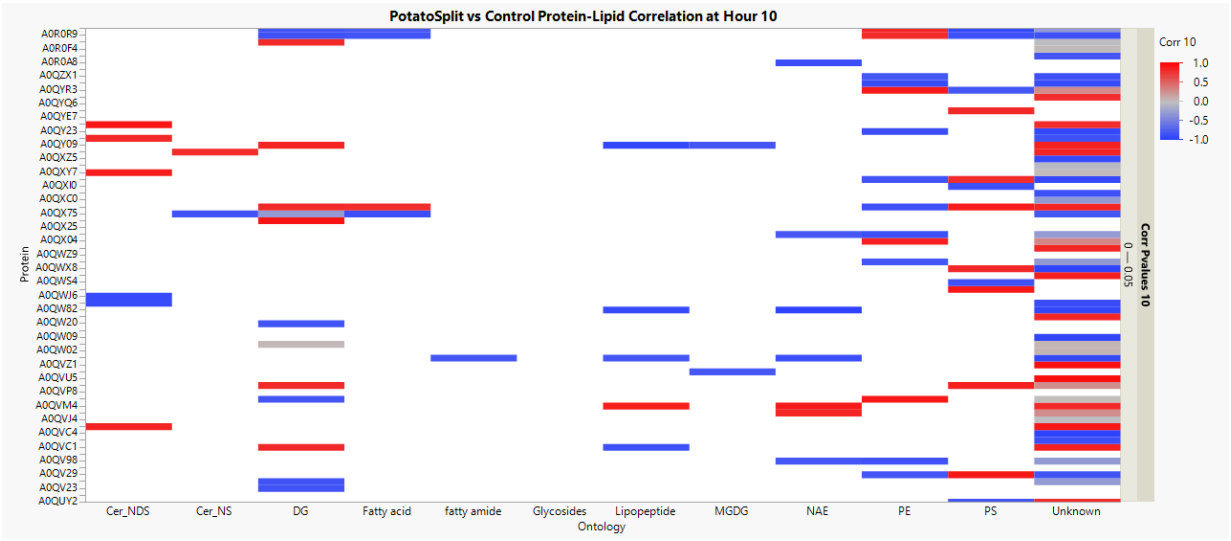
Appendix Figure 14. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



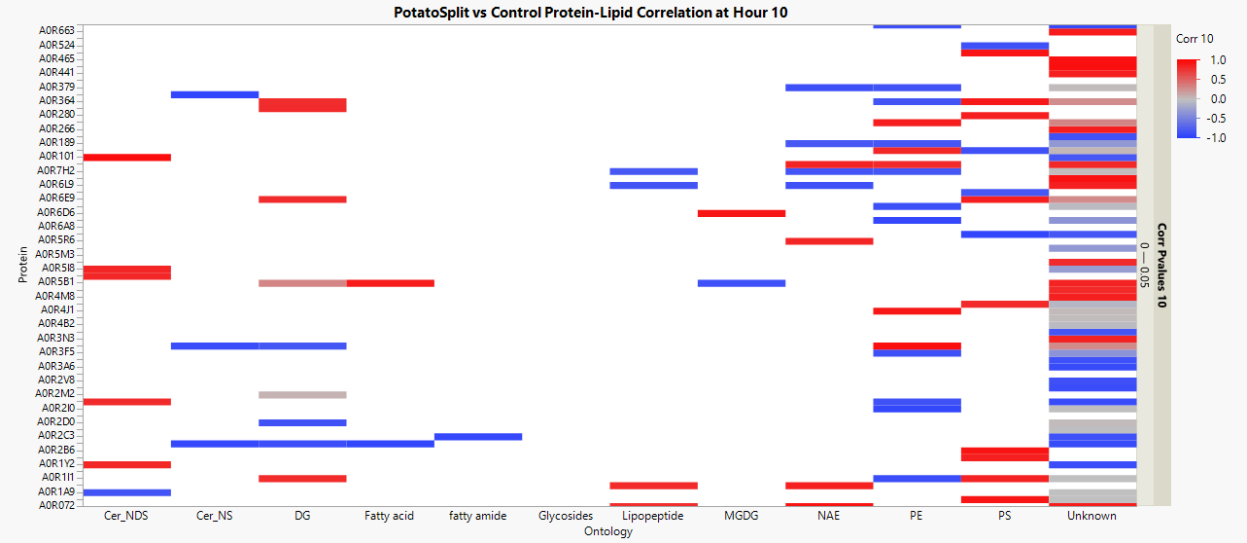
Appendix Figure 15. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



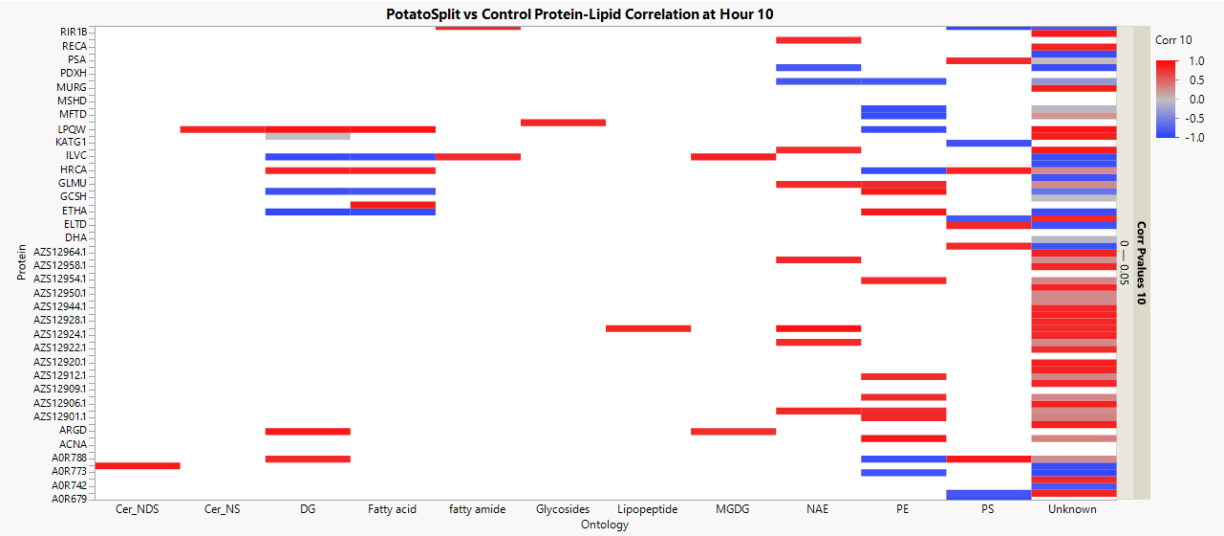
Appendix Figure 16. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



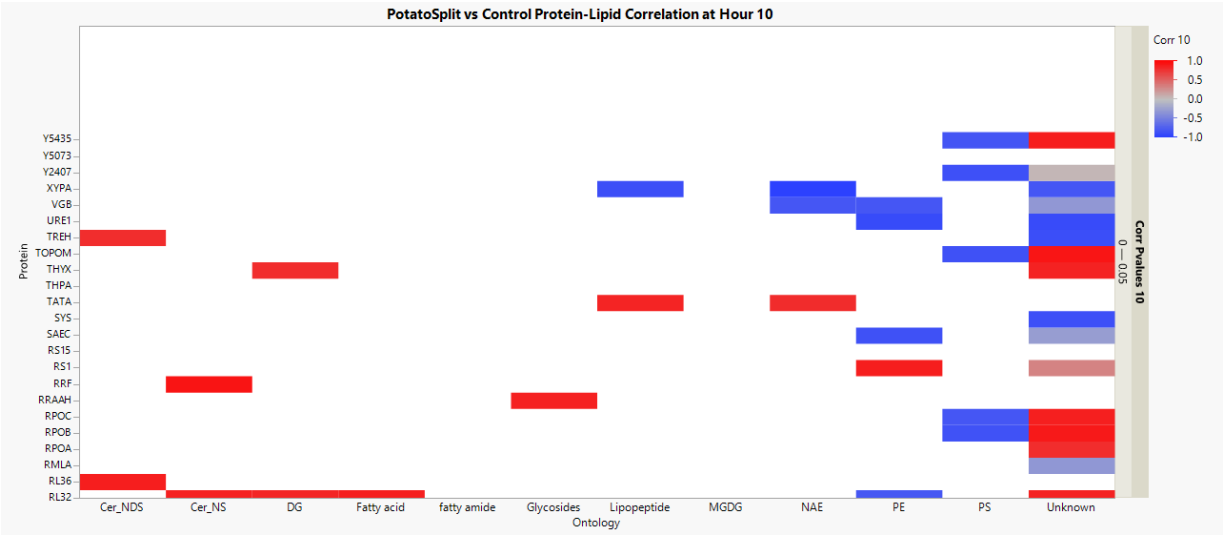
Appendix Figure 17. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



Appendix Figure 18. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



Appendix Figure 19. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



Appendix Figure 20. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for PotatoSplit vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

Zalkecks Compared Against the Control Data

Appendix Table 3. Zalkecks MS2 acquired protein results from a Metaboanalyst linear model with covariate adjustments, when compared against the control.

	ZALKECKS PROTEIN ID	LOG(FC)	P.VALUE
COMPARISON TO CONTROL	QAU03332.1	1.4744	0.00011098
	QAU03331.1	1.4289	0.0001797
	QAU03330.1	1.3952	0.00025468
	QAU03293.1	1.3613	0.00035858
	QAU03427.1	1.3392	0.00044685
	QAU03299.1	1.3307	0.00048573
	QAU03398.1	1.3284	0.00049682
	QAU03435.1	1.3276	0.00050049
	QAU03414.1	1.327	0.00050368
	QAU03461.1	1.3267	0.000505
	QAU03399.1	1.3261	0.0005079
	QAU03424.1	1.3256	0.00051049
	QAU03411.1	1.3235	0.00052137
	QAU03268.1	1.3231	0.00052334
	QAU03417.1	1.3222	0.00052754
	QAU03307.1	1.322	0.00052878
	QAU03420.1	1.3208	0.0005349
	QAU03462.1	1.3192	0.00054361
	QAU03419.1	1.3181	0.00054925
	QAU03422.1	1.3174	0.00055326
	QAU03294.1	1.3165	0.00055792
	QAU03290.1	1.3162	0.0005596
	QAU03308.1	1.3156	0.00056262
	QAU03350.1	1.3135	0.00057409
	QAU03423.1	1.312	0.00058264
	QAU03432.1	1.3055	0.00062057
	A0QRH0	1.3021	0.0006409
	QAU03302.1	1.3012	0.00064661
	QAU03373.1	1.3	0.00065397
	QAU03351.1	1.2997	0.00065624
	QAU03258.1	1.2936	0.00069579
	QAU03391.1	1.2929	0.00070021
	A0QV98	-1.2919	0.00070723
	QAU03395.1	1.2893	0.00072487
	QAU03321.1	1.2867	0.00074315
	QAU03363.1	1.2836	0.00076494
	QAU03460.1	1.2835	0.00076568

Appendix Table 3. Continued.

COMPARISON TO CONTROL	QAU03377.1	1.279	0.00079935
	QAU03402.1	1.2766	0.00081745
	QAU03359.1	1.2727	0.00084838
	QAU03393.1	1.2694	0.00087541
	QAU03364.1	1.2658	0.00090497
	QAU03329.1	1.2636	0.00092434
	QAU03445.1	1.2542	0.0010094
	QAU03439.1	1.2335	0.0012223
	QAU03459.1	1.2313	0.0012463
	QAU03326.1	1.2205	0.0013758
	QAU03272.1	1.2204	0.0013767
	QAU03404.1	1.2188	0.0013968
	QAU03251.1	1.2116	0.0014913
	QAU03450.1	1.2103	0.0015086
	QAU03347.1	1.2061	0.0015675
	QAU03458.1	1.2057	0.0015733
	COAD	-1.2054	0.0015772
	QAU03426.1	1.204	0.0015969
	QAU03406.1	1.192	0.0017788
	QAU03323.1	1.191	0.0017945
	QAU03360.1	1.1895	0.0018176
	QAU03449.1	1.1892	0.0018239
	QAU03372.1	1.1826	0.0019337
	QAU03446.1	1.1757	0.0020544
	QAU03401.1	1.1591	0.0023754
	QAU03338.1	1.1477	0.0026216
	QAU03412.1	1.1248	0.0031895
	A0QQJ9	-1.1232	0.0032336
	QAU03281.1	1.1197	0.0033314
	QAU03441.1	1.1144	0.0034837
	A0QPZ0	-1.1107	0.0035928
	QAU03387.1	1.1106	0.0035976
	A0R5P8	-1.1099	0.0036183
	QAU03354.1	1.1088	0.0036495
	A0QW20	-1.1076	0.003687
	QAU03369.1	1.106	0.0037367
	QAU03339.1	1.1059	0.0037393
	QAU03405.1	1.1052	0.003761
	QAU03409.1	1.1045	0.0037852
	QAU03431.1	1.1041	0.0037975
	QAU03451.1	1.1029	0.0038362
	A0QTU1	-1.102	0.0038643
	QAU03442.1	1.1007	0.0039071

Appendix Table 3. Continued.

COMPARISON TO CONTROL	QAU03457.1	1.0995	0.0039458
	QAU03319.1	1.0991	0.0039575
	QAU03317.1	1.0984	0.0039835
	QAU03356.1	1.0971	0.0040258
	QAU03376.1	1.0968	0.0040353
	QAU03352.1	1.0959	0.0040649
	QAU03448.1	1.0949	0.0040993
	QAU03397.1	1.0947	0.0041051
	QAU03348.1	1.0919	0.0042028
	QAU03367.1	1.0901	0.0042637
	QAU03452.1	1.0874	0.0043615
	QAU03368.1	1.0854	0.0044344
	QAU03324.1	1.0834	0.0045057
	A0QRU5	-1.0827	0.0045337
	NDP	-1.0807	0.0046098
	QAU03463.1	1.0802	0.0046259
	QAU03342.1	1.0802	0.0046265
	QAU03447.1	1.0764	0.0047737
	QAU03269.1	1.0681	0.0051062
	A0R723	1.0603	0.0054416
	A0R3Q2	-1.0592	0.0054902
	A0QZB3	-1.0582	0.0055335
	QAU03335.1	1.0553	0.0056621
	A0QW04	-1.0443	0.0061839
	QAU03257.1	1.0376	0.0065237
	A0QZY1	-1.0192	0.0075399
	A0QSX3	1.0186	0.0075762
	PUP	-1.0175	0.0076393
	A0QVC3	-0.99724	0.0089382
	QAU03444.1	0.99659	0.0089828
	A0R0D8	0.99643	0.0089937
	A0QVC1	-0.99441	0.0091341
	HIS3	-0.99397	0.0091652
	QAU03265.1	0.99123	0.009359
	QAU03346.1	0.98887	0.0095292
	QAU03349.1	0.98622	0.0097232
	QAU03429.1	0.9823	0.010017
	QAU03365.1	0.98072	0.010138
	QAU03425.1	0.97879	0.010287
	QAU03278.1	0.97851	0.010308
	QAU03312.1	0.97821	0.010331
	A0R4L6	-0.97755	0.010384
	QAU03320.1	0.97348	0.010707

Appendix Table 3. Continued.

COMPARISON TO CONTROL	RIR1B	0.97202	0.010825
	QAU03403.1	0.97052	0.010948
	QAU03296.1	0.9664	0.011291
	QAU03410.1	0.9636	0.01153
	QAU03408.1	0.95721	0.012091
	A0QR48	-0.94508	0.013224
	A0R717	0.93334	0.01441
	A0QTF6	-0.93319	0.014425
	A0QSL0	-0.92496	0.015311
	A0QRN7	-0.91825	0.01607
	A0QT13	-0.91211	0.016791
	CH602	0.9058	0.017563
	A0QQK0	-0.90301	0.017914
	A0R3W1	-0.89065	0.019545
	A0QRZ7	-0.88543	0.02027
	A0QS33	-0.88524	0.020297
	A0QYT1	-0.88049	0.02098
	A0R6V7	-0.87466	0.021845
	A0QQ02	-0.87071	0.022448
	A0R763	-0.8706	0.022464
	QAU03358.1	0.87017	0.022532
	A0QYK0	0.86959	0.022621
	LYSX	-0.86796	0.022877
	A0R5R1	-0.86741	0.022963
	QAU03353.1	0.86355	0.023578
	A0R401	-0.86295	0.023675
	QAU03454.1	0.86168	0.023882
	QAU03361.1	0.8613	0.023943
	A0QZ03	-0.85916	0.024296
	A0R2Z2	-0.85797	0.024493
	A0QYU2	-0.85486	0.025016
	QAU03274.1	0.85432	0.025108
	A0QSZ6	0.85232	0.025451
	SECA1	0.85126	0.025634
	A0QWF7	0.84978	0.025892
	A0QQW2	-0.84656	0.026459
	A0R2T2	-0.84317	0.02707
	A0R6S9	-0.84268	0.027159
	QAU03413.1	0.84265	0.027165
	A0QYD4	0.84146	0.027383
	A0QTS4	-0.8404	0.027577
	A0R7D8	-0.8386	0.027911

Appendix Table 3. Continued.

