UNIVERSITY OF CALIFORNIA SAN DIEGO

Lost in Translation: A Post-translational Modification-inclusive Analysis of Infectious Diseases

A dissertation submitted in partial satisfaction
of the requirements for the degree Doctor of Philosophy

in

Biomedical Sciences

by

Jacob M. Wozniak

Committee in charge:

Professor David J. Gonzalez, Chair
Professor John Guatelli
Professor Elizabeth Komives
Professor Victor Nizet
Professor Eugene Yeo

2020

The dissertation of Jacob M. Wozniak is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California San Diego

2020

DEDICATION

This dissertation is dedicated to my family, friends and mentors who have made a lasting impact on my life. Without their help, this work would not have been possible.

Mentors. To my mentors who have guided this incipient voyage of my scientific career. To my PI, David Gonzalez, who taught me to always strive for excellence and push the boundaries of what is known, but to have fun while doing it. To my post-doc mentors, John Lapek and Igor Wierzbicki, who not only taught me the majority of the experimental techniques included in this dissertation, but also how to think critically and communicate clearly. To the other faculty and mentors (particularly from the UCSD Pharmacology Department) that I had the pleasure of interacting with during my dissertation work, who are too numerous to mention by name, but nonetheless have contributed to my development as a scientist. Thank you.

Friends. To my friends, new and old, who helped keep my morale high throughout graduate school. From the weekend camping getaways to explore the Californian wilderness to flying halfway around the world to visit Japan, these experiences have made me forever grateful. Having an enriched life outside of work has made long hours in the lab all the more manageable. Thank you.

Family. To my parents, who raised me to be a curious individual and to always question the world around me. To my brothers, Giles and Lucas, who I grew up alongside and inspire me every day with their own achievements. To my girlfriend, Alice, who's unconditional love and willingness for adventure kept my spirits high when things in lab were low. Thank you.

EPIGRAPH

“We should not trust the masses who say only the free can be educated, but rather the lovers of wisdom who say that only the educated are free”
----Epictetus

“Sucking at something is the first step to becoming sorta good at something.”
----Jake the Dog

“Real artists ship.”
----Steve Jobs

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

A portion of the material within is reprinted from published and or submitted manuscripts where the dissertation author was the primary author.

Chapter 2, in part, is a reprint of material submitted to *Cell Reports*, 2020 (*manuscript in revision*), Jacob M. Wozniak, Julieta Aguilar., Dominic McGrosso, Igor H. Wierzbicki, Eri Nakatani-Webster, Michael R. Dores, Katrin Schilcher, Anvesh Macherla, Diana Dehaini, Xiaoli Wei, Ronnie H. Fang, JoAnn Trejo, Brian J. Werth, Abhinav Nath, Ross Corriden, Liangfang Zhang, Alexander R. Horswill, and David J. Gonzalez. The dissertation author was the primary author of this paper.

Chapter 3, in part, is a combination of material as it appears in *PeerJ*, 2019, Jacob M. Wozniak and David J. Gonzalez, and *PLOS Neglected Tropical Diseases***,** 2020, Jacob M. Wozniak, Tatiana Araújo Silva, Diane Thomas, Jair L. Siqueira-Neto, James H. McKerrow, David J. Gonzalez, Claudia M. Calvet. The dissertation author was the primary author of these papers.

Chapter 4, in part, is a reprint of material submitted to *Cell*, 2020 (*manuscript in press*), Jacob M. Wozniak, Joshua Olson, JR Caldera, Robert H. Mills, Marvic Carrillo-Terrazas, Chih-Ming Tsai, Fernando Vargas, Pieter C. Dorrestein, George Y. Liu, Victor Nizet, George Sakoulas, Warren Rose and David J. Gonzalez. The dissertation author was the primary author of this paper.

VITA

2015 - B.S. in Biochemistry and Molecular Biology, The Pennsylvania State University

2020 - Ph.D. in Biomedical Sciences, University of California San Diego

Publications

1. Rojony, R., Danelishvili, L., Campeau, A., **Wozniak, J.M**., Gonzalez, D.J., and Bermudez, L.E. (2020). Exposure of Mycobacterium abscessus to Environmental Stress and Clinically Used Antibiotics Reveals Common Proteome Response among Pathogenic Mycobacteria. **Microorganisms** *8*.

2. **Wozniak, J.M.**, Silva, T.A., Thomas, D., Siqueira-Neto, J.L., McKerrow, J.H., Gonzalez, D.J.*, Calvet, C.M.* (2020) Molecular Dissection of Chagas Induced Cardiomyopathy Reveals Central Disease Associated and Druggable Signaling Pathways. **PLoS Negl Trop Dis** *14*, e0007980,

3. Lin, Y.*, **Wozniak, J.M.***, Grimsey, N.J.*, Girada, S., Patwardhan, A., Molinar-Inglis, O., Smith, T.H., Lapek, J.D., Gonzalez, D.J., and Trejo, J. (2020). Phosphoproteomic analysis of protease-activated receptor-1 biased signaling reveals unique modulators of endothelial barrier function. **Proc Natl Acad Sci U S A** 117, 5039-5048.

4. Mills, R.H., **Wozniak, J.M.**, Vrbanac, A., Campeau, A., Chassaing, B., Gewirtz, A., Knight, R., and Gonzalez, D.J. (2020). Organ-level protein networks as a reference for the host effects of the microbiome. **Genome Res** 30, 276-286.

5. Gauglitz, J.M., Aceves, C.M., Aksenov, A.A., Aleti, G., Almaliti, J., Bouslimani, A., Brown, E.A., Campeau, A., Caraballo-Rodriguez, A.M., Chaar, R.*, et al.* (2020). Untargeted mass spectrometry-based metabolomics approach unveils molecular changes in raw and processed foods and beverages. **Food Chem** *302*, 125290.

6. Grainger, S., Nguyen, N., Richter, J., Setayesh, J., Lonquich, B., Oon, C.H., **Wozniak, J.M.**, Barahona, R., Kamei, C.N., Houston, J.*, et al.* (2019). EGFR is required for Wnt9a-Fzd9b signalling specificity in haematopoietic stem cells. **Nat Cell Biol** *21*, 721-730.

7. Rojony, R., Martin, M., Campeau, A., **Wozniak, J.M.**, Gonzalez, D.J., Jaiswal, P., Danelishvili, L., and Bermudez, L.E. (2019). Quantitative analysis of Mycobacterium avium subsp. hominissuis proteome in response to antibiotics and during exposure to different environmental conditions. **Clin Proteomics** *16*, 39.

8. Korandla, D.R., **Wozniak, J.M.**, Campeau, A., Gonzalez, D.J., and Wright, E.S. (2019). AssessORF: combining evolutionary conservation and proteomics to assess prokaryotic gene predictions. **Bioinformatics**.

9. Gao, N.J., Al-Bassam, M.M., Poudel, S., **Wozniak, J.M.**, Gonzalez, D.J., Olson, J., Zengler, K., Nizet, V., and Valderrama, J.A. (2019). Functional and Proteomic Analysis of Streptococcus pyogenes Virulence Upon Loss of Its Native Cas9 Nuclease. **Front Microbiol** *10*, 1967.

10. **Wozniak, J.M.**, and Gonzalez, D.J. (2019). PTMphinder: an R package for PTM site localization and motif extraction from proteomic datasets. **PeerJ** *7*, e7046.

11. Lapek, J.D., Jr.*, Jiang, Z.*, **Wozniak, J.M.**, Arutyunova, E., Wang, S.C., Lemieux, M.J., Gonzalez, D.J., and O'Donoghue, A.J. (2019). Quantitative Multiplex Substrate Profiling of Peptidases by Mass Spectrometry. **Mol Cell Proteomics** *18*, 968-981.

12. Al-Bassam, M.M., Kim, J.N., Zaramela, L.S., Kellman, B.P., Zuniga, C., **Wozniak, J.M.**, Gonzalez, D.J., and Zengler, K. (2018). Optimization of carbon and energy utilization through differential translational efficiency. **Nat Commun** *9*, 4474.

13. Lapek, J.D.*, Jr., Mills, R.H.*, **Wozniak, J.M.**, Campeau, A., Fang, R.H., Wei, X., van de Groep, K., Perez-Lopez, A., van Sorge, N.M., Raffatellu, M.*, et al.* (2018). Defining Host Responses during Systemic Bacterial Infection through Construction of a Murine Organ Proteome Atlas. **Cell Syst** *6*, 579-592 e574.

14. McDonald, D., Hyde, E., Debelius, J.W., Morton, J.T., Gonzalez, A., Ackermann, G., Aksenov, A.A., Behsaz, B., Brennan, C., Chen, Y.*, et al.* (2018). American Gut: an Open Platform for Citizen Science Microbiome Research. **mSystems** *3*.

15. Markmiller, S., Soltanieh, S., Server, K.L., Mak, R., Jin, W., Fang, M.Y., Luo, E.C., Krach, F., Yang, D., Sen, A.*, et al.* (2018). Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. **Cell** *172*, 590-604 e513.

16. Lapek, J.D.*, Jr., Lewinski, M.K.*, **Wozniak, J.M.**, Guatelli, J., and Gonzalez, D.J. (2017). Quantitative Temporal Viromics of an Inducible HIV-1 Model Yields Insight to Global Host Targets and Phospho-Dynamics Associated with Protein Vpr. **Mol Cell Proteomics** 16, 1447-1461

ABSTRACT OF DISSERTATION

Lost in Translation: A Post-translational Modification-inclusive Analysis of Infectious Diseases

By

Jacob M. Wozniak

Doctor of Philosophy in Biomedical Sciences

University of California San Diego, 2020

Professor David J. Gonzalez, Chair

The overall theme of this dissertation is the use post-translational modification (PTM)-tolerant approaches to investigate infectious diseases. Chapter 1 contains background information regarding quantitative proteomics, PTMs and infectious disease, particularly the anti-microbial resistant pathogen, *Staphylococcus aureus*. This information is provided to introduce the reader with the fundamentals of the biology and the main techniques used during the doctoral studies. The following chapters describe primary author works (published or under review) completed by the author of this dissertation.

Chapter 2 describes the discovery and characterization of modified microproteins produced by *Staphylococcus aureus*. These microproteins were discovered using an

unbiased peptidogenomic approach and characterized using *in vitro* and *in vivo* assays. One of the new microproteins functions similarly to previously described virulence factors, while to other appears to have distinct mechanisms of action in the skin. The further investigation of these microproteins may provide a more comprehensive understanding of the host-pathogen interaction.

Chapter 3 introduces a computational tool, PTMphinder, created by the author of the dissertation. This tool enables the high-throughput localization of phospho-sites in full length proteins and the extraction of flanking sequences. The utility of this tool is demonstrated by applying it to gain a more comprehensive view of the *in vivo* host response to infection in a chronic Chagas disease model. Kinase-substrate pairs and drug targets are predicted from phospho-proteomic data, providing numerous hypotheses to further interrogate.

Chapter 4 aims to set a new standard in the biomarker research field. By employing a multi-omic approach, >10,000 features were quantified from >200 *S. aureus* bacteremia (SaB) patient samples, including abundant post-translational modifications (PTMs). A model was constructed from the multi-omic data which provides the best predictive ability for mortality from any infection to date. Further, the host response to infection was detailed and used to predict treatment strategies, which were substantiated using animal models. Overall, Chapter 4 details a comprehensive molecular view of the early host response to infection.

Lastly, Chapter 5 details experiments to further develop the research described in Chapter 4. This includes the development of a multi-marker immunoassay-based tool for

the prediction of SaB outcomes and experiments to further understand detailed mechanisms of the host response to *S. aureus* bacteremia.

Chapter 1 - General Introduction

Systems Biology and Quantitative Proteomics

Conventional approaches, such as western blot and molecular biology, have been the foundation for answering biological questions. However, these methods only interrogate a few targets and even simple stimuli are known to induce drastic molecular alterations in complex biological milieux. Thus, systems level technologies have emerged as an effective tool to unbiasedly assess various disease conditions. These include approaches such as genomics, the complete sequencing and comparison of organisms' genome, and transcriptomics, the measurement of the RNA transcripts in a sample, often produced in response to a stimulus. While these nucleic acid-based analyses have provided a wealth of data that has driven research across all realms of biomedical inquiry, from cancer genomics to microbiome meta-transcriptomics, they possess an inherent flaw in that they fail to provide direct measurements of the primary bioactive molecules: proteins and metabolites. Thus, these somewhat incomplete methods are challenged when answering particular questions, such as when molecules function in a specific environmental context (eg. functional molecules secreted into serum) or are influenced by the presence of post-translational modifications (PTMs). These areas of biological interest are more readily accessible by mass spectrometry (MS)-based approaches, such as metabolomics and proteomics.

Throughout history, MS-based approaches have made significant contributions to biomedical research(1, 2). With the complete sequencing of thousands of organisms over the past 20 years, bottom up MS-based proteomics has become increasingly useful(3-5). Recent developments in analytical instruments and protein quantification methods(6-9) have only exponentially increased the value of MS applied to biology. While there are a

multitude of instrument types and quantification strategies, a few have emerged as workhorses for proteomics applications. These include various mass analyzers (eg. ion-trap and Orbitrap) and quantification strategies (eg. label-free quantification (LFQ) and isobaric tags). As these are the primary tools utilized in this dissertation, a brief overview of the benefits and drawbacks of each approach is provided below.

Mass analyzers function to directly detect molecular ions inside mass spectrometers. These analyzers can be grouped into low-resolution (eg. ion-trap) and high-resolution (eg. Orbitrap) based on their discriminating power of similar weight ions. Ion-traps typically have a resolution of ~2000, while Orbitraps can achieve a resolution of >100,000 depending on the duty cycle time. However, this increased resolution comes with the drawback of a reduced scan rate. Therefore, combination strategies, utilizing the high-resolution of the Orbitrap in conjunction with the rapid scan rate of ion-trap, have been employed to a great success. This co-operation of both the Orbitrap and ion-trap is unique to the architecture of the Fusion line of Thermo Fisher instruments (ie. Orbitrap Fusion, Lumos and Eclipse instruments), which are the primary instruments utilized in this dissertation. While this combination of mass analyzers is widely used for standard proteomic applications, specific focus areas, such as glyco-proteomics, benefit from only taking high-resolution scans in the Orbitrap, which enables enhanced PTM assignment and localization(10, 11).

Another factor to consider when designing proteomics experiments is the method of quantification used. Quantification via MS can be either relative, where differences in abundance are compared between samples, or absolute, where the total abundance of a protein (eg. ug/ml) is determined. The vast majority of proteomic applications use relative

quantification as the latter requires the design, purchase and optimization of synthetic standards for every quantified target, drastically increasing cost and reducing throughput. Relative quantification strategies can be split into three groups: label-free quantification (LFQ), metabolic labeling (eg. Stable isotope labeling with amino acids in cell culture (SILAC)), and chemical labeling (eg. tandem-mass tags (TMT)). LFQ is the cheapest and quickest workflow for relative quantification, but it suffers greatly from missing values across distinct MS runs, limiting the throughput and statistical power. SILAC strategies benefit from early sample mixing in the workflow, which reduces the impact of sample preparation errors; however, this approach is generally restricted to cell-culture experiments, has a high cost association, and increases sample complexity, resulting in less efficient peptide identification. Finally, chemical labeling approaches, such as TMTs, offer a solution to the drawbacks associated to other quantification strategies. While there is still a high cost associated with purchase of the TMT labels, this is balanced by the increased throughput achieved via multiplexing up to 16 samples in a single MS run. Chemical labeling-based approaches also drastically reduce the number of missing values across samples and can be readily applied to any biological source material, included human samples. Importantly, due to the isobaric nature of the chemical tags, there is minimal increase in overall sample complexity. Therefore, the development of chemical labeling strategies, paired with improved instrumentation, is likely one of the greatest advances in proteomic quantification over the past ten years.

To demonstrate the benefit of chemical labeling over LFQ analyses, we applied both approaches to human serum samples, a notoriously difficult sample to analyze via MS given the large dynamic range of proteins and prevalent PTMs. Overall, both methods

detected a similar number of unique peptides (LFQ: 6,919, TMT: 6,597; **Figure 1-1A**), but the TMT method quantified a larger number of proteins (LFQ: 888, TMT: 1,261; **Fig 1-1B**), while encompassing the majority of the proteins from the LFQ method (~77%). Comparing the quantitation values from both experiments, we found that the fold-changes of infected vs. healthy patients from the distinct workflows were significantly correlated (Pearson Correlation: 0.713, **Fig 1-1C**), validating each approach. However, due to the proper design of the TMT experiments and the high number of missing values associated with LFQ, the TMT method had a drastically higher number of comparable proteins (ie. detected in at least three samples for each condition) between healthy and infected patients (**Fig 1-1D**). Therefore, TMT-based quantification is utilized for the majority of experiments in this dissertation.

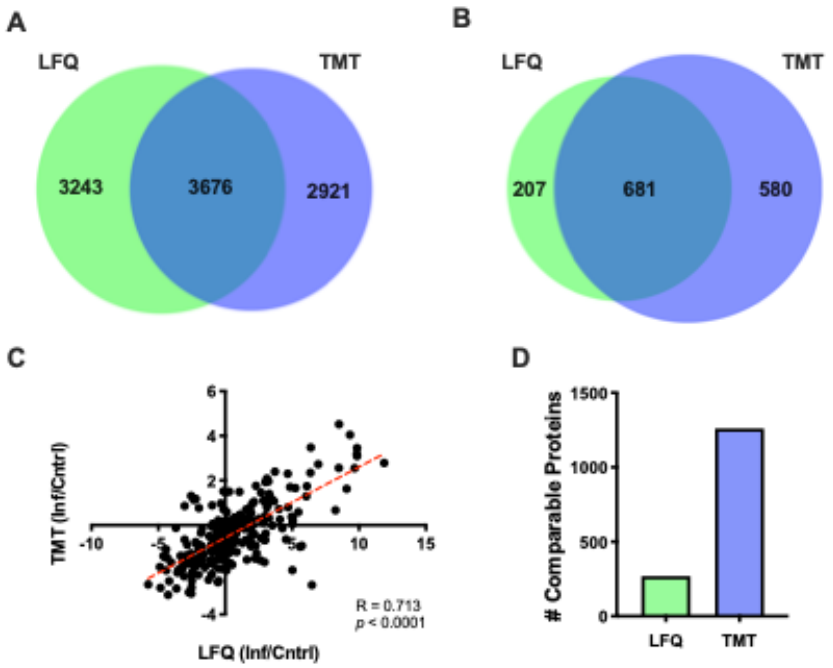


Figure 1-1 *Comparison of LFQ and TMT-based serum proteomics.*

Venn diagrams of unique peptides (**A**) and proteins (**B**) detected in each experiment. (**C**) Scatter plot of $\log_2$(infected/healthy) quantitation values of each protein detected in each experiment. (**D**) Number of comparable proteins detected in each experiment.

Post-translational Modifications (PTMs)

The human genome contains tens of thousands of genes that encode for proteins, the functional unit of the cell. Protein activity can be modulated by many factors such as expression level, alternative-splicing, localization and post-translational modifications (PTMs). Common PTMs range from small molecules (phosphorylation(12, 13) and acetylation(14, 15)) to peptide features (ubiquitination(16, 17)) to sugars moieties (glycosylation(18, 19)). When PTMs are taken into consideration, the complexity of the proteome exponentially increases to over one million predicted protein products. Consequently, the analysis of protein PTMs has led to the deeper understanding of biological systems and development of therapies for a wide-range of maladies from cancer(20-22) to infectious diseases(23, 24) to auto-immune disorders(22, 25). Thus, assessing the PTM status of proteins under various physiological conditions and disease states has the potential to facilitate new biological discoveries.

While many studies interrogate the PTM status of a single or a few proteins via conventional techniques (i.e. western blot), PTMs are rarely singular events and a simple stimulus can induce a massive number of alterations in protein modifications. In addition, conventional studies are limited by a lack of reliable antibodies for specific PTM site detection. To address these gaps, mass spectrometry-based proteomics has emerged as a powerful tool to simultaneously analyze the PTM status of thousands of proteins at tens of thousands of sites(26, 27). These approaches have led to many discoveries, from complicated signaling networks induced during infection(23, 24) to the identification of motifs recognized by particular kinases(28-30). With the further development of such

technology, one can expect the depth of proteomic-based PTM analyses to increase substantially.

To date, there are a number of tools available that allow users to analyze PTMs within proteomic data(29, 31, 32). These include both tools focused on increasing PTM identification and quantification, such as open database and PTM-tolerant searches and tools focused on increasing PTM interpretability, such as motif enrichment analyses and kinase predictions. While useful, many of these applications require significant preprocessing of raw data and substantial programming knowledge to implement. Thus, for biologists with limited programming experience, a gap in expertise and software exists to fully analyze the PTMs within proteomic data sets. These gaps in knowledge can lead to the misinterpretation or overlooking of biologically relevant data. Therefore, the development of new and easy to use computational tools for the analysis of PTM-containing proteomic data can empowers researchers to fully analyze and interpret PTMs derived from proteomic data.

Infectious Disease and *Staphylococcus aureus*

Antibiotic resistance is one of the greatest contemporary challenges facing mankind, showing little sign of slowing without direct and precise intervention. Antibiotic treatments target essential metabolic functions of bacterial pathogens, applying high selective pressures that result in the development of resistance(33). Over and improper prescription of broad-spectrum antibiotics combined with patient non-compliance with therapeutic regimens further exacerbates this issue(34). Thus, bacterial pathogens are developing multi-drug resistance at an alarming rate(33). Over the past 60 years, highly

infectious, multidrug-resistant bacterial isolates that pose pressing threat levels across the United States (US) have emerged(33, 35, 36). An estimated one million Americans are afflicted with healthcare-associated (HA) bacterial infections each year(37), most of which are from antibiotic-resistant strains, resulting in up to 100,000 deaths and an economic burden of $21 to $34 billion annually(38). In addition, some bacteria are migrating from their typical HA niches and have become community-associated (CA), which are often more virulent than their HA counterparts.

One the deadliest human microbial pathogens is *Staphylococcus aureus.* Infections can range from minor skin infections, readily treatable with over-the-counter antibiotics, to life-threatening *S. aureus* bacteremia (SaB), requiring intensive, intravenous therapy to properly treat. Severe disease states such as SaB carry alarming mortality rates ranging from 20 - 30%, which can increase in at risk populations and if the microbe is resistant to antimicrobials (eg. methicillin-resistant (MRSA)). However, even in infections caused by methicillin sensitive strains, clinical failure is common, revealing significant shortcomings in predictive power of standard antimicrobial susceptibility testing. Therefore, it is unsurprising that *S. aureus* has remained a prevalent pathogen and significant burden on society despite being highly studied for decades. Further, CA-MRSA strains have emerged, which can cause more severe skin infections in younger, previously healthy individuals. This increased virulence has largely been attributed to the expression of additional virulence factors, such as the phenol-soluble modulins (PSMs)(39, 40).

MS has been readily applied to the field of infectious disease(41-43).The unbiased nature and precise quantification of MS-based approaches, presents these techniques as

having the potential to guide the construction and prioritization of testable hypotheses. In order to combat the spread of problematic microbes, it is imperative to explore new avenues of antibacterial therapies; strategies that not only focus on the pathogen but also the host response to infection(44). Critical host responses modulated by interaction with an invading pathogen must be elucidated, not only as single entities, but how they function as a dynamic, concerted network. Merging the fields of analytical chemistry and microbiology can result in unprecedented views of microbial pathogenesis, ultimately leading to the enhancement of infectious disease treatments. This dissertation describes a number of unique vantage points of infectious diseases obtained through MS approaches including: 1) the functions of microproteins produced by CA-MRSA, 2) a global phospho-proteomic analysis of *Tryponosoma cruzi* infected murine hearts and 3) a multi-omic analysis of SaB patient serum.

Chapter 2 – Discovery and Characterization of Modified S. aureus Microproteins

Introduction

Staphylococcus aureus is a commensal organism that persistently colonizes the skin and upper respiratory tract of an estimated 30% of humans(45-47). *S. aureus* can readily cause diseases ranging from skin lesions to invasive endocarditis, particularly during times of immunodeficiency(48). With the emergence of community-associated, methicillin-resistant *S. aureus* (CA-MRSA), a steady rise in the infection rate of healthy individuals lacking traditional risk factors has been observed(49, 50). Thus, substantial effort has been placed upon understanding the basis of *S. aureus* pathogenicity. The discovery of new virulence factors that contribute to the pathogenesis of *S. aureus* would represent a critical step towards the development of therapeutics against this prevalent human pathogen. To date there remains an unmet need for deeper comprehension of the molecular dynamics that govern the extraordinary virulence of *S. aureus*.

Microproteins are peptide effector molecules that originate from small open reading frames (smORFs)(51). It is important to note that, while microproteins are peptidic by nature, they are distinct from functional peptides derived from larger proteins via cleavage events. The study of microproteins and their importance in biological systems has gained momentum in recent years(51, 52). In the realm of bacteria, the functions of microproteins range from signaling molecules to potent toxins(40, 53). One family of toxin microproteins, the phenol soluble modulins (PSMs), are well characterized *S. aureus* virulence factors with roles ranging from host cell lysis(40) to biofilm formation(54). Interestingly, the PSMs are robustly produced by CA-MRSA relative to other MRSA strains, and often retain a formylated initiator methionine (fMet)(40, 55). It is predicted that bacteria produce a number microproteins in culture. However, smORFs that encode

for microproteins are often overlooked for two reasons: 1) their size is typically smaller than the lower limits of annotation for many genomic algorithms, or 2) they may be embedded within larger ORFs(56, 57).

In this work, we define the microproteins and endogenous peptides detected in *S. aureus* TCH1516 (CA-MRSA) culture extracts and functionally characterize a bioactive subset. We report the high confidence identification of 57 peptides, encompassed by 115 peptido-forms. This study focuses on the characterization of two newly discovered microproteins, *S. aureus* microprotein (SAM1) and 2 (SAM2), within this cohort. We report that the SAMs are regulated by the accessory gene regulator (*agr*) and that SAM1 has immuno-stimulatory and cytotoxic properties similar to the PSMs. Utilizing quantitative proteomics, we found that SAM2 affects host cells through a non-canonical mechanism, which appears to be disruption of host keratin networks. Finally, we show that the SAMs enhance *S. aureus* virulence and/or survival through distinct mechanisms both *in vitro* and *in vivo*. Overall, this study expands the current mechanistic understanding of microprotein virulence factors and offers a resource detailing the *S. aureus* peptidome that can be further explored.

Microbial pathogenesis is a complex process that involves crosstalk between an array of host and pathogen factors. The understanding of how bacterial microproteins contribute to virulence is limited relative to larger protein and transcript studies. The human pathogen, *Staphylococcus aureus,* is known to produce a diverse array of virulence factors that contribute to its pathogenicity. However, many virulence factors, including microproteins and protein cleavage products, remain uncharacterized. To address these gaps, a peptidogenomic workflow was applied to *S. aureus* culture

supernatants. Overall, we identified 57 high-confidence, non-redundant peptides detected as approximately 115 peptido-forms. Within this group, we discovered two microproteins, SAM1 and SAM2, which are regulated by the classical *S. aureus* accessory gene regulator system. We find that SAM1 appears to act as a canonical cytolysin, reminiscent of the phenol-soluble modulins (PSMs), while SAM2 possesses unique bioactivity, the perturbation of keratin networks. Finally, we show that both SAMs can enhance *S. aureus* pathogenesis *in vitro* and *in vivo* through distinct mechanisms. This work highlights the value of studying the peptidome, a largely unexplored molecular space, and its contribution to bacterial virulence.