COMPARISON TO CONTROL	A0QRZ6	-0.83527	0.028539
	A0QSA8	-0.83432	0.028719
	A0QTF1	0.83337	0.028901
	A0QW71	0.83274	0.029023
	A0R5U7	-0.83153	0.029257
	A0QR94	0.83008	0.02954
	RRAAH	-0.82704	0.030141
	A0R474	-0.82625	0.030299
	A0QS54	-0.8243	0.030693
	QAU03343.1	0.82262	0.031034
	A0QWI4	-0.8207	0.031429
	A0R783	-0.8129	0.033076
	A0R479	0.81112	0.033462
	A0QSJ0	-0.80924	0.033875
	A0QPM2	0.80854	0.034028
	A0QR49	0.80832	0.034078
	A0QNZ3	-0.80784	0.034185
	A0R2B6	-0.80628	0.034531
	A0QYU5	-0.80342	0.035178
	A0R3Y8	-0.80148	0.035622
	A0R5I8	-0.80123	0.035679
	A0QQ62	0.80107	0.035715
	MTF2	-0.79545	0.037032
	A0R4C5	0.79398	0.037383
	A0QW16	-0.79292	0.037639
	A0QTR8	-0.79277	0.037674
	A0QUA2	-0.7921	0.037836
	FMT	-0.79164	0.037946
	A0QRN4	-0.78895	0.038604
	CH10	0.78842	0.038735
	A0QNT6	-0.78739	0.038992
	RL5	0.7854	0.039488
	A0R2L4	-0.78469	0.039667
	DPRP	0.78458	0.039693
	A0R461	-0.78208	0.040329
	A0R1A9	-0.77943	0.041012
	A0QUX4	-0.77821	0.041328
	A0QZ34	-0.77603	0.041901
	A0R4I6	-0.77446	0.042316
	A0R0A9	-0.77245	0.042857
	RPOB	0.77065	0.043343
	A0QQX6	0.77011	0.04349

Appendix Table 3. Continued.

COMPARISON TO CONTROL	A0QX97	-0.76926	0.04372
	A0QXD0	-0.7663	0.044539
	A0QWP9	-0.76192	0.04577
	A0QTW7	-0.76028	0.046239
	A0QWK5	0.76008	0.046297
	A0R6I9	-0.75781	0.04695
	A0R477	0.75763	0.047005
	A0QX01	0.75743	0.047062
	A0QQ04	-0.75742	0.047064
	A0QRZ3	-0.75612	0.047445
	CH601	0.75512	0.047739
	A0R782	0.75487	0.047814
	LERK	0.75386	0.04811
	A0QPZ4	-0.75125	0.048892
	A0R3L9	-0.75096	0.048978
	A0R506	-0.75034	0.049166
	A0QR80	0.74992	0.049293
	A0QTQ4	-0.74975	0.049343

Appendix Table 4. Zalkecks ANOVA Simultaneous Component Analysis (ASCA) significant lipid results when compared against the control to model phenotype and time effects and their interaction.

	PROTEIN ID	LEVERAGE	SPE	PROTEIN ID	LEVERAGE	SPE
PHENOTYPE	QAU03427.1	0.003384	2.96E-31	QAU03367.1	0.002243	1.48E-31
	QAU03398.1	0.00333	1.48E-31	QAU03452.1	0.002231	2.96E-31
	QAU03414.1	0.003323	1.48E-31	QAU03368.1	0.002223	2.96E-31
	QAU03461.1	0.003321	2.96E-31	QAU03324.1	0.002215	2.96E-31
	QAU03399.1	0.003318	2.96E-31	A0QRU5	0.002212	2.96E-31
	QAU03424.1	0.003316	1.48E-31	NDP	0.002204	2.96E-31
	QAU03411.1	0.003305	2.96E-31	QAU03463.1	0.002202	2.96E-31
	QAU03307.1	0.003298	2.96E-31	QAU03269.1	0.002153	2.96E-31
	QAU03422.1	0.003275	2.96E-31	A0QZB3	0.002113	2.96E-31
	QAU03423.1	0.003248	2.96E-31	QAU03335.1	0.002102	1.48E-31
	QAU03302.1	0.003195	2.96E-31	A0QW04	0.002058	0
	QAU03351.1	0.003187	2.96E-31	QAU03257.1	0.002032	2.96E-31
	QAU03258.1	0.003158	1.48E-31	A0QZY1	0.00196	2.96E-31
	QAU03395.1	0.003137	2.96E-31	A0QSX3	0.001958	2.96E-31
	QAU03321.1	0.003124	2.96E-31	PUP	0.001954	0
	QAU03363.1	0.003109	2.96E-31	A0QVC3	0.001877	1.85E-31
	QAU03445.1	0.002968	2.96E-31	HIS3	0.001864	7.40E-32
	QAU03439.1	0.002871	1.48E-31	QAU03429.1	0.001821	1.85E-31
	QAU03459.1	0.002861	0	QAU03365.1	0.001815	7.40E-32
	QAU03326.1	0.002811	2.96E-31	QAU03425.1	0.001808	1.85E-31
	QAU03404.1	0.002803	1.48E-31	QAU03312.1	0.001806	2.96E-31
	QAU03251.1	0.00277	2.96E-31	QAU03320.1	0.001788	3.70E-32
	QAU03450.1	0.002764	2.96E-31	QAU03403.1	0.001777	2.96E-31
	QAU03347.1	0.002745	0	QAU03296.1	0.001762	2.96E-31
	QAU03406.1	0.002681	2.96E-31	QAU03410.1	0.001752	1.85E-31
	QAU03449.1	0.002668	2.96E-31	A0QR48	0.001685	1.85E-31
	QAU03446.1	0.002608	2.96E-31	A0QRN7	0.001591	7.40E-32
	QAU03401.1	0.002535	2.96E-31	A0QT13	0.00157	1.85E-31
	QAU03338.1	0.002486	2.96E-31	CH602	0.001548	2.96E-31
	QAU03412.1	0.002387	2.96E-31	A0QQK0	0.001539	2.96E-31
	QAU03441.1	0.002343	2.96E-31	A0R3W1	0.001497	3.70E-32
	QAU03387.1	0.002327	2.96E-31	A0QS33	0.001479	3.70E-32
	A0R5P8	0.002324	1.48E-31	A0QYT1	0.001463	2.96E-31
	QAU03354.1	0.00232	2.96E-31	A0QQ02	0.001431	1.85E-31
	A0QW20	0.002315	1.48E-31	A0R763	0.00143	7.40E-32
	QAU03369.1	0.002308	0	A0QYK0	0.001427	7.40E-32
	QAU03339.1	0.002308	2.96E-31	A0R5R1	0.00142	2.96E-31
	QAU03405.1	0.002305	2.96E-31	QAU03353.1	0.001407	2.96E-31

Appendix Table 4. Continued.

PHENOTYPE	QAU03409.1	0.002302	1.48E-31	A0R401	0.001405	1.85E-31
	QAU03451.1	0.002295	2.96E-31	QAU03454.1	0.001401	1.85E-31
	QAU03442.1	0.002286	2.96E-31	QAU03361.1	0.0014	2.96E-31
	QAU03319.1	0.00228	2.96E-31	A0QZ03	0.001393	2.96E-31
	QAU03356.1	0.002271	2.96E-31	A0R2Z2	0.001389	2.96E-31
	QAU03352.1	0.002266	2.96E-31	SECA1	0.001367	1.85E-31
	QAU03448.1	0.002262	0	A0QQW2	0.001352	1.85E-31
	QAU03397.1	0.002261	2.96E-31	A0R2T2	0.001342	7.40E-32
TIME	A0QPE1	0.002063	3.0502	A0QNQ9	0.001467	4.0186
	A0QPW0	0.002042	1.2054	A0R379	0.001453	1.2917
	A0R478	0.00197	4.4556	A0R1B0	0.001441	2.1952
	A0R5T7	0.001963	0.072714	A0QR03	0.001441	0.72382
	A0QUI9	0.001943	2.3145	A0QV29	0.00144	2.5255
	XYPA	0.001883	2.7582	A0R2V3	0.001439	2.6934
	A0QPG7	0.001845	2.1627	A0QX32	0.001439	0.041484
	A0QT08	0.001834	0.016771	Y5790	0.001438	1.7479
	A0QU91	0.001825	1.6719	A0QTS9	0.001433	2.099
	A0R2P2	0.001824	2.0622	A0QVZ5	0.001432	0.17402
	A0QPX3	0.001821	1.0301	A0QUW4	0.00143	1.3216
	A0QQY2	0.001818	2.9186	A0R221	0.001429	1.3118
	A0QTP1	0.001815	0.24282	A0QT42	0.001427	0.11405
	A0QZ96	0.001789	6.4502	A0R242	0.001418	5.1666
	A0QQ63	0.001784	2.7306	A0QSZ4	0.001414	0.29717
	A0QUM7	0.001778	2.293	A0QTV7	0.00141	1.4037
	GLNE	0.001769	1.7884	A0R5G9	0.001409	1.3203
	ECCE1	0.001765	0.39871	A0QYH8	0.001408	2.8417
	A0QVX4	0.001759	0.71278	A0R4Y7	0.001405	0.40103
	A0QQ51	0.001757	2.5066	A0R4L2	0.001399	2.5984
	A0R0M4	0.001749	0.005263	GLGE	0.001392	0.14593
	A0R3I2	0.001743	0.41518	A0R3L4	0.001391	0.48527
	A0R2J4	0.001705	4.2266	A0QYS2	0.00139	2.6013
	RSMH	0.001665	2.759	A0QPF7	0.001387	0.63247
	HRCA	0.001645	1.7531	A0R101	0.001384	1.4405
	A0QQD7	0.001636	0.69306	A0R020	0.001379	0.25481
	A0QRB1	0.00163	1.006	A0R3P4	0.001377	2.1988
	A0QZA2	0.001608	5.3117	A0R4N5	0.001376	0.75434
	A0R2E7	0.001606	1.7867	A0R3H8	0.001368	4.8501
	A0QT19	0.001599	0.35972	AFTC	0.001363	0.67867
	DPPRS	0.001583	2.7123	A0QU92	0.001362	1.8924
	A0R1H2	0.001563	0.30701	A0QQ61	0.001352	0.213
	A0QQZ4	0.001562	0.44173	A0QT70	0.001346	0.50132
	RL14	0.001547	3.2463	A0QQ64	0.001345	0.38058

Appendix Table 4. Continued.

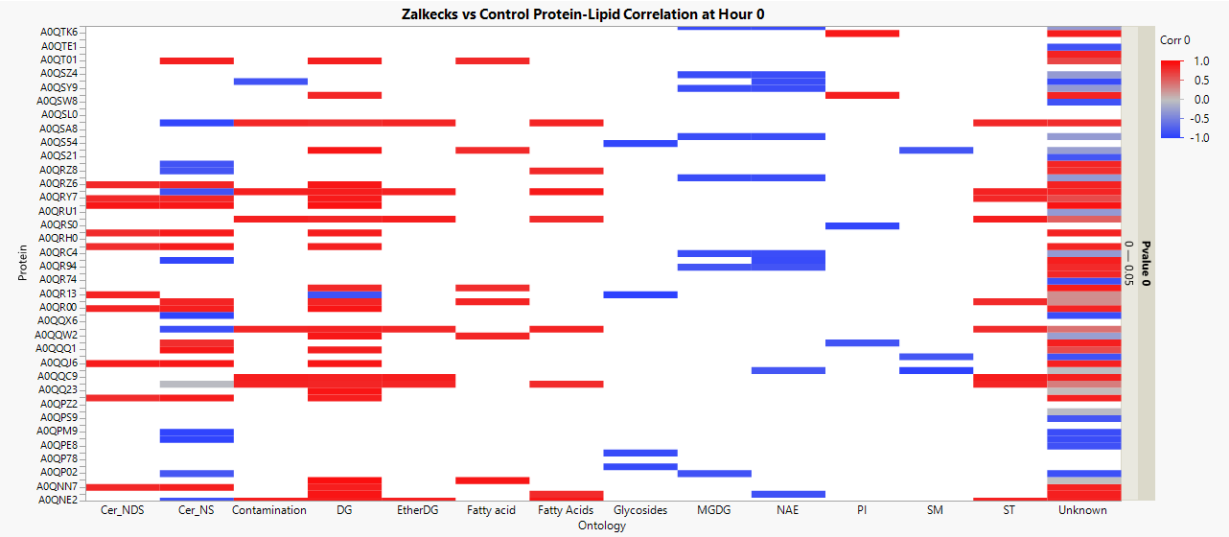
TIME	A0QX90	0.001529	0.51232	A0QUW5	0.001345	2.311
	A0R3B7	0.001516	0.62614	DXR	0.001344	0.82866
	A0QNI4	0.001515	0.10832	CAPP	0.001342	0.80567
	Q2YHI9	0.001515	0.17413	A0R2J0	0.00134	0.92059
	A0QZZ9	0.001514	2.2394	A0R0D5	0.001336	0.82891
	A0R2V4	0.001512	3.3101	A0QPY5	0.001332	2.2111
	Y3950	0.001505	2.3245	AFTD	0.001331	5.0886
	A0QT10	0.001499	1.7219	PAFA	0.001329	0.11602
	A0QW43	0.001493	0.3042	A0QXB1	0.001324	2.611
	A0QQR3	0.001486	0.35065	A0QWS9	0.001324	1.0558
	A0QTT5	0.001481	0.2361	A0QR14	0.001323	0.26127
	A0R084	0.001478	0.049533	A0R2V5	0.001322	2.1254
	TRES	0.001468	0.27615			
INTERACTION	PNP	0.004556	0.051213	QAU03310.1	0.00233	2.1497
	A0R1D1	0.004013	0.97008	A0QZI0	0.002322	2.1427
	A0QVZ5	0.003955	0.15348	QAU03394.1	0.002314	2.1357
	A0QR79	0.003953	0.17303	A0QQU8	0.002304	0.4202
	A0R2R7	0.003838	0.18516	ALR	0.002301	0.78281
	A0R218	0.003836	1.0777	A0R0S3	0.002292	0.41803
	A0QS85	0.003599	0.36445	A0R1H5	0.002291	0.28051
	EGTB	0.003581	0.65295	Y1165	0.002291	0.24954
	A0QVL1	0.003484	0.96512	QAU03367.1	0.002288	0.3482
	A0R6D6	0.003345	0.30268	QAU03376.1	0.002286	0.39633
	A0QZL1	0.003294	0.35764	A0QZQ6	0.00227	0.13383
	A0QYB0	0.003252	0.000172	QAU03451.1	0.002267	0.46972
	A0QSB1	0.003217	0.28837	SEPF	0.002267	1.0035
	A0QSZ3	0.003135	0.53668	A0QYU5	0.002259	0.43063
	RS16	0.003104	0.11093	A0QYT4	0.002257	0.008648
	A0R6I9	0.003095	0.84146	ATPG	0.002249	0.073132
	A0QS91	0.003059	0.070697	A0QYD8	0.002249	0.002992
	ECCD3	0.003057	0.17809	QAU03346.1	0.002248	0.006265
	A0QQX4	0.003007	0.008631	QAU03257.1	0.002245	0.12339
	A0QX14	0.002974	1.5379	QAU03338.1	0.002241	0.052184
	A0R614	0.002928	0.02403	QAU03320.1	0.002241	9.44E-06
	A0QTK6	0.00288	0.56325	QAU03353.1	0.002232	0.35221
	A0R6A9	0.002827	1.44	PKS5	0.002223	0.40535
	A0QXS0	0.002823	0.43525	A0R0W7	0.002221	0.002201
	A0QV26	0.00282	0.61768	A0R1A4	0.00222	0.31049
	WHIB2	0.002812	1.3698	Q2M5K3	0.002219	0.016619

Appendix Table 4. Continued.