Methods

Peptide Extraction.

Peptides were extracted from *S. aureus* TCH1516, LAC, or RN9120 alone or in the presence of 1.2 mg/L LL-37 or lipopolysaccharide using 1-butanol and ethyl acetate. The collected extracts were dried, resuspended in 5% acetonitrile/5% formic acid and analyzed via the standard MS2 method described below.

Bacterial Growth Inhibition Assays.

S. pyogenes strain M3 was grown overnight (O/N), diluted 1:100 and grown for 8 hours (hrs) prior to experiments. The bacteria were then suspended to an OD_{600} of 0.4 and diluted 1:10 in 10% BHI in RPMI media containing the α-hemolysin signal peptide derivative (50 µg/ml – TTLLLGSILMNPVANA) or a matched DMSO control. Growth curves were generated by measuring OD_{600} every 15 minutes (min) for 12 hrs using a plate reader.

Monte Carlo Simulations.

Monte Carlo simulations of peptides in 100 mM NaCl were performed using the MCPep webserver, specifying a 40 Å membrane consisting of a 30 Å hydrophobic region flanked by two 5 Å polar regions. A 15% charged lipid was used to mimic the lipid composition of human keratinocytes (58, 59). Three replicates consisting of 500,000 Monte Carlo steps were performed for each peptide in each condition (membrane and aqueous).

Construction of S. aureus Mutants.

Accessory Gene Regulator (agr). The TCH1516 *agr*::tetM deletion mutant was generated by using phage φ11 to transduce the *agr*::tetM mutation from *S. aureus* strain RN6911 into TCH1516. TCH1516 *agr*::tetM has a complete deletion of the *agr* locus, which was confirmed by polymerase chain reaction (PCR) and sequencing using the primers KS111 and KS112 (**Table S1**).

S. aureus Microprotein 1 (sam1). To construct the *sam1* deletion plasmid, about 1000 base-pair (bp) regions flanking the *sam1* gene were amplified from *S. aureus* TCH1516 genomic DNA with primer pairs KS19/KS21 (for upstream homologous region; HR up) and KS20/KS22 (for downstream homologous region; HR down) (**Table S1**). The products were column purified, digested with SalI/XhoI (HR up), KpnI/XhoI (HR down) and ligated into pJB38 linearized with SalI and KpnI. The constructed plasmid was electroporated into *E. coli* DC10B cells and positive clones were selected on LB plates containing 100 µg/ml ampicillin at 37 ºC. The plasmid was isolated and directly electroporated into electrocompetent *S. aureus* TCH1516 cells. Cells carrying the plasmid were selected on TSA plates containing 10 µg/ml chloramphenicol at 30 ºC. Individual

colonies were then streaked on TSA chloramphenicol (10 µg/ml) plates incubated at 42 ºC to select for homologues recombination into the chromosome. To promote a second round of recombination, single colonies were sub-cultured in TSB at 30 ºC for 3 days without addition of antibiotics. The cultures were then diluted 10^{-6} and plated on TSA containing 0.2 µg/ml anhydrotetracycline (ATc) to select for plasmid loss. Colonies were screened for resistance to chloramphenicol (10 µg/ml) and chloramphenicol sensitive colonies were verified by PCR and sequencing using primers KS31, KS32, KS35 and KS36 (**Table S1**).

S. aureus Microprotein 1/2 (sam1/2). To construct the Δ*sam1/2* double mutant, a similar approach was undertaken. Briefly, 1000 bps up (JMW10 and JMW11) and downstream (JMW3 and JMW12) of the *sam2* gene were amplified from *S. aureus* TCH1516 *Δsam1* genomic DNA (**Table S1**). PCR products were column purified and ligated together utilizing a Gibson Assembly protocol according the manufacturer instructions. Gibson Assembly products were amplified and KpnI/EagI restriction sites were incorporated using primers JMW13 and JMW14 (**Table S1**). Assemblies were ligated into the pKOR1-MCS vector and transformed into *E. coli* NEB5α. Cells carrying the plasmids were selected for on THB plates containing 100 µg/ml Ampicillin. The resultant plasmids were subsequently isolated, confirmed via PCR and sequencing and transformed into *E. coli* DC10B. Cells carrying the plasmid were selected for on THB plates containing 100 µg/ml Ampicillin and 30 µg/ml Chloramphenicol. The plasmid was again isolated and electroporated into electrocompetent *S. aureus* TCH1516 Δ*sam1* cells. Cells carrying the plasmid were selected on THB plates containing 20 µg/ml chloramphenicol at 37 ºC. Plasmid containing cells were grown for 24 hours at 43 ºC to

promote plasmid integration into the chromosome. Serial dilutions of 10^{-3} to 10^{-7} were prepared and plated on THB plates containing 20 µg/ml chloramphenicol and incubated at 43 ºC O/N. Large colonies were selected and inoculated at 30 ºC O/N to promote double-crossover events. The cultures were then diluted 10^{-1} to 10^{-5} and plated on THB plates containing 1 µg/ml ATc to select for plasmid loss. Colonies were screened for resistance to chloramphenicol (20 µg/ml) and chloramphenicol sensitive colonies were verified by PCR and sequencing using JMW7, JMW8, JMW9, JMW15 and JMW16 (**Table S1**).

Quantative Reverse Transcription (qRT)-PCR.

To determine *sam1* and *sam2* expression in TCH1516 wild-type (WT) and the isogenic *agr*::tetM mutant, three independent cultures were grown in TSB to OD_{600} of 4. Cells were pelleted and incubated briefly with RNA protect Bacterial Reagent. Cells were lysed with lysostaphin for 30 min at 37 ºC and then incubated with proteinase K for another 30 min at 37ºC. RNA was purified using the RNeasy Plus Mini Kit and residual chromosomal DNA was removed using the Turbo DNA free kit according to the manufacturer's protocol. DNase-treated RNA (2 µg) was used as a template to generate cDNA with the High-Capacity Reverse Transcription Kit. qRT-PCR was performed using SsoAdvanced Universal SYBR Green Supermix, 350 nM of each primer (**Table S1**) and 10 ng cDNA. The expression was normalized to that of DNA gyrase (*gyrB*) and the fold change was calculated according to the Pfaffl analysis method (60).

Neutrophil Isolation.

Human research participants were consenting volunteers as approved under our local Institutional Review Board (IRB) protocol. Primary neutrophils were isolated from

venous blood collected from healthy volunteers, with heparin used as an anticoagulant. Blood was subjected to density gradient separation using Polymorphprep density gradient medium according to the manufacturer's protocol. Subsequently, neutrophils were collected, counted, and resuspended in PBS at a concentration of 1 x 10^7 cells/ml prior to use in experiments.

Neutrophil Chemotaxis.

Corning Transwell Permeable Supports with 3.0 μm-pore polycarbonate membranes were used to perform neutrophil migration assays. The transwell inserts were removed and 650 μl of Hank's Balanced Salt Solution (HBSS) + Ca^{2+}/Mg^{2+} alone or with 2 μg/ml PSMα3, mock-SAM, SAM1 or SAM2 was added to the lower wells. For the inhibitor assays, neutrophils in HBSS + Ca^{2+}/Mg^{2+} were pre-incubated in siliconized tubes for 20 min with PBP-10 (10 μM). Following this incubation, neutrophils (1 x 10^6 cells in 150 μl media) were added to upper wells and allowed to migrate for 1 hr. After migration, the inserts were removed and the cells in the lower wells were lysed by addition of 72 μl PBS containing 1% Triton X-100 for 10 min. 180 μl from each well was collected and added to a flat-bottom plate containing 20 μl 10 mM N-Methoxysuccinyl-Ala-Ala-Pro-Val p-nitroanilide. Following a 30 min incubation at room temperature (RT), absorbance at 405 nm was measured in a plate reader to quantify relative cell migration.

HaCaT LDH Release.

HaCaT cells were grown to 80-90% confluence and plated at 2 x 10^4 cells per well for 48 hrs. For synthetic microprotein experiments, cells were pretreated with 20 mM Hepes and 10 μg/ml of BSA, 10% FBS or no supplement for 1 hr before adding PSMα3, mock-SAM, SAM1 or SAM2 for 4 hrs. For inducible expression experiments, plasmid-

containing strains were grown in THB media supplemented with ATc (25 ng/ml) and chloramphenicol (10 µg/ml) for 16 hrs. Supernatants from the resulting cultures were passed through filter units and added to HaCaT cells for 6 hrs. HaCaT cell lysis was measured by release of lactate dehydrogenase (LDH) activity using an LDH Cytotoxicity Assay Kit according to manufacturer instructions.

HaCaT Microscopy.

For confocal microscopy experiments, circular glass coverslips that had been ethanol-dipped and dried were placed in 12-well polystyrene tissue culture plates. Coverslips were percolated with 300 µl fibronectin for 15 min at RT. HaCaT cells were plated on top of the coverslips (2 x 10^5 cells/well) and grown at 37 °C. After 48 hrs the cells were 80-90% confluent and pretreated with 20 mM Hepes and 10 µg/ml of BSA for an hour. PSMα3, mock-SAM, SAM1, SAM2 were added to the cells at 10 µg/ml for 4 hrs. Cells were then placed on ice, washed with PBS and stained. For micrographs in Figure 3, cells were stained with DAPI, SYTOX-Green, and Alexa Fluor 555 Phalloidin. For micrographs in Figure 4, cells were fixed with 4% paraformaldehyde for 15 min, blocked with 1% BSA + 0.1% triton in 1X PBS, and stained with Rabbit Polyclonal anti-Wide Spectrum Cytokeratin antibodies in 1% BSA O/N at 4 °C. The next day, the primary antibody was aspirated, the cells were washed with 1X PBS, and incubated in the dark with Donkey anti-Rabbit IgG H&L (Alexa Fluor 647) for 1 hr at RT. Cells were washed again and mounted to microscopy slides using Flouromount with DAPI. Microscopic analysis was repeated at least three times and a representative image from each treatment are shown. Quantification of keratin area was performed by measuring the average area of keratin per cell of four replicates (ten cells each) using ImageJ.

Construction of sam1 Inducible Vector.

To create an ATc-inducible *sam1* plasmid (pKAS28), the *sam1* gene was amplified from TCH1516 genomic DNA using the primers KS120 and KS121 which contain KpnI and SacI sites at their 5' ends (**Table S1**). The resulting PCR product was ligated into pRMC2 (linearized with KpnI and SacI). The plasmid was electroporated into *E. coli* DC10B and positive clones were selected on LB ampicillin (100 µg/ml) and verified by PCR and sequencing using primer KS88 and KS89. The plasmid pKAS28 and the empty vector control pRMC2 were electroporated into TCH1516 carrying the chromosomally integrated pLL29 plasmid which confers resistance to tetracycline.

HaCaT Infections.

Standard Infection. HaCaT cells were grown in 12-well plates to 80-90% confluency for 48 hrs without antibiotics. *S. aureus* TCH1516 WT and Δ*sam1* containing pRMC2 and Δ*sam1* containing pKAS28 were grown in the presence of 25 ng/ml ATc (maintained for the duration of the experiment) at 37 °C O/N prior to experiments. On the day of infection, bacterial cultures were back-diluted 1:100 and incubated for 1 hr at 37 °C. HaCaT cells were infected at an MOI of 1 for 4 hrs. At the 4 hr time point, Triton X-100 (final concentration: 0.1%) was added and colony forming units (CFUs) were quantified by serial dilution.

Intracellular Infection. Infections were repeated as above with the following additions. After the infection stage, extracellular bacteria were killed by addition of 20 µg/ml lysostaphin and 50 µg/ml gentamicin for 1 hr. Antibiotics were then washed away and the intracellular infection was allowed to persist for 2 and 4 hrs (4 and 6 total hrs after the addition of bacteria, respectively). For EGTA supplementation experiments, 1 mM

EGTA was supplemented following antibiotic removal. At each time point, extracellular CFUs were quantified from the supernatants and intracellular CFUs were quantified after lysing the HaCaT cells with 0.1% Triton X-100 in PBS. CFUs were determined by serial dilution.

Mouse Skin Lesion Experiments.

Four-day Lesion Size Experiment. Six-week-old female BALB/c mice were shaved to remove the hair on their back. 100 µl of solution containing 6 x 10^9 CFUs/ml of *S. aureus* Δ*agr* (RN9120) supplemented with DMSO, PSMα3, SAM1 or SAM2 to a final concentration of 10 µg/µl was then injected into the dorsal skin. The lesion of each mouse was monitored and reported as the area of the visibly affected area on day four post-challenge (quantified using ImageJ). The mice were then euthanized, and the affected skin was carefully collected and processed for bacterial CFU enumeration with sterile ceramic beads. The sample size for each group was 5 mice. This experiment was repeated twice for a total of 10 replicates.

Seven-day Lesion Penetration Experiment. The above infection experiments were repeated for a week-long experiment with the following changes. First, only DMSO and SAM2 were used as supplements. Second, in addition to *S. aureus* RN9120, we included a strain of WT *S. aureus* (RN6607). Third, mouse weight was measured each day and, on the final day of the experiment, both the lesion and the underlying tissues were photographed and processed for bacterial CFU enumeration with sterile ceramic beads. Percent penetrated CFUs is reported (equation: CFUs/g in underlying tissue/ total CFUs/g recovered). Finally, relative inflammation was quantified by averaging the redness of the

pixels (equation: red / (red + blue + green)) of two lines drawn through the photographs of the underlying tissue using ImageJ. The sample size for each group was 5 mice.

Microprotein Affinity Purification

HaCaT cells (TMT experiment) were grown in 15 cm dishes to 80-90% confluency and HL-60 cells (reductive demethylation experiment) were grown in T175 flasks to 1 x 10^6 cells/ml prior to experiments. Cells were washed with 1X PBS and lysed in a buffer composed of n-Dodecyl β-D-maltoside (DDM) 1% w/v, 100 mM NaCl, 10 mM sodium fluoride (NaF), 10 mM β-glycerophosphate, 10 mM sodium pyrophosphate (NaPP), 2 mM sodium orthovanadate ($NaVO_4$) and 1X cOmplete EDTA-free protease inhibitor cocktail. Insoluble membranes were removed by centrifugation at 13,000 x g for 10 min at 4 °C. Affinity purification was performed using a Biotinylated Protein Interaction Pull-Down Kit according to manufacturer instructions. Beads were prepared as either empty, beads-only controls, or with biotinylated PSMα3, SAM1 or SAM2. Prey proteins were eluted according to manufacturer instructions and processed for quantitative MS analysis.

Western Blot.

Proteins were isolated from the microprotein affinity purification experiments described above. Equal volumes of eluent from beads-only and SAM1 beads were subject to western blot analysis along with 10 µg of input material. Proteins were separated on a poly-acrylamide gel and transferred on to a nitrocellulose membrane. The nitrocellulose membrane was then blocked in 5% milk for 1 hr at RT before being washed and stained with Rabbit Polyclonal anti-BAP37 (PHB2) antibodies (1:500 in 5% milk) shaking O/N at 4 °C. The next day, membranes were washed and stained with Goat anti-Rabbit IgG, HRP-linked antibodies (1:2000 in 5% milk) for 1 hr at RT. Membranes were

then washed and incubated in the dark with enhanced chemiluminescent substrate reagents for 1 min at RT. Light emission was measured using X-ray film developed in a dark room. Quantification of protein from three independent pulldowns was performed using ImageJ.

Protein Digestion and Labeling.

Protein Digestion. Proteins were isolated from HaCaT cells treated with microproteins for 0, 2, 4, 6 and 8 hrs or the microprotein affinity purification experiments described above. For microprotein treatment experiments, cells were washed with 1X PBS and lysed in a buffer containing: 75 mM NaCl, 3% sodium dodecyl sulfate (SDS), 1 mM NaF, 1 mM β-glycerophosphate, 1 mM $NaVO_4$, 1 mM NaPP, 1 mM phenylmethylsulfonyl fluoride, 1X complete EDTA-free protease inhibitor cocktail and 50 mM HEPES(23). Proteins were reduced and alkylated with dithiothreitol (DTT) and iodoacetamide (IAA), respectively (61), then methanol/chloroform precipitated. Proteins were re-solubilized in 1 M urea in 50 mM HEPES, pH 8.5. and digested in a two-step process (LysC and Trypsin) before being desalted with C18 Sep-Paks (62). *TMT Labeling.* Samples were labeled with TMT 10-plex reagents (6, 8) for multiplexed quantitative proteomics. For microprotein treatment experiments, TMT reagents 126 and 131 were reserved for “bridge channels”, and the remaining reagents were used to label digests in random order. Bridge channels consisted of 25 µg of each digest pooled together and then re-aliquoted into 50 µg portions for labeling. The bridge served as a means to control for experimental variation between mass spectrometry experiments. For microprotein affinity purification experiments, peptides were labeled directly following desalting to minimize sample loss. Labeling was conducted for 1 hr at RT and was

quenched by addition of 9 µl of 5% hydroxylamine. Samples were then acidified by addition of 50 µl of 1% TFA and pooled. The pooled, multiplex samples were desalted with C18 Sep-Paks as described above.

Reductive Dimethylation Labeling. Reductive dimethylation was conducted for beads-only and SAM1 microprotein affinity purification samples as previously described (62). Briefly, heavy (H) and light (L) formaldehyde (F) and sodium cyanoborodeuteride (R) solutions were prepared or purchased as follows; FH: 20% formaldehyde-d2, FL: 20% formaldehyde, RH: 5 M NaCNBd3 in 1 M NaOH, RL: 5 M NaCNBH3 in 1 M NaOH, and reagent buffer: 0.227 M Na2HPO4, 0.026 M citric acid, pH 5.39. Heavy and light wash solutions were made by combining 20 µl of FL + 12 µl of RL + 968 µl of reagent buffer. Immediately following digestion with trypsin, peptides were loaded onto C18 stage-tips, labeled by washing with the reagents made above, desalted, and pooled into appropriate duplex experiments prior to standard LC-MS2 analysis.

Basic pH Reverse-phase Liquid Chromatography (bRPLC) Fractionation.

Fractionation was carried out by bRPLC (63) with fraction combining as previously described (23, 62). Briefly, samples were solubilized in 110 µl of 5% formic acid in 5% acetonitrile and 100 µl was separated on a 4.6 mm x 250 mm C18 column on an Ultimate 3000 HPLC. The resultant 96 fractions were combined into 24 distinct fractions and dried prior to multiplexed LC-MS2/MS3 analysis.

LC-MS Analysis

Standard MS2 Method. Peptides were resupsended in 5% acetonitrile/5% formic acid and analyzed on an Orbitrap Fusion Tribrid mass spectrometer with an in-line Easy-nLC 1000 System. All data acquired were centrioded. Samples were loaded onto a 30

cm in-house pulled and packed glass capillary column (I.D. 100 µm, O.D. 350 µm). The column was packed with 0.5 cm of 5 µm C4 resin followed by 0.5 cm of 3 µm C18 resin, with the remainder of the column packed with 1.8 µm of C18 resin. Following sample loading, peptides were eluted using a gradient ranging from 11-30% acetonitrile in 0.125% formic acid over 85 min at a flow rate of 300 nL/minute and heating the column to 60 °C. Electrospray ionization was assisted by the application of 2,000 V of electricity through a T-junction connecting the column to the nLC.

MS1 spectra were acquired in data dependent mode with a scan range of 500-1200 m/z and a resolution of 60,000. Automatic gain control (AGC) was set to 2×10^5 with a maximum ion inject time was 100 ms and a lower threshold for ion intensity of 5×10^4. Ions selected for MS2 analysis were isolated in the quadrupole at 0.5 Th. Ions were fragmented using collision-induced dissociation (CID) with a normalized collision energy of 30% and were detected in the linear ion trap with a rapid scan rate. AGC was set to 1×10^4 and a maximum inject time of 35 ms.

Multiplexed MS3 Method. The multiplexed MS3 method settings are the same as the standard MS2 method above with the following additions. MS3 analysis was conducted using the synchronous precursor selection (SPS) option to maximize TMT quantitation sensitivity (9). Up to 10 MS2 ions were simultaneously isolated and fragmented with high energy collision induced dissociation using a normalized energy of 50%. MS3 fragment ions were analyzed in the Oribtrap at a resolution of 6×10^4. The AGC was set to 5×10^4 using a maximum ion injection time of 150 ms. MS2 ions 40 m/z below and 15 m/z above the MS1 precursor ion were excluded from MS3 selection.

Peptide Identification by Proteome Discoverer.

Peptidomic Experiments. Raw files were processed using Proteome Discoverer 2.1. MS2 data were queried against two databases: the TCH1516 UniProt database (downloaded from UniProt on 12/30/2015) and a six-frame translated database generated from the TCH1516 genome (downloaded from EBI on 08/03/2016) using the Sequest algorithm (64). The six-frame translated database was made in R, utilizing the “marray” and “seqinr” packages (65). A decoy search was also conducted with sequences in reverse order (3, 66, 67). For MS1 spectra, a mass tolerance of 50 ppm was used and for MS2 spectra a 0.6 Da tolerance was used. Oxidation of methionine (+15.99492 Da) and formylation of the peptide N-terminus (+27.995 Da) were included as variable modifications. Enzyme digest specificity was set to ‘unspecific’ to account for all endogenous peptide cleavage sites. Data were filtered to a 1% peptide and protein level false discovery rate using the target-decoy strategy (66). For subsequent label-free quantitation (LFQ) of microproteins, raw files were processed as described above except the data were searched against the TCH1516 reference proteome with the SAMs appended. MS1 area under the curve (AUC) were used for quantitation.

Proteomic Experiments. Resultant data files were processed using Proteome Discoverer 2.1. MS2 data were queried against the Uniprot human database (downloaded: 05/2017) using the Sequest algorithm (64). A decoy search was also conducted with sequences in reversed order (3, 66, 67). For MS1 spectra, a mass tolerance of 50 ppm was used and for MS2 spectra a 0.6 Da tolerance was used. Static modifications included TMT 10-plex reagents (+229.162932 Da) or dimethylation (+28.031300 Da) on lysine and peptide n-termini and carbamidomethylation of cysteines (+57.02146 Da). Variable oxidation of methionine (+15.99492 Da) was also included in

the search parameters. Data were filtered to a 1% peptide and protein level false discovery rate using the target-decoy strategy (3, 66, 67). For TMT experiments, reporter ion intensities were extracted from MS3 spectra for quantitative analysis. For reductive dimethylation experiments, relative abundances of peptides in the duplexed samples were quantified as the ratio of MS1 peak areas of heavy and light versions of the same peptide. Protein-level quantitation values were calculated by summing signal to noise values for all peptides per protein meeting the specified filters. Data were normalized in a two-step process as previously described (68). First, the values for each protein were normalized to the pooled bridge channel value (microprotein treatment experiment) or an *in-silico* average "bridge channel" of all the samples in the experiment (microprotein affinity purification experiments). Then, the values were normalized to the median of each reporter ion channel (for TMT) or sample MS1 AUC (for reductive dimethylation), as well as the entire dataset.

Results

Endogenous Peptides Detected in *S. aureus* Supernatants

To investigate the microproteins and peptides secreted by *S. aureus* TCH1516, a mass spectrometry (MS) based peptidogenomic approach was employed (**Fig 2-1A**). In an effort to capture microproteins that may not present as annotated ORFs, the TCH1516 reference genome was six-frame translated, creating a database that contains all possible ORFs with at least six amino acids. Using the TCH1516 Uniprot database and the custom six-frame translated database to match spectra, a total of 57 peptides represented by 115 peptido-forms were identified at a peptide and protein level false discovery rate of < 1%.

Categorizing peptides based on their subcellular location confirmed the detection of peptides derived primarily from membrane and secreted proteins with minimal cytosolic contamination (**Fig 2-1B**). Most of the peptides identified correspond to known virulence factors, with hemolysins and PSMs being highly abundant (**Fig 2-1C**). All previously characterized PSMs were detected, except for PSM-mec, which is not encoded in the TCH1516 reference genome(55). Among the other virulence factors, a high number of peptides from accessory gene regulator D (*agrD*), immunodominant antigens A and B, and various IgG binding proteins were detected. While some of these variants were previously described(69), most of them, to our knowledge, have never been reported. The majority of these peptides contain the classical SPaseI "AXA" motif, and are presumably derived from the signal sequences that target these proteins to the membrane or extracellular space(70). However, the full N-terminal signal sequences for any of these proteins were not observed, suggesting additional proteolytic processing following SPaseI cleavage. Interestingly, most of the detected cleavages occur immediately after a lysine residue, which may play a role in cleavage specificity. While primarily functioning to target proteins for the membrane or extracellular space, alternative functions for signal peptides have been proposed(71, 72). Additionally, it has been reported that other peptide fragments produced by *S. aureus* possess antimicrobial effects(73, 74). With this in mind, we hypothesized that cleaved signal peptides could retain some bioactivity. We tested this with the α-hemolysin signal peptide derivative, and found that it inhibited the growth of *Streptococcus pyogenes* strain M3 in a liquid culture assay at a similar concentration as other *S. aureus* antimicrobial peptide fragments(73). Thus, we conclude that these set

of endogenous merit further studies to elucidate the precise cleavage mechanisms and any additional biological functions of these putative signal peptide derivatives.

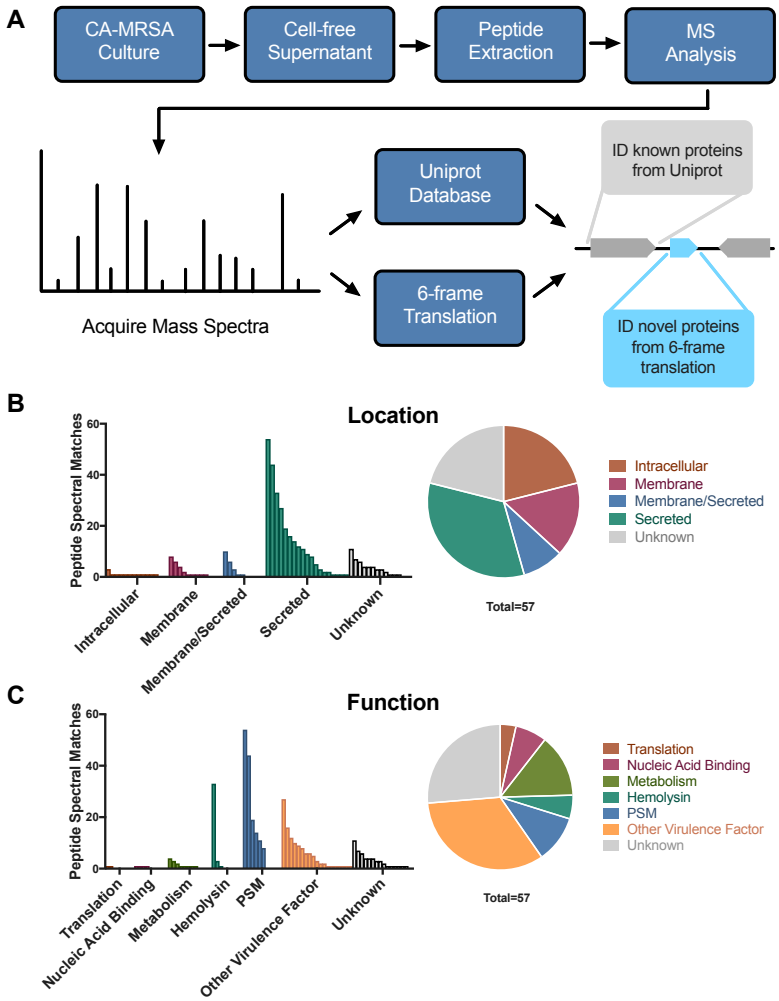


Figure 2-1 *Overview of Peptides Secreted by S. aureus.*

(**A**) Workflow for the identification of known and novel peptides secreted by *S. aureus* TCH1516. (**B**) Manual annotation of the locations of the peptides identified in this experiment. (**C**) Manual annotation of the functions of the peptides identified in this experiment.

Characterization of Novel *S. aureus* Microproteins

In addition to the repertoire of known microproteins and peptides derived from larger proteins, two previously unidentified microproteins were identified only in the six-frame translated database. These microproteins retain their formylated initiator methionine and are similar in length to the α-type PSMs. However, their genetic locus is distal from the α-type PSMs. We named these peptides *S. aureus* microprotein (SAM) 1 and 2. In total, 11 and 7 high-confidence spectra were detected from SAM1 and SAM2, respectively. These assignments were verified by manual annotation of representative spectra. Based on number of spectral matches, we estimate these peptides are produced in culture to a similar degree as the β-type PSMs(75).

An initial genetic analysis was carried out to investigate the evolutionary conservation of the SAMs. Basic local alignment search tool (BLAST) analysis revealed that the SAMs are restricted to the Aureus-Epidermidis group of *Staphylococci* (**Fig 2-2A**). While SAM1 is found in at least eight distinct species, SAM2 is only found in three, all on a single branch, suggesting more recent evolution. Notably, the strains encoding SAM2 are highly similar, with the distinction of *S. aureus* from *S. argenteus* and *S. simiae* only formally defined over the past 5 years. Thus, it appears that the SAMs are primarily encoded by more pathogenic *Staphylococci* (eg. *S. aureus* and *S. epidermidis*) rather than clinically irrelevant species (eg. *S. carnosus* and *S. equorum*). Furthermore, generating frequency plots uncovered residues in SAM1 (**Fig 2-2B**) and SAM2 (**Fig 2-2C**) that are fully conserved across all the encoding *Staphylococcal* species, potentially signifying that these residues are essential for microprotein function.

Mapping the SAMs to the TCH1516 reference genome shows that they are located proximal to each other, separated by a single lipase gene (**Fig 2-2D**). The SAMs contain comparable hydrophobic side chains, share 33.3% sequence homology, and a positive net charge. Performing multiple sequence alignments of the SAMs and α-type PSMs demonstrates the similarity of their primary sequences (**Fig 2-2E**). Furthermore, Monte Carlo simulations predict that the SAMs have amphipathic features and moderate helical propensity, which would assume a largely helical structure upon binding to a lipid bilayer (**Fig 2-2F**). These traits highlight the relatedness of the SAMs to the PSMs and indicate they may function in a similar manner.

The SAMs are Regulated by *agr*

The *agr* of *S. aureus* is a quorum sensing system that controls the expression of genes, such as the PSMs and other crucial virulence factors, depending on local *S. aureus* cell density(53). We hypothesized our newly identified microproteins may also be under the control of *agr*. By repeating our peptidomic experiments using an additional wild-type (WT) CA-MRSA strain (LAC) and an *agr*-deficient strain (RN9120), we observed that SAM1 was expressed at a similar level in both of the WT strains but completely absent in the *agr*-deficient strain (**Figure 2-2G**). To validate these findings, we performed qRT-PCR on WT TCH1516 and an isogenic *agr*-deficient strain and found that the expression of both SAM mRNA are drastically reduced in the absence of *agr* (**Figure 2-2H**). This observation is consistent in other strains of *S. aureus* that were tested. Regulation by *agr* lends further support to the idea that these microproteins are functionally important, particularly in times of high bacterial cell density. Based on the

above observations, synthetic copies of the SAMs and PSMα3 were obtained to characterize and compare the effects of these microproteins in parallel. A synthetic, formylated “mock-SAM” microprotein with a scrambled sequence was utilized as a negative control.

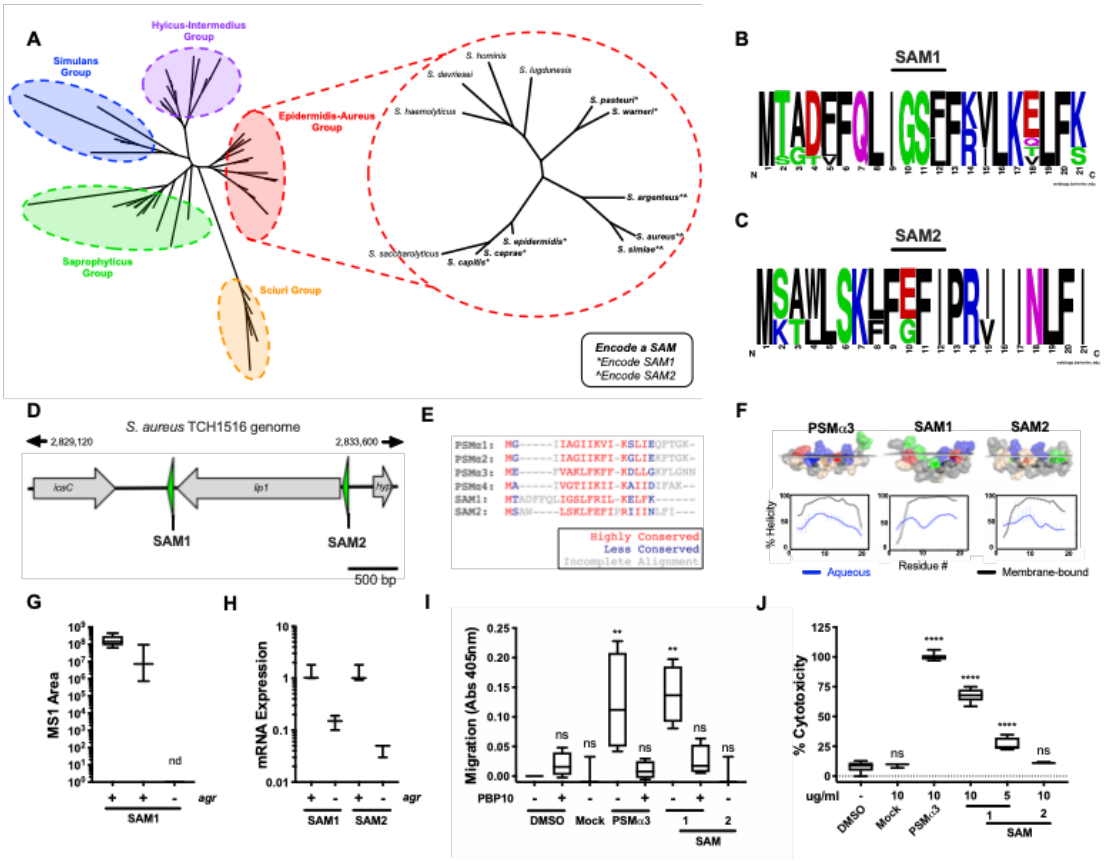


Figure 2-2 *Initial Characterization of Novel Microproteins*

(**A**) Conservation of SAMs across the *Staphyloccocus* phylogenetic tree. (**B**) SAM1 gene multiple sequence alignment across various staphylococcal species. (**C**) SAM2 gene multiple sequence alignment across various staphylococcal species. (**D**) Newly annotated SAM1 and SAM2 loci within the *S. aureus* TCH1516 genome (*icaC* – intercellular adhesion protein C; *lip1* – lipase 1; *hyp* – hypothetical protein). (**E**) Multiple sequence alignment of α-type PSM and SAM protein sequences. (**F**) *Top Panel:* Comparison of molecular structures of PSMα3 and the SAMs derived from Monte Carlo simulations in the presence or absence of a keratinocyte-like lipid bilayer. Aliphatic hydrophobic residues are shown in gray, aromatic hydrophobic residues in tan, neutral polar residues in green, positively charged residues in blue, and negatively charged residues in red. *Bottom Panel*: Predicted mean helicity per-residue of free peptides (blue) and in the presence of keratinocyte-like membrane (black). (**F**) Relative SAM1 MS1 intensity in *S. aureus* TCH1516, LAC and RN9120 supernatants (left to right; nd - not detected). (**G**) Relative SAM mRNA levels in TCH1516 WT (+) and *agr*::tetM mutant (-). Both genes are normalized to the reference gene *gyrB.* (**H**) Migration of neutrophils treated with Mock-SAM, PSMα3 or SAMs with or without pretreatment with the FPR2 inhibitor, PBP-10. (**I**) LDH release assay of HaCaT cells treated with Mock-SAM, PSMα3 or the SAMs. For all graphs, results are expressed as the median +/- the IQR (box) and max/min values (whiskers) of at least 3 independent experiments. Significance was determined by one-way ANOVA with Dunnett's multiple comparison test (****$p < 0.0001$; **$p < 0.01$; ns - not significant).

SAM1 Stimulates Neutrophil Migration and Lyses Human Keratinocytes *in vitro*

Neutrophils are highly abundant at the site of infection and function by recognizing, phagocytosing and eradicating pathogens by regulating an effective immune response(76, 77). It has been proposed that *S. aureus* actively recruits neutrophils via microprotein secretion for subsequent lysis and induction of tissue damage, a process that can be mediated by the PSMs (78). Therefore, we hypothesized the SAMs may also contribute to the excessive influx of neutrophils at the site of infection. Upon examination, SAM1 stimulated the migration of primary human neutrophils to a comparable extent as PSMα3 (**Figure 2-2I**). No appreciable migration was observed for SAM2 or the mock-SAM. Neutrophils express both a formyl peptide receptor, fMet-Leu-Phe receptor (FPR1), which recognizes various formylated peptides, and FPR1-related formyl peptide receptor 2 (FPR2), which is the major receptor for PSMs(78, 79). We observed that SAM1-induced migration was completely abolished when neutrophils were pre-treated with the selective FPR2 inhibitor PBP-10 (**Figure 2-2I**), suggesting that SAM1 induces neutrophil migrations in an FPR2 dependent manner.

Keratinocytes represent the major cell type of the human epidermal skin layer and comprise the first line of defense against invading pathogens(80). *S. aureus* in particular is a leading cause of skin and soft tissue infections (SSTIs) in the United States(36). Therefore, the toxicity of the SAMs was assessed in a keratinocyte cell line via an LDH release assay. We found that SAM1 treatment induced LDH release in a concentration-dependent manner while no lysis was observed for the mock-SAM or SAM2 (**Fig 2-2J**). The above results demonstrate that SAM1 appears to function in a similar manner to the α-type PSMs, while SAM2 may possesses distinct mechanisms of action.

SAM1's Cytolytic Ability is Inhibited by Serum and Enhances *S. aureus* Virulence

To confirm the cytolytic activity of SAM1, confocal fluorescence microscopy using SYTOX green staining was employed. We found that both PSMα3 and SAM1 induce high amounts of membrane permeability (**Fig 2-3A**), supporting the results in **Figure 2-2J**. As expected, mock-SAM did not show any significant alteration to the keratinocyte membranes. Recent studies on the α-type PSMs demonstrate that they are sequestered and inhibited by serum lipoproteins, suggesting they have an intracellular role(81, 82). Thus, we were curious if the presence of serum would also dampen the effect of SAM1 on host cells. We found that including serum in the treatment media with SAM1 drastically reduced host cell lysis (**Figure 2-3B**), indicating that the SAMs may also function intracellularly or at a body site devoid of serum lipoproteins (ie. the skin). Again, PSMα3 was used as a positive control and demonstrated the expected lytic phenotypes.

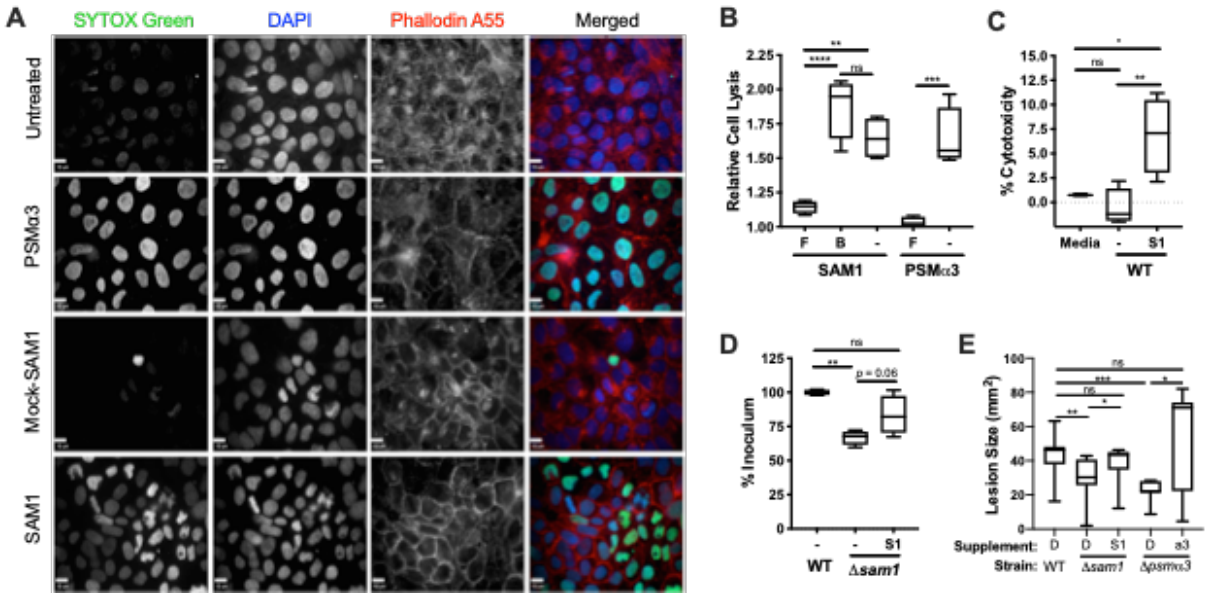


Figure 2-3 *SAM1 Lyses Keratinocytes and Contributes to S. aureus Virulence Phenotypes*

(**A**) Confocal microscopy images of HaCaT cells treated with PSMα3, Mock-SAM, and SAM1. (**B**) Relative cell lysis of HaCaT cells treated with SAM1 in the presence of FBS, BSA or no supplement (-). (**C**) Percent cytotoxicity of HaCaT cells treated with bacterial media or culture supernatants from WT *S. aureus* cells containing pRMC2-empty (-) or pRMC2-SAM1 (S1) induced with anhydrotetracycline. (**D**) CFUs recovered from *in vitro* infections of HaCaT cells infected with WT or Δ*sam1* containing an empty vector (-), or Δ*sam1* containing pRMC2-SAM1 (S1) *S. aureus* induced with anhydrotetracycline. (**E**) Lesion sizes from mouse skin infections with WT, Δ*sam1*, Δ*psmα3* bacteria supplemented with SAM1, PSMα3 or a vehicle control (DMSO). For all graphs, results are expressed as the median +/- the IQR (box) and max/min values (whiskers) of at least 4 independent replicates. Animal experiments contained 5 replicates. For B, C and D, significance was determined by one-way ANOVA with Dunnett's multiple comparison test (***p < 0.001; **p < 0.01; *p < 0.05; ns - not significant). For E, significance was determined by a Student's t-test (***p < 0.001; **p < 0.01; *p < 0.05; ns - not significant). Welch's correction was used for comparisons with unequal variance.

Given the above results, we made constructs that would allow the study of SAM1 produced from *S. aureus* in the context of other secreted protein/peptides, rather than a synthetic version of the peptide. Utilizing the inducible expression vector pRMC2(83), strains of WT *S. aureus* carrying either an empty vector or a SAM1-containing plasmid were generated. We found that inducing SAM1 expression resulted in significantly more LDH release from human keratinocytes than the empty vector or media treated controls (**Figure 3C**). Subsequently, we generated an isogenic deletion strain for SAM1 to determine whether loss of this factor would reduce *S. aureus* virulence. We challenged this strain in an *in vitro* skin infection model and found that it had reduced fitness relative to the WT strain (**Figure 3D**). This reduction in bacterial CFUs could be rescued using the inducible vector constructed above, supporting SAM1's role in this process (**Figure 3D**). The above results indicate that the data generated from the synthetic microprotein were not simply features of the purified molecule and that SAM1 may function during infection.

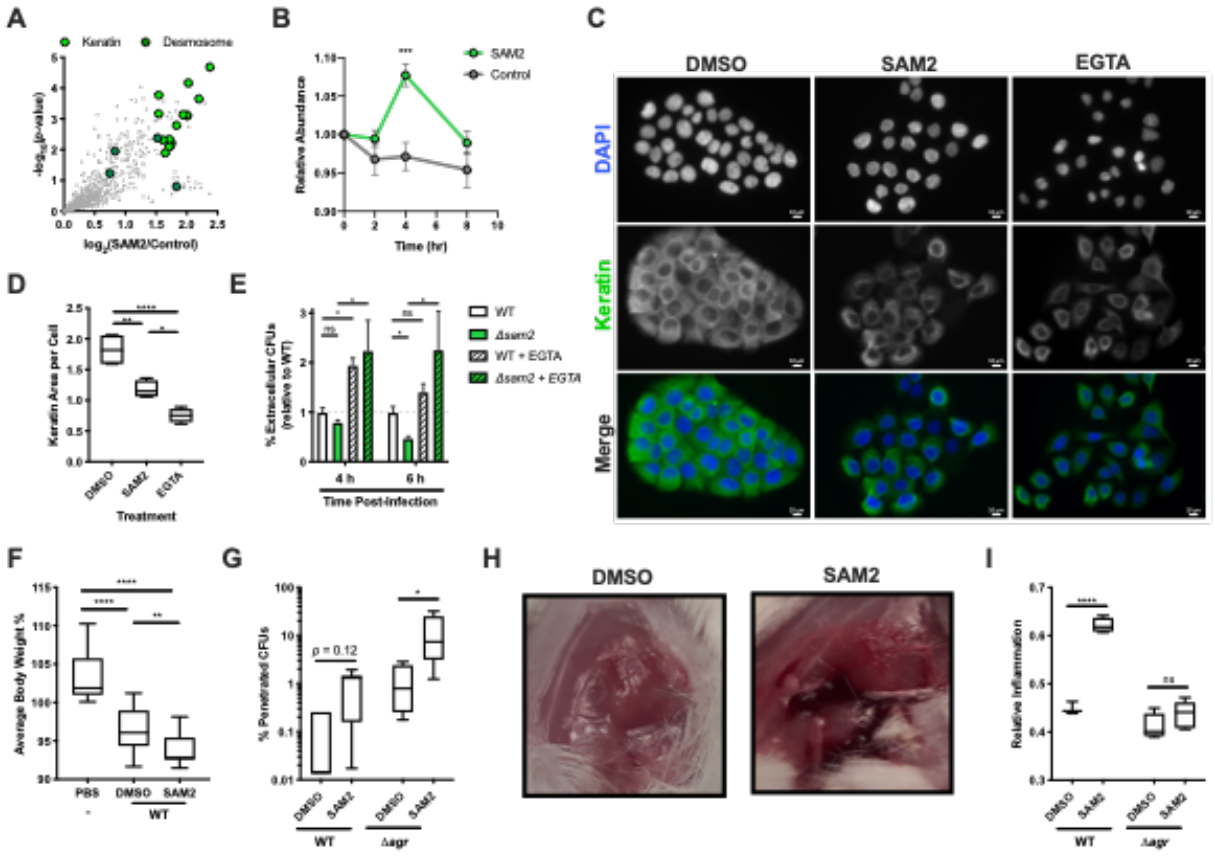
Next, we aimed to assess the impact of the SAMs *in vivo*. Therefore, we performed murine skin infection experiments using a isogenic deletion strains that were supplemented with microproteins as previously described(69). WT *S. aureus* and PSMα3 isogenic deletion strains were used as controls. We found that deletion of SAM1 resulted in significantly smaller lesion sizes compare to WT controls, similar to the effects observed upon deletion of PSMα3. As expected, supplementation of each deletion strain with its corresponding microprotein were able to rescue the lesion size defect. Overall, these results demonstrate a significant contribution of SAM1 to *S. aureus* pathology *in vitro* and *in vivo*.

SAM2 Interacts with and Disrupts Keratin Networks *in vitro*

Given that we could not assign a function to SAM2 using standard assays, quantitative proteomic techniques were employed to gain insight into potential mechanisms of action. First, we characterized the host protein interactome of SAM2 using a biotinylated microprotein as bait (**Table S3**). We found that SAM2 interacted highly with keratin intermediate filaments as well as related desmosome proteins (**Figure 4A**). As controls, we assessed the interactions of host proteins with PSMα3 and SAM1 and noted that host interactome of each microprotein is highly specific (**Table S3**, **Figure S4A**). It was observed that PSMα3 interacts significantly with proteasome and Rab proteins (**Figure S4B**), while SAM1 primarily interacts with prohibitin (PHB) proteins and karyopherins (**Figures S4C, G-H**). In parallel, we assayed the temporal protein dynamics of keratinocytes treated with SAM2 for a 2, 4 and 8 hours (**Table S4**). Supporting our interactome results, it was found that SAM2 significantly perturbs keratin protein expression (**Figure 4B**). Again, we also performed these experiments using PSMα3 and SAM1 and determined that each microprotein differentially altered the host proteome (**Table S4, Figure S4D**) in ways that validate the interaction data (**Figures S4E-F**). Taken together, these unbiased analyses corroborate each other and suggest a specific mechanism of action for SAM2, disruption of the keratin network. Furthermore, these results highlight the differential effects of *S. aureus* microproteins on host cells and offer PSMα3 and SAM1 interaction candidates for subsequent biological validation.

Figure 2-4 *SAM2 Disrupts Keratin Networks and Enhances S. aureus Penetration into Underlying Tissue from Skin Lesion.*

(**A**) Volcano plot of potential host interactors of SAM2 depicting the $\log_2$(fold-change) for each protein on the x axis and the $-\log_{10}$(p-value) for each protein on the y axis. Comparisons were performed for SAM2 vs. SAM1, PSMα3 and a beads-only bait (Control). A Student's t-test was used to calculate p-values and Welch's correction was used for comparisons with unequal variance. Keratin (light green) and desmosome proteins (dark green) are highlighted. (**B**) Relative abundance of keratin proteins during the treatment of HaCaT cells with SAM2. (**C**) Confocal microscopy of HaCaT cells treated with DMSO, SAM2 and EGTA. (**D**) Relative keratin area per cell of HaCaT cells treated with DMSO or SAM2 compared using a Student's t-test. (**E**) Relative extracellular CFUs recovered from *in vitro* intracellular infections of HaCaT cells infected with WT, Δ*sam1*, and Δ*sam1/2 S. aureus* and Δ*sam1/2 S. aureus* infections treated with EGTA. Results are expressed as the mean +/- the standard error of the mean of at least 5 independent replicates and compared using a two-way ANOVA with Tukey's multiple comparison test. (**F-K**) Mice were infected intradermally with WT or Δ*agr S. aureus* supplemented with DMSO or SAM2. A PBS mock infection was used as a control. (**F**) Body weight of mice during the WT infection time course. Results are expressed as the mean +/- the standard deviation of 5 independent replicates and compared using a two-way ANOVA with Tukey's multiple comparison test. (**G**) Average body weight for mice on each day of the WT infection time course compared using a paired t-test. (**H**) Percent of bacterial CFUs that penetrated the underlying tissue compared using a Mann-Whitney test. Representative photographs of underlying tissue from infections supplemented with (**I**) DMSO or (**J**) SAM2. (**K**) Relative inflammation of underlying tissue compared using a two-way ANOVA with Tukey's multiple comparison test. Unless otherwise noted, results are expressed as the median +/- the IQR (box) and max/min values (whiskers) of at least 4 independent replicates (****$p < 0.0001$; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; ns - not significant).


A
Keratin
Desmosome
-log10(p-value)
log2(SAM2/Control)
B
SAM2
Control
Relative Abundance
Time (hr)
C
DMSO
SAM2
EGTA
DAPI
Keratin
Merge
D
Keratin Area per Cell
DMSO
SAM2
EGTA
Treatment
E
% Extracellular CFUs (relative to WT)
WT
Δsam2
WT + EGTA
Δsam2 + EGTA
4 h
6 h
Time Post-Infection
F
Average Body Weight %
PBS
DMSO
SAM2
WT
G
% Penetrated CFUs
p = 0.12
DMSO
SAM2
WT
Δagr
H
DMSO
SAM2
I
Relative Inflammation
DMSO
SAM2
WT
Δagr

To validate the mechanism of SAM2, we interrogated any visible effects that SAM2 had on the host keratin network. Using a pan-keratin fluorescent antibody, we observed that treatment with SAM2 induced perturbations of the keratin network compared to a vehicle treated control (**Figure 4C**). This resulted in a significant decrease in the average keratin area per cell (**Figure 4D**). The observed disruption is similar to an effect observed from EGTA treatment (**Figure 4C**), a known keratin perturbing agent(84, 85). These microscopy results substantiate our proteomics findings and further support the mechanism of SAM2.

SAM2 Contributes to *S. aureus* Escape from Host Keratinocytes *in vitro* and Dissemination from Skin Lesions *in vivo*

As mentioned above, keratinocytes play a major barrier role in defending against invading pathogens. We hypothesized that the disruption of host keratin by SAM2 could facilitate *S. aureus* penetration through this barrier. Therefore, we tested whether or not SAM2 contributes to *S. aureus* invasion into, or escape from, keratinocytes using an intracellular infection assay. Using an isogenic deletion strain of SAM2, we observed that there was no difference in keratinocyte invasion, but strains lacking SAM2 could not escape host cells as readily (**Figure 4E**). This defect in exiting host cells came along with a slight, albeit not statistically significant, reduction in host cell cytotoxicity, suggesting it is a rather subtle process. As the single Δ*sam1* mutant was able to escape to the same extent as the WT strain, we believe this process mediated through the interaction of SAM2 with the keratin-desmosome network. To interrogate this hypothesis further, we intentionally disrupted the keratin network via EGTA treatment (**Figure 4E**) and found that

it rescued the escape defect, providing additional evidence that the keratin network contributes to containment of *S. aureus* intracellularly.

The *in vitro* infection results above, paired with our initial *in vivo* microprotein supplementation experiments using SAM2 (**Figure S3**), demonstrate that this microprotein may contribute to bacterial penetration through the epidermal layer into the underlying tissue. To test this hypothesis directly, we repeated the microprotein supplementation experiment with SAM2 and assessed mouse weight during the course of infection. Both a WT and Δ*agr* strain were used in these experiments. At the end of the time course, both the lesions, as well as the underlying tissue, were imaged and processed for CFU enumeration. We found that *S. aureus* supplemented with SAM2 induced a more significant reduction in body weight than their vehicle control counter parts (**Figure 4F**). For nearly every day of the time course, the weight of the SAM2 supplemented mice was noticeably less than the control mice (**Figure 4G**). Quantifying CFUs in both the lesion and the underlying tissue revealed that, when *S. aureus* is supplemented with SAM2, bacterial cells can penetrate into the mouse tissues more effectively (**Figure 4H**), likely resulting in the observed exacerbation of weight loss. To support this hypothesis, we photographed the underlying tissue from these experiments (**Figures 4I and J**) and observed an increase in damage and inflammation (**Figure 4K**). Overall, these results lend further support to SAM2 acting upon keratin networks and demonstrate that the increased production of this peptide at the lesion can enhance the pathogenicity of *S. aureus*.