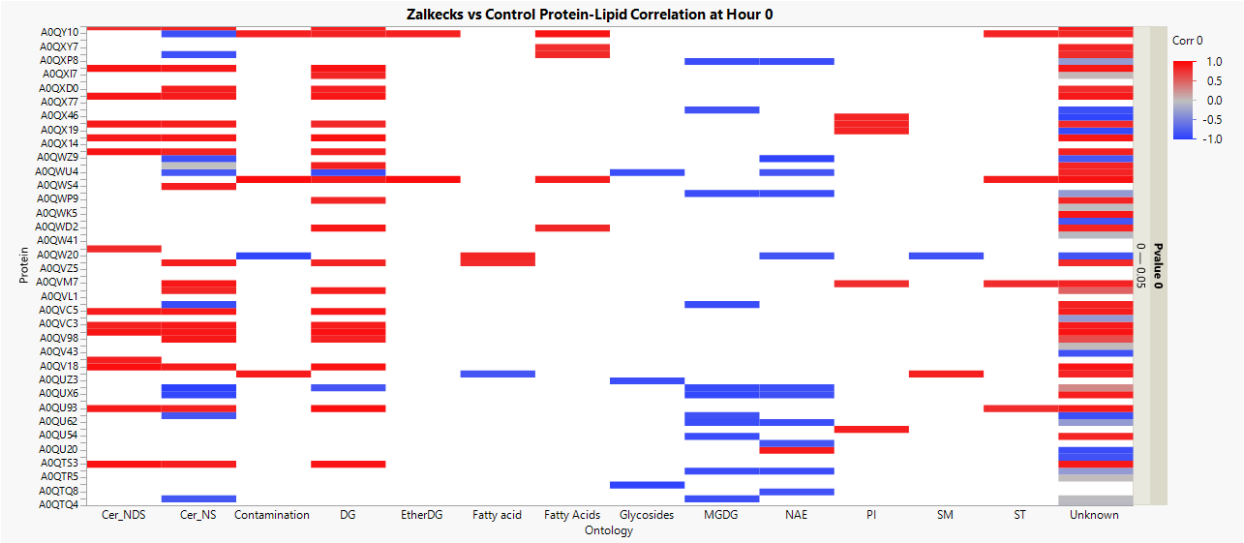
INTERACTION	ATPA	0.00279	0.2452	A0QWG0	0.002213	0.001672
	A0QXI7	0.002781	0.81492	QAU03352.1	0.002199	0.53208
	A0R2E9	0.002729	1.626	QAU03365.1	0.002196	0.009344
	PSA	0.002715	0.006068	A0QSX9	0.002191	0.39954
	A0QPM2	0.002712	0.61813	LERK	0.002187	0.74892
	Y2731	0.00271	0.006928	A0R2Q4	0.002186	0.049426
	A0R1E4	0.002683	0.52824	RRF	0.002185	0.12043
	A0R028	0.002681	0.025526	A0R5W6	0.002184	0.026498
	A0QZX8	0.002678	0.002087	A0QZX1	0.00218	0.015151
	A0QVT1	0.002674	0.22348	QAU03369.1	0.00218	0.65327
	Y6073	0.002673	0.63282	ENO	0.002173	0.051999
	A0QS43	0.002668	0.74382	QAU03349.1	0.002171	0.024685
	GCSH	0.002658	0.070788	QAU03444.1	0.002171	0.046169
	ACDH2	0.002648	0.018206	A0R6D2	0.002171	0.049822
	PUP	0.002641	0.29051	A0R2I1	0.002168	0.81072
	A0QWI4	0.002628	0.52793	QAU03457.1	0.002166	0.62285
	A0QYY3	0.002606	0.005439	A0QPE0	0.002155	0.39288
	Y4692	0.002605	0.055481	A0QZY1	0.002142	0.29923
	ETHA	0.002596	5.18E-05	QAU03358.1	0.002133	0.19097
	A0QVB6	0.002554	0.19251	A0R594	0.002125	0.51783
	A0R0C8	0.002541	0.012192	A0R0N9	0.002123	0.064787
	A0R716	0.002513	0.000287	QAU03405.1	0.002122	0.76463
	GPGS	0.002501	0.51831	A0R5R4	0.002109	0.22546
	A0QRC4	0.002491	0.13172	A0R678	0.002106	0.07554
	SAHH	0.002486	0.26439	QAU03387.1	0.002104	0.85724
	A0QTE1	0.002475	0.80978	A0QW02	0.002101	0.007691
	A0QQC9	0.002471	0.32159	A0QWS3	0.002101	0.16218
	GYRB	0.002459	0.053167	A0QS63	0.002098	9.66E-05
	Q3L887	0.002457	0.23192	QAU03339.1	0.002097	0.8267
	A0QQF8	0.002455	0.16118	QAU03265.1	0.002094	0.078296
	A0R5B0	0.002442	0.12674	QAU03319.1	0.002092	0.77137
	A0QYH0;tr	0.002441	0.89977	A0QQJ6	0.002089	0.13312
	QAU03335.1	0.002434	0.051836	A0R003	0.002088	0.083287
	QAU03408.1	0.002428	0.093928	QAU03317.1	0.002087	0.77335
	A0R4N3	0.002426	0.078372	A0QQ23	0.002084	0.50373
	MIMA	0.002419	0.061812	A0R0A8	0.002075	1.004
	A0R206	0.002411	2.2253	A0QVA0	0.002072	2.6104
	QAU03447.1	0.002408	0.1398	QAU03312.1	0.002069	0.057211
	A0R613	0.002405	0.026692	QAU03446.1	0.002059	0.16978
	QAU03348.1	0.002381	0.24075	QAU03372.1	0.002053	0.49215
	RNJ	0.002374	0.29174	RS12	0.002046	1.4157
	PYRF	0.002372	1.4257	Q2M5K4	0.002042	0.011644

Appendix Table 4. Continued.

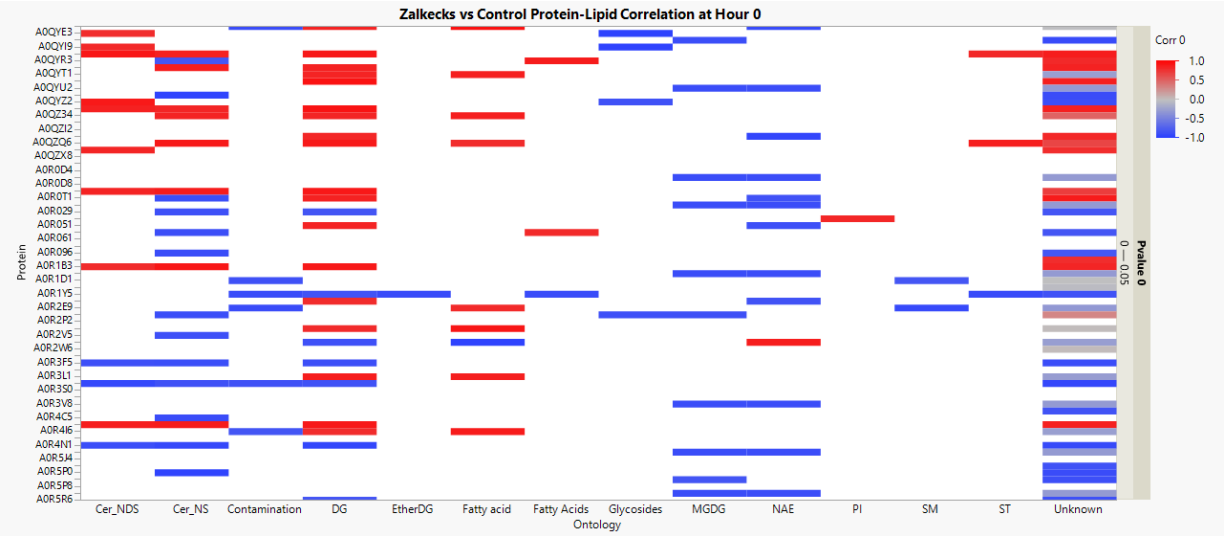
INTERACTION	A0R4N1	0.002367	0.63276	APT	0.002038	0.28293
	DCDA	0.002358	0.064968	A0QX15	0.002031	0.13515
	A0QNZ8	0.002351	0.004273	A0QSU3	0.002031	0.038611
	QAU03452.1	0.00235	0.25286	CH602	0.002028	0.2617
	QAU03356.1	0.002343	0.31773	A0QX21	0.002022	0.045265
	QAU03448.1	0.002341	0.30737	A0R477	0.002011	0.18732
	QAU03368.1	0.002341	0.25222	QAU03354.1	0.002009	1.0718
	QAU03361.1	0.002341	0.5395	A0QU62	0.002008	1.0223
	QAU03463.1	0.002339	0.22634	A0R4S8	0.002007	0.29025
	A0QXM6	0.002331	0.6737	GLYA	0.002006	0.16654
	A0R269	0.002331	0.17377	A0R6G4	0.002003	0.001038



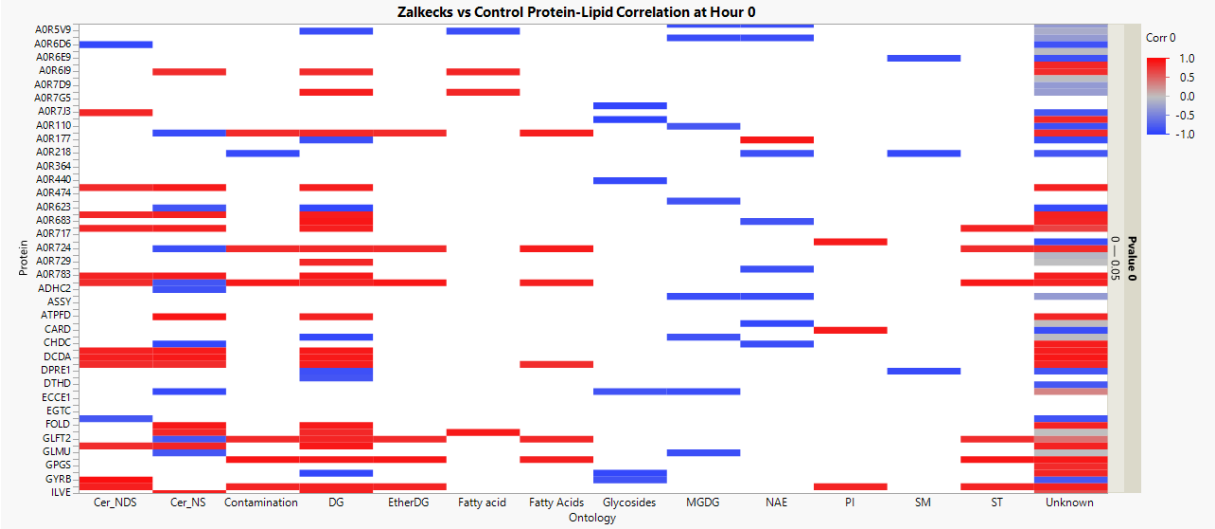
Appendix Figure 21. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



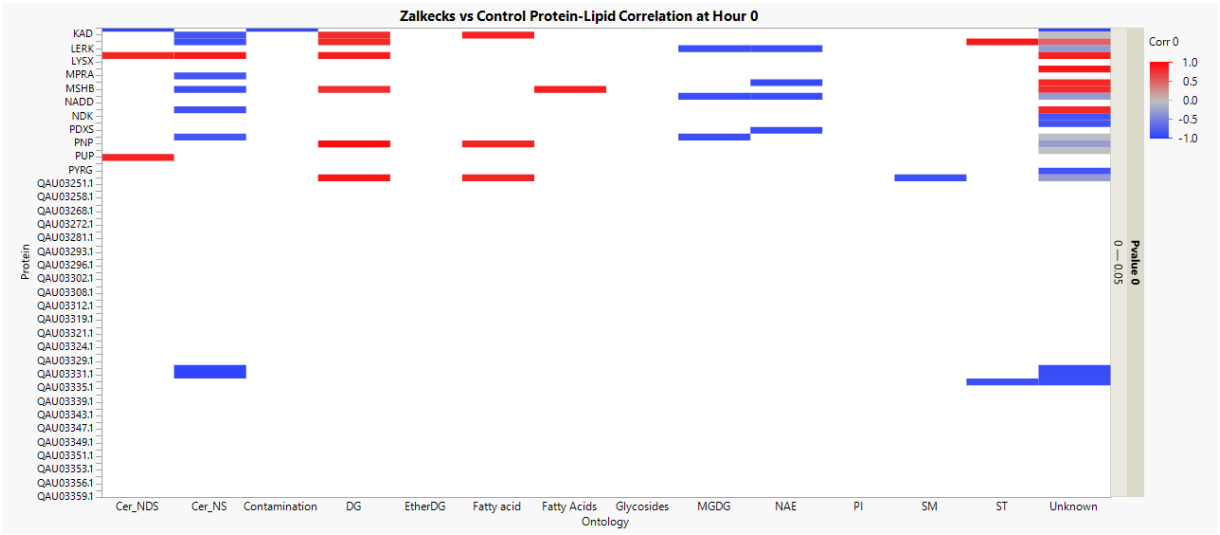
Appendix Figure 22. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



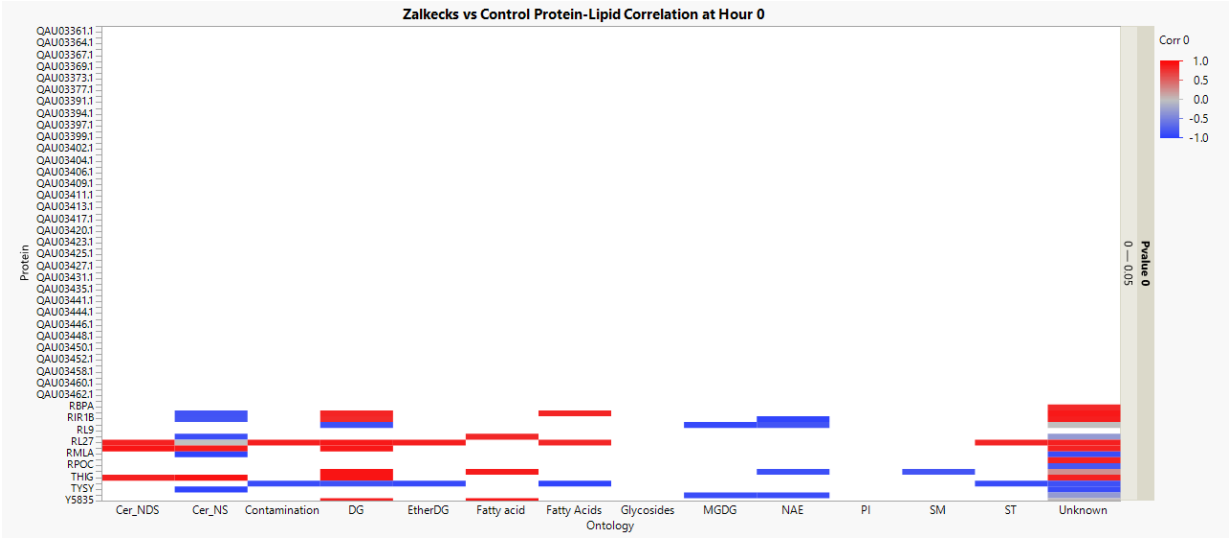
Appendix Figure 23. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



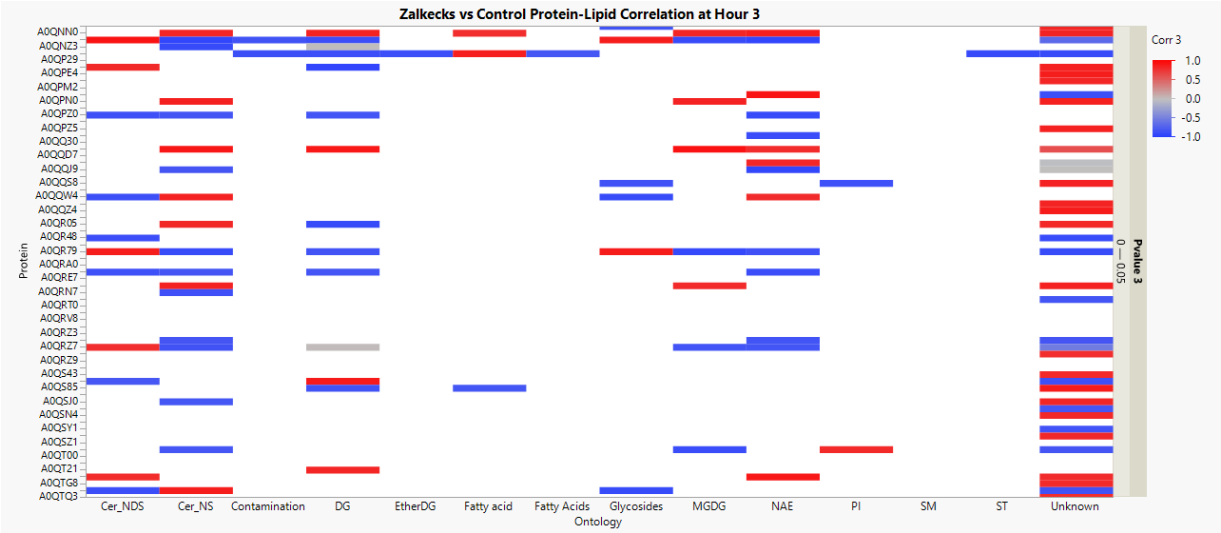
Appendix Figure 24. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



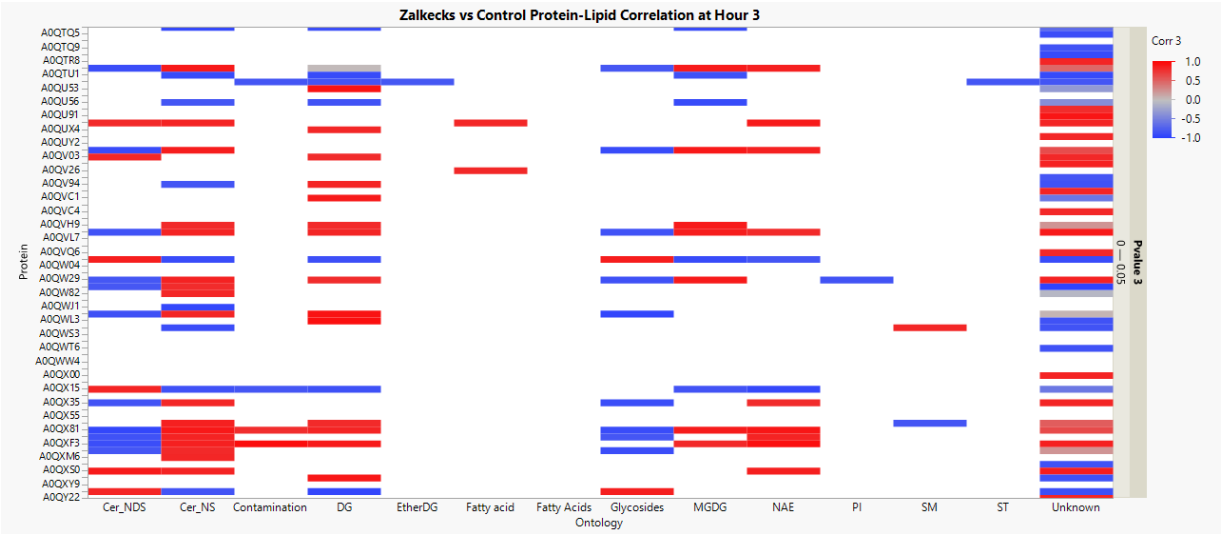
Appendix Figure 25. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



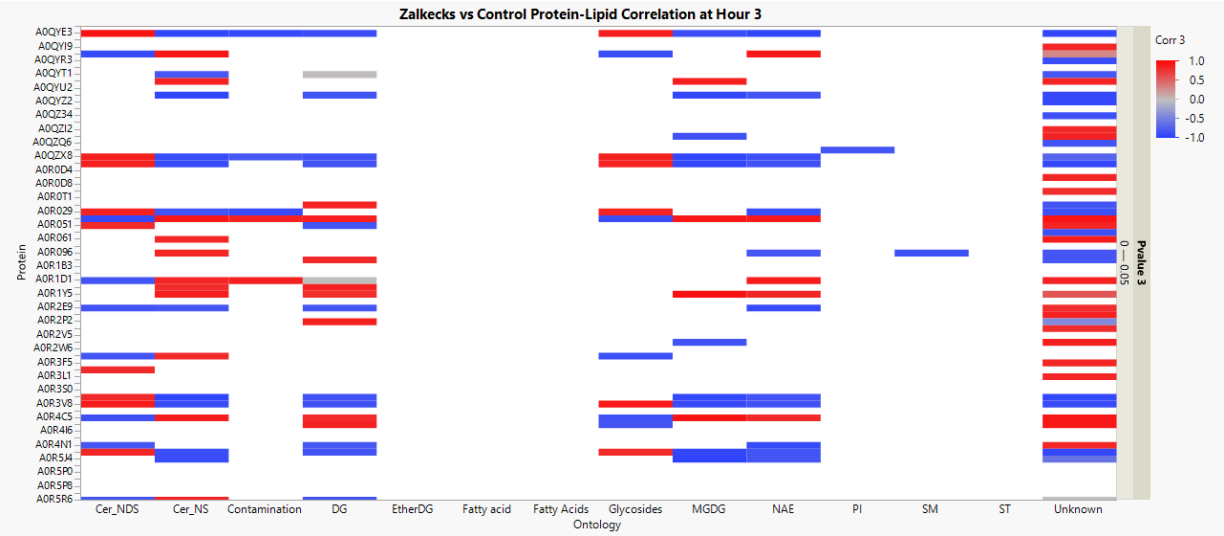
Appendix Figure 26. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 0, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



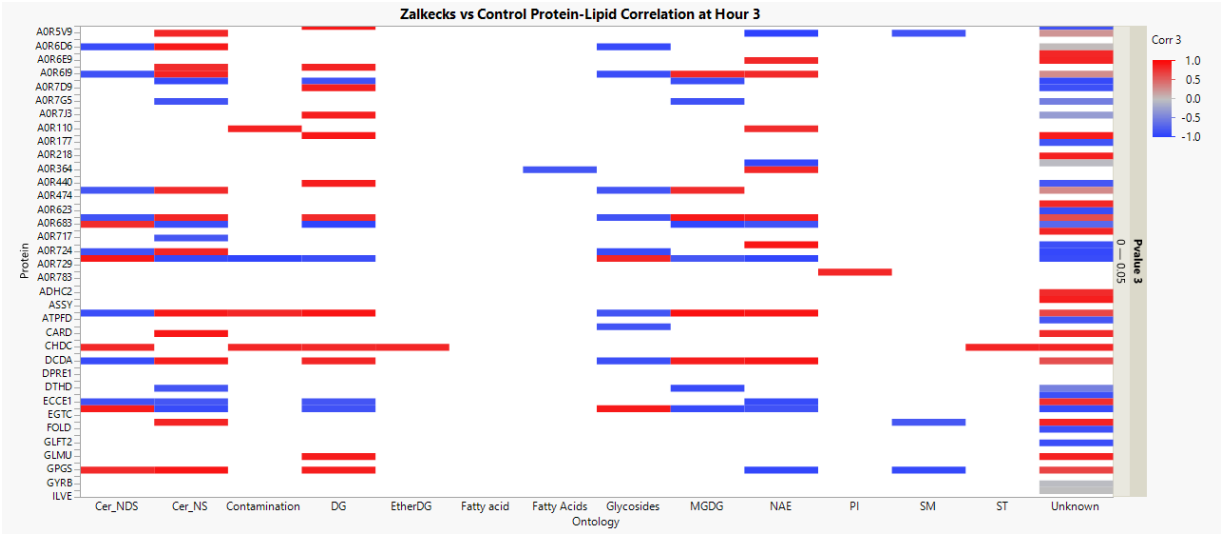
Appendix Figure 27. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



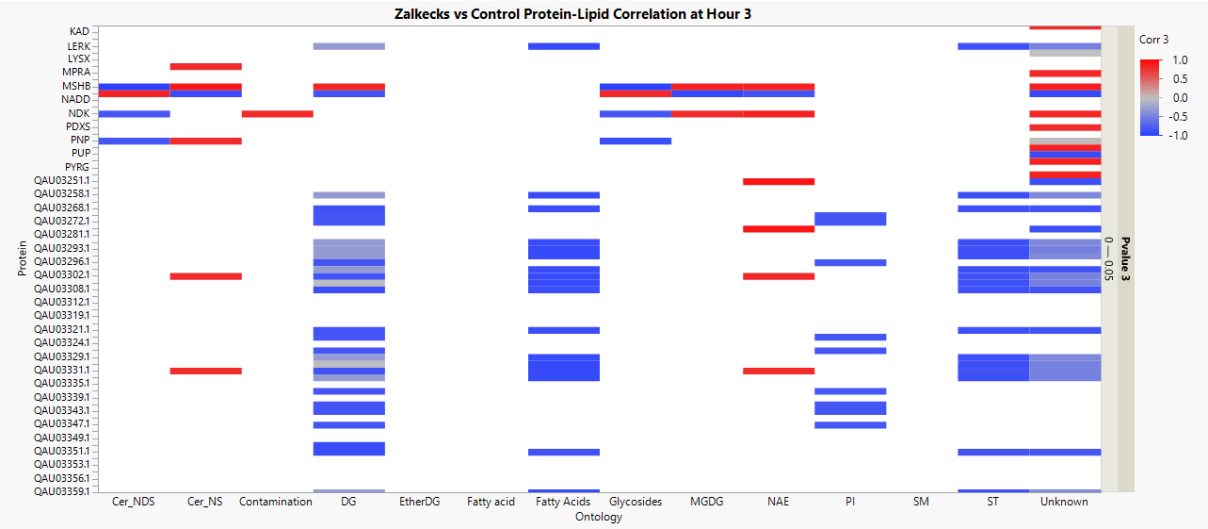
Appendix Figure 28. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



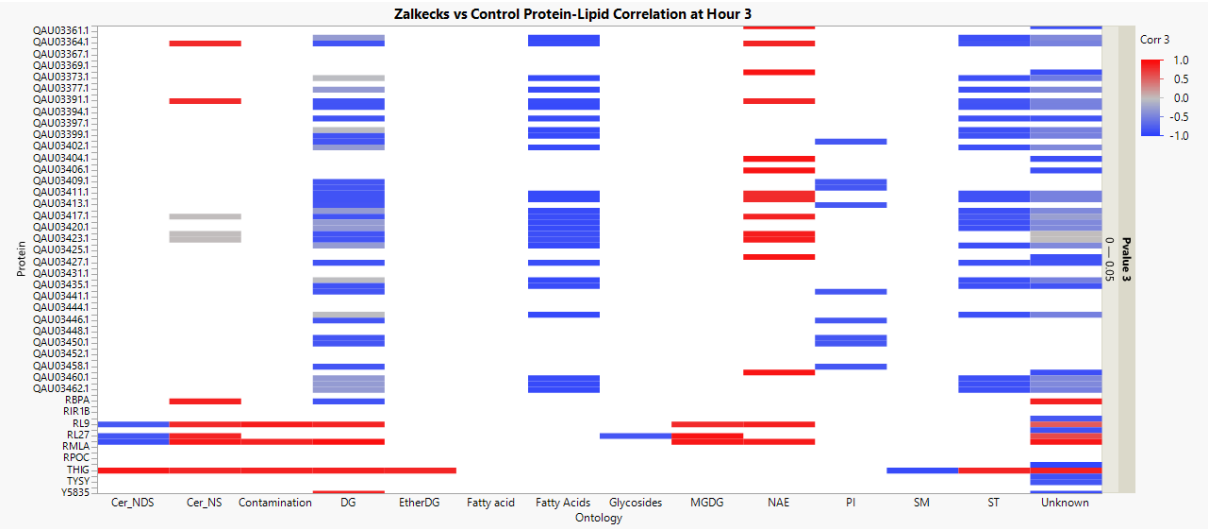
Appendix Figure 29. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



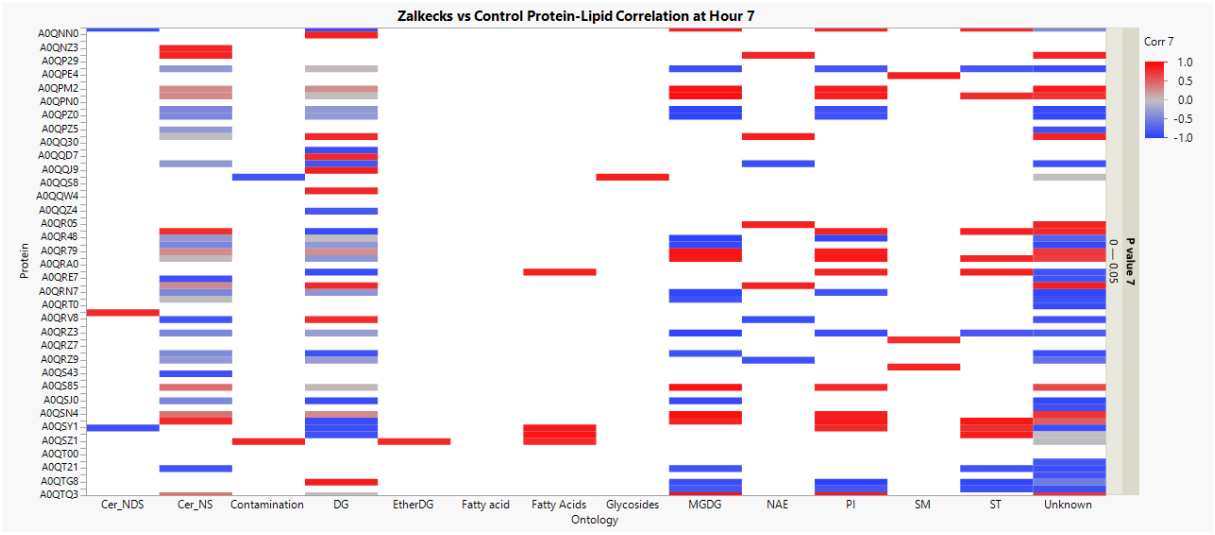
Appendix Figure 30. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



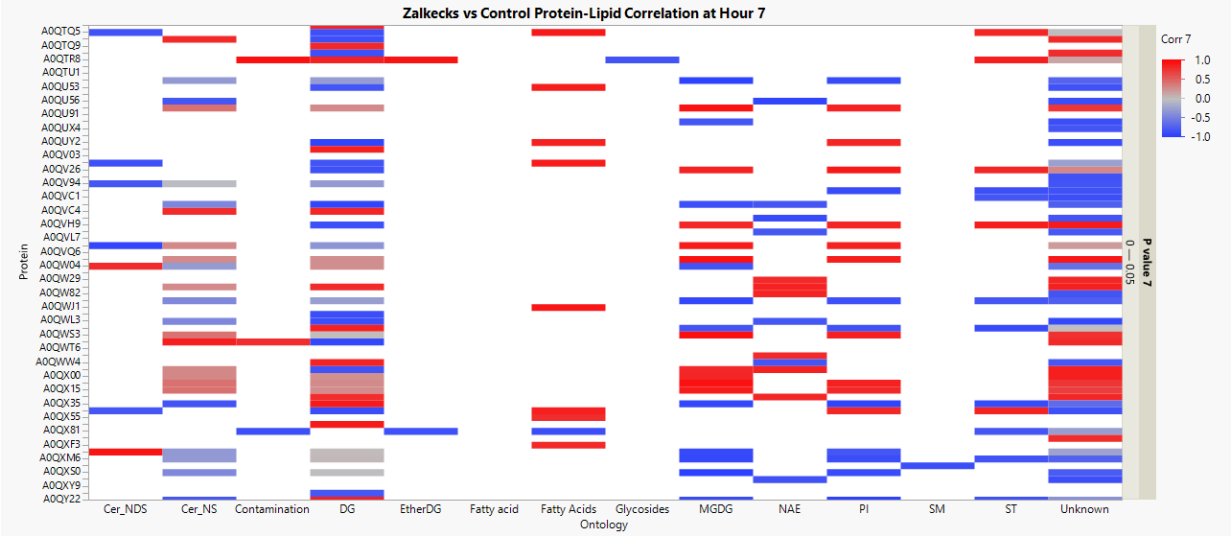
Appendix Figure 31. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



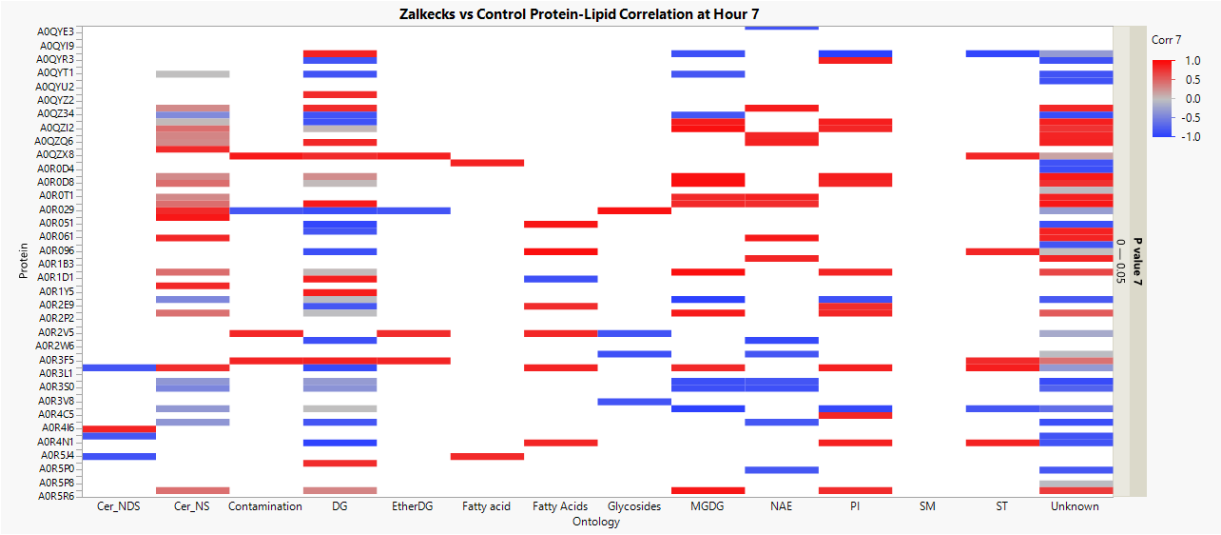
Appendix Figure 32. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 3, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



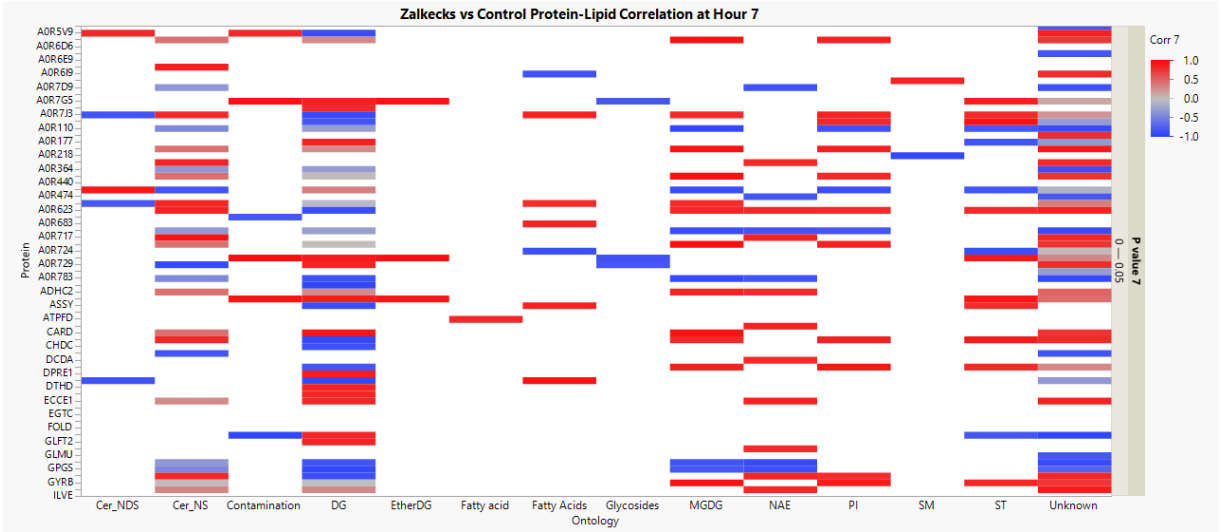
Appendix Figure 33. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



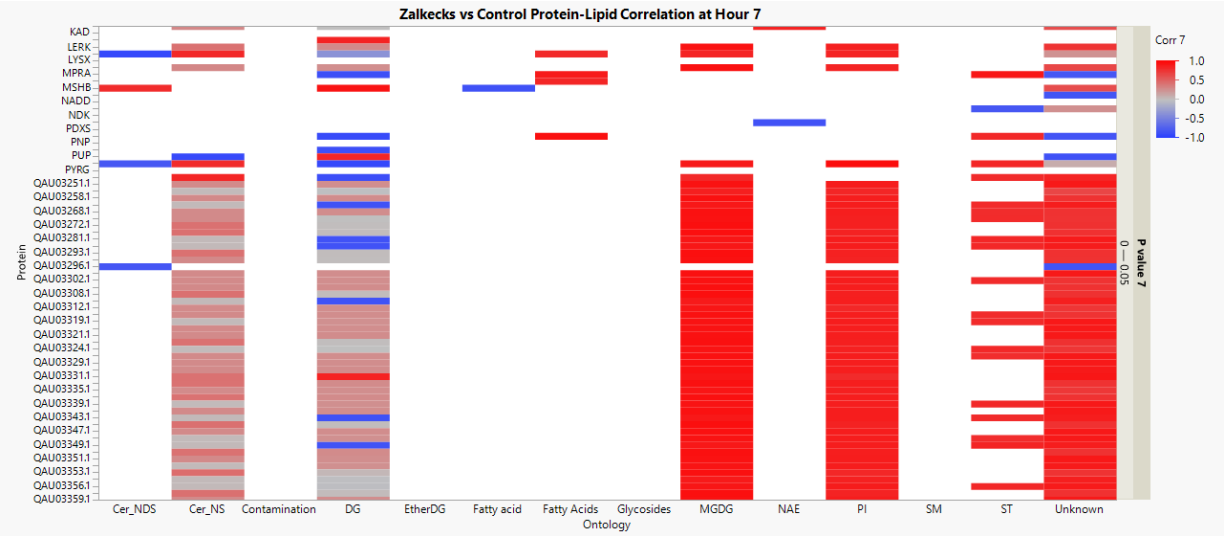
Appendix Figure 34. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