Discussion

Our results highlight the potential of applying peptidogenomics to discover new bioactive molecules in a prevalent, highly-studied pathogen, *S. aureus*. The majority of peptides identified in this study originated from known, secreted virulence factors, such as PSMs(40), hemolysins and variant forms of *agrD*(69). Many of the endogenous peptides that were detected are derived from larger protein ORFs that possess the SPaseI “AxA” cleavage motif(70). It is unclear why these putative signal peptide fragments were consistently identified when the remainder of the signal sequences were not detected. High concentrations of signal peptides are known to impair proper protein secretion in *Escherichia coli*(86). If signal peptides are no longer needed following cleavage or are actively removed to ensure proper protein secretion, why do these peptides remain readily detectable? While it is possible that post-translational modifications may have made it difficult to match spectra to the N-terminus of the signal peptides, we cannot rule out the possibility that these C-terminal signal peptide fragments retain bioactivity and are stably expressed for a purpose. As we have determined that the α-hemolysin signal peptide derivative has antimicrobial effects, this dataset offers additional peptides to follow up with biological assays. Further studies are needed to fully assess the function of these molecules in the context of *S. aureus* pathology.

In addition to peptides that mapped to predicted ORFs, the discovery of two novel, conserved microproteins, SAM1 and SAM2, originating from previously unannotated smORFs is reported. These factors were only identified when using a six-frame translated genome database to match mass spectra. We believe this approach results in a more comprehensive survey of complex systems when compared to only using a predicted-

ORF database and will yield intriguing results when applied to other pathogens. Of note, these two new microproteins highly resemble the α-type PSMs in their primary sequence, predicted secondary structures and regulation by *agr*.

Stemming from this data, we investigated whether the SAMs possess bioactivity against human cells. While we could not assign a function to SAM2 with our initial assays, it was observed that SAM1 stimulated neutrophil migration and lysed human keratinocytes. We further found that the SAM1 induction of neutrophil migration functions through an FPR2-dependent mechanism, similar to the previously reported *S. aureus* microproteins(78). This data demonstrates redundant mechanisms for neutrophil recruitment by a variety of microproteins produced by *S. aureus*. Given that the same receptor recognizes all of these bacterial factors, this mechanism may be exploited in the development of targeted therapeutics. Additionally, we found that SAM1 readily caused lysis of human keratinocytes, in both a synthetic form and when overexpressed in culture, and that this activity could be inhibited by serum. These data argue that SAM1 functions in areas where there is no serum, such as on the skin or within human cells. Furthermore, the increased expression of SAM1 during *in vitro* infections could rescue an observed Δ*sam1* deficiency in propagation, indicating it is important for *S. aureus* pathogenesis. We validated these results in animal studies where supplementing skin infections with SAM1 resulted in the exacerbation of *S. aureus* virulence phenotypes. These results constitute the current knowledge pertaining to SAM1 bioactivity and suggests that this microprotein can effectively enhance *S. aureus* pathology.

As we could not define any SAM2 bioactivity through standard *in vitro* experiments, we employed a quantitative proteomics approach to investigate its mechanism(s) of

action. We found that SAM2 could interact with and perturb the keratin networks *in vitro*, which appears to facilitate the escape of *S. aureus* from human keratinocytes. We followed up these studies by demonstrating that the supplementation of SAM2 in skin infections *in vivo* resulted in increased bacterial penetration into the underlying tissue and an exacerbation of disease burden. Together, these studies suggest that SAM2 can disrupt keratin networks in the skin, enabling dissemination of the bacteria from the lesion. To our knowledge, this is the first *S. aureus* virulence factor reported to function in this manner. It has not escaped our notice that this mechanism may play a role in the transition of *S. aureus* from a commensal microbe, asymptomatically carried by up to 30% of the population, into an invasive pathogen. Thus, the further interrogation of this process can lead to the development of therapeutics to prevent *S. aureus* dissemination from its residence on the skin, or from superficial lesions, into the organism, effectively impeding this crucial step in pathogenesis.

Distinct and diverse functions appear to be a hallmark of the PSM-family as different members and variants have been linked to cytotoxicity(40, 87), immune stimulation(40, 87), antimicrobial effects(73, 74), and biofilm formation(54). This dogma is further supported by the distinct bioactivity and mechanisms found for the SAMs. Thus, our study paves the way for additional experimentation to fully assess the importance of the SAMs during infection and to further delineate mechanisms of action. Additional questions into the basic biology regarding these microproteins, as well as their clinical relevance, can be raised. For example, what are the production levels of these microproteins across different *S. aureus* strains? Are they solely dependent on *agr* or are there other factors influencing their expression? Is their secretion dependent on the same

machinery as the PSMs(88)? In the realm of clinical importance, what other virulence phenotypes do these microproteins manifest? Can we block the activity or secretion of these peptides as an anti-virulence therapeutic strategy? Are the interactions of PSMα3 and SAM1 with their respective host proteins clinically relevant? Further studies into these microproteins should be performed to address these questions, among others. In conclusion, our investigations into the *S. aureus* peptidome is highlighted by the intriguing bioactivity of two novel *S. aureus* microproteins, SAM1 and SAM2. The additional characterization of these molecules, as sole entities and in conjunction with other virulence factors, can contribute to a better understanding of *S. aureus* pathogenesis and lead to the development of new therapeutics.

Chapter 2, in part, is a reprint of material submitted to *Cell Reports*, 2020 (*manuscript in revision*), Jacob M. Wozniak, Julieta Aguilar., Dominic McGrosso, Igor H. Wierzbicki, Eri Nakatani-Webster, Michael R. Dores, Katrin Schilcher, Anvesh Macherla, Diana Dehaini, Xiaoli Wei, Ronnie H. Fang, JoAnn Trejo, Brian J. Werth, Abhinav Nath, Ross Corriden, Liangfang Zhang, Alexander R. Horswill, and David J. Gonzalez. The dissertation author was the primary author of this paper.

Chapter 3 – A Computational Tool to Analyze PTM-containing Proteomic Data and Applications to a Chronic Chagas Disease Model

Introduction

As mentioned above, a key advantage of MS-based proteomics is the ability to detect protein post-translational modifications (PTMs) and localize them to specific amino acid residues. These approaches have led to many significant findings in a wide range of biological disciplines, from developmental biology to cancer and infectious diseases. However, PTMs in proteomic data are often ignored in standard search algorithms and there is a current lack of tools available to connect raw PTM site information to biologically meaningful results in a high-throughput manner. Furthermore, many of the available tools require significant programming knowledge to implement.

Here, we describe the R package *PTMphinder*, which enables a more complete analyses of PTM-containing proteomic data. To demonstrate functionality of this tool, we apply it to an *in vivo* infection model, hearts isolated from mice with chronic Chagas disease. We utilize *PTMphinder* to localize phospho-sites within their respective full-length proteins and extract flanking amino acid sequences. These extracted sequences are readily input into additional computational tools to identify enriched motifs, predict causal kinases and suggest potential therapeutics.

Chagas disease is the manifestation of an infection by the protozoan parasite *Trypanosoma cruzi*. First described in 1909 by a Brazilian physician, Carlos Chagas(89), this disease is a significant health concern, particularly in areas with low socio-economic status. Facilitated by human population flow, Chagas disease has spread out of endemic areas into more developed countries, with more than 100,000 cases in Europe(90) and 200,000 - 300,000 in the United States(91, 92) reported. Infected insect vectors and congenital transmission are the most common means of disease spread, accounting for

up to ~96% of recorded cases (70% insect vectors, 26% congenital); blood transfusion, organ transplantation and consumption of contaminated foods also contribute to *T. cruzi* dissemination(93). Historically overlooked, Chagas disease is classified as a neglected tropical disease by the world health organization(94) and is estimated to result in a global economic burden of > $7 billion (USD) per year(95). Thus, Chagas disease is a major human health concern that causes significant morbidity and mortality worldwide.

The progression of Chagas disease can be classified into two phases, the acute phase and the chronic phase(96, 97). The acute phase is asymptomatic in most cases, lasts approximately 1 - 2 months and usually resolves spontaneously(96). However, if left untreated, patients can remain chronically infected, resulting in critical health concerns later in life(96). These delayed adverse effects occur in approximately 30% of the infected individuals and include cardiac and visceral involvement, with cardiomyopathies being the most severe and frequent manifestation(96, 97). Interstitial fibrosis of the heart is thought to be a major determinant factor for the pathogenesis of Chagas disease(93). In fact, even after successfully lowering parasite loads with the current standard of therapy (ie. benznidazole), patients with advanced cardiomyopathies remained under high disease burden(97). The reason for this is currently unclear, but suggestions have ranged from auto-immune responses(98, 99) to dormant, low-proliferating forms of *T. cruzi* that are resistant to anti-trypanosomals(100). Regardless, directing therapies against fibrotic phenotypes of heart, in combination with trypanocidal agents, have great potential to effectively treat this disease.

Chronic Chagas disease in the heart is driven by an intense inflammatory response and excessive immune infiltration(93, 101). Cytokines and chemokines, secreted by both

cardiomyocytes(102) and invading immune cells(103), stimulate a wound-healing response (eg. extracellular matrix (ECM) deposition) from fibroblasts to repair the damaged tissue(93). Specifically, transforming growth factor β (TGF-β)(104), tumor necrosis factor-α (TNF-α)(102), and interferon gamma (IFN-γ)(103) are central to the immune response and pathology of Chagas disease. These effectors induce a myriad of downstream signaling cascades, resulting in diverse functional outcomes from apoptosis(105) to accumulation of ECM(104, 106). The potential intracellular signaling pathways include c-Jun N-terminal kinase (JNK)(107) and p38(108). Recent reports from *in vitro* models of *T. cruzi*-host cell interaction demonstrated that cardiac fibroblasts display an increase in phosphorylation of p38 and c-Jun after infection(109). Transcriptomic analysis showed that *T. cruzi* infection upregulates the JUNB gene and results in translocation of JunB to nuclei of primary human cardiomyocytes(107). In line with this hypothesis, treatment of mice with genistein(110), a tyrosine kinase inhibitor, lowered TAK1 and JNK activity and decreased cardiac fibrosis in a hypertension model(111), suggesting this pathway is associated with cardiac remodeling. Further, SP600125(112), a canonical JNK inhibitor, is well tolerated by mice and provides protective effects for hearts in damaging conditions(113). Despite these developments, Chagas disease progression results in alterations of other signaling pathways(108, 114) that should not be overlooked as therapeutic targets. (23). A phospho-proteomic approach applied to Chagas disease models in the heart has not been attempted to our knowledge and thus has the potential to deepen the understanding of global signaling pathways affected by chronic *T. cruzi* infections.

In this study, we apply a phospho-proteomic workflow to interrogate chronic Chagas disease progression. This investigation constitutes a foundational examination of the global phospho-signaling response to *T. cruzi* in the primary affected organ, the heart. Our analyses uncover both known and previously uncharacterized alterations in total protein abundance and phosphorylation status. As expected, we captured the classical induction of IFN-mediated signaling pathways(103) and repression of mitochondrial function(115, 116). Of significance, our unbiased approach identified new players that may have a role in disease progression including Immunity Related GTPase M (IRGM) 1 and 2 and the immune-associated, guanylate binding proteins (GBPs). In addition to total protein abundance changes, we uncovered a vast signaling network of plasma membrane and intermediate filament proteins with perturbed phosphorylation status following infection. These include new targets such as Striated Muscle Enriched Protein Kinase (SPEG), Tensin 1, BCL2 Associated Athanogene 3 (BAG3), Sorbin and SH3 domain-containing protein (SORBS) 1/2 and myosin-family proteins in addition to the previously described p38 axis. Further, we applied bioinformatics to predict active kinases, supporting the involvement of JNK and identification of new activated (DYRK2 and AMPKA2) and repressed (casein kinase family) kinases in the host response to *T. cruzi* infection. Finally, through the creation of a druggable disease network, we propose a number of FDA approved drugs that may be repurposed for the treatment of Chagasic cardiomyopathy. Overall, this study reveals new signaling pathways modulated during chronic Chagas disease that expand the understanding of molecular mechanisms of pathogenesis and inform rational drug design.

Methods

PTMphinder

Data Sources. The example Human proteome used in this study was downloaded from Uniprot on 12/12/18 (https://www.uniprot.org/proteomes/UP000005640). The example shotgun proteomics data used in this study was published previously(23). The rmotix package(30) was downloaded from GitHub (https://github.com/omarwagih/rmotifx).

PTM Localization Methodology. All software was written in the R programming language. The algorithm for PTM localization and motif extraction receives the experimentally determined peptide sequence, total number of modifications, and potential PTM locations and scores as input. The program then extracts the represented proteins from a parsed proteome database, scans the full protein sequences for the detected peptides, and outputs the location in the full protein along with the flanking sequences. The extracted data is formatted into a new data table that can be exported or analyzed further within the R framework.

PTM Localization Validation. For initial validation of the package, a pseudo-dataset (**Supplemental Table 1**) was created using Microsoft Excel to match the format of the phindPTMs function input. Second, a previously published phospho-proteomic dataset was used to validate the use of the package on real-world data. The second dataset contained PTM information in the form of ptmRS output, which has to be modified prior to input into phindPTMs. The data was reformatted to match the phindPTMs function input using the text-to-column and find/replace functions in Microsoft Excel. This reformatted

input table is provided as an example file with the *PTMphinder* package (phindPTMs_Input_Example.csv).

Assessment of Time of Motif Extraction. The time of motif extraction was measured for manual extraction and using *PTMphinder* with a previously published data set(23). First, the time of extraction per peptide was measured when manually extracting flanking sequences for five randomly selected peptides using standard spreadsheet and text editor software. Next, the time of extraction per peptide was measured when using *PTMphinder* within R. The starting point for each method was the same input file (phindPTMs_Input_Example.csv) and protein database (Human_Uniprot_Parsed_Example.txt).

Code Availability. An open source version of *PTMphinder* is freely available for download from GitHub (https://github.com/jmwozniak/PTMphinder).

Chronic Chagas Disease Model and Assessment

Animals. Six week old female C57Bl/6 mice, weighing 16 - 18 g were used for the animal experiments. Mice were housed and kept in a conventional room at 20 to 24 °C under a 12 hour (hr)/12 hr light/dark cycle. The animals were provided with sterilized water and chow ad libitum. All procedures involving animals were approved by the Institutional Animal Care and Use Committee at UCSD (Protocol #14187 - Jair L. Siqueira Neto).

Chagas Disease Infection Model. T. cruzi Sylvio X10/4 was maintained in C2C12 myoblasts culture. After 5-7 days of passage of cells and parasites, trypomastigotes released in the supernatant were collected after centrifugation for 15 minutes (min) at 3300 rpm, re-suspended in Dulbecco's Modified Eagle Medium (DMEM), and used to

infect mice by intraperitoneal injection with $1x10^6$ trypomastigote form of *T. cruzi*/mouse. The infected groups were age and sex matched with uninfected controls and kept under the same conditions. The general health of the mice was evaluated weekly for one year. At 1-year post infection, the development of heart disease was monitored by electrocardiography at Seaweed Canyon Cardiovascular Physiology Laboratory, Institute for Molecular Medicine, UCSD. After ECG analysis of the mice, heart tissue was collected for further analysis.

Surface ECG. Adult mice were anesthetized with isoflurane (5% induction, 1-1.5% maintenance in 100% oxygen) and placed on a warming pad (35 °C – 37 °C). Needle electrodes made of 27-gauge needles were inserted subcutaneously into each of the four limbs and the chest area. Simultaneous standard ECG leads I, II, and chest leads were recorded at a frequency response of 0.05–500 Hz. The signal was digitized and recorded at 2,000 Hz on LabChart (ADinstruments).

Histology. Upon euthanasia, the hearts were collected from the mice and cut in half in the sagittal orientation, placed in cryomolds, embedded in Tissue-Tek (O.C.T., Sakura Finetek) and snap frozen in liquid nitrogen. Samples were sectioned in a cryostat, fixed in buffered formalin and stained with hematoxylin and eosin or Sirius Red/Fast Green to stain collagen fibers. The slides were scanned using Nanozoomer Slide Scanner (Hamamatsu Photonics, NJ, USA) and images were obtained through NDP viewer software (Hamamatsu Photonics, NJ, USA).

Histopathology analysis. Levels of inflammation and fibrosis were quantified as previously described(117). Briefly, 5 random images of mouse hearts (10X magnification) were obtained from each animal, comprising most of the heart section area. Lymphocyte

nuclei was segmented through the Particle Analyzer Image processing plugin from FIJI software(118), and lymphocyte nuclei were counted. To measure fibrosis, the red staining of collagen fibers was segmented through color thresholding, and the area fraction of collagen was measured in the generated binary image after segmenting.

Proteomic LC-MS^2/MS^3 Analysis

Proteomics analyses were conducted as described above with the following alterations.

Tissue Lysis. Heart tissue was homogenized in a buffer consisting of 3% SDS, 75 mM NaCl, 1 mM β-glycerophosphate, 1 mM NaF, 1 mM Na_3VO_4, 10 mM $Na_4P_2O_7$, 1 mM phenylmethanesulfonyl fluoride and 1X Roche cOmplete mini EDTA free protease inhibitor in 50 mM HEPES, pH 8.5. Homogenization was conducted using a bead beater 3X for 1 min each time with a 1 min rest in between each session. Homogenates were sonicated for 5 min in a water bath sonicator to ensure complete lysis. Lysate supernatants were transferred to a new tube and any remaining cellular debris was pelleted via centrifugation (5 min, 16,000 x *g*, 4 °C). The resulting, clarified supernatant was processed for downstream analysis.

Enrichment of Phospho-peptides. Phospho-peptides were enriched by TiO_2 beads as previously described(119). Briefly, the following buffers were made. Binding buffer: 2 M lactic acid, 50% acetonitrile; wash buffer: 50% acetonitrile/0.1% trifluoracetic acid; elution buffer: 50 mM KH_2PO_4, pH 10. TiO_2 beads were washed (once with binding buffer, once with elution buffer and twice with binding buffer). Peptides were resuspended in binding buffer, mixed with beads at a ratio of 1:4 (peptides to beads) and vortexed at RT for 1 hr. Beads were then washed three times with binding buffer, followed by three times

with wash buffer. Phospho-peptides were eluted from the beads with elution buffer (two 5 min incubation while vortexing). Enriched peptides were desalted with solid-phase extraction columns then lyophilized and stored at −80 °C until they were labeled for quantitation.

Data Processing and Analysis. Variable phosphorylation of serine, threonine and tyrosine residues were also included in the search parameters. Phospho-peptides were normalized the same with the following changes. First, phospho-peptides were summed to the unique peptide level rather than protein level. Second, phospho-peptide abundance was normalized to their respective total protein abundance from the standard proteomic experiments. Pi score(120) was used as a final measure of significance. Gene ontology(121, 122) and functional protein association networks(123) were used to identify enriched groups of related proteins. PTMphinder(124) was used to extract protein locations and flanking motifs and motif-x(30, 31) was used to identify enriched motifs. GPS(125) was used to predict active kinases and their corresponding sites on target proteins.

Results

PTMphinder

Format a Uniprot Database. The parseDB function reformats a proteome database for subsequent input into the phindPTMs and extractBackground functions (**Fig 3-1**). This function extracts the protein ID and full protein sequence from the provided database and creates a data table with those two columns, respectively. The resultant data table can be further analyzed with R using various functions or written to a file and opened in a text

or spreadsheet editor. This function is capable of handling two of the most common proteome databases used, Uniprot and RefSeq, as well as homemade databases provided the protein ID is the only value in ".fasta" description lines. Users can specify their database sources using the database source option ("db_source"). We recommend that users remove any redundant, reverse or contaminant sequences from the database prior to parsing. However, for convenience, we have included an option to remove any entries that start with "reverse" or "contaminant" if the filter option ("filt") is set to TRUE. We have provided both an unformatted and parsed Human Uniprot database as examples.

Localize PTM sites and Extract Motifs. Many proteomic workflows export PTM data that is difficult to manipulate and format properly for subsequent analysis. The phindPTMs function reads in experimentally acquired proteomic data, as well as a parsed reference database, to localize the detected PTMs in their respective proteins and extract relevant motifs. Users should re-format the proteome database used to search the mass spectrometry data with the above parseDB function prior to running phindPTMs. The experimentally acquired proteomic data must have the following six columns: Identifier, Protein_ID, Peptide_Seq, Total_Sites, PTM_Loc, and PTM_Score. Column headers need to be spelled exactly as above for proper import of data. For peptides with multiple modifications, the PTM_Locs and PTM_Scores should be separated by a semi-colon (";"). phindPTMs returns a data table with the following eight columns: Identifier, Protein_ID, Pep_Loc, Prot_Loc, Score, Flank_Seq, Ambiguity, Prot_Seq. The location of the PTMs in the full protein (Prot_Loc) can be further analyzed in the context of previous knowledge regarding that site using tools specific to the PTM of interest, such as

PhosphoSitePlus(126) or KinaseNet (http://www.kinasenet.ca/) in the case of phosphorylation events. The flanking sequences of each modified site (Flank_Seq) can be used as input into the a motif-finding tool (such as motif-x, see below) for subsequent motif enrichment analysis.

Validation of the phindPTMs Function. To validate the results of the phindPTMs function, two approaches were undertaken. First, a hypothetical pseudo-dataset was generated, which contained a variety of PTMs and scores that resemble the output from common proteomics software. This data set was input into phindPTMs function and the results were manually verified. We found that phindPTMs correctly localized and extracted the flanking sequences of 100% of the modified sites contained within the pseudo-dataset. Next, a real-world, phospho-proteomic data set was analyzed using the phindPTMs function. Again, manual validation found that the software correctly localized and extracted the flanking sequences of 100% of the modified sites. These results demonstrate that users can confidently trust the output of the package using data containing various PTMs and from real-world experiments.

Extract Proteome-specific Background Motifs. In order to properly perform motif enrichment analyses, a set of background motifs needs to be provided for statistical reasons. The motifs around every occurrence of a particular amino acid are commonly used as background for this purpose. However, many motif tools have limited proteome backgrounds to choose from(29, 31) or simply use 10,000 random sites(30). Using another species background or a small background set may result in abberent enrichment results. By using the extractBackground function contained in this package, users can generate their own background datasets from any proteome database. Furthermore, this

function gives the user the flexibility to specify the width and central character of the background motifs. This tool simply reads in a parsed proteome database (from the parseDB function above), the central character and the width of the desired motifs. The output can be directly input into a motif enrichment tool, such as motif-x, as a list of background sequences (see below).

Integration with Existing Tools. As with any new software, intergration with existing tools drastically increases its applicability to a variety of problems and accessibility to different users. Therefore, we designed the output of our tool to interface directly with existing tools for motif enrichment analysis, such as motif-x. With minimal reformatting, a user can simply use the output from the phindPTMs and extractBackground functions as input into motif-x (**Figure 1**). Fully-documented code to implement this workflow with the example files is provided as an ".R" script. This example uses real-world phospho-proteomic data previously published from our lab(23). Through implementing this package, a user can replicate the results reported in the previous publication in a simpler, more high-throughput manner.

Comparison with Existing Tools. While there is no open-source tool to our knowledge that performs the functions of *PTMphinder*, we compared the efficiency of *PTMphinder* to the manual extraction of motifs. We found that using *PTMphinder* resulted in a decrease in motif extraction time of over two orders of magnitude compared to manually extracting the sequences. Additionally, *PTMphinder* provides the location of the modified site in the full-length protein and reduces the potential of human error, effectively enhancing research reproducibility.

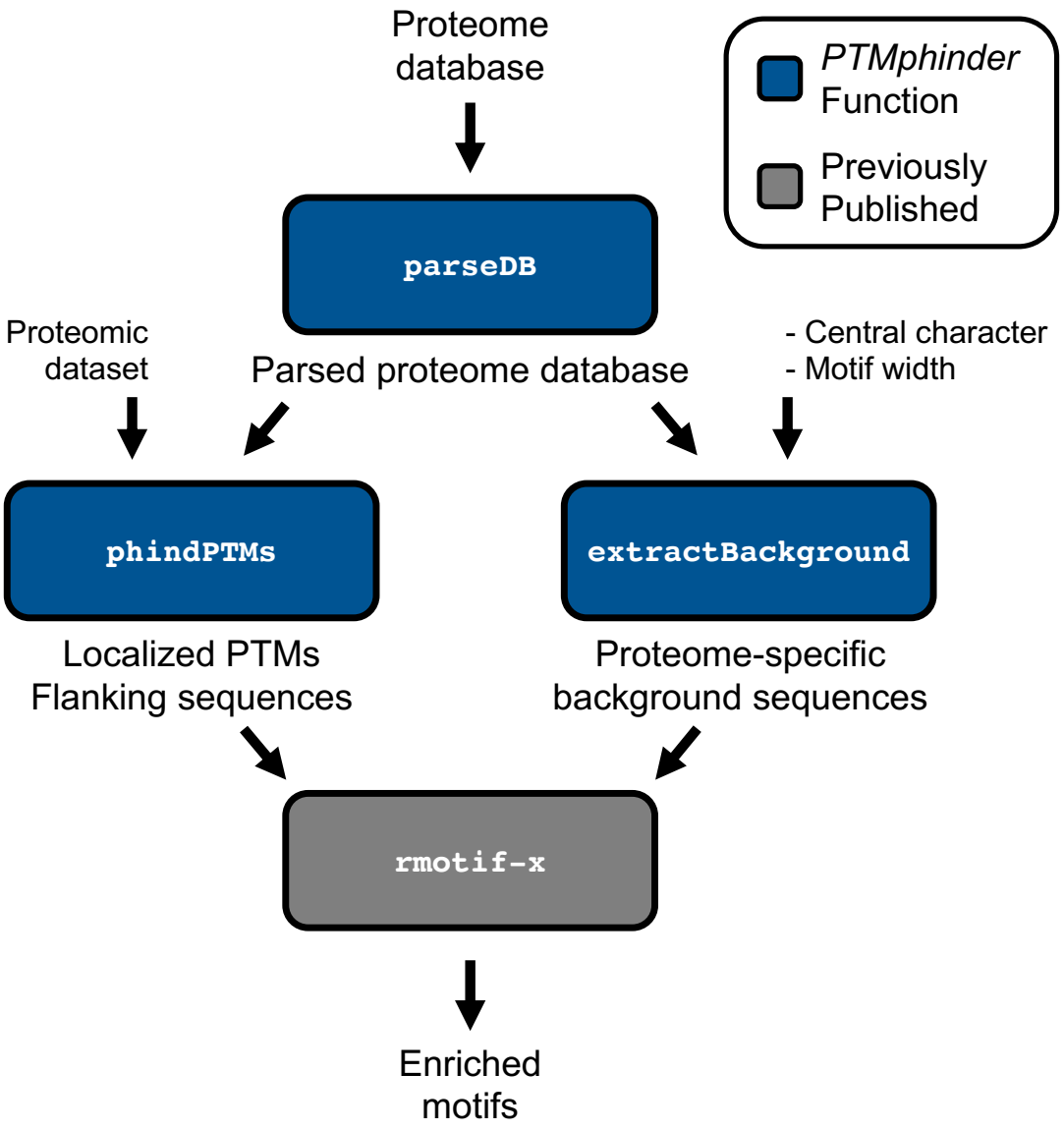


Figure 3-1 *Workflow of PTMphinder R Package*

Blue boxes indicate *PTMphinder* functions while grey bocks indicate functions outside of *PTMphinder*. Pipeline simply requires 1) the user's proteomic dataset, 2) a proteome database and 3) additional, minor user-defined parameters.