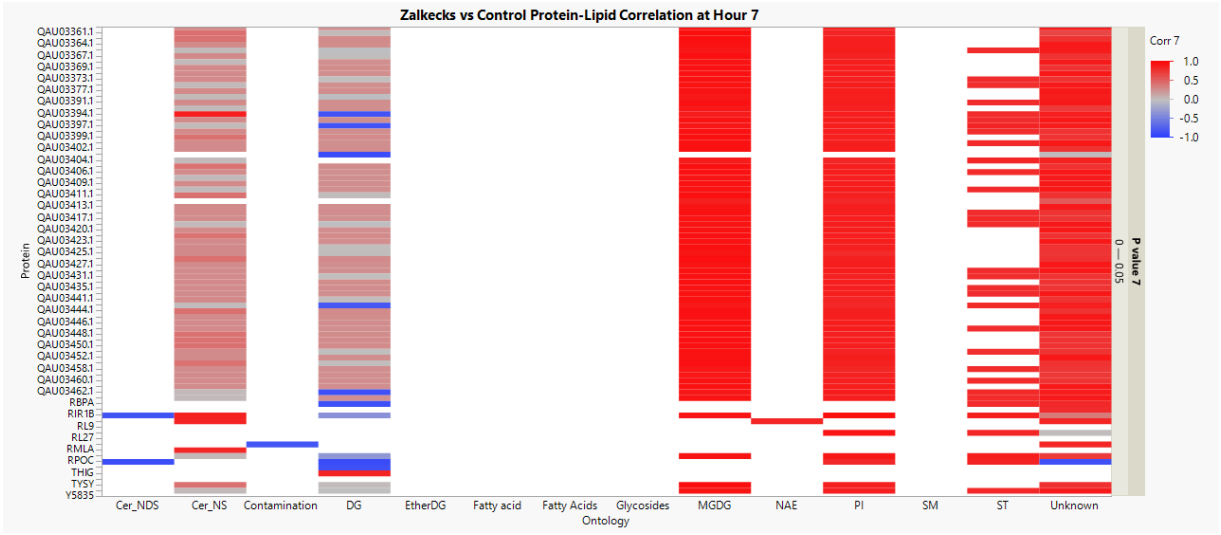
Appendix Figure 35. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



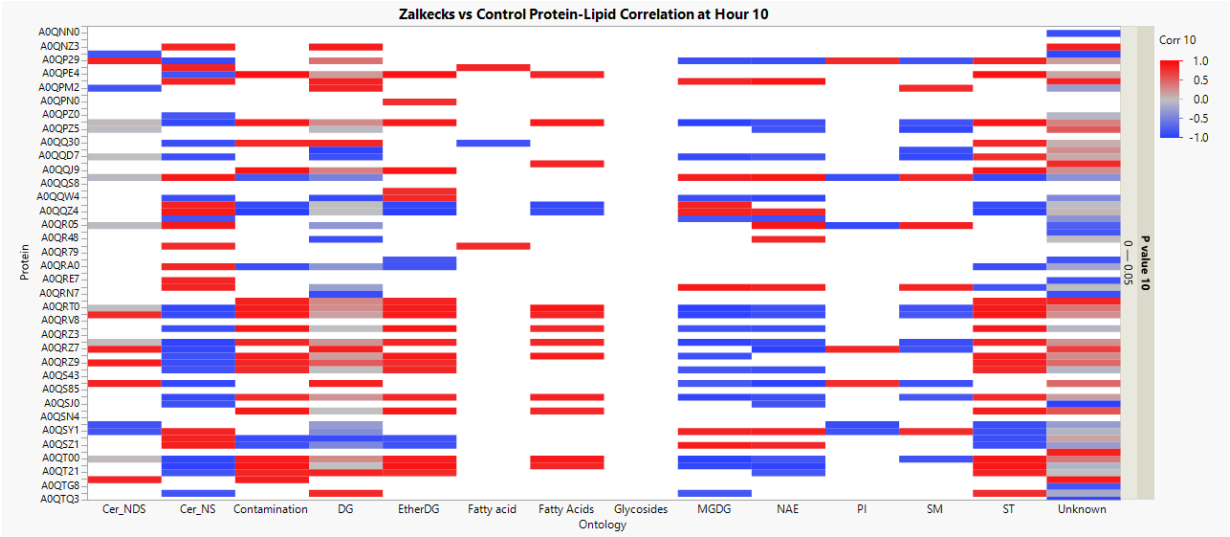
Appendix Figure 36. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



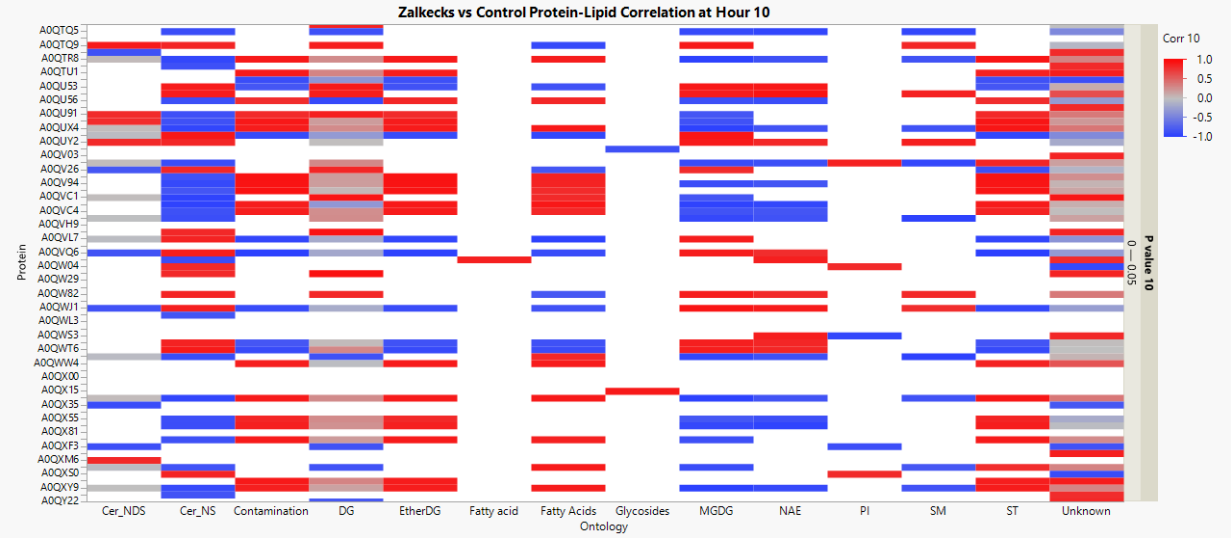
Appendix Figure 37. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



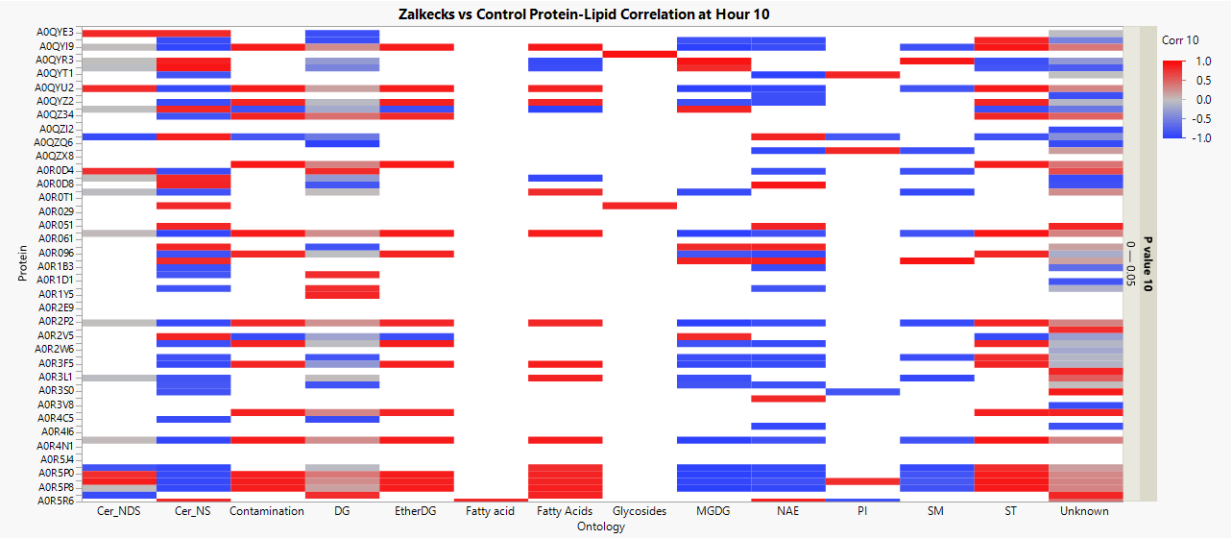
Appendix Figure 38. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 7, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



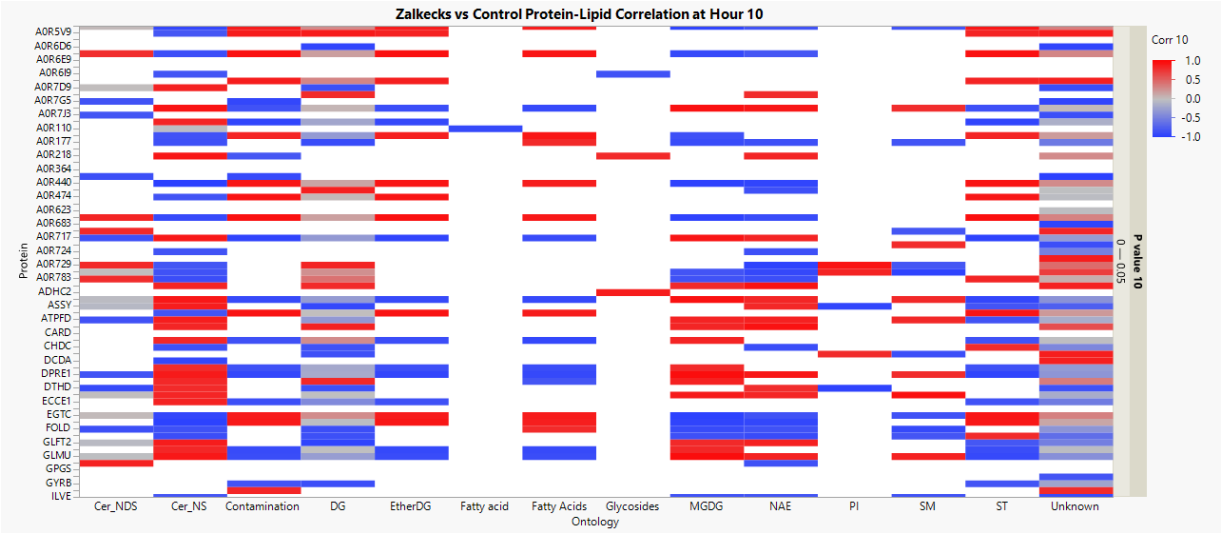
Appendix Figure 39. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



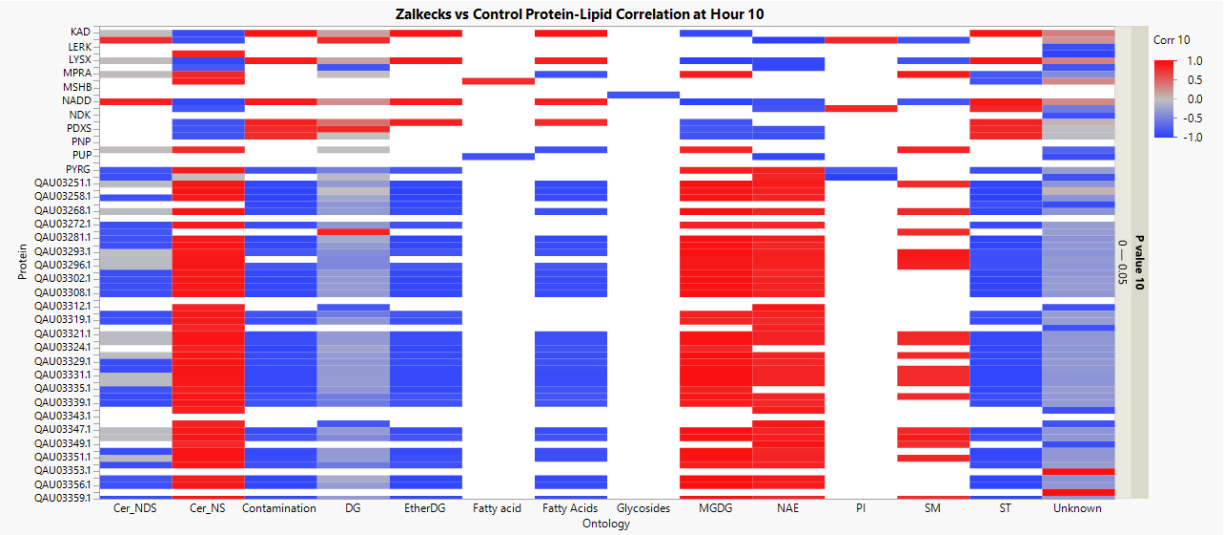
Appendix Figure 40. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



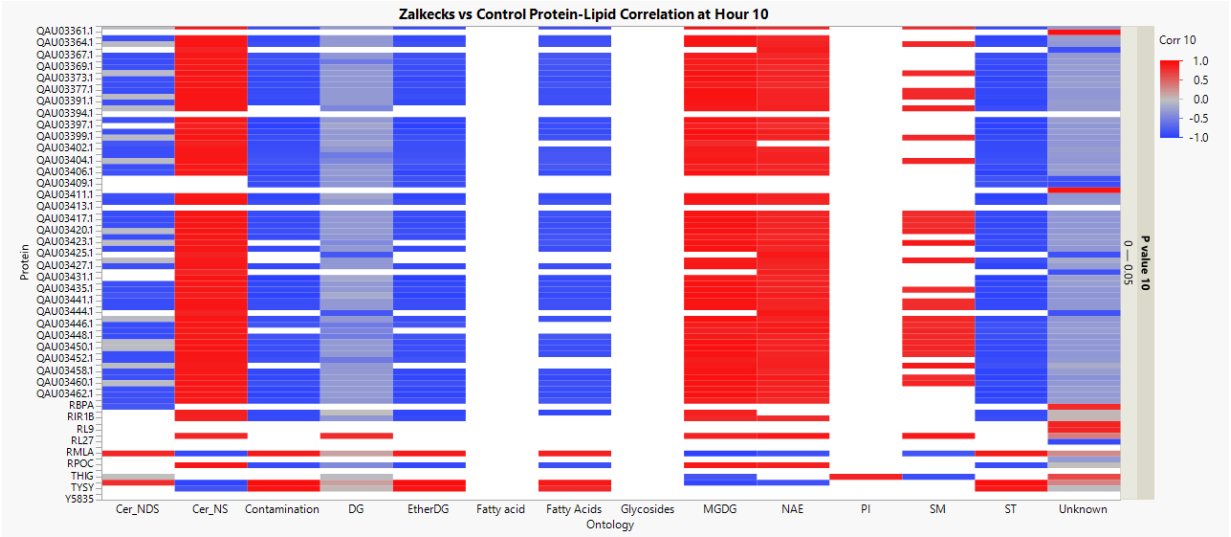
Appendix Figure 41. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



Appendix Figure 42. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



Appendix Figure 43. Heatmap of the correlation of all proteins and all lipids, that have a p-value below 0.05 at Hour 10, for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.



Appendix Figure 44. Heatmap of the correlation of all proteins and lipids, that have a p-value below 0.05 at Hour 10 for Zalkecks vs Control samples. Red shows a positive correlation and blue shows a negative correlation between the groups.

Zalkecks Compared Against PotatoSplit Data

Appendix Table 5. Zalkecks MS2 acquired protein results from a Metaboanalyst linear model with covariate adjustments, when compared against PotatoSplit.

	ZALKECKS VS POTATOSPLIT PROTEIN	LOG(FC)	P.VALUE
COMPARISON TO POTATOSPLIT	QAU03332.1	1.4744	0.00011402
	AZS12919.1	-1.466	0.00012473
	AZS12912.1	-1.4463	0.00015369
	QAU03331.1	1.4049	0.00023622
	QAU03330.1	1.3952	0.00026096
	QAU03293.1	1.3613	0.00036702
	QAU03427.1	1.3392	0.00045704
	AZS12921.1	-1.3321	0.00048995
	QAU03299.1	1.3307	0.00049668
	QAU03398.1	1.3284	0.00050798
	QAU03435.1	1.3276	0.00051172
	QAU03414.1	1.327	0.00051497
	QAU03461.1	1.3267	0.00051632
	QAU03399.1	1.3261	0.00051927
	QAU03424.1	1.3256	0.00052191
	QAU03411.1	1.3235	0.000533
	QAU03268.1	1.3231	0.00053501
	QAU03417.1	1.3222	0.00053929
	QAU03307.1	1.322	0.00054055
	QAU03420.1	1.3208	0.00054679
	AZS12901.1	-1.3202	0.00055012
	QAU03462.1	1.3192	0.00055566
	QAU03419.1	1.3181	0.00056141
	QAU03422.1	1.3174	0.00056549
	QAU03294.1	1.3165	0.00057024
	QAU03290.1	1.3162	0.00057195
	QAU03308.1	1.3156	0.00057503
	QAU03350.1	1.3135	0.00058672
	QAU03423.1	1.312	0.00059542
	QAU03432.1	1.3055	0.00063405
	QAU03302.1	1.3012	0.00066058
	QAU03373.1	1.3	0.00066807
	QAU03351.1	1.2997	0.00067038
	QAU03412.1	1.2989	0.00067552
	QAU03258.1	1.2936	0.00071065
	QAU03391.1	1.2929	0.00071515
	QAU03395.1	1.2893	0.00074025

Appendix Table 5. Continued.

COMPARISON TO POTATOSPLIT	QAU03321.1	1.2867	0.00075886
	QAU03363.1	1.2836	0.00078104
	QAU03460.1	1.2835	0.00078179
	AZS12950.1	-1.2821	0.0007923
	QAU03377.1	1.279	0.00081606
	QAU03402.1	1.2766	0.00083448
	AZS12954.1	-1.2744	0.00085187
	QAU03359.1	1.2727	0.00086595
	QAU03393.1	1.2694	0.00089345
	QAU03364.1	1.2658	0.00092351
	QAU03329.1	1.2636	0.00094323
	QAU03445.1	1.2542	0.0010298
	QAU03439.1	1.2335	0.0012461
	QAU03459.1	1.2313	0.0012706
	QAU03326.1	1.2205	0.0014021
	QAU03272.1	1.2204	0.001403
	QAU03404.1	1.2188	0.0014235
	QAU03251.1	1.2116	0.0015194
	QAU03450.1	1.2103	0.001537
	AZS12909.1	-1.2093	0.001551
	QAU03347.1	1.2061	0.0015969
	QAU03458.1	1.2057	0.0016028
	QAU03426.1	1.204	0.0016267
	QAU03406.1	1.192	0.0018113
	AZS12905.1	-1.1911	0.0018254
	QAU03323.1	1.191	0.0018273
	QAU03360.1	1.1895	0.0018507
	QAU03449.1	1.1892	0.0018571
	QAU03372.1	1.1826	0.0019686
	QAU03446.1	1.1757	0.0020911
	QAU03401.1	1.1591	0.0024167
	QAU03338.1	1.1477	0.0026663
	A0QNT6	-1.1452	0.0027254
	QAU03281.1	1.1197	0.0033857
	QAU03441.1	1.1144	0.00354
	A0QT98	-1.1117	0.0036217
	QAU03387.1	1.1106	0.0036553
	AZS12907.1	-1.1104	0.0036614
	AZS12964.1	-1.1094	0.0036919
	AZS12906.1	-1.1091	0.0036994
	QAU03354.1	1.1088	0.0037079
	QAU03369.1	1.106	0.0037962
	QAU03339.1	1.1059	0.0037989

Appendix Table 5. Continued.

COMPARISON TO POTATOSPLIT	QAU03405.1	1.1052	0.0038208
	QAU03409.1	1.1045	0.0038453
	AZS12961.1	-1.1044	0.0038474
	QAU03431.1	1.1041	0.0038578
	QAU03451.1	1.1029	0.003897
	QAU03442.1	1.1007	0.0039687
	QAU03457.1	1.0995	0.0040079
	QAU03319.1	1.0991	0.0040198
	QAU03317.1	1.0984	0.0040461
	AZS12923.1	-1.0981	0.0040556
	QAU03356.1	1.0971	0.004089
	AZS12928.1	-1.097	0.0040915
	AZS12958.1	-1.0969	0.0040937
	QAU03376.1	1.0968	0.0040986
	AZS12924.1	-1.0966	0.0041045
	QAU03352.1	1.0959	0.0041286
	QAU03448.1	1.0949	0.0041634
	QAU03397.1	1.0947	0.0041693
	QAU03348.1	1.0919	0.0042681
	QAU03367.1	1.0901	0.0043298
	AZS12920.1	-1.0879	0.0044123
	AZS12953.1	-1.0875	0.0044265
	QAU03452.1	1.0874	0.0044288
	QAU03368.1	1.0854	0.0045026
	QAU03324.1	1.0834	0.0045748
	QAU03463.1	1.0802	0.0046964
	QAU03342.1	1.0802	0.004697
	QAU03447.1	1.0764	0.004846
	A0QQ23	1.0704	0.0050886
	AZS12955.1	-1.0694	0.0051309
	QAU03269.1	1.0681	0.0051824
	QAU03335.1	1.0553	0.0057447
	AZS12922.1	-1.0389	0.0065505
	QAU03257.1	1.0376	0.006616
	A0QRD2	-1.0184	0.0076942
	A0QX91	-1.0159	0.0078434
	AZS12944.1	-0.99722	0.0090573
	QAU03444.1	0.99659	0.0091008
	A0R0D8	0.99643	0.0091119
	AZS12948.1	-0.99615	0.0091318
	AZS12910.1	-0.99167	0.0094493
	QAU03265.1	0.99123	0.0094808
	A0QWF7	0.99096	0.0095003

Appendix Table 5. Continued.

COMPARISON TO POTATOSPLIT	A0R782	0.99076	0.0095147
	QAU03346.1	0.98887	0.0096527
	QAU03349.1	0.98622	0.0098486
	A0QT00	-0.98596	0.0098681
	QAU03429.1	0.9823	0.010145
	QAU03365.1	0.98072	0.010267
	QAU03425.1	0.97879	0.010418
	QAU03278.1	0.97852	0.010439
	QAU03312.1	0.97822	0.010463
	AZS12940.1	-0.97789	0.010488
	QAU03320.1	0.97348	0.010842
	A0QW02	-0.9733	0.010857
	QAU03403.1	0.97052	0.011085
	A0R723	0.96984	0.011141
	A0R2T2	-0.96872	0.011235
	QAU03296.1	0.9664	0.011431
	QAU03410.1	0.9636	0.011672
	QAU03408.1	0.95721	0.012239
	AZS12927.1	-0.95421	0.012514
	A0QPM2	0.94931	0.012973
	A0QTF6	-0.94789	0.013109
	A0R6A6	-0.9438	0.013508
	RS15	0.94127	0.013761
	A0R2B7	-0.93238	0.014679
	A0QXM6	-0.92782	0.015171
	A0QPE5	-0.92255	0.015758
	A0QRB9	-0.92243	0.015772
	A0R0X1	0.91322	0.016846
	A0QNQ9	-0.91153	0.01705
	A0R717	0.90017	0.018477
	A0R5Y9	-0.8989	0.018643
	A0R5I4	-0.89765	0.018809
	A0QUF4	-0.89475	0.019195
	A0QSL0	-0.8854	0.02049
	A0R5G6	0.88534	0.020499
	A0QVD5	-0.884	0.020691
	RL11	0.88373	0.020729
	A0QZE4	-0.87641	0.021806
	AZS12973.1	-0.87206	0.022469
	A0QT68	-0.87062	0.022692
	QAU03358.1	0.87017	0.022763
	A0QU53	-0.87007	0.022778

Appendix Table 5. Continued.

COMPARISON TO POTATOSPLIT	A0QWP9	-0.86879	0.022979
	AZS12918.1	-0.86865	0.023001
	A0QV19	-0.86829	0.023058
	A0R3A4	0.8648	0.023614
	A0QSK8	0.86444	0.023673
	AZS12935.1	-0.86364	0.023803
	QAU03353.1	0.86355	0.023818
	A0R401	-0.863	0.023906
	A0QQ02	-0.86262	0.023969
	QAU03454.1	0.86168	0.024123
	QAU03361.1	0.8613	0.024184
	A0QZ08	0.8611	0.024218
	A0QVP0	-0.86022	0.024363
	A0QRH0	0.85833	0.024678
	A0QVC1	-0.85619	0.025039
	QAU03274.1	0.85432	0.025358
	A0QW21	-0.8537	0.025464
	A0R3V8	-0.85275	0.025628
	RS14Z	0.8485	0.026373
	A0R267	-0.84828	0.026411
	AZS12917.1	-0.84665	0.026703
	A0R4V8	-0.84493	0.027012
	A0R555	-0.84448	0.027094
	A0R7D8	-0.84442	0.027104
	A0R5G2	-0.84421	0.027142
	QAU03413.1	0.84265	0.027429
	A0R5B1	-0.84234	0.027485
	A0QQF5	-0.83832	0.028232
	A0QQA3	-0.8346	0.02894
	A0QTF1	0.83337	0.029176
	A0R1A7	-0.83033	0.029771
	SAHH	-0.82815	0.030202
	A0QQJ9	-0.82813	0.030206
	A0QPZ4	-0.82761	0.030311
	A0R2S4	-0.82587	0.03066
	A0QPN0	-0.8253	0.030774
	QAU03343.1	0.82262	0.031322
	A0QUW5	-0.8079	0.034479
	CH10	0.80727	0.03462
	A0QPZ0	-0.80468	0.035204
	A0QQY9	-0.80214	0.035785
	ATPG	0.80185	0.035853

Appendix Table 5. Continued.

COMPARISON TO POTATOSPLIT	A0R6M1	-0.79849	0.036635
	Y6286	0.79808	0.036733
	A0R1A9	-0.79029	0.038608
	RSMH	-0.78907	0.038908
	A0R1D1	-0.78398	0.040187
	A0QPY2	-0.78389	0.040209
	A0QY09	0.78346	0.040318
	DDL	-0.78052	0.041074
	A0R292	0.77689	0.042024
	A0QRU5	-0.77528	0.042452
	A0R4Z0	-0.7752	0.042472
	A0QWS4	0.77154	0.04346
	A0R5Y1	0.76794	0.044447
	SYDND	-0.76666	0.044804
	A0QR49	0.76643	0.044868
	A0QWN0	0.7639	0.04558
	MNMA	0.76155	0.046247
	A0R5M3	0.76139	0.046293
	A0R720	-0.76113	0.046367
	A0QWK5	0.75706	0.047549
	A0R5W6	-0.756	0.04786
	A0QQK8	-0.75509	0.048129
	MSHC	-0.75157	0.049181
	A0QS21	-0.75002	0.049651
	Y5073	0.7497	0.04975

Appendix Table 6. Zalkecks ANOVA Simultaneous Component Analysis (ASCA) significant lipid results when compared against PotatoSplit to model phenotype and time effects and their interaction.

	PROTEIN ID	LEVERAGE	SPE	PROTEIN ID	LEVERAGE	SPE
PHENOTYPE	AZS12912.1	0.003841	0	A0R0D8	0.001823	0
	QAU03331.1	0.003625	0	QAU03265.1	0.001804	0
	QAU03293.1	0.003403	0	A0QWF7	0.001803	3.70E-32
	QAU03435.1	0.003237	0	A0R782	0.001803	3.70E-32
	QAU03461.1	0.003232	0	QAU03349.1	0.001786	3.70E-32
	QAU03399.1	0.00323	0	A0QT00	0.001785	3.70E-32
	QAU03424.1	0.003227	0	QAU03429.1	0.001772	3.70E-32
	QAU03417.1	0.003211	0	QAU03365.1	0.001766	0
	QAU03420.1	0.003204	0	QAU03278.1	0.001758	3.70E-32
	QAU03462.1	0.003196	0	QAU03312.1	0.001757	3.70E-32
	QAU03419.1	0.003191	0	AZS12940.1	0.001756	3.70E-32
	QAU03422.1	0.003187	0	QAU03320.1	0.00174	0
	QAU03294.1	0.003183	0	A0QW02	0.00174	0
	QAU03308.1	0.003179	0	QAU03403.1	0.00173	0
	QAU03350.1	0.003169	0	QAU03296.1	0.001715	7.40E-32
	QAU03432.1	0.00313	0	QAU03410.1	0.001705	7.40E-32
	QAU03302.1	0.003109	0	QAU03408.1	0.001683	0
	QAU03391.1	0.00307	0	AZS12927.1	0.001672	7.40E-32
	QAU03395.1	0.003053	0	A0R6A6	0.001636	3.70E-32
	QAU03321.1	0.00304	0	RS15	0.001627	0
	QAU03363.1	0.003026	0	A0R2B7	0.001597	7.40E-32
	AZS12950.1	0.003019	0	A0QXM6	0.001581	3.70E-32
	QAU03402.1	0.002993	0	A0QPE5	0.001563	0
	AZS12954.1	0.002983	0	A0QRB9	0.001563	7.40E-32
	QAU03459.1	0.002784	0	A0QNQ9	0.001526	0
	QAU03326.1	0.002736	0	A0R717	0.001488	7.40E-32
	QAU03251.1	0.002696	0	A0R5Y9	0.001484	3.70E-32
	AZS12909.1	0.002686	0	A0R5I4	0.00148	7.40E-32
	QAU03347.1	0.002671	0	A0QSL0	0.00144	3.70E-32
	QAU03458.1	0.00267	0	A0R5G6	0.001439	7.40E-32
	QAU03426.1	0.002662	0	A0QVD5	0.001435	7.40E-32
	AZS12905.1	0.002605	0	RL11	0.001434	0
	QAU03323.1	0.002605	0	AZS12973.1	0.001397	3.70E-32
	QAU03360.1	0.002599	0	A0QT68	0.001392	3.70E-32
	QAU03372.1	0.002568	0	QAU03358.1	0.001391	3.70E-32
	QAU03401.1	0.002467	0	A0QU53_2	0.00139	3.70E-32
	QAU03281.1	0.002302	0	A0QWP9	0.001386	7.40E-32
	AZS12907.1	0.002264	0	AZS12918.1	0.001386	0
	AZS12964.1	0.00226	0	A0R3A4	0.001373	7.40E-32

Appendix Table 6. Continued.

PHENOTYPE	QAU03369.1	0.002247	0	A0QSK8	0.001372	7.40E-32
	QAU03339.1	0.002246	0	AZS12935.1	0.00137	3.70E-32
	QAU03409.1	0.00224	0	QAU03353.1	0.001369	7.40E-32
	QAU03451.1	0.002234	0	A0R401	0.001368	0
	QAU03319.1	0.002219	0	A0QQ02	0.001367	7.40E-32
	QAU03317.1	0.002215	0	QAU03454.1	0.001364	0
	QAU03356.1	0.00221	0	QAU03361.1	0.001362	0
	AZS12928.1	0.00221	0	A0QZ08	0.001362	3.70E-32
	AZS12958.1	0.00221	0	A0QVP0	0.001359	0
	AZS12924.1	0.002208	0	A0QVC1	0.001346	0
	QAU03352.1	0.002206	0	QAU03274.1	0.00134	7.40E-32
	QAU03348.1	0.002189	0	A0QW21	0.001338	7.40E-32
	AZS12920.1	0.002173	0	A0R3V8	0.001335	3.70E-32
	QAU03452.1	0.002172	0	RS14Z	0.001322	0
	QAU03368.1	0.002163	0	A0R267	0.001322	3.70E-32
	QAU03463.1	0.002143	0	AZS12917.1	0.001316	7.40E-32
	QAU03342.1	0.002143	0	A0R4V8	0.001311	7.40E-32
	QAU03447.1	0.002128	0	A0R7D8	0.00131	3.70E-32
	AZS12955.1	0.0021	0	A0R5G2	0.001309	7.40E-32
	AZS12922.1	0.001982	0	QAU03413.1	0.001304	0
	A0QRD2	0.001905	0	A0R5B1	0.001303	3.70E-32
	QAU03444.1	0.001824	3.70E-32			
TIME	A0QSW8	0.002541	0.32816	MYCP1	0.001461	1.5328
	A0R3H8	0.002467	0.62043	A0QR03	0.001459	1.7965
	A0QYJ2	0.002058	0.4865	A0QPD6	0.001459	1.9396
	A0R0M4	0.00199	0.69067	A0QTT6	0.001458	0.71659
	A0QUW4	0.001937	0.26215	A0QSA9	0.001453	5.4285
	A0QTT5	0.001859	0.16453	A0QNR5	0.00143	1.4483
	A0QTP1	0.001842	0.49345	CBXPD	0.001426	0.3218
	AFTC	0.001816	0.35263	A0R588	0.001419	0.26526
	A0R2S9	0.001809	1.1975	A0QNR6	0.001417	0.25527
	A0R240	0.001808	1.8459	A0R4L2	0.001412	5.0352
	A0R4K5	0.001795	1.2	A0QQ51	0.001405	0.80646
	A0QSY9	0.001794	0.64615	MFS55	0.001404	2.5646
	A0R430	0.001778	0.30874	ECCB1	0.001394	0.79474
	A0QU91	0.001746	0.38583	A0R3C9	0.001391	2.8128
	ECCE1	0.001738	0.77504	MUTT1	0.001385	1.3828
	A0R2J0	0.001712	0.11599	A0QPX4	0.001385	1.3709
	A0QZ58	0.001671	0.25436	DGTL1	0.001384	0.23946
	A0R221	0.00166	2.6951	A0R5Q0	0.001383	0.27888
	SECA2	0.001656	1.2054	A0R529	0.001383	0.30994
	ACYLT	0.001648	0.020037	HTPX	0.001382	0.22752

Appendix Table 6. Continued.