Phospho-proteomics Applied to a Chagas Disease Model

Mouse Model of Chagas Disease. To ensure that our infection model properly reflects Chagas disease symptoms, we performed electrocardiography and histological analyses of the infected hearts (**Figure 1A**). We found that mice infected with *T. cruzi* presented heart arrhythmias (**Fig 1B-C**), a lower heart rate (**Fig 1D**) and AV block (**Fig 1C, E**), which mimics human Chagasic hearts(127) and are in accordance with previous mouse models of chronic *T. cruzi* infection using different strains of the parasite(128). The extended QT interval observed in previous mouse infections(128) was also captured in our model (**Fig 1F**). These phenomena usually occur in parallel with interstitial fibrosis and inflammation(93). Therefore, we performed histological analyses and confirmed that these mice displayed intense cardiac inflammation (**Fig 1G**), quantified by % area of cell nuclei (**Fig 1I**), and increased collagen deposition (red staining, **Fig 1H**) in the heart tissue. In agreement with human clinical symptoms(129), quantification of the fibrotic area demonstrated a 3-fold increase in the area occupied by collagen in cardiac sections from *T. cruzi* infected animals (**Fig 1J**). Together, this data validates the establishment of an *in vivo* model of chronic Chagas disease cardiomyopathy that shows clinical symptoms of heart disease similar to those observed in human patients, providing a strong foundation for subsequent analyses.

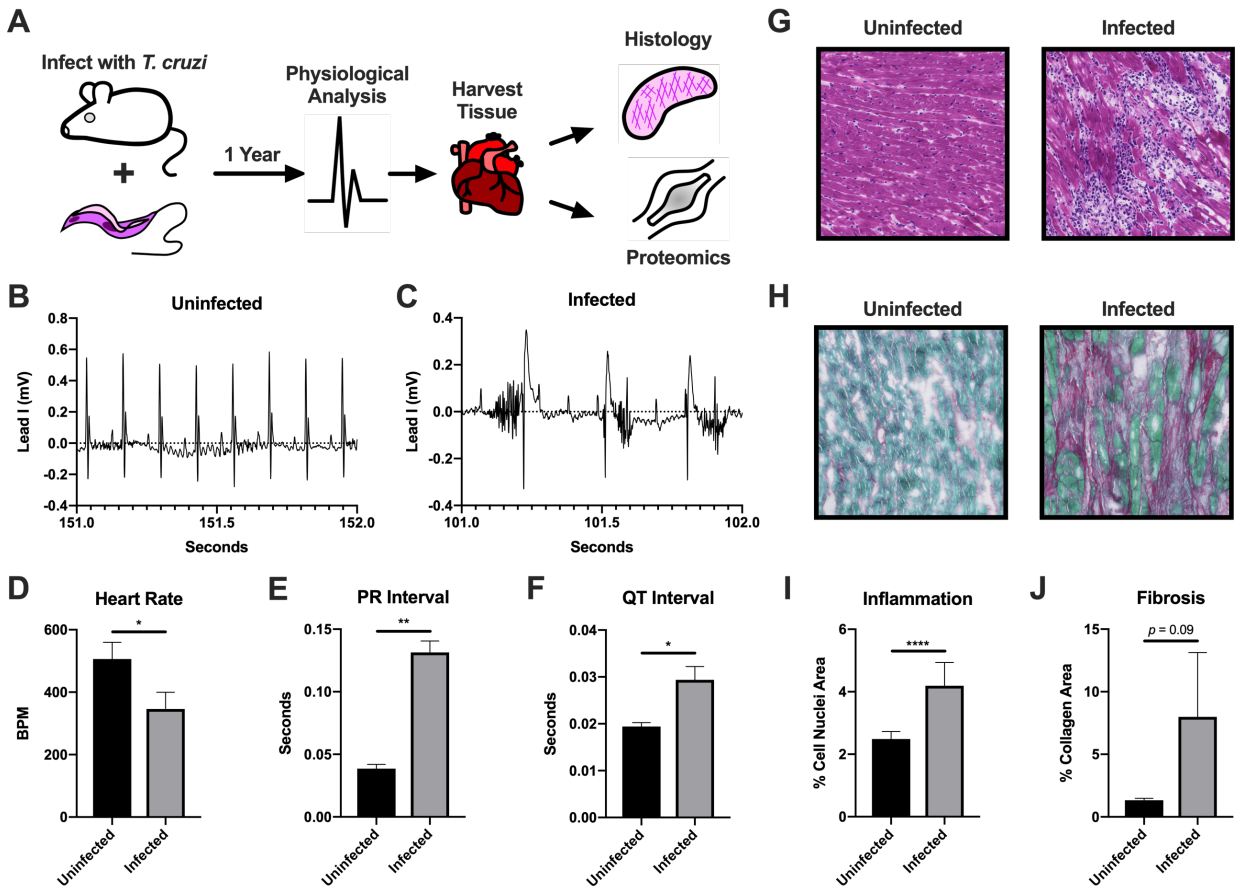


Figure 3-2 *Chronic Chagas Cardiomyopathy Model*

(**A**) Schematic of model infection and analyses. (**B**) Representative one-second ECG trace from uninfected mouse. (**C**) Representative one-second ECG trace from infected mouse. (**D**) Average heart rates from uninfected and infected animals during ECG analysis. (**E**) Average PR intervals from uninfected and infected animals during ECG analysis. (**F**) Average QT intervals from uninfected and infected animals during ECG analysis. (**G**) Representative images of cell nuclei staining in hearts of uninfected and infected animals. Bar=100μm. (**H**) Representative images of collagen staining in hearts of uninfected and infected animals. Bar=100μm. (**I**) Quantitation of areas of cell nuclei. (**J**) Quantitation of areas of collagen staining. For all graphs, significance was determined using an unpaired Student's t-test (*$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$).

General Overview of Phospho-proteomic Results. The applied phospho-proteomic workflow (**Fig 2A**) identified and quantified 4,622 proteins (**Table S1**) and 6,807 unique phospho-peptides at a false discovery rate (FDR) of < 1%. Our analyses revealed a high overlap of phospho-peptides that were matched to total protein abundances (**Fig S1A**). Thus, we were able to normalize 4,526 phospho-peptides to their respective total protein levels (**Table S2**). Subsequent phospho-analyses were performed on the protein-normalized, phospho-peptide dataset. We found that the biological replicates of both the proteomic and phospho-proteomic data had low inter-replicate variation (coefficient of variation (CV) < 15% for all replicates; **Fig S1B**) and high correlation (**Fig S1C-D**), endorsing the reproducibility of the Chagas disease model and proteomic methods. In line with previous TiO_2-enriched phospho-proteomic studies (119, 130), we primarily detected serine phosphorylation (81%), followed by threonine (16%) and tyrosine (3%) residues (**Fig 2B**). The majority of the detected phospho-peptides had only one phosphorylation event (80%) but a notable fraction possessed 2 or more events (20%) (**Fig 2C**). Overall, we found significant changes (pi score < 0.05)(120) in 394 proteins (238 increased, 156 decreased) and 353 phospho-proteins (133 increased, 159 decreased, 61 both) (**Fig 2D, Fig S1E-F**). There was low overlap between the proteins found to be significantly altered by total abundance and phosphorylation events, suggesting distinct pathway involvement (**Fig 2E**). The phospho-sites quantified in the experiments were primarily evenly distributed across the total protein length, with a slight bias for protein C-termini (**Fig 2F**). These broad analyses demonstrate high quality results and that chronic Chagas disease induces a vast rearrangement in host proteins and phospho-sites in the heart.

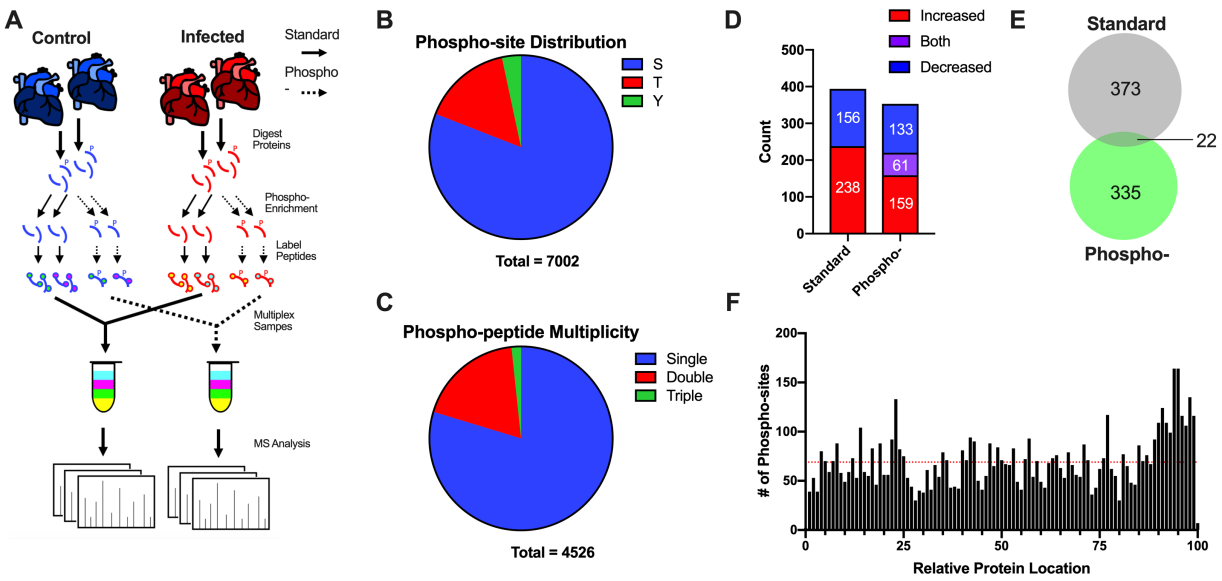


Figure 3-3 *Overview of Dual Proteome/Phospho-proteome Analyses*

(**A**) Schematic of phospho-proteomic analysis workflow. (**B**) Distribution of phospho-sites detected in this experiment. (**C**) Multiplicity of phospho-peptides detected in this experiment. (**D**) Numbers of significantly altered (pi score < 0.05) proteins and phospho-proteins. (**E**) Venn diagram of significantly altered proteins and phospho-proteins. (**F**) Relative protein locations of phospho-sites detected in this experiment.

Chronic Chagas Disease Induces an IFN-mediated Immune Response and Represses Mitochondrial Proteins. We focused our initial analysis on the total protein abundance changes following chronic Chagas disease. We found that total levels of 394 proteins were significantly altered during infection (**Fig 2D, Fig S1E**). To gain a deeper understanding of the functional outcome of this proteome perturbation, significantly altered proteins were subject to functional protein association network and gene ontology (GO) analyses using String-DB(123) and DAVID(121, 122), respectively. We observed that the altered proteins formed a highly interconnected network consisting of two primary clusters of proteins with increasing and decreasing abundance (**Fig 3A**). Interestingly, proteins altered in the opposite directions were located within the primary clusters, suggesting that an increase of abundance of one protein may be related to the decreased abundance of another. GO analysis revealed proteins with increased abundance were primarily secreted glycoproteins involved in generating an effective immune response against protozoan invaders (**Fig 3B; Table S3**). These changes appeared to be primarily driven by IFN-mediated signaling pathways, supporting previous studies(103). Lysosomal proteins and antigen processing pathways were also increased (**Fig 3B**), again consistent with an IFN-driven immune response(131, 132). In contrast, proteins with decreased abundances were primarily derived from the mitochondria (**Fig 3C**). Specifically, transmembrane proteins associated with electron transport and redox reactions were the most enriched among the decreased proteins. Peroxisomal proteins and lipid metabolic processes were also significantly reduced in the infected mice (**Fig 3C**). While it should be noted that GO analysis are not experimental and need to be validated further, the total

protein level alterations during Chagas disease suggest a strong immune response and suppression of general host metabolism.

To validate the proteomic studies, we cross-examined our data against the current knowledge regarding the impact of *T. cruzi* infections on host cells (**Fig S2A-B**). It has been reported that chronic Chagas disease is associated with increased chemokine (eg. CCL2(133) and CCL5(134)) and adhesion molecule (eg. ICAM1 and VCAM1)(135) expression in the heart leading to an infiltration of pro-inflammatory immune cells. In the experiments herein, we found a significant increase in both VCAM1 and ICAM1 (**Fig S2A**) but the chemokines CCL2 or CCL5 were not identified. However, we did identify CCL8, which had an ~6-fold increase in expression (**Fig 3D**). *T. cruzi* infections have also been associated with increased STAT1(136), caspase(137) and MHC class I(138) expression, all of which were detected in our data (**Fig S2A**). These results suggest highly inflammatory conditions and excessive immune cell influx. To strengthen this notion, the data was mined for proteins that are primarily expressed in immune cells and should be absent in normal heart tissue to use as markers for immune invasion. We found an increase in expression of a number of immune-cell enriched markers such as: Klra2 (killer-cells), LSP1 and Fas (lymphocytes), TGTP1 (T-cells), MHC class II (antigen presenting cells) (**Fig S2B**) in our heart samples, suggesting a notable association of these cells with the heart tissue.

To put our study in the context of the current understanding of *T. cruzi* infections and heart disease, we performed an in-depth comparison with previous -omic studies. First, a recent genome-wide association study (GWAS) identified numerous single-nucleotide polymorphisms (SNPs) with significant relationships to Chagas

cardiomyopathy in humans(139). Comparing our list of significant proteins to the GWAS findings, we identified a number of proteins (**Fig S2C**) that were detected in the distinct experiments. We found proteins with both increased expression (SMAP2, RBP1, COL14A1 and TGM2) and decreased expression (THBD and GNG7) to have SNPs associated with Chagas-induced cardiomyopathy. The concurrent associations of SNP variation and expression level of these proteins with Chagas disease offers support for their distinct involvement in disease progression.

A survey of the literature showed that no prior proteomic studies of Chagasic hearts had been reported. However, recent transcriptomic analyses have focused on 1) organ-level host responses to *T. cruzi* infection(140, 141) and 2) the effects of a genetically-induced cardiomyopathy(142). These host responses may provide a means to validate the current proteomic results and be a useful comparison to our dataset. We hypothesized that a similar modulation of gene expression may be present in various organs affected by *T. cruzi* as well as distinct cardiomyopathy models. Indeed, we observed a notable correlation of protein levels from our study with significantly altered transcripts in both Chagasic hearts ($R = 0.630$, $p < 0.001$; **Fig S2E**) and placentas ($R = 0.598$, $p < 0.001$; **Fig S2F**). Further, we found that the correlation of gene expression between Chagasic hearts and genetically driven, dilated cardiomyopathy ($R = 0.417$, $p < 0.001$; **Fig S2G**) and heart failure ($R = 0.425$, $p < 0.001$; **Fig S2H**) was also significant, albeit with lower correlation values. Strikingly, 80-90% of the significant genes identified in both studies were altered in the same direction (**Fig S2E-H**), suggesting there are similar mechanisms involved in these distinct models. As expected, this proportion was highest for proteins (this study) and transcripts from Chagasic hearts (>96% agreement;

Fig S2E). Together, the above analyses demonstrate the concordance of our results with previous studies, supporting the validity of our findings.

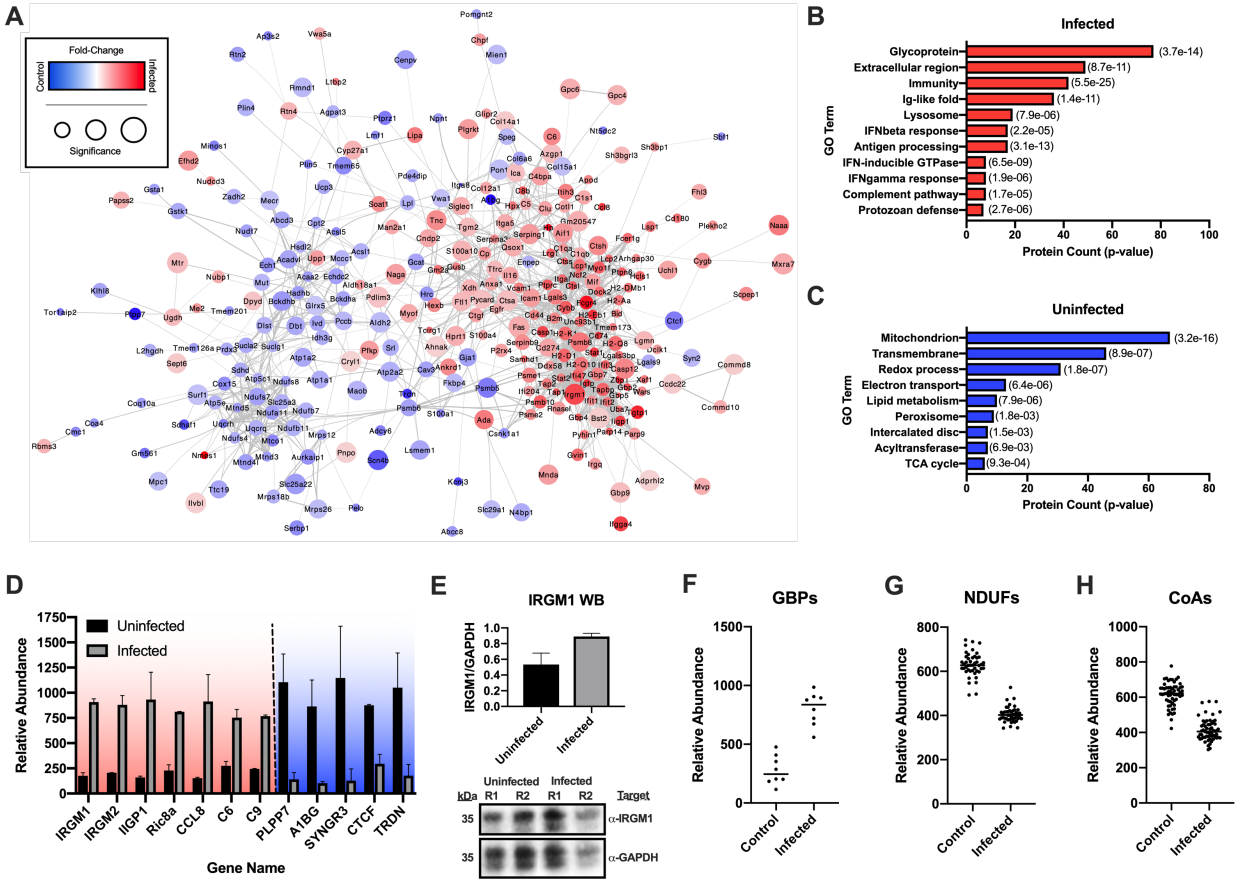


Figure 3-4 *Total Proteome Perturbations Demonstrate Induction of Host Immune Response and Suppression of Mitochondrial Proteins*

(**A**) Functional protein association network of all significantly altered proteins detected in the standard proteomic workflow. (**B**) GO results of significantly increased proteins. (**C**) GO results of significantly decreased proteins. (Significance of enrichment is shown in parenthesis to the right of the bars). (**D**) Selected increased and decreased proteins detected in the standard proteomic workflow. Significance is noted in reference to the pi score cutoffs for respective significance thresholds (* - $\alpha < 0.05$; ** - $\alpha < 0.01$; *** - $\alpha < 0.001$; **** - $\alpha < 0.0001$). (**E**) Western blot validation of IRGM1 expression (R1 – replicate 1, R2 – replicate 2). (**F**) Differentially expressed GBPs detected in the standard proteomic workflow. (**G**) Differentially expressed NDUFs detected in the standard proteomic workflow. (**H**) Differentially expressed CoA-containing enzymes detected in the standard proteomic workflow. For differentially expressed families of proteins (F-H), a paired T-test was used (*** - $\alpha < 0.001$; **** - $\alpha < 0.0001$).

In addition to previously known associations to *T. cruzi* infections, our unbiased, systems-level assessment found a number of proteins that have not yet been interrogated in the context of Chagas disease. Interesting proteins were ranked by pi-score, a metric that combines both fold-change and *p*-value. For example, we found additional, immune-related proteins increased in response to *T. cruzi* infection including: IRGM1/2, IIGP1, CCL8, C6, C9 (**Fig 3D-E**), among others. We also noted an increase in all of the guanylate-binding proteins (GBPs) detected in our analysis (**Fig 3F**). On the other hand, we detected a decrease in abundance of proteins with no prior association to Chagas disease or immune responses. The most significantly altered among these include: PLPP7, A1BG, SYNGR3, CTCF and TRDN (**Fig 3D**). As mentioned above, our GO pathway analysis highlighted a decrease in mitochondrial proteins related to redox and electron transport (**Fig 3C**). Specifically, we noted a decrease in nearly all of the NADH dehydrogenases (NDUFs) (**Fig 3G**) and coenzyme A-containing proteins (**Fig 3H**) detected in our experiments. These findings are in-line with previous reports of decreased mitochondrial function(115) and identify the major protein families affected. These analyses demonstrate that chronic Chagas disease drives an IFN-mediated immune response and suppresses mitochondrial pathways in murine hearts. Our global proteomic data are substantiated by comparisons with previous studies and western blot validation. These results also highlight new proteins that may be involved in Chagas disease progression.

The Phospho-proteome of Chagasic Hearts. While total protein abundances provide useful information, many cellular signaling pathways are mediated via phosphorylation, with minimal alterations in total abundance. Therefore, we employed a

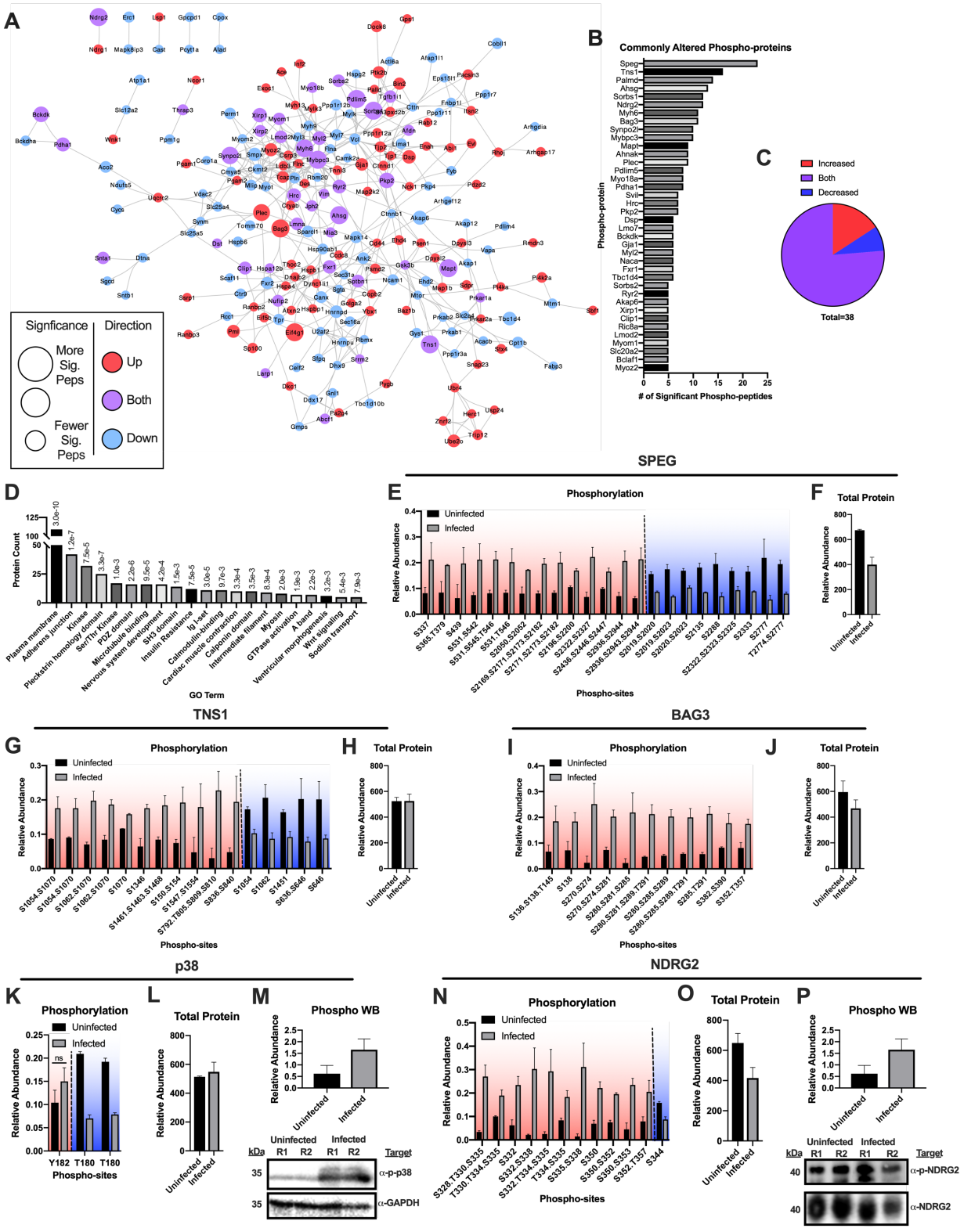
phospho-proteomic approach to dissect signaling pathways affected by chronic Chagas disease. We found that phospho-sites on 353 phospho-proteins were significantly altered following infection. The majority (83%) of these phospho-proteins had only increased (159) or decreased (133) phosphorylation, but a subset had both increased and decreased phospho-sites (61) (**Fig 2D**). Creating a functional protein association network(123) (**Fig 4A**), we observed that proteins with opposite directions of phosphorylation abundance were highly interconnected, indicating that the increase of phosphorylation of one protein is associated with a decrease in phosphorylation of related proteins, and vice-versa. While the majority of phospho-proteins possessed only one significantly altered phospho-peptide, ~40 proteins had five or more significantly altered phospho-peptides (**Fig 4B**). Interestingly, most of these proteins had both increasing and decreasing phosphorylation sites (**Fig 4C**) and were centrally located in our network, putting forth their roles as key signaling hubs in Chagas-induced cardiomyopathy.

Chronic Chagas Disease Alters Membrane Protein and Microtubule Phosphorylation. To determine the global signaling pathways altered during chronic Chagas disease, we performed GO analysis on all of the phospho-proteins with significantly altered phospho-sites (**Fig 4D; Table S4**). We found highly significant enrichments for proteins localized to the plasma membrane, adherens junctions and intermediate filaments. As expected, we noted significant enrichments for many proteins involved in intracellular signaling such as kinases, calmodulin-binding proteins, and GTPases. Protein domains that govern signaling pathways from the plasma membrane through the cytoskeleton, such as pleckstrin-homology, PDZ and SH3 domains, were also enriched in this dataset. These results indicate that chronic Chagas disease has the most

significant effects on intracellular signaling pathways bridging the plasma membrane and the cytoskeleton.

Figure 3-5 *Phospho-proteome Interrogation Reveals Alterations in Membrane and Cytoskeletal Protein Phosphorylation and Activation of p38.*

(**A**) Functional protein association network of all proteins with significantly altered phospho-peptides. (**B**) Bar graph of proteins with 5 or more significantly altered phospho-peptides. (**C**) Pie chart depicting the phospho-peptide direction shifts of proteins with 5 or more significantly altered phospho-peptides. (**D**) GO results of all proteins with significantly altered phospho-peptides. Graph depicts the number of phospho-proteins in the respective ontology with enrichment *p*-values displayed above the bar. (**E**) Bar graphs of significantly altered SPEG phospho-peptides. (**F**) Bar graph of total SPEG protein. (**G**) Bar graphs of significantly altered TNS1 phospho-peptides. (**H**) Bar graph of total TNS1 protein. (**I**) Bar graphs of significantly altered BAG3 phospho-peptides. (**J**) Bar graph of total BAG3 protein. (**K**) Bar graphs of significantly altered p38 phospho-peptides. (**L**) Bar graph of total p38 protein. (**M**) Western blot validation of p38 phosphorylation status. (**N**) Bar graphs of significantly altered NDRG2 phospho-peptides. (**O**) Bar graph of total NDRG2 protein. (**P**) Western blot validation of NDRG2 phosphorylation status. For western blot figures, R1 – replicate 1 and R2 – replicate 2. Significance is noted in reference to the pi score cutoffs for respective significance thresholds (* - $\alpha < 0.05$; ** - $\alpha < 0.01$; *** - $\alpha < 0.001$; **** - $\alpha < 0.0001$)


A
Signficance
More Sig. Peps
Fewer Sig. Peps
Direction
Up
Both
Down
B
Commonly Altered Phospho-proteins
Phospho-protein
# of Significant Phospho-peptides
C
Increased
Both
Decreased
Total=38
SPEG
D
Protein Count
GO Term
E
Phosphorylation
Uninfected
Infected
Relative Abundance
Phospho-sites
F
Total Protein
TNS1
G
Phosphorylation
H
Total Protein
BAG3
I
Phosphorylation
J
Total Protein
p38
K
Phosphorylation
ns
L
Total Protein
M
Phospho WB
kDa
R1
R2
Target
α-p-p38
α-GAPDH
NDRG2
N
Phosphorylation
O
Total Protein
P
Phospho WB
α-p-NDRG2
α-NDRG2

Specific proteins of interest were prioritized based on 1) the number and significance of altered phospho-sites, 2) their involvement in plasma membrane and cytoskeletal signaling, and 3) their relationship to previously described Chagas disease processes or other cell stresses. These include: striated muscle preferentially expressed protein kinase (SPEG), Tensin 1, BCL2 Associated Athanogene 3 (BAG3), Sorbin and SH3 domain-containing protein (SORBS) 1/2 and a number of myosin-related proteins (MYH6, MYBPC3, MYO18A, MYL2, MYOM1, MYOZ2). SPEG had the highest number of significantly altered phospho-peptides (23 phospho-peptides; **Fig 4B, E-F**) followed by Tensin 1 (16 phospho-peptides; **Fig 4B, G-H**). Both of these proteins had phospho-sites that increased or decreased in abundance, suggesting distinct kinases/phosphatase regulation of these proteins. BAG3 contained 11 significantly altered phospho-peptides that were all increased (**Fig 4B, I-J**). SORBS1/2 (**Fig S3A-D**) and a number of myosin proteins (**Fig S3E-P**) also had multiple, significant phosphorylation shifts in response to *T. cruzi* infection. The fact that we captured multiple members of these protein families argues for their importance in the host response to *T. cruzi*.

In accordance with previous *in vitro* studies with cardiac fibroblasts(109) and colon cells(108), we detected an increase in p38 Y182 phosphorylation following chronic *T. cruzi* infection in mice (**Fig 4K-M**). Interestingly, we also detected a significant decrease in singly phosphorylated p38 T180 peptides **(Fig 4K**). Residues T180 and Y182 of p38 are sequentially phosphorylated in response to pro-inflammatory conditions(143). The observed decrease in singly phosphorylated p38 T180 peptides may be a result of an increase in the doubly phosphorylated species, which was not detected in our phospho-proteomic data. To clarify these findings, we performed western blot of p38 Y182 (**Fig**

4M), confirming its increased phosphorylation status. We also detected increased phosphorylation of N-Myc Downstream-Regulated Gene 2 (NDRG2) (**Fig 4N-P**), which can act upstream of p38(144, 145). These results further support the activation of this pathway in the host response to infection. Overall, the above analyses highlight global pathways and specific proteins whose phosphorylation status is modulated during chronic Chagas disease.

Prediction of Kinase Activity During Chronic Chagas Disease Reveals Increased JNK/DYRK and Decreased CK2 activity. Detecting alterations in phospho-site abundances is crucial to understanding the ultimate outcomes of various signaling processes. However, linking phospho-site abundance to specific kinases provides more mechanistic insight into the underlying biology and can identify potential drug targets. Therefore, in an effort to identify kinase-substrate interactions with increased and decreased activity during Chagas disease progression, we undertook a two-pronged approach.

First, we used the group-based prediction (GPS) tool(125) to predict the kinases that target the significant phosphorylation sites detected in our experiments. This analysis predicted substrates for 125 kinases within our dataset (**Table S5**), with AGC-group kinases having the most unique members (25 kinases; **Fig S4A**) and CMGC-group kinases targeting the most sites (422 sites; **Fig S4B**). We then compared which kinases were predicted for phosphorylation sites with increased vs. decreased abundance (**Fig 5A; Fig S4C-F**). We found that more site-kinase pairs were predicted to have increased activity in the infected samples than in the controls (**Fig S4C**). Stratifying predicted kinases by kinase groups and families suggests that there is a significant increase in

activity from the CAMK, CMGC and TKL groups (**Fig S4D**) and MAPK, PKC, MAPKAPK, DYRK and LISK families (**Fig S4E**). Individual kinases with increased activity include: IKKB, CDC28, MNK1, AMPKA2, MARK1, LIMK, ERK1, MAPK2K2, JNK2 and DYRK2 (**Fig 5A, Fig S4F**). In contrast, the kinases with the most decreased activity are VRK2, CSNK1D, TTBK1 and ROCK. Interestingly, we identified kinases that are predicted to target phospho-sites with both increased and decreased abundance, such as FRAP, PIM1, PAK1 and RAF, suggesting that the specificity of these kinases may be modulated as opposed to overall activity.

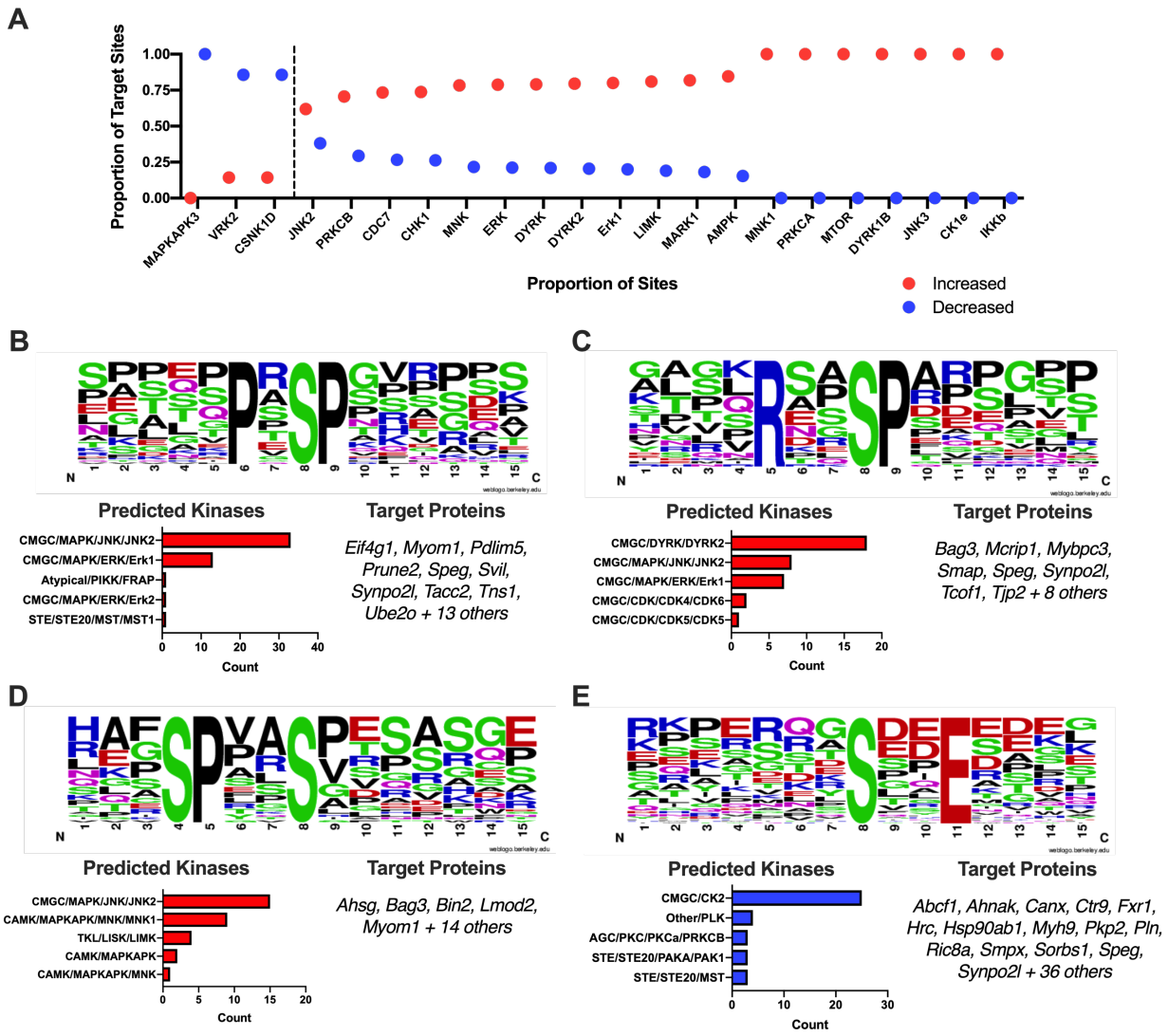


Figure 3-6 *Bioinformatic Kinase Prediction Uncovers Stimulation of JNK and DYRK2 and Suppression CK2 Activities.*

(**A**) Kinases with significant proportions (Chi-squared p-value < 0.05) of differentially expressed phospho-sites. (**B**) Logo of enriched motif "PxS$_p$P" (top) with bar graph of the number of predicted kinase sites (bottom left) and list of target proteins (bottom right). (**C**) Logo of enriched motif "RxxS$_p$P" (top) with bar graph of the number of predicted kinase sites (bottom left) and list of target proteins (bottom right). (**D**) Logo of enriched motif "SPxxS$_p$" (top) with bar graph of the number of predicted kinase sites (bottom left) and list of target proteins (bottom right). (**E**) Logo of enriched motif "S$_p$xxE" (top) with bar graph of the number of predicted kinase sites (bottom left) and list of target proteins (bottom right).

Second, we extracted the flanking sequences of each phospho-site using the *PTMphinder* R package(124) and used the motif-x algorithm(29, 30) to identify enriched peptide motifs that were either increased or decreased during infection (**Table S6**). Through this analysis we found that phospho-sites within the motifs: PxS_pP (**Fig 5B**), $RxxS_pP$ (**Fig 5C**) and $SPxxS_p$ (**Fig 5D**) were increased following infection and phospho-sites within the S_pxxE motif (**Fig 5E**) were decreased following infection. We then linked the phospho-sites within these motifs to their predicted kinases from the GPS analysis to create putative kinase-substrate interactions (**Figs 5B-E; Table S7**). This dual prediction approach provides further evidence for an activation of JNK/p38 family kinases and downstream MNK kinases, but also uncovers unexpected kinases such as DYRK2 (increased activity) and CK2 (decreased activity). Overall, both known (eg. JNK) and previously unknown (eg. DYRK2, CK2) kinases were predicted to be differentially activated in response to *T. cruzi* infection. We also identified their respective substrates for further validation as functional players in Chagas disease progression.

The Druggable Network of Chagasic Hearts. To elucidate druggable signaling pathways that can be further explored for their therapeutic potential, we linked the significant proteomic, phospho-proteomic and kinase prediction results to known drug targets within the DrugBank database(146, 147) (**Table S8**). Visualizing these druggable proteins using functional protein association analysis, we found that they formed a highly interconnected network with the predicted kinases at the center, closely followed by phospho-proteins, and the standard proteins residing in the distal regions (**Fig 6A**). Gene ontology analysis revealed that the most druggable pathways are mitochondria, secreted proteins, kinases and cell-cell adherens junctions (**Fig 6B; Table S9**), all of which are

intimately linked to *T. cruzi* infections and heart disease. Strikingly, more than 80% of the proteins in our network are known interactors with FDA approved drugs (**Fig 6C**); approximately 35% of which are linked to the intended, on-target effects (**Fig 6D**). To identify drugs that could broadly impact multiple proteins altered during chronic *T. cruzi* infections, putative therapeutics were ranked based on the total number of targets present in the network (**Fig 6E**). We found that Fostamatinib, Artenimol, metals bound by secreted immune proteins (zinc and copper) and NADH may have a strong influence on the outcomes of *T. cruzi* infections. Of note, acetylsalicylic acid (aspirin), emerged as one of the most influential drugs in our network, and has been demonstrated to be beneficial for the host during *T. cruzi* infections (148, 149), specifically at later stages of the disease(148).

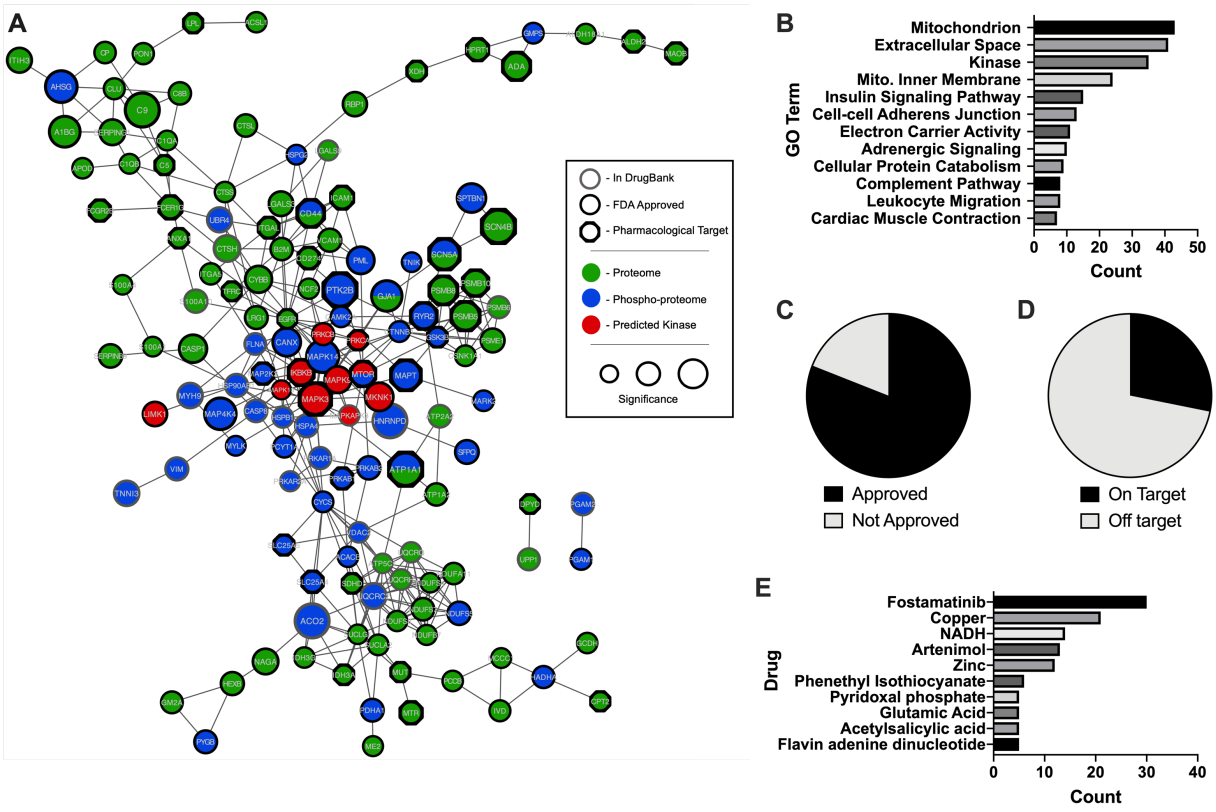


Figure 3-7 *Druggable Network of Chagasic Hearts*

(**A**) Functional protein association network of significantly altered proteins, phospho-proteins and predicted kinases that are known interactors of drugs from DrugBank. (**B**) GO analysis of proteins in the druggable network. Pie charts of the (**C**) approval status and (**D**) intended target status of drugs that interact with proteins in the druggable network. (**E**) Bar chart of drugs with 5 or more targets in the druggable network.

Discussion

Cardiomyopathy is the most common and severe pathological result of chronic Chagas disease(93, 96, 101). While studies focusing on individual host factors have provided insight into specific mechanisms of pathogenesis, proteins typically work in co-regulated networks with significant involvement of PTMs, which are often overlooked in standard assays. Therefore, a systems-level analysis would represent a significant step toward understanding the signaling pathways and molecular mechanisms underlying disease progression. Previous studies applying proteomics to investigate Chagas disease have mostly been focused on serum, looking for disease progression markers using two-dimensional gel electrophoresis prior to mass spectrometry in rats(150) and human patients(151, 152). Also, proteomic and phospho-proteomic approaches have been targeted toward the parasite itself, describing its response to nutritional stress(153) and proteins modulated during the differentiation process(154, 155). However, a large-scale, phospho-proteomic analysis of Chagasic hearts, one of the most affected organs, has not been performed. To address this gap in knowledge, we performed a multiplexed, quantitative phospho-proteomic analysis of an *in vivo* murine model of Chagas-induced cardiomyopathy. We find that our model possesses common clinical symptoms of Chagasic hearts(127, 128), indicating that it accurately reflects human biology. While this manuscript describes findings based on two, independent, biological replicates, we believe that the agreement in our replicates (Fig S1), particularly in the proteins/phospho-sites displayed in the figures, and comparisons to previous studies (Fig S3), largely substantiate the dataset.

We compared our proteomic expression results to previously published studies including conventional assays (eg. western blot and histology) as well as GWAS and transcriptome analyses. Together this helped validate our proteomic and prioritize interesting proteins for analysis. From this meta-analysis, we found that the host response to *T. cruzi* is significantly correlated across distinct organs and that Chagasic hearts follow gene expression profiles similar to other cardiomyopathies. We also found a number of proteins that were differentially expressed in our proteomic data and possessed SNP-associations to Chagas disease progression. Notable among these is retinol-binding protein 1 (RBP1), which was significantly increased upon infection and linked to Chagas cardiomyopathy through GWAS studies(139). Rbp1 functions in the uptake and storage of retinol (vitamin A)(156), a molecule with anti-oxidant(157) and immune-related(158) activities. Further, isotretinoin, a naturally-occurring derivative of retinol(159) has been shown to have promising trypanocidal effects in the nanomolar concentration range(160). Due to the high structural similarity of retinol and isotretinoin, it seems plausible that increased Rbp1 could also increase intracellular isotretinoin concentrations, acting as a host defense mechanism against *T. cruzi* or chronic inflammation. Treatment of *T. cruzi* infections with vitamin C has been shown to have myocardial protective effects(161) and retinol can help defend against many forms of cardiovascular disease(157), however, treatment of *T. cruzi* infections with either retinol or isotretinoin has not yet been tested.

Other interesting groups of proteins include immune-related proteins, with both described and previously undescribed associations with Chagas disease. For example, we verified previous reports of an upregulation of the adhesion molecules VCAM1 and ICAM1(135) and the chemokine CCL8(141) which is thought to enable excessive immune

infiltration. Previous studies identified CCL5 as a primary contributor to immune influx(134). Both CCL5(162) and CCL8(163) can interact with CCR1, CCR3 and CCR5 on target cells inducing immune recruitment to the site of inflammation. Their co-expression increase suggests a redundant mechanism in the chemotaxis and could result in excessive immune infiltration.

A striking finding in our standard proteomic data was the increase in nearly the entire family of guanylate binding proteins (GBPs). Previous studies have indicated an increase of expression of GBP2 and GBP6 following *T. cruzi* infection in both the heart and the placenta(140, 141), but not in response to genetic cardiomyopathies(130). However, our findings include the additional family members: GBP2B, GBP4, GBP5, GBP7 and GBP9. GBPs are induced by IFN-γ as a host defense against invading pathogens(164, 165) and are thought to function as a complex, supporting the precise co-regulation observed in the present study. Of note, GBP1, along with GBP2, 4 and 5, has been shown to localize to *T. gondii* containing vacuoles(166). Further, it has been shown that mice deficient in GBP2 are highly susceptible to *T. gondii* infections(167). While GBP1 was demonstrated to not co-localize with *T. cruzi*(166), the alternate GBPs detected in our experiments might be involved in host recognition of *T. cruzi*. This data demonstrates a mechanism of host response to *T. cruzi* that is conserved in multiple organs and highlights the additional proteins that are co-regulated in the heart. These results warrant further investigation into which specific GBPs functionally respond to *T. cruzi* and the ultimate consequences of this interaction.

In contrast to the proteins with increased abundance mentioned above, we detected a strong signature of mitochondrial proteins being significantly downregulated

following infection. The dysfunction in mitochondria following *T. cruzi* infections is well known(115), but poorly understood. Previous studies have exhibited *T. cruzi* co-localizing with mitochondria via their single flagella(168). Mitochondrial dysfunction is described to start in the acute phase of the infection, continuing to a reduction of oxidative phosphorylation capacity in the heart of mice during chronic infection(169). In addition, a decline in the activities of the respiratory complex III and antioxidant enzymes (MnSOD and GPX) as well as in GSH contents has been observed in chronic Chagas patients(170). With respect to chronic *T. cruzi* mouse infections, the current study, supported in part by previous transcriptome analyses(141), have demonstrated a decrease in many mitochondrial proteins. Our analyses have identified NDUFs and CoA-containing enzymes as primary contributors to this observed mitochondrial signature. NDUFs and CoA-containing enzymes are involved in many metabolic functions including electron transport and fatty acid synthesis. Intriguingly, it has been demonstrated that *T. cruzi* actively scavenges host long-chain fatty acids to promote its intracellular survival, a mechanism that was further validated by showing reduced parasite proliferation in fibroblasts lacking *de novo* triacylglycerol biosynthesis(171). The downregulation of NDUFs and CoA enzymes in the heart may reflect a host effort to effectively starve out the parasite. Unfortunately, this decrease in mitochondrial function can also contribute to cardiomyopathy progression(172). By identifying the specific proteins involved, it may be possible to target therapeutic approaches that enhance myocyte function and survival while simultaneously starving the parasite. As such, our studies reinforce previously described interplays of host-parasite lipid metabolism and highlight specific host proteins likely to be involved.

In addition to total protein expression, our analysis captured the first shotgun, phospho-proteomic analysis of the interaction of *T. cruzi* with host cells *in vivo*. We found drastic alterations in host phosphorylation status following infection as evidenced by significant changes of phospho-site abundance in 353 unique phospho-proteins. Functional protein association analysis(123) highlighted potential key players, which sat central to the network with large numbers of phosphorylation changes (up to 23 distinct phospho-peptides significantly altered). Specifically, SPEG, Tensin 1, BAG3, SORBS1/2 and a number of myosin-related proteins (MYH6, MYBPC3, MYO18A, MYL2, MYOM1, MYOZ2) emerged as integral components of this network. Of note, all of these proteins have been linked to proper heart function and/or implicated in the development of cardiomyopathies(173-182). Thus, their differential phosphorylation in response to *T. cruzi* may be mechanistically related to the development of severe disease states.

Chronic Chagas disease drives severe inflammation and fibrosis in the heart, and our phospho-proteomic analyses found substantial evidence that the primary implicated intracellular signaling pathways are c-Jun N-terminal kinase (JNK) and p38. First, we found elevated phosphorylation of p38 at Y182, a known activation site. Second, we detected an increase in phosphorylation of NDRG2, a cell-stress related protein. Phosphorylation of p38 has been described for *in vitro* models of *T. cruzi* infection in colon cells (18) and cardiac fibroblasts(109). p38 was also found increased during *T. cruzi* acute infection(183), but our data shows for the first time that p38 might be implicated in chronic Chagas cardiomyopathy in a model that displays clinical symptoms of heart failure. Overexpression of NDRG2 has been shown to increase p38 phosphorylation(145), suggesting it can function upstream. The observed increase in phosphorylation of NDRG2

provides further evidence for activation of this pathway. Finally, our dual kinase-substrate prediction analysis reported an increase in JNK activity, particularly at phospho-sites within the motifs PxS_pP and $SPxxS_p$. Of note, many of the highly altered phospho-proteins of interest mentioned above possess predicted JNK target sequences, suggesting downstream involvement in this signaling pathway. Together, this data supports a central role of JNK and p38 signaling pathways in Chagas disease progression and identifies new components both up- and downstream.

In addition to activation of JNK and p38 pathways above, our analysis predicted an activation of IKKB, MNK1 and DYRK2 and suppression of CK2, CSNK1D and VRK2. Activation of IKKB and MNK1 is logical given the highly inflammatory conditions and activation of JNK/p38 described above(184, 185). IKKB phosphorylates the NFKB Inhibitor Alpha in the NFKB complex, causing dissociation of the inhibitor and activation of NFKB in response to pro-inflammatory stimuli(186). MNK1 is a MAPK activated protein (MAPKAP) and functions downstream of p38(187). Upon activation, MNK1 phosphorylates downstream substrates, primarily involved in translational regulation, including eIF4E and eIF4G(187, 188). While we didn't detect any spectra assigned to eIF4E, we detected significant increases in phosphorylation of eIF4G at both S1231 and S1238, further reinforcing activation of this pathway. DYRK2 functions to regulate cell cycle and proliferation(189), apoptosis(190) and organization of the cytoskeleton(191). We predicted DYRK2 to phosphorylate a number of interesting proteins from our phospho-hits including SPEG, BAG3, NDRG2, MYBPC3 and MYO18A. The potential for DYRK2 to phosphorylate these central proteins and also affect cytoskeletal organization make it an interesting finding for further interrogation. Finally, we noted an intriguing

downregulation of CK2, CSNK1D and VRK2 activity, all kinases related to casein kinase families. VRK2 has been previously shown to downregulate JNK activity(192, 193). Further, CSNK1D has been hypothesized to be phosphorylated by p38 at a regulatory site, reducing its activity(194). Thus, the inactivation of these kinases further supports JNK and p38 activation in our model. On the other hand, there are conflicting reports on the impact of CK2 on JNK activity in response to inflammation(195-198). It is important to note that, while CK2 activity was predicted to decrease, this finding was primarily linked to sites within $S_p$xxE motifs. A nearly equivalent number of other CK2-predicted sites were found to be increased; however, they showed no bias toward the $S_p$xxE motif. These results demonstrate a decrease in CK2 activity against specific sites, suggesting the specificity of CK2 may be altered rather than a global decrease in activity.

We conclude our study by identifying FDA approved drugs that could be repurposed for the treatment of Chagas disease in the chronic phase. This analysis indicates that therapeutic targeting of the central kinases highlighted herein (such as JNK, p38 and IKKB) could be an effective means to influence a variety of downstream proteomic and phospho-proteomic alterations. Fostamatinib, a kinase inhibitor recently approved for rheumatoid arthritis and immune thrombocytopenic purpura(199), has affinity for all these targets and thus represents a means to simultaneously disrupt these key signaling pathways. On the other hand, targeting a particular group of downstream effectors (eg. complement proteins, NDUFs, and the proteasome) could be implemented as a more focused therapy. For example, zinc supplementation has been shown to reduce complement activity in age-related macular degeneration(200). Effectively blocking complement deposition could reduce the disease burden associated with high

inflammatory conditions of the Chagasic heart. Interestingly, we identified a large number of antibody-based, anti-inflammatory therapeutics (Efalizubmab, Natalizumab, etc.) that also interact with complement proteins (C1QA and C1QB), as well as EGFR and immune adhesion molecules (ICAM1, ITGAL). These targeted therapies may be useful in reducing the inflammatory conditions of the heart, particularly when used in combination with zinc or other immuno-modulatory agents. As with any immune suppressing drugs, there is an increased risk of infection for patients taking the medication(201). Further, excessive zinc can also lead to immune dysfunction(202). Determining which approach would be most effective at reducing disease burden, while simultaneously preserving normal cardiac function and immune system health, would represent a critical step toward developing new therapeutic options for Chagas cardiomyopathies. A thorough assessment of the impact of the identified pharmacological agents on disease progression should be pursued to build further evidence for the roles of these proteins and tease out precise mechanisms of action.

Overall, this study presents a comprehensive view of the molecular underpinnings of Chagasic hearts. We identify specific host pathways and proteins that respond to *T. cruzi* infections at both the raw protein and phospho-site abundance levels. As expected, we found a predominantly IFN-driven immune response as evidenced by infiltration of immune cells and activation of the JNK and p38 pathways. We also found a downregulation of many mitochondrial proteins which may be linked to the known mitochondrial defects observed in *T. cruzi* infections. Our phospho-proteomic analyses revealed potential up- and downstream mediators of JNK, p38 and NDRG2 signaling and identified additional kinase families that may be activated or repressed. Finally, we

highlight and discuss the potential repurposing of FDA approved drugs for the reduction of Chagas disease burden. Together, these studies provide a systems-level overview of Chagasic hearts and sets the groundwork for future studies to validate the functional consequence of these alterations.

Chapter 3, in part, is a combination of material as it appears in *PeerJ*, 2019, Jacob M. Wozniak and David J. Gonzalez, and *PLOS Neglected Tropical Diseases***,** 2020, Jacob M. Wozniak, Tatiana Araújo Silva, Diane Thomas, Jair L. Siqueira-Neto, James H. McKerrow, David J. Gonzalez, Claudia M. Calvet. The dissertation author was the primary author of these papers.

Chapter 4 – Multi-omic Serum Analysis Reveals Early Predictive and Pathogenic Molecular Signatures for S. aureus Bacteremia

Introduction

Staphylococcus aureus bacteremia (SaB) causes significant mortality in a wide range of patient populations. Due to medical advances that support increasingly immunocompromised hosts, invasive surgical procedures and indwelling medical devices, underlying risk factors for serious *S. aureus* infections are expanding(47). Overall mortality rates for SaB range from 20 - 30%(203-205), however SaB patients display a highly heterogeneous array of disease severity and patient outcomes(206, 207). Whereas some patients clear the infection promptly on first-line antibiotic therapy, others fail to properly eradicate the pathogen. Extended bacteremia leads to dysregulation of the host immune response, which is highly correlated with patient mortality(208-210). This heterogeneity in SaB disease progression complicates the determination of optimal treatment strategies. Current standard of care is to administer broad-spectrum antibiotics while awaiting pathogen susceptibilities to guide treatment decisions; however, blood cultures are not always attainable, and it may take days to fully deduce antibiotic susceptibilities. Any delay in time-to-diagnosis and initiation of appropriate therapy exacerbates patient mortality, especially in sepsis(211, 212). Furthermore, in the case of the first-line antibiotic vancomycin, while drug resistance is rare, clinical failure is common, revealing significant shortcomings in predictive power of standard antimicrobial susceptibility testing(213). Here we explore if the host response measured within hours of clinical presentation can accurately predict mortality risk, which could ultimately help clinicians optimize more appropriate and personalized therapeutic regimens.

Previously, our group identified robust immunological biomarkers for SaB mortality, e.g. elevated interleukin-10(208, 210) and prolonged bacteremia, e.g. reduced

interleukin 1β(210), which were subsequently corroborated in independent studies(209, 214). While these findings represent starting points for ‘precision medicine’ to individualize intensity of SaB monitoring and therapy based on predicted patient outcome, the host response to infection is highly complex and extends beyond just immunological parameters. For example, lactic acidosis and acute kidney injury are associated with increased SaB mortality, and are potential markers of mitochondrial dysfunction, reflecting profound effects on host metabolism(215-217). Thus, a comprehensive and unbiased assessment of host factors altered in response to SaB may elucidate additional key features that can help predict patient outcomes and guide development of novel therapeutics. In this context, critical host responses modulated by interaction with an invading pathogen must be elucidated, not only as single entities, but how they function as a dynamic, concerted network.

Here, we construct a comprehensive bioanalytical snapshot of SaB patient serum collected immediately upon clinical presentation, providing the earliest view possible of the *in vivo* human response to infection. Using metabolomics and multiplexed quantitative proteomics in tandem, we analyzed samples from two cohorts totaling >200 individuals, including healthy and hospitalized uninfected controls. By conducting multiple rounds of biomarker analyses, we defined new features with strong individual predictive value that can be further used concert to increase predictive confidences. The depth of our analysis was increased through advanced computational methods, which identified prevalent post-translational modifications (PTMs) and inferred underlying cytokine signaling networks. These techniques pinpointed glycopeptides as more precise biomarkers and uncovered widespread carbamylation on abundant serum proteins in patients who succumbed to

infection. Ultimately, these data provide a starting point for the development of a rapid clinical test for identifying patients at high mortality risk. Further investigation on these findings would allow clinicians to identify patients who need intensified monitoring and therapy when it is most critical for a successful outcome, rather than responding post-hoc to failures in standard of care.

Methods

Patient and Isolate Identification and Collection. Patients were identified with SaB for study inclusion by electronic notification of blood cultures growing *S. aureus*, identified by Matrix Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF, Bruker Scientific LLC, Billerica, MA, USA). Methicillin-resistance was identified using GeneXpert® test (Cepheid, Sunnyvale, CA, USA). Patients were included if at least two positive blood cultures were identified, or one positive culture was congruent with a clinical diagnosis of SaB from an Infectious Diseases Physician Specialist. This study did not analyze consecutive samples from SaB patients, but rather outcomes of death (n=76) and survival (n=99) were selected from the SaBIR biobank for multi-omic serum analysis. The other subject groups included non-hospitalized, non-infected healthy volunteers (n=15), and hospitalized, non-infected patients (n=10) at UW Health identified through the electronic medical record.

Patient serum samples were obtained on the same day of initial presentation of SaB, before antibiotic therapy initiation and often within 1 hour of blood culture. The samples were stored at -80 °C until analysis.

Clinical Measurements and Outcomes. Patient electronic medical records were reviewed to collect basic demographics. The mean age was 58.7±15.5 years and 49.1% of patients were male. In the SaB patient group, 33.2% were infected with MRSA and 66.8% had MSSA bacteremia, identified as above and confirmed by routine antimicrobial susceptibility testing in the clinical microbiology laboratory. Total duration of bacteremia included cases of persistent bacteremia (consecutive days of positive blood cultures) and in-hospital microbiologic relapse defined as recurrence of a positive blood culture after the first negative culture while receiving appropriate antibiotic. The median duration of bacteremia duration was 2 days with an interquartile range of 1-4 days. All included patients received appropriate antimicrobial therapy for the treatment of MSSA (anti-staphylococcal β-lactam or vancomycin/daptomycin where needed for β-lactam allergic patients) and MRSA bacteremia (vancomycin or daptomycin).

Serum Metabolite Extraction. All steps for this protocol are to be done on ice. Serum samples (100 µl) were thawed for 30 mins, then 400 µl of prechilled extraction solvent (100% MeOH with 1µM sulfamethazine as a synthetic standard) was added to each sample. Samples were mix using a vortexer for 2 minutes (min) then incubated at -20 °C for 20 min to aid in protein precipitation. Samples were centrifuged the samples at 16,000 x g for 15 min to pellet the protein precipitate. The supernatant was transferred into 96-Well DeepWell, dried using centrifugal low-pressure system and stored at -80 °C once dry.

Metabolomic LC-MS^2 Analysis. Metabolomic LC-MS^2 was performed on a Bruker Daltonics® Maxis qTOF mass spectrometer (Bruker, Billerica, MA USA) with a Thermo Scientific UltiMate 3000 Dionex UPLC (Fisher Scientific, Waltham, MA USA). Metabolites

were separated using a Phenomenex C_{18} core shell (50 x 2 mm, 1.7 µm particle size) UHPLC column fitted with a C_{18} guard cartridge. The mobile phase solvents (solvent A, water-0.1% formic acid; solvent B, acetonitrile-0.1% formic acid) were run at a flow rate of 0.5 ml/min and chromatographic separation was achieved using the following elution gradient: 0 to 1 min 5% B, 1 to 10 min a linear increase from 5 to 100% B, 10 to 12 min held at 100% B, 12 to 12.5 min a linear decrease from 100 to 5% B, and 12.5 to 13 min maintained at 5% B. The mass spectrometer was calibrated twice daily using Tuning Mix ES-TOF (Agilent Technologies). For accurate mass measurements, lock mass internal calibration used a wick saturated with hexakis (2,2-difluoroethoxy)phosphazine (Synquest Laboratories, m/z 622.0289) located within the source. Ions were generated using the following parameters: nebulizer gas pressure, 2 Bar; Capillary voltage, 3,500 V; ion source temperature, 200°C; dry gas flow, 9 l/min; spectra rate acquisition, 3 spectra/s. Full scan MS spectra (m/z 50 – 1500) were acquired in the qTOF and the top five most intense ions in a particular scan were fragmented using a ramped collision induced dissociation (CID) energy from 10 - 50 eV. Data dependent automatic exclusion protocol was used so that an ion was fragmented when it was first detected, then twice more, but not again unless its intensity was 2.5x the first fragmentation. This exclusion method was cyclical, being restarted after every 30 seconds.

Metabolite Molecular Networking and Identification by GNPS. Metabolomics data files were converted to the .mzXML format using the Bruker Data Analysis software and uploaded to GNPS(218) through the MassIVE server (MSV000083593). Molecular networking was optimized as previously described(219) to an estimated false discovery rate of 1%. The data was filtered by removing all MS^2 fragment ions within +/- 17 Daltons

(Da) of the precursor m/z. MS^2 spectra were window filtered by choosing only the top 6 fragment ions in the +/- 50 Da window throughout the spectrum. The precursor ion mass tolerance was set to 0.05 Da and an MS^2 fragment ion tolerance of 0.05 Da. A network was then created where edges were filtered to have a cosine score above 0.59 and more than 6 matched peaks. Spectra were searched against the spectral libraries contained within GNPS. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least 6 matched peaks.

Area under the curve feature abundances were calculated to produce a metabolome bucket-table with the mzMine software(220). MzMine modules were used with the following settings. Peak mass detection: 1E3 MS^1 noise level, 1E2 MS^2 noise level. Chromatogram deconvolution: Local minimum search algorithm: 0.2 min minimum retention time (RT) range, 3 min ratio of peak top/edge, 0.05 – 0.5 min peak duration range, 0.05 Da m/z range for MS^2, and 0.2 min RT range for MS^2. Isotopic peak grouper: 0.05 m/z tolerance, 0.1 min RT tolerance, maximum charge 4. Join aligner: 0.01 m/z tolerance, 0.3 min RT tolerance, 75% weight for m/z, 25% weight for RT, 2 minimum peaks per row. Gap filling: 20% Intensity tolerance, 0.01 m/z tolerance, 0.2 min RT tolerance. Peak filter: Area 1E3 – 1E12. The abundances of each feature in the final bucket-table were normalized first by abundance of the internal standard within each sample and next by the total ion intensity of each sample.

PTM-tolerant Proteomics Workflow. Proteomics analyses were conducted as described above with the following alterations. For high-resolution spectra (fourteen 10-plexes), ions were fragmented using higher-energy collision-induced dissociation (HCD)

with a normalized collision energy of 30% and were detected in the Orbitrap with a resolution of 3 x 10^4. High resolution MS^2 data from proteomic experiments were submitted to molecular networking via GNPS as described above. Overrepresented mass shifts, as determined by the total number of network edges corresponding to each mass shift, were selected as modifications to include in a PTM-tolerant search. MS^2 data were queried against a focused human serum proteome database (proteins detected in standard search, 1,088 entries) using Byonic. A decoy search was also conducted with sequences in reversed order (3, 66, 67). For MS^1 spectra, a mass tolerance of 50 ppm was used and for MS^2 spectra a 0.05 Da tolerance was used. Static modifications included TMT 10-plex reagents (+229.162932 Da) on peptide n-termini and carbamidomethylation of cysteines (+57.02146 Da). Variable modifications were specified using Modification Fine Control. Variable modifications included: deamidation (+0.984016 Da) of asparagine and glutamine, oxidation (+15.99492 Da) of methionine, tryptophan and histidine, formylation (+27.994915 Da) of lysine, dioxidation (+31.989829 Da) of tryptophan, carbamylation (+43.005814) of lysine and arginine and dihydroxyimidazolidine (+72.021129 Da) of arginine. Spectra that contained low-mass glycan reporter ions as determined by the IMP-glycan reporter node were submitted to a glyco-peptide search with the following modification parameters. Static modifications included: TMT 10-plex reagents (+229.162932 Da) on peptide n-termini and lysines and carbamidomethylation of cysteines (+57.02146 Da). Variable modifications included: oxidation (+15.99492 Da) of methionine and glycosylation (57 common human N-glycans – various Da) of asparagine. Reporter ion intensities for modified peptides were summed to the unique peptide level then normalized as above. PTMs were localized in the context

of the total protein length and flanking sequences were extracted using the *PTMphinder* R package(124).

Mouse Model of SaB. 8-week old female CD1 mice were used for all animal experiments. Mice were treated before infections as follows or with vehicle controls. Hyperthyroid mice were given I.P. injections of 100 µg thyroxine (T4) once daily for the three days prior to infection. Hypothyroid mice were given drinking water containing hypothyroid treatment (1% (wt/vol) sodium perchlorate and 0.1% (wt/vol) methimazole) for three weeks prior to infections. Adiponectin mice were given I.P. injections of 1 mg/kg AdipoRon one day prior infection, then injected daily with AdipoRon for the duration of the experiment. Mice were then I.V. infected with *S. aureus* LAC (high dose (Fig 7B): $1x10^8$ CFUs, low dose (Fig 7C-H): $5x10^7$ CFUs) and survival was monitored every 12 hours. For CFU burden experiments, mice were treated and infected as above, then euthanized 48 hours post-infection and organs were harvested for quantitation of bacterial burden.

Results

Overview of Multi-omic SaB Patient Serum Analysis. We employed a multi-omic approach to gain a comprehensive view of the SaB host-pathogen interaction (**Fig 4-1A**). First, a small discovery cohort was analyzed by standard multiplexed proteomic techniques to assess and optimize the ability of serum protein biomarkers to predict SaB patient phenotypes. This initial analysis identified 1,405 proteins with a false discovery rate (FDR) <1%, and yielded biomarker candidates associated with various disease features including infection, mortality and duration of bacteremia (**Table S1**). To verify these initial markers and deepen our comprehension of the molecular features related to

this disease state, an expanded cohort was analyzed through both proteomic and metabolic approaches (**Fig 4-1A**). In the extended cohort, standard proteins were identified by a conventional database search. Metabolomics data were interrogated for small molecules using MZmine(221) and Global Natural Products Social Networking Molecular Networking (GNPS)(218). Modified proteins, including glycoproteins, were identified by combining molecular networking with a PTM-tolerant database search through Byonic(222). Importantly, all of the modified protein identifications were derived from high-resolution MS^2 scans. Overall, 1,088 proteins, 5,280 metabolomic features and 6,700 modified peptides were quantified in this experiment. The resultant >10,000 molecular features were statistically analyzed with binary comparisons to identify biomarkers, and through clustering and network based approaches to define disease associations within our primary sample groups (Control groups: NN – Non-hospital, Non-infected, HN – Hospital, Non-infected; Infection groups: HS –Hospital, Survival, HM – Hospital, Mortality).

An initial hierarchical clustering analysis of the standard proteomics data showed clear segregation of the control samples from the infected samples (**Fig 4-1B**). In contrast, SaB survival and mortality groups were highly intermixed, indicating that the differences between death and survival in SaB infections were subtle. Nevertheless, stratification of death and survival groups was observed in the clustered dendrogram, indicating the potential within this data to probe for mortality biomarkers. As a proof-of-principle for distinguishing disease states, we first selected a highly discriminating protein for predicting infection, rather than mortality. SERPINA5 emerged as a top hit (**Fig 4-1C**), with a receiver operator characteristic (ROC) curve area under the curve (AUC) of 0.9891

(**Fig 4-1D**). This protein had a higher AUC than the commonly used clinical marker for infection, C-reactive protein (CRP)(223), in the current dataset (AUC = 0.9691) as well as values reported in the literature (maximum AUC = 0.92)(224-226). This example demonstrates the power of unbiased proteomics for biomarker discovery.

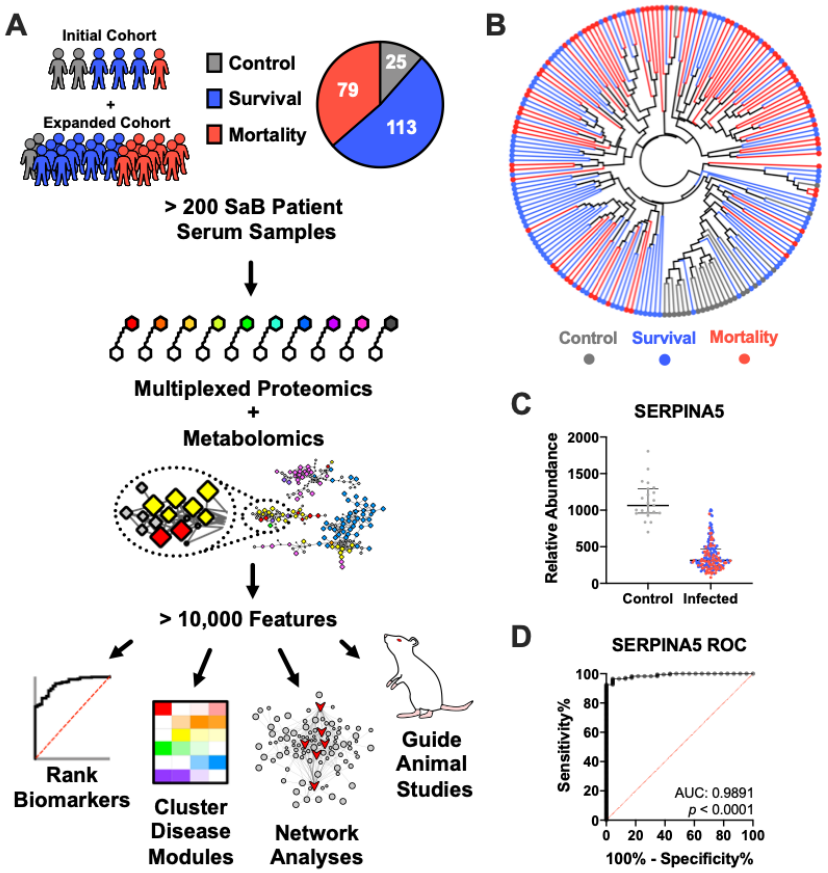


Figure 4-1 *Multiomic Analysis of SaB Patient Serum*

(**A**) Workflow for multi-omic analysis of SaB patient serum. (**B**) Unbiased hierarchical clustering based on Pearson correlation values for proteins detected across all samples. (**C**) Relative protein abundance of SERPINA5 in control and infected samples. (**D**) ROC curve of SERPINA5 abundance discriminating between control and infected samples.

Definition of High-confidence Biomarkers to Predict SaB Patient Mortality. Our first goal was to define high-confidence biomarkers to predict SaB patient mortality. To rank biomarkers, we took an ensemble feature selection (EFS) approach(227), which applied multiple feature selection algorithms, then aggregated and ranked the results. This method can avoid biases associated with individual feature selection algorithms(228). This workflow was applied to the two primary datasets and used to rank the top biomarkers (Proteomics – **Fig 4-2A**; Metabolomics – **Fig 4-2D**). The highest ranked protein biomarkers were fetuin B (**Fig 4-2B**), heparin cofactor II (encoded by the SERPIND1 gene, **Fig S1A**), and insulin-like growth factor binding protein 3 **(**IGFBP-3, **Fig S1B**), all with decreased serum levels. The decrease in serum fetuin B was also captured in the initial cohort, despite a low number of mortality samples analyzed (**Fig S1C**). Our top biomarkers with increased serum levels were insulin-like growth factor binding protein 2 (IGFBP-2, **Fig 4-2C**), adiponectin (encoded by the ADIPOQ gene, **Fig S1D**) and EGF-containing fibulin-like extracellular matrix protein 1 (EFEMP1, **Fig S1E**). Applying a similar approach to the metabolomics data, we found that all of the top-ranked biomarkers were unidentified MS features (**Fig 4-2D**). However, these molecules showed considerable predictive utility (**Fig 4-2E and F**), similar to our top-ranked protein biomarkers (ROC AUC = ~0.75; p-value < 0.0001).

The ensemble feature selection approach ensures that the top-ranked biomarkers are not correlated to one another and therefore could be used in combination for the enhanced prediction of SaB patient mortality(229). Using the top two markers from both workflows significantly enhanced the predictive power relative to the individual markers alone (**Fig 4-2G**). To support the proteomics workflow, we validated fetuin B (**Fig 4-2H**)

and other top biomarkers (**Fig S1F-G**) in a subset of our samples using enzyme-linked immunosorbent assays (ELISAs) (Fetuin B AUC = 0.8945). Our results indicate that patients with low fetuin B levels (< 2.2 μg/ml) had significantly reduced survival compared to patients with high concentrations of fetuin B (**Fig 4-2I**).

An important factor to consider when performing biomarker analyses is the influence of confounding factors on the proposed biomarkers(230). To investigate this issue, we performed a metadata-wide association assessment for every protein and metabolite detected in our experiments (28 total metadata variables). Using an appropriate statistical test for each variable type (ie. Mann-Whitney U test for binary categorical metadata, Kruskall-Wallis test for >2 categorical metadata and Pearson correlation for continuous metadata), the association of any particular feature with every measured metadata variable can be determined. Through this approach, we found that all top biomarkers, both up and downregulated, are predominantly associated with infection and mortality, with minimal associations to other metadata (Proteomics - **Fig 4-2J, S2A-G**; Metabolomics – **Fig S2H-M**). The next most common metadata associated with the top biomarkers are dialysis and serum creatinine levels, both related to kidney function(231) and mortality in bacteremia(232, 233). This analysis verified that there was negligible influence of typical confounding variables (such as age, gender, hospitalization, etc.) on the top SaB mortality-related biomarkers. Overall, the above results provide a comprehensive list of biomarkers associated with SaB mortality risk.

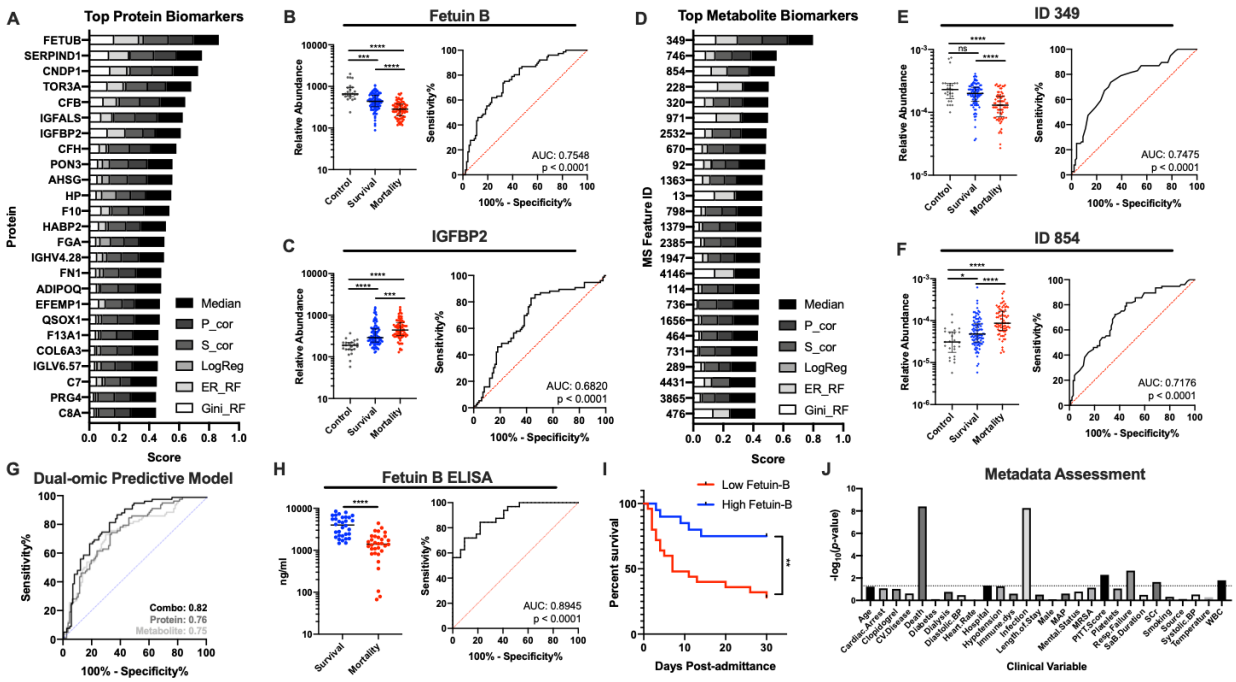


Figure 4-2 *Definition of High-confidence Biomarkers for the Prediction of SaB Patient Mortality*

(**A**) Top 25 EFS ranked protein biomarkers for discriminating patient survival vs. mortality. (**B**) Relative protein abundance and ROC curve of Fetuin B discriminating between survival and mortality samples. (**C**) Relative protein abundance and ROC curve of IGFBP2 discriminating between survival and mortality samples. (**D**) Top 25 EFS ranked metabolite biomarkers for discriminating patient survival vs. mortality. (**E**) Relative protein abundance and ROC curve of metabolite ID 349 discriminating between survival and mortality samples. (**F**) Relative protein abundance and ROC curve of metabolite ID 854 discriminating between survival and mortality samples. (**G**) Dual-omic ROC curve for discriminating between survival and mortality samples (**H**) Absolute protein abundance and ROC curve of Fetuin B discriminating between survival and mortality samples measured via ELISA. (**I**) Survival curves of Fetuin B high (> 2.2 µg/ml) and low (< 2.2 µg/ml). (**J**) Metadata assessment of Fetuin B protein levels measured via proteomics.