TIME	ATPFD	0.001648	0.19221	PSD	0.001378	0.17414
	A0R7H4	0.001641	2.0289	A0QR14	0.001373	1.6731
	A0R6P9	0.001628	0.83951	A0QT70	0.00137	0.51363
	A0QVU5	0.001599	0.73312	A0QR13	0.001368	0.36661
	A0R4N7	0.001569	0.93681	A0R7J0	0.001368	2.5812
	A0QTL1	0.001565	1.1643	A0R069	0.001367	3.7915
	A0QS42	0.001552	0.13607	A0QSZ4	0.001359	0.46115
	A0R5H3	0.001543	0.079428	A0R448	0.001358	3.081
	TRES	0.001532	0.86276	A0R2I2	0.001349	0.56186
	THTR	0.001524	0.58902	A0R020	0.001345	0.68747
	CON__P17690	0.001523	5.2477	GLGE	0.001338	0.041854
	A0QTQ9	0.001522	0.76591	A0QUI9	0.001336	0.044292
	A0QQ64	0.001518	0.87756	A0R2P2	0.001332	4.5261
	HOA2	0.001517	0.30284	A0QU11	0.001332	0.46818
	ATPG	0.001501	1.0759	Y1603	0.001327	0.04379
	A0R2Q0	0.001484	1.5474	A0QXA1	0.001326	0.63603
	Y4692	0.001479	0.50282	A0R4N4	0.001324	0.63595
	A0QX90	0.001469	4.5284	A0R1H2	0.001322	2.5579
	A0QQG1	0.001469	0.18851	A0QQZ2	0.001316	0.41891
	A0R436	0.001468	0.33069	A0QT19	0.001314	3.2449
INTERACTION	A0QRZ4	0.004082	0.61181	QAU03442.1	0.002257	0.3595
	A0QV26	0.003772	0.15378	A0QPW1	0.002252	0.032765
	Y5073	0.003595	0.39872	QAU03449.1	0.002245	0.13196
	GLMU	0.003539	0.24511	QAU03281.1	0.00224	0.65465
	A0QZX4	0.003406	1.6228	QAU03320.1	0.002235	0.24755
	A0QQF0	0.003401	0.3248	QAU03444.1	0.002231	0.080344
	A0R595	0.003322	9.58E-07	A0QV98	0.002223	0.10456
	A0QQI4	0.003112	0.20106	QAU03360.1	0.002221	0.19379
	A0R407	0.003095	0.025446	QAU03365.1	0.002219	0.16113
	A0R2C3	0.003084	0.59216	QAU03349.1	0.002213	0.11542
	MSHC	0.003071	0.24396	QAU03323.1	0.002212	0.18604
	RS11	0.003069	0.042971	A0QWK9	0.002209	1.952
	A0QTK6	0.003054	0.79458	A0QW13	0.002198	0.078601
	A0QQS8	0.003052	0.001662	A0QXD2	0.002191	0.078297
	TOPON	0.003015	1.5895	QAU03406.1	0.002191	0.018907
	A0QS01	0.003005	0.04836	A0QSX3	0.00219	0.34465
	A0QYC5	0.002997	0.29078	A0QU18	0.002187	0.62303
	A0QS43	0.002996	0.29372	QAU03446.1	0.002182	0.013096
	A0QVZ5	0.002979	0.045158	A0QSH0	0.002174	0.23267
	A0R4S3	0.002927	0.055884	AZS12958.1	0.002173	0.55222
	A0QPZ1	0.002899	0.31746	QAU03265.1	0.002173	0.044688
	A0QYH0;tr	0.002896	0.60524	AZS12964.1	0.002173	0.73452

Appendix Table 6. Continued.

INTERACTION	A0QRB4	0.002823	0.047561	QAU03347.1	0.002172	0.1128
	A0QTY3	0.002808	0.054386	QAU03324.1	0.002171	0.38952
	A0R5H5	0.002795	0.013244	AZS12906.1	0.002171	0.73924
	A0QS52	0.002681	0.90204	A0QZR9	0.002169	0.26247
	A0QWJ6	0.002646	0.000519	A0QY29	0.002164	0.66994
	KGD	0.002643	0.45984	A0R2P7	0.002163	0.26747
	DDL	0.002634	0.11933	AZS12961.1	0.002162	0.6963
	A0R365	0.002629	0.14021	A0QZX9	0.002147	1.8963
	A0QTQ5	0.002629	0.003792	A0QQB4	0.002142	0.051599
	A0QT14	0.002617	0.3618	QAU03450.1	0.002136	0.15573
	A0R783	0.002554	2.2755	QAU03312.1	0.002135	0.061527
	A0R4I6	0.002542	0.58883	A0R0B7	0.002127	0.93552
	QAU03348.1	0.002535	0.001459	A0QQX8	0.002125	0.59412
	A0QTL8	0.002524	0.15742	A0R0A8	0.002123	0.018558
	QAU03447.1	0.002524	0.023972	QAU03251.1	0.002122	0.011598
	QAU03356.1	0.00252	0.001371	A0R2V8	0.002121	0.012035
	A0R614	0.002516	5.05E-05	A0R5W6	0.00212	0.04808
	QAU03448.1	0.002515	0.000799	BFRB	0.002116	0.12408
	QAU03452.1	0.002507	0.000524	AZS12924.1	0.002115	0.74854
	A0R2M2	0.002505	0.30966	QAU03401.1	0.002113	0.005151
	QAU03335.1	0.002502	0.090009	QAU03361.1	0.002109	1.5051
	QAU03368.1	0.002497	0.000508	QAU03445.1	0.002098	0.033267
	QAU03463.1	0.002487	0.002316	AZS12928.1	0.002091	0.84646
	QAU03376.1	0.002483	0.011547	A0R0D8	0.002089	0.000165
	QAU03451.1	0.002482	0.026859	QAU03425.1	0.002087	0.019873
	A0R4C3	0.002478	0.29962	RL10	0.002084	0.50036
	QAU03367.1	0.002472	0.004731	3O1D	0.002083	0.12001
	A0QPL0	0.002471	0.57804	A0QWS5	0.002081	0.001942
	A0R716	0.002454	0.13198	A0QW30	0.002078	0.23837
	A0R1Y2	0.002432	1.0825	AZS12923.1	0.002076	0.92911
	QAU03369.1	0.002431	0.086081	A0QNY8	0.002073	0.39409
	QAU03352.1	0.002425	0.046008	TOPOM	0.002067	0.73034
	A0R0I3	0.002419	0.47741	A0QS36	0.002067	0.000862
	QAU03457.1	0.00241	0.076416	TSAD	0.002048	0.022892
	A0QNZ8	0.002401	0.10967	QAU03353.1	0.002047	1.161
	A0QTR2	0.002398	0.097654	A0R7J4	0.002047	0.040714
	A0QXD0	0.002396	0.62252	QAU03329.1	0.002035	0.002499
	A0R2V1	0.002395	2.1626	A0QPG3	0.002033	0.015525
	QAU03405.1	0.002392	0.13279	QAU03364.1	0.002029	0.041932
	QAU03387.1	0.00239	0.17352	A0R477	0.002026	0.30955
	QAU03339.1	0.002377	0.16095	A0QRU7	0.002023	0.4023
	GYRB	0.002375	0.32005	QAU03359.1	0.002021	0.057124
	A0QV23	0.002373	2.1508	A0QSL0	0.002018	0.078575

Appendix Table 6. Continued.

INTERACTION	A0QXI7	0.002363	1.4312	A0QYF6	0.002006	0.2588
	QAU03319.1	0.002361	0.13812	GLMM	0.002001	0.000238
	QAU03317.1	0.002357	0.13932	QAU03358.1	0.001997	0.83531
	A0QNQ9	0.002355	0.40949	AZS12905.1	0.001995	0.78158
	QAU03257.1	0.00235	0.025284	A0QYR2	0.00199	0.1441
	DOP	0.002343	0.034777	ARGD	0.001989	0.14988
	AZS12907.1	0.002325	0.29937	QAU03326.1	0.001987	0.46403
	QAU03354.1	0.002325	0.28536	QAU03377.1	0.001986	0.031799
	QAU03408.1	0.002323	0.67152	A0R0Z9	0.001983	0.00088
	A0R2A4	0.002319	0.027197	QAU03429.1	0.001973	0.018742
	A0QWU9	0.002318	0.10438	QAU03426.1	0.001972	0.006199
	QAU03441.1	0.002313	0.37071	A0QRN8	0.001971	0.12649
	A0QT21	0.002313	1.6621	A0QYT3	0.001967	0.10432
	QAU03338.1	0.002307	0.094242	A0QNE2	0.00196	0.047559
	A0R1Y8	0.002306	0.65478	QAU03321.1	0.001954	0.013098
	A0QSB1	0.002279	0.11203	A0QT77	0.001954	1.2141
	QAU03431.1	0.002275	0.35242	QAU03363.1	0.001952	0.074
	QAU03397.1	0.002274	0.26029	A0QRD4	0.001951	0.38451
	A0QQF9	0.002274	0.13608	A0QYB0	0.00194	0.005948
	QAU03372.1	0.002272	0.03139	HUTU	0.00194	0.40357
	QAU03346.1	0.002266	0.18002			

Appendix Table 7. All lipids found in both bacteriophages Zalkecks and PotatoSplit – both significant and unsignificant results from raw MS data

	LIPID NAMES	LIPID ONTOLOGY GROUP
SHARED BETWEEN ZALKECKS AND POTATOSPLIT	C12H11N6O17PS8	Cer_NS
	Cer 57:4;2O	Cer_NS
	C24H38N16O	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	1,2-DG(35:0)	DG
	UNPD60739	DG
	DG 39:5	DG
	DG 39:5	DG
	DG 39:5	DG
	DG 39:5	DG
	DG 39:5	DG
	DG 39:5	DG
	DG 39:5	DG
	DG 39:5	DG
	DG 39:5	DG
	DG 41:11	DG
	DG 41:11	DG
	DG 41:11	DG
	DG 41:11	DG
	DG 41:11	DG

Appendix Table 7. Continued.

SHARED BETWEEN ZALKECKS AND POTATOSPLIT	DG 64:17	DG
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl)-3,8-dimethyl-4-[2-(5-oxo-2,5-dihydro furan-3-yl)ethyl]-decahydronaphthalen -1-yl acetate	EtherMGDG
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl)-3,8-dimethyl-4-[2-(5-oxo-2,5-dihydro furan-3-yl)ethyl]-decahydronaphthalen-1-yl acetate	EtherMGDG
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl)-3,8-dimethyl-4-[2-(5-oxo-2,5-dihydro furan-3-yl)ethyl]-decahydronaphthalen-1-yl acetate	EtherMGDG
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl)-3,8-dimethyl-4-[2-(5-oxo-2,5-dihydro furan-3-yl)ethyl]-decahydronaphthalen-1-yl acetate	EtherMGDG
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl)-3,8-dimethyl-4-[2-(5-oxo-2,5-dihydro furan-3-yl)ethyl]-decahydronaphthalen-1-yl acetate	EtherMGDG
	8-cyclohexyl-4,5-dihydroxy-4a-(1-hydroxy-4-methyl-7-phenylheptyl)-3,8-dimethyl-4-[2-(5-oxo-2,5-dihydro furan-3-yl)ethyl]-decahydronaphthalen-1-yl acetate	EtherMGDG
	Pipericine	fatty amide
	Pipericine	fatty amide
	Pipericine	fatty amide
	triciribine	Glycosides
	2,4,12-Octadecatrienoic acid isobutylamide	NAE
	5-{[(diaminomethylidene)amino]methyl}-6-[3,4-dihydroxy-3a,6,9a,11a-tetramethyl-6-(3-methylbutyl)-7,10-dioxo-1H,2H,3H,3aH,4H, 5H,5aH, 6H,7H,8H, 9H,9aH,10H,11H,11aH-cyclopenta[a] phenanthren-1-yl]-4-hydroxy-2-methylhept-2-enoic acid	PC
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE

Appendix Table 7. Continued.

SHARED BETWEEN ZALKECKS AND POTATOSPLIT	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 32:1|PE 11:0_21:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 34:1|PE 21:0_13:1	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE

Appendix Table 7. Continued.

SHARED BETWEEN ZALKECKS AND POTATOSPLIT	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE 35:0	PE
	PE-Cer(d35:7)	PE-Cer
	Gitogenin	Spirostans
	Gitogenin	Spirostans
	Gitogenin	Spirostans
	Gitogenin	Spirostans
	Gitogenin	Spirostans
	ST 28:1;O	ST
	ST 28:1;O	ST
	ST 28:1;O	ST
	ST 28:1;O	ST
	UNPD180606	Unknown
	2-amino-N1,N9-bis[11-hydroxy-2,5,9-trimethyl-1,4,7,14-tetraoxo-6,13-bis(propan-2-yl)-1H,2H,3H,4H,5H,6H,7H,9H,10H,13H,14H,16H,17H,18H,18aH-pyrrolo[2,1-i]1-oxa-4,7,10,13-tetraazacyclohexadecan-10-yl]-4-methyl-3-oxo-3H-phenoxazine-1,9-dicarboximidic acid	Unknown
	C69H104N4O4S7	Unknown
	C37H78N2O8	Unknown
	C37H78N2O8	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown

Appendix Table 7. Continued.

SHARED BETWEEN ZALKECKS AND POTATOSPLIT	C22H50N29O4PS17	UNKNOWN
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	C22H50N29O4PS17	Unknown
	pyridin-3-ylmethyl N-[(2S,3S,5S)-5-[[(2S)-2-[[(6-ethylpyridin-2-yl)methyl-methylcarbamoyl]amino]-3-methylbutanoyl]amino]-3-hydroxy-1,6-diphenylhexan-2-yl]carbamate	Unknown
	pyridin-3-ylmethyl N-[(2S,3S,5S)-5-[[(2S)-2-[[(6-ethylpyridin-2-yl)methyl-methylcarbamoyl]amino]-3-methylbutanoyl]amino]-3-hydroxy-1,6-diphenylhexan-2-yl]carbamate	Unknown
	C26H70N18O4	Unknown
	C26H70N18O4	Unknown
	C26H70N18O4	Unknown
	C26H70N18O4	Unknown
	C26H70N18O4	Unknown
	C26H70N18O4	Unknown
	C26H70N18O4	Unknown
	C26H70N18O4	Unknown
	C26H70N18O4	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_ Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown

Appendix Table 7. Continued.

SHARED BETWEEN ZALKECKS AND POTATOSPLIT	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	UNKNOWN
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	RIKEN P-VS1 ID-9000 from Mouse_Feces_WT_N_Ctr	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C36H81N5O9S	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown

Appendix Table 7. Continued.

SHARED BETWEEN ZALKECKS AND POTATOSPLIT	C22H68N21O8P	UNKNOWN
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	C22H68N21O8P	Unknown
	RIKEN P-VS1 ID-11326 Mouse_ApoEKO_N_F1EPA	Unknown
	RIKEN P-VS1 ID-11326 Mouse_ApoEKO_N_F1EPA	Unknown
	RIKEN P-VS1 ID-11326 Mouse_ApoEKO_N_F1EPA	Unknown
	C33H83N15S3	Unknown
	C42H71N5O4S6	Unknown
	C34H94N19O9PS2	Unknown
	C38H99N15O10S2	Unknown
	24-{[3-cyclohexyl-5-(4-hydroxyoxan-4-yl) phenyl] methyl}-17-[2-(3,3-dimethyloxiran-2-yl)-2-hydro xyethyl]-2-hydroxy-6-(3-hydroxypropyl)-4,8,12,13,17-pentamethyl-8-[2-(methylamino) ethyl]-22,26-diaza-heptacyclo [14.11.1.1??,??. 0?,??.0?,?.0??,??.0??,??] nonacosa-16(28),20,23 (29),24-tetraene-7,15-dione	Unknown
	C23H28N4OS20	Unknown

Appendix Table 8. All proteins found in both bacteriophages Zalkecks and PotatoSplit – both significant and unsignificant results from raw MS data

	PROTEIN NAMES
SHARED BETWEEN ZALKECKS AND POTATOSPLIT	GLFT2
	A0QR79
	A0QNE2
	A0QTE1
	A0QTU1
	A0QV98
	A0QVC4
	A0QVZ5
	A0QV26
	A0QVC1
	A0QVC3
	A0QVM7
	A0QW20
	A0QWD2
	A0QWS4
	A0R072
	A0QWZ9
	A0QX00
	A0R0F4
	A0QXF3
	A0R2C6
	A0R678
	A0R783
	RL29
	A0R2V8
	A0R364
	A0R3A6
	ILVC
	AFTC
	A0R3F5
	TREH
	A0R3Q2
	A0QZX1
	A0R0D5
	RMLA
	A0R0T1
	RPOC
	A0R2P2
	A0QRZ7
	RIR1B

Appendix Table 8. Continued.