PTM-tolerant Analysis of Serum Samples Enables the Identification of Disease-associated PTMs. Serum is a notoriously difficult sample to analyze via proteomics(234, 235), attributable to a large dynamic range of protein concentrations and high numbers of PTMs, such as glycosylation. Thus, standard proteomic searches fail to identify greater than 90% of the spectra acquired from mass spectrometry (MS)-based shotgun proteomics of serum(236). We hypothesized that if we could predict highly abundant protein PTMs, we could perform a PTM-inclusive database search and match more spectra per sample. To identify the most wide-spread protein modifications in the serum proteome data, we used GNPS-based molecular networking to group similar spectra that differ by regular mass shifts. Overall, we were able to network >80% of the MS/MS spectra (**Fig 4-3A**), suggesting that many of the peptides identified in the standard database search may also be present as modified variants. Displaying the total number of mass shift occurrences in a histogram, we can observe highly abundant PTMs present in our data (**Fig 4-3B**). These include expected artifacts such as oxidation of methionine (+15.99), carbamidomethylation of cysteine (+57.02), and deamidation (+0.984), but also unanticipated modifications such as carbamylation (+43.005), dioxidation (+31.99), and formylation (+27.99). Of note, the glycan moieties fucose (+146.06), hexose (+162.05), and sialic acid (+291.1) were also highly abundant. This suggests that the peptides captured in our MS analysis are rich in PTMs, complicating identification through traditional strategies.

To address this issue and take these modifications into account, we employed a PTM-tolerant search strategy. First, we focused the serum proteome database, including only proteins that were matched in the initial, standard search. Then, we used the Byonic

search algorithm(222), whose Modification Fine Control feature allows specifying high numbers of modifications and can also match intact glycopeptides. Including the top predicted modifications and common human plasma glycans, we achieved a doubling of the serum spectral match rate (**Fig 4-3C**). Low-weight modifications ultimately called in the PTM-tolerant search correlated highly with the number of edges of the corresponding mass shifts from the initial GNPS analysis (**Fig 4-3D**). Further, the vast majority (> 85%) of glycosylation sites identified have been previously reported in Uniprot (**Fig 4-3E**), authenticating their identification as true glycopeptides. The distribution of mass errors from the modified peptides identified in the PTM-tolerant search was nearly identical to the standard search (**Fig 4-3F**), suggesting there is minimal loss in spectral identification quality. We also found that the total peptide-spectrum matches (PSMs) and unique peptides per protein were highly correlated between the standard search and the PTM-tolerant workflow (**Fig 4-3G-H**). One of the initial benefits noted from the PTM-tolerant search was an increased number of unique peptides detected for low-abundant proteins. Proteins with the smallest number of unique peptides in the standard search gained a higher proportion of unique peptides in the PTM-tolerant search than those proteins that had more unique peptides originally detected (**Fig 4-3I**). Gene ontology (GO) analysis(237) revealed that the majority of proteins with boosted unique peptides were immunoglobulins (**Fig 4-3J**), but a number of intracellular proteins also showed large gains in unique peptides (**Fig 4-3K**). While some of these proteins had minimal abundance changes among groups, there are others that demonstrate significant associations to hospitalization (**Fig 4-3L**) and infection (**Fig 4-3M**). Our initial results demonstrate that predicting highly abundant modifications can inform subsequent search

strategies, yielding higher spectrum match rates and increased confidence in low-abundant proteins while also identifying disease-associated PTMs.

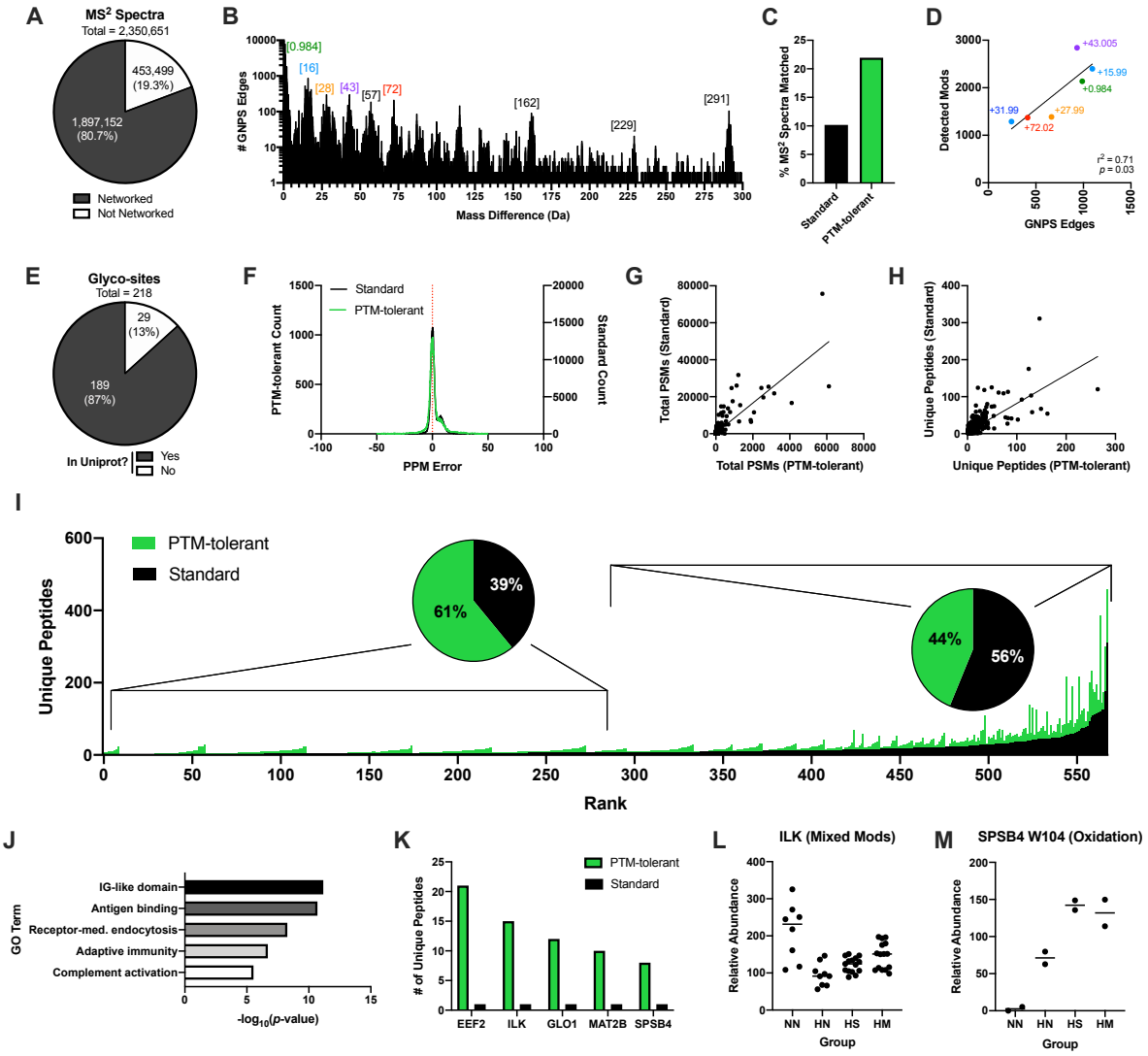


Figure 4-3 *PTM-tolerant search enables deeper proteomic analysis and identification of disease-relevant PTMs*

(**A**) Pie chart of MS^2 spectra incorporated into GNPS network. (**B**) Edge histogram from GNPS with top mass shifts highlighted. (**C**) Percent of MS^2 spectra matched in standard and PTM-tolerant workflows. (**D**) Correlation of GNPS edges and PTMs detected in PTM-tolerant workflow. (**E**) Proportion of detected glyco-sites present in Uniprot. (**F**) MS^1 mass errors for standard and PTM-tolerant database searches. Correlations of total PSMs (**G**) and unique peptides (**H**) per protein detected in the standard and PTM-tolerant database searches. (**I**) Unique peptides detected in the standard and PTM-tolerant database search ranked by number of unique unmodified peptides then number of unique modified peptides. Pie charts depict unique peptide proportions of top and bottom 50% of proteins detected in the standard and PTM-tolerant workflows. (**J**) GO analysis of proteins with bottom 50% of unique peptides in the standard search. (**K**) Proteins with the largest gain in unique peptides detected in the PTM-tolerant search. (**L**) Relative abundance of modified ILK peptides detected in PTM-tolerant search. (**M**) Relative abundance of dioxidation of SPSB4 104W detected in PTM-tolerant search.

We reasoned that some modified peptides might represent high quality biomarkers for predicting infection or SaB mortality. Due to the incompatibility of the higher number of missing values for the modified peptides and the EFS approached utilized above, we ranked biomarkers simply on Mann Whitney U test *p*-values. Interestingly, the top biomarkers for infection and mortality were both glycan PTMs of alpha-2-HS-glycoprotein (AHSG), also known as fetuin A (**Fig 4-4A-B**). The support for these two glycans is strong as evidenced by >50 unique PSMs detected for each peptide, some of which have Byonic scores of greater than 400 (corresponding to an FDR <0.1%)(222). Individually, these biomarkers demonstrated higher ROC curve AUC values than our top unmodified biomarker identified in the standard search (0.9981 vs. 9891 for infection and 0.8066 vs. 0.7548 for mortality, **Fig 4-4A-B**). Unmodified fetuin A was also a top biomarker for both infection and mortality (**Fig 4-2A**), although the observed fold-change was higher for the glycans than the total protein (**Fig 4-4C**). Again, we assessed the relation of these biomarkers to the collected metadata and primarily found associations with infection and mortality (**Fig 4-4D**). Together, this suggests that these glycans may yield better predictive value than the total protein alone. When used in concert with our top unmodified protein and metabolite biomarkers (nine total molecular features), these PTMs further enhanced predictive power (AUC = 0.92, **Fig 4-4E**). To our knowledge, this approach generated the best model, based on AUC and *n*, for predicting mortality from any infection using patient-derived biomarkers.

The above PTM analysis was performed on the normalized raw abundances for each modified peptide without consideration for the change in total protein levels. However, the modified peptide abundances can also be normalized to the total protein

level to investigated potential divergent regulation of protein and PTM. Overall, we found that the majority of modified peptides had a positive correlation to their respective total protein levels. Correspondingly, when we compare the fold-changes of the modified peptide abundance to the protein normalized values for infection and mortality changes, we observe highly similar results (**Fig4-4F-G**); however, there were some exceptions. To better understand which modifications deviated from their total protein level, we first filtered the protein-normalized PTMs with significant alterations in any of the primary sample groupings (ie. NN, HN, HS, HM) using ANOVA ($p < 0.05$). Then, we clustered the significantly altered features using the K-means algorithm. Most striking among these was a cluster of modifications that showed a stark increase in abundance only in the mortality samples (**Fig 4-4H**). Interestingly, nearly half (46%) of the modified peptides in this cluster were assigned to only two proteins: albumin and serotransferrin (**Fig 4-4I**). Moreover, the modifications on these peptides were primarily carbamylation and formylation (**Fig 4-4I**). Comparing the relative changes in the modified peptide to the total protein abundance for albumin (**Fig 4-4J**), it was observed that, while total albumin levels dropped upon infection and were reduced further in mortality patients, albumin was modified at a higher level in mortality patients compared to all other groups (**Fig 4-4J**). Modifications on serotransferrin demonstrated a similar trend (**Fig 4-4K**). Together, this analysis enabled a deeper interrogation of serum-derived proteomics data and linked multiple, distinct PTMs, including specific glycosylation and carbamylation sites, to increased mortality in SaB patients.

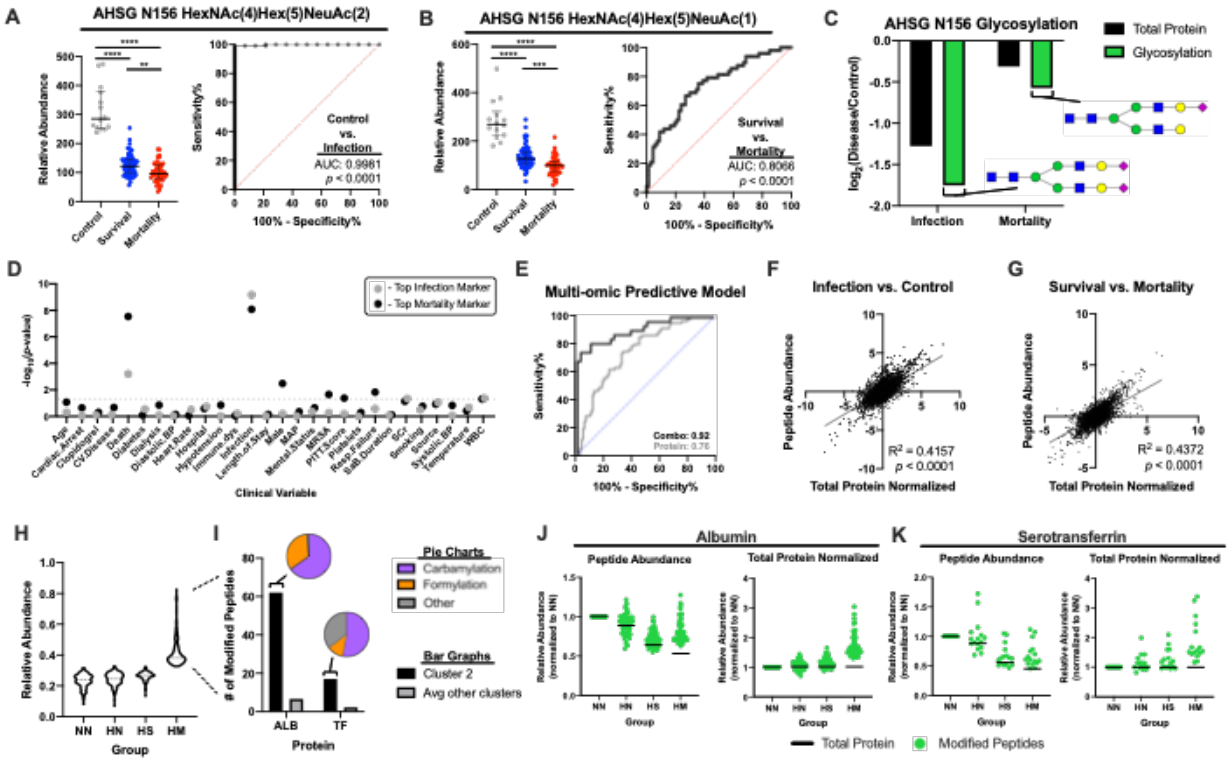


Figure 4-4 *Mortality Associated PTM Signatures in SaB Patient Serum*

(**A**) Relative peptide abundance and ROC curve of AHSG N156 HexNAc(4)Hex(5)NeuAc(2) discriminating between control and infected samples. (**B**) Relative peptide abundance and ROC curve of AHSG N156 HexNAc(4)Hex(5)NeuAc(1) discriminating between survival and mortality samples. (**C**) Relative fold-changes of total AHSG protein and glycosylations of N156 for infection and mortality samples. Depicted glycan structures were inferred from the monosaccharide compositions in accordance with common serum glycans. (**D**) Metadata assessment of top modified biomarkers for infection and mortality. (**E**) Multi-omic ROC curve for discriminating between survival and mortality samples. Scatter plot of fold-changes comparing (**F**) control vs. infected and (**G**) survival vs. mortality. (**H**) Relative abundance of modified peptides assigned to mortality-specific expression cluster (Control groups: NN – Non-hospital, Non-infected, HN – Hospital, Non-infected; Infection groups: HS –Hospital, Survival, HM – Hospital, Mortality). (**I**) Distribution of peptide counts and modifications types of albumin (ALB) and serotransferrin (TF) detected in PTM tolerant workflow. (**J**) Albumin mortality associated PTM plot depicting modified peptide abundance (left) and modified peptide abundance normalized to total protein levels (right). (**K**) Serotransferrin mortality-associated PTM plot depicting modified peptide abundance (left) and modified peptide abundance normalized to total protein levels (right).

Unbiased Clustering of SaB Disease Modules. While our primary goal was to define biomarkers for the prediction of SaB mortality, our cohort represents the largest serum-focused bioanalytic experiment of human *S. aureus* infection to date. Therefore, we sought to further understand the effects of *S. aureus* infection on the human serum landscape using our two primary datasets, the standard proteomics and metabolomics findings. Performing a similar clustering approach used for the PTM analysis above, the proteomics data was grouped into 6 clusters (**Fig 4-5A**) and the metabolomics data grouped into 7 clusters (**Fig 4-5D**).

Simply observing the proteomic clusters, we found some interesting trends in the expression profiles (**Fig 4-5A**). Specifically, proteomics cluster 2 (C2 - orange) captured the host response to infection, regardless of mortality status. Within this cluster we found C-reactive protein and serum-amyloid proteins 1 and 2, along with other major components of the acute-phase response (**Fig 4-6A-E**). Other interesting clusters include clusters 4, 5 and 6, which showed increases (cluster 5) and decreases (clusters 4 and 6) in the mortality group making them prime clusters for further investigation. Clusters 1 and 3 were largely healthy control-specific and hospital control-specific, respectively, but did not demonstrate a strong association with infection or mortality. For the sake of clarity, the mortality-associated proteomics clusters were renamed according to their expression direction and magnitude compared to control samples (C4: pMortality-, C5: pMortality+, C6: pMortality--; p signifies “proteomics”).

To gain a deeper understanding of the crosstalk of proteins between clusters, we submitted all of the clustered proteins to a protein association network analysis on String-DB(123) (**Fig 4-5B**). Interestingly, we found the largest number of connections between

proteins within the pMortality+ and pMortality-- clusters (**Fig 4-5C**), despite the fact that they change in opposite directions relative to the control patients. This suggests that proteins that increase in expression may be related to the decrease in expression of another protein, and vice-versa.

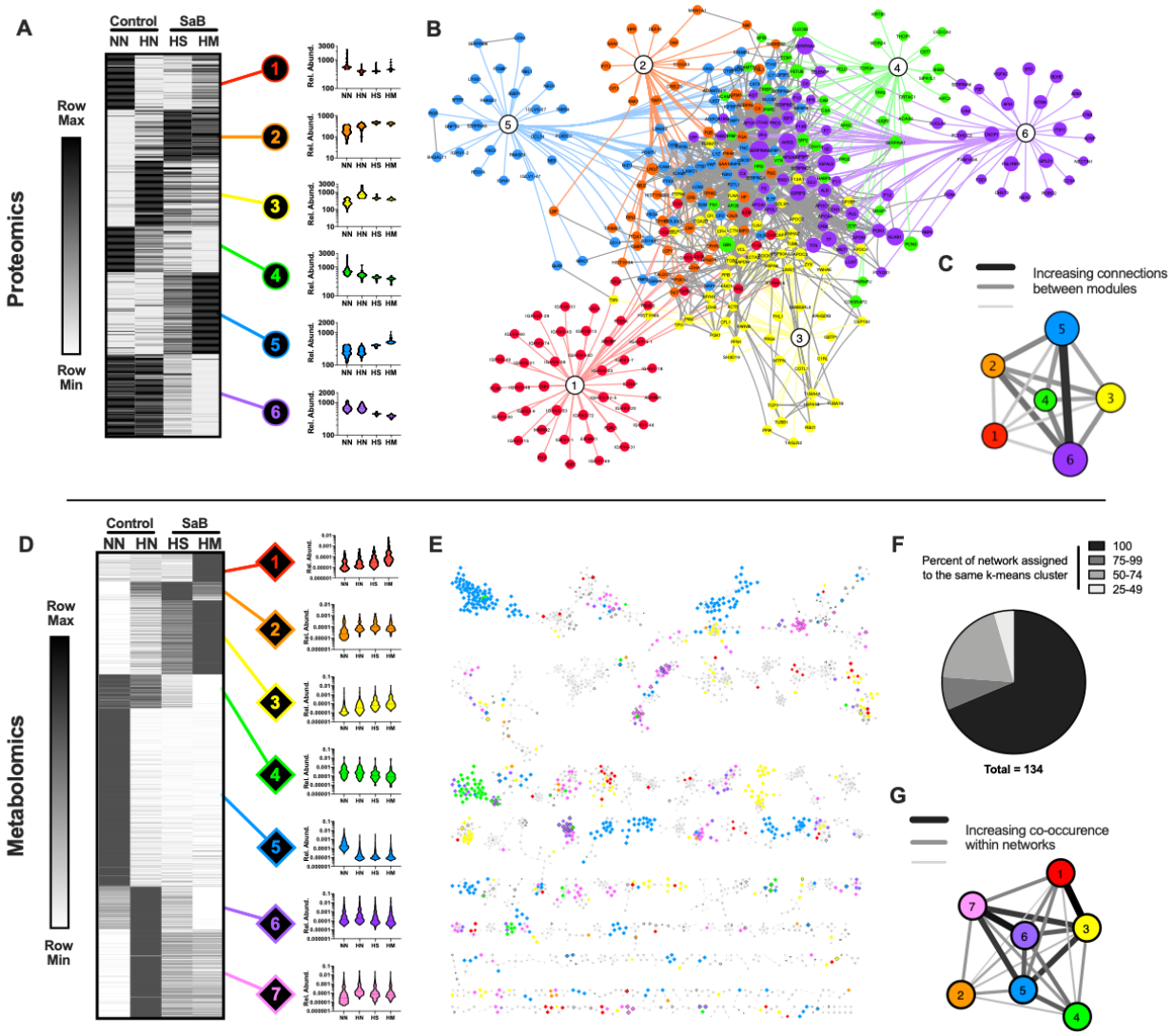


Figure 4-5 *Clustering of MS Data into Disease-relevant Modules*

(**A**) K means clustered heatmap, (**B**) protein association network, and (**C**) module cross-talk network of all significantly altered proteins (ANOVA $p < 0.05$) across the four primary groups (Control groups: NN – Non-hospital, Non-infected, HN – Hospital, Non-infected; Infection groups: HS –Hospital, Survival, HM – Hospital, Mortality). In B and C, nodes are colored according to cluster designations in A. (**D**) K means clustered heatmap, (**E**) molecular networking overview, (**F**) within network co-regulation pie chart, and (**G**) module cross-talk network of all significantly altered metabolites (ANOVA $p < 0.05$) across the four primary groups. In E and G, nodes are colored according to cluster designations in D.

Due to their prominent association with SaB mortality, we delved further into the major groups of proteins within the mortality-associated clusters using GO analysis (**Fig 4-6F-H**). While the pMortality- cluster had few, low-significant enrichments (**Fig 4-6F**), the pMortality+ and pMortality-- clusters had multiple, highly significant protein groups enriched. Cluster pMortality+ is dominated by extracellular matrix (ECM) and insulin-like growth factor binding proteins (IGFBPs) and has a moderate enrichment for tumor necrosis factor (TNF)/ interleukin-1 (IL1) response (**Fig 4-6G**). The ECM adhesion proteins ICAM1 and VCAM1 have previously been shown to be elevated in SaB patients, particularly those with endocarditis(238), and TNF can be used as mortality biomarkers in humans(208). However, the striking enrichment for IGFBPs in this cluster represents a novel finding in our data. In contrast, pMortality-- was highly enriched for protease inhibitors, complement/coagulation cascade members and lipoproteins (**Fig 4-6H**). A decrease of lipoproteins is well described in septic patients(239) and the reduction in complement/coagulation proteins is consistent with the activation of these proteins by proteolytic cleavage. We also noted that a subset of IGFBPs were present within this cluster and possessed some of the most significant ANOVA *p*-values, which is particularly interesting given the upregulation of other family members observed in pMortality+. The divergent regulation of distinct IGFBP family members is discussed further below.

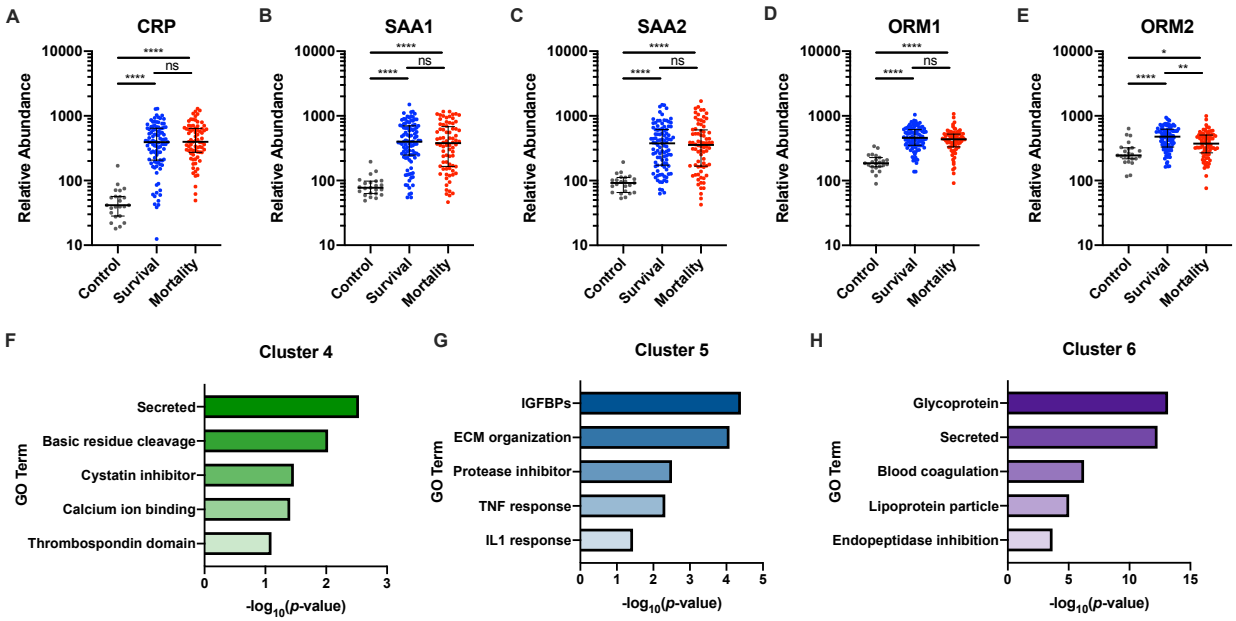


Figure 4-6 *Detailed Analysis of Proteomic SaB Disease Modules*

(**A-E**) Individual plots for major acute-phase reactant proteins contained within proteomics infection-associated cluster 2 from Fig 4A including: (**A**) CRP, (**B**) SAA1, (**C**) SAA2, (**D**) ORM1, and (**E**) ORM2. (**F-H**) GO analysis of proteomics mortality-associated clusters: pMortality- (**F**), pMortality+ (**G**) and pMortality-- (**H**).

Similar to the proteomics clusters, biological associations could be defined for our metabolite clusters based on their expression profiles