SHARED BETWEEN ZALKECKS AND POTATOSPLIT	A0QRZ8
	A0QRU1
	A0QRZ6
	A0QYE7
	A0QXY9
	RPOB
	A0QSW8
	LEXA
	GCSH
	NADD
	ATPFD
	MSHB
	GLMU
	KATG1
	A0QQX6
	A0R6E6
	A0QXP8
	A0R5I8
	A0R5P8
	A0R5R6
	A0R6D6
	A0R6E9
	A0QND7
	A0R753
	A0R7J4
	A0QPZ0
	A0QPZ5
	A0QQD7
	A0QX35
	A0QYR3
	A0QU53
	A0QW82
	A0QU93
	A0QX19
	A0QXY7
	A0QSZ4
	A0QUX6
	A0QUY2

REFERENCES

6545 Quadrupole Time-of-Flight LC/MS, LC/Q-TOF MS | Agilent. (n.d.). Retrieved February 26, 2022, from https://www.agilent.com/en/product/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-instruments/quadrupole-time-of-flight-lc-ms/6545-lc-q-tof

Abedon, S. T. (2012). Bacterial 'immunity' against bacteriophages. *Bacteriophage*, *2*(1), 50. https://doi.org/10.4161/BACT.18609

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature 2003 422:6928*, *422*(6928), 198–207. https://doi.org/10.1038/nature01511

Ali, J., Rafiq, Q., & Ratcliffe, E. (2019). A scaled-down model for the translation of bacteriophage culture to manufacturing scale. *Biotechnology and Bioengineering*, *116*(5), 972–984. https://doi.org/10.1002/BIT.26911

Barák, I., & Muchová, K. (2013). The Role of Lipid Domains in Bacterial Cell Processes. *International Journal of Molecular Sciences*, *14*(2), 4050. https://doi.org/10.3390/IJMS14024050

Beckmann, N., & Becker, K. A. (2021). Ceramide and Related Molecules in Viral Infections. *International Journal of Molecular Sciences*, *22*(11). https://doi.org/10.3390/IJMS22115676

Bertin InstrumentsPrecellys Evolution tissue homogenizer, the most advanced homogenizer | Bertin Instruments. (n.d.). Retrieved January 31, 2022, from https://www.bertin-instruments.com/product/sample-preparation-homogenizers/precellys-evolution-homogenizer/?gclid=Cj0KCQiA0eOPBhCGARIsAFIwTs7i5HNhkEUY4M5ovhUXvteHBKKd-b_k78TcQ06NZ3IqS-aR0T6jkU8aAnWyEALw_wcB

Bertozzi Silva, J., Storms, Z., & Sauvageau, D. (2016). Host receptors for bacteriophage adsorption. *FEMS Microbiology Letters*, *363*(4), 2. https://doi.org/10.1093/FEMSLE/FNW002

Boundless. (2020, August 15). *21.2B: The Lytic and Lysogenic Cycles of Bacteriophages - Biology LibreTexts*. https://bio.libretexts.org/Bookshelves/Introductory_and_General_Biology/Book%3A_General_Biology_(Boundless)/21%3A_Viruses/21.2%3A_Virus_Infections_and_Hosts/21.2B%3A_The_Lytic_and_Lysogenic_Cycles_of_Bacteriophages

Chung, I. Y., Jang, H. J., Bae, H. W., & Cho, Y. H. (2014). A phage protein that inhibits the bacterial ATPase required for type IV pilus assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(31), 11503–11508. https://doi.org/10.1073/PNAS.1403537111/SUPPL_FILE/PNAS.201403537SI.PDF

Cronan, J. E., & Thomas, J. (2009). Bacterial Fatty Acid Synthesis and its Relationships with Polyketide Synthetic Pathways. *Methods in Enzymology*, *459*(B), 395. https://doi.org/10.1016/S0076-6879(09)04617-5

DAVID Functional Annotation Bioinformatics Microarray Analysis. (n.d.). Retrieved February 26, 2022, from https://david.ncifcrf.gov/

Desbois, A. P., & Smith, V. J. (2010). Antibacterial free fatty acids: activities, mechanisms of action and biotechnological potential. *Applied Microbiology and Biotechnology*, *85*(6), 1629–1642. https://doi.org/10.1007/S00253-009-2355-3

DJ, S., LA, H., & CE, C. (2017). Lipidomics in translational research and the clinical significance of lipid-based biomarkers. *Translational Research : The Journal of Laboratory and Clinical Medicine*, *189*, 13–29. https://doi.org/10.1016/J.TRSL.2017.06.006

Etienne, G., Laval, F., Villeneuve, C., Dinadayala, P., Abouwarda, A., Zerbib, D., Galamba, A., & Daffé, M. (2005). The cell envelope structure and properties of Mycobacterium smegmatis mc2155: is there a clue for the unique transformability of the strain? *Microbiology*, *151*(6), 2075–2086. https://doi.org/10.1099/MIC.0.27869-0

Fatty acid amides | Cyberlipid. (n.d.). Retrieved March 22, 2022, from http://cyberlipid.gerli.com/lipids/fatty-acid-amides/

Fujiwara, N., Naka, T., Ogawa, M., Yamamoto, R., Ogura, H., & Taniguchi, H. (2012). Characteristics of Mycobacterium smegmatis J15cs strain lipids. *Tuberculosis*, *92*(2), 187–192. https://doi.org/10.1016/J.TUBE.2011.10.001

Glycoside | biochemistry | Britannica. (n.d.). Retrieved March 22, 2022, from https://www.britannica.com/science/glycoside

Gordillo Altamirano, F. L., & Barr, J. J. (2019). Phage therapy in the postantibiotic era. *Clinical Microbiology Reviews*, *32*(2). https://doi.org/10.1128/CMR.00066-18

Guerrero-Bustamante, C. A., Dedrick, R. M., Garlena, R. A., Russell, D. A., & Hatfull, G. F. (2021). Toward a phage cocktail for tuberculosis: Susceptibility and tuberculocidal action of

mycobacteriophages against diverse Mycobacterium tuberculosis Strains. *MBio*, *12*(3). https://doi.org/10.1128/MBIO.00973-21

Guo, Z., Lin, H., Ji, X., Yan, G., Lei, L., Han, W., Gu, J., & Huang, J. (2020). Therapeutic applications of lytic phages in human medicine. *Microbial Pathogenesis*, *142*, 104048. https://doi.org/10.1016/J.MICPATH.2020.104048

Hadas, H., Einav, M., Fishov, I., & Zaritsky, A. (1997). Bacteriophage T4 Development Depends on the Physiology of its Host Escherichia Coli. *Microbiology*, *143*(1), 179–185. https://doi.org/10.1099/00221287-143-1-179

Hampton, H. G., Watson, B. N. J., & Fineran, P. C. (2020). The arms race between bacteria and their phage foes. *Nature 2020 577:7790*, *577*(7790), 327–336. https://doi.org/10.1038/s41586-019-1894-8

Haq, I. U., Chaudhry, W. N., Akhtar, M. N., Andleeb, S., & Qadri, I. (2012). Bacteriophages and their implications on future biotechnology: a review. *Virology Journal 2012 9:1*, *9*(1), 1–8. https://doi.org/10.1186/1743-422X-9-9

Hatfull, G. F. (2018a). Mycobacteriophages. *Microbiology Spectrum*, *6*(5), 76–78. https://doi.org/10.1128/MICROBIOLSPEC.GPP3-0026-2018

Hatfull, G. F. (2018b). Mycobacteriophages. *Microbiology Spectrum*, *6*(5), 76–78. https://doi.org/10.1128/MICROBIOLSPEC.GPP3-0026-2018

Hatfull, G. F., Brinkman, F., & Parkhill, J. (2008). Bacteriophage genomics This review comes from a themed issue on Genomics Edited by. *Current Opinion in Microbiology*, *11*, 447–453. https://doi.org/10.1016/j.mib.2008.09.004

Hedrick VE, LaLand MN, Nakayasu ES, Paul LN. Digestion, Purification, and Enrichment of Protein Samples for Mass Spectrometry. Curr Protoc Chem Biol. 2015 Sep 1;7(3):201-222. doi: 10.1002/9780470559277.ch140272. PMID: 26331527.

Jacobs-Sera, D., Marinelli, L. J., Bowman, C., Broussard, G. W., Guerrero Bustamante, C., Boyle, M. M., Petrova, Z. O., Dedrick, R. M., Pope, W. H., Modlin, R. L., Hendrix, R. W., & Hatfull, G. F. (2012). On the nature of mycobacteriophage diversity and host preference. *Virology*, *434*(2), 187–201. https://doi.org/10.1016/J.VIROL.2012.09.026

Jeckelmann, J. M., & Erni, B. (2020). Transporters of glucose and other carbohydrates in bacteria. *Pflugers Archiv : European Journal of Physiology*, *472*(9), 1129–1153. https://doi.org/10.1007/S00424-020-02379-0

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modelling, prediction and analysis. *Nature Protocols*, *10*(6), 845. https://doi.org/10.1038/NPROT.2015.053

Kobayashi, K., Kondo, M., Fukuda, H., Nishimura, M., & Ohta, H. (2007). Galactolipid synthesis in chloroplast inner envelope is essential for proper thylakoid biogenesis, photosynthesis, and embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(43), 17216. https://doi.org/10.1073/PNAS.0704680104

Kutter, E., de Vos, D., Gvasalia, G., Alavidze, Z., Gogokhia, L., Kuhl, S., & Abedon, S. (2010). Phage Therapy in Clinical Practice: Treatment of Human Infections. *Current Pharmaceutical Biotechnology*, *11*(1), 69–86. https://doi.org/10.2174/138920110790725401

Loewenstein, Y., Raimondo, D., Redfern, O. C., Watson, J., Frishman, D., Linial, M., Orengo, C., Thornton, J., & Tramontano, A. (2009). Protein function annotation by homology-based inference. *Genome Biology*, *10*(2), 207. https://doi.org/10.1186/GB-2009-10-2-207/FIGURES/2

Ma, L., Green, S. I., Trautner, B. W., Ramig, R. F., & Maresso, A. W. (2018). Metals Enhance the Killing of Bacteria by Bacteriophage in Human Blood. *Scientific Reports*, *8*(1). https://doi.org/10.1038/S41598-018-20698-2

Mahony, J., Casey, E., & Sinderen, D. van. (2020). The Impact and Applications of Phages in the Food Industry and Agriculture. *Viruses 2020, Vol. 12, Page 210*, *12*(2), 210. https://doi.org/10.3390/V12020210

Mancuso, F., Shi, J., & Malik, D. J. (2018). High Throughput Manufacturing of Bacteriophages Using Continuous Stirred Tank Bioreactors Connected in Series to Ensure Optimum Host Bacteria Physiology for Phage Production. *Viruses 2018, Vol. 10, Page 537*, *10*(10), 537. https://doi.org/10.3390/V10100537

Masuda, S., Harada, J., Yokono, M., Yuzawa, Y., Shimojima, M., Murofushi, K., Tanaka, H., Masuda, H., Murakawa, M., Haraguchi, T., Kondo, M., Nishimura, M., Yuasa, H., Noguchi, M., Oh-Oka, H., Tanaka, A., Tamiaki, H., & Ohta, H. (2011). A Monogalactosyldiacylglycerol Synthase Found in the Green Sulfur Bacterium Chlorobaculum tepidum Reveals Important Roles for Galactolipids in Photosynthesis. *The Plant Cell*, *23*(7), 2644. https://doi.org/10.1105/TPC.111.085357

Masukawa, Y., Narita, H., Sato, H., Naoe, A., Kondo, N., Sugai, Y., Oba, T., Homma, R., Ishikawa, J., Takagi, Y., & Kitahara, T. (2009). Comprehensive quantification of ceramide species in human stratum corneum. *Journal of Lipid Research*, *50*(8), 1708. https://doi.org/10.1194/JLR.D800055-JLR200

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, *405*(2), 442–451. https://doi.org/10.1016/0005-2795(75)90109-9

Max Planck Institute of Biochemistry. (2008). *MaxQB - MaxQuant Database*. http://maxqb.biochem.mpg.de/mxdb/

Max Planck Institute of Biochemistry. (n.d.). *MaxQuant*. Retrieved February 26, 2022, from https://www.maxquant.org/

Meinicke, P., Lingner, T., Kaever, A., Feussner, K., Göbel, C., Feussner, I., Karlovsky, P., & Morgenstern, B. (2008). MetaboAnalyst: Metabolite-based clustering and visualization of mass spectrometry data using one-dimensional self-organizing maps. *Algorithms for Molecular Biology*, *3*(1). https://doi.org/10.1186/1748-7188-3-9

Molecular Devices SpectraMax Plus 384 Microplate Reader | GMI - Trusted Laboratory Solutions. (n.d.). Retrieved January 31, 2022, from https://www.gmi-inc.com/product/molecular-devices-spectramax-plus-384-microplate-reader/?keyword=&leadsource=googleads&gclid=Cj0KCQiA0eOPBhCGARIsAFIwTs5WkX9v2qVg75kpf8d6LOZUu8hje4ocBM0ghKk94P5pVjsLUMGaPGwaAn09EALw_wcB

Mycolicibacterium smegmatis (strain ATCC 700084 / mc(2)155) (Mycobacterium smegmatis). (n.d.). Retrieved February 26, 2022, from https://www.uniprot.org/proteomes/UP000000757

Nagel, T. E., Chan, B. K., de Vos, D., El-Shibiny, A., Kang'ethe, E. K., Makumi, A., & Pirnay, J.-P. (2016). The Developing World Urgently Needs Phages to Combat Pathogenic Bacteria. *Frontiers in Microbiology*, *0*(JUN), 882. https://doi.org/10.3389/FMICB.2016.00882

Nanospray Flex™ Ion Sources. (n.d.). Retrieved January 31, 2022, from https://www.thermofisher.com/order/catalog/product/ES071

Oliveira, C., Faoro, H., Alves, L. R., & Goldenberg, S. (2017). RNA-binding proteins and their role in the regulation of gene expression in Trypanosoma cruzi and Saccharomyces cerevisiae. *Genetics and Molecular Biology*, *40*(1), 22. https://doi.org/10.1590/1678-4685-GMB-2016-0258

Oliveira, H., Sampaio, M., Melo, L. D. R., Dias, O., Pope, W. H., Hatfull, G. F., & Azeredo, J. (2019). Staphylococci phages display vast genomic diversity and evolutionary relationships. *BMC Genomics 2019 20:1*, *20*(1), 1–14. https://doi.org/10.1186/S12864-019-5647-8

Ongenae, V., Briegel, A., & Claessen, D. (2021). Cell wall deficiency as an escape mechanism from phage infection. *Open Biology*, *11*(9). https://doi.org/10.1098/RSOB.210199

Phages DB. (n.d.-a). *Details for Cluster A phages*. Https://Phagesdb.Org/Clusters/A/.

Phages DB. (n.d.-b). *The Actinobacteriophage Database | Cluster C Phages*. Retrieved September 11, 2021, from https://phagesdb.org/clusters/C/

Pierce™ Peptide Desalting Spin Columns. (n.d.). Retrieved January 31, 2022, from https://www.thermofisher.com/order/catalog/product/89851

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., & Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Research*, *46*(W1), W200–W204. https://doi.org/10.1093/NAR/GKY448

Q Exactive™ Plus Hybrid Quadrupole-Orbitrap™ Mass Spectrometer. (n.d.). Retrieved January 31, 2022, from https://www.thermofisher.com/order/catalog/product/IQLAAEGAAPFALGMBDK?SID=srch-srp-IQLAAEGAAPFALGMBDK

RIKEN Center for Sustainable Resource Science : Metabolome Informatics Research Team. (2020). *CompMS | MS-DIAL*. http://prime.psc.riken.jp/compms/msdial/main.html

Samaddar, S., Grewal, R. K., Sinha, S., Ghosh, S., Roy, S., & das Gupta, S. K. (2016). Dynamics of mycobacteriophage-mycobacterial host interaction: Evidence for secondary mechanisms for host lethality. *Applied and Environmental Microbiology*, *82*(1), 124–133. https://doi.org/10.1128/AEM.02700-15/SUPPL_FILE/ZAM999116810SO1.PDF

Sapkota, A. (2020, December 28). *Bacteriophage- Definition, Structure, Life Cycles, Applications, Phage Therapy*. https://microbenotes.com/bacteriophage/

Shahnazari, S., Namolovan, A., Klionsky, D. J., & Brumell, J. H. (2011). A role for diacylglycerol in antibacterial autophagy. *Autophagy*, *7*(3), 331. https://doi.org/10.4161/AUTO.7.3.14045

Shimamura, M. (2012). Immunological functions of steryl glycosides. *Archivum Immunologiae et Therapiae Experimentalis*, *60*(5), 351–359. https://doi.org/10.1007/S00005-012-0190-1/FIGURES/5

Shop ACQUITY UPLC BEH C18 Columns | 2.1mm X 50mm | 186002350 | Waters. (n.d.). Retrieved February 26, 2022, from https://www.waters.com/nextgen/us/en/shop/columns/186002350-acquity-uplc-beh-c18-column-130a-17--m-21-mm-x-50-mm-1-pk.html

Söding, J., Biegert, A., & Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research*, *33*(Web Server issue), W244. https://doi.org/10.1093/NAR/GKI408

SoftMax Pro Software, Microplate Data Acquisition & Analysis Software | Molecular Devices. (n.d.). Retrieved January 31, 2022, from https://www.moleculardevices.com/products/microplate-readers/acquisition-and-analysis-software/softmax-pro-software#gref

Storms, Z. J., Arsenault, E., Sauvageau, D., & Cooper, D. G. (2010). Bacteriophage adsorption efficiency and its effect on amplification. *Bioprocess and Biosystems Engineering 2010 33:7*, *33*(7), 823–831. https://doi.org/10.1007/S00449-009-0405-Y

UltiMate™ 3000 RSLCnano System. (n.d.). Retrieved January 31, 2022, from https://www.thermofisher.com/order/catalog/product/ULTIM3000RSLCNANO

Vecino, X., Rodríguez-López, L., Rincón-Fontán, M., Cruz, J. M., & Moldes, A. B. (2021). Nanomaterials synthesized by biosurfactants. *Comprehensive Analytical Chemistry*, *94*, 267–301. https://doi.org/10.1016/BS.COAC.2020.12.008

Walsh, D., & Mohr, I. (2011). Viral subversion of the host protein synthesis machinery. *Nature Reviews Microbiology 2011 9:12*, *9*(12), 860–875. https://doi.org/10.1038/nrmicro2655

Wenk, M. R. (2005). The emerging field of lipidomics. *Nature Reviews Drug Discovery 2005 4:7*, *4*(7), 594–610. https://doi.org/10.1038/nrd1776

Wenk, M. R. (2010). Lipidomics: New Tools and Applications. *Cell*, *143*(6), 888–895. https://doi.org/10.1016/J.CELL.2010.11.033

White, H. E., & Orlova, E. v. (2019). Bacteriophages: Their Structural Organisation and Function. *Bacteriophages - Perspectives and Future*. https://doi.org/10.5772/INTECHOPEN.85484

Zhang, W., & Wu, Q. (2020). Applications of phage-derived RNA-based technologies in synthetic biology. *Synthetic and Systems Biotechnology*, *5*(4), 343. https://doi.org/10.1016/J.SYNBIO.2020.09.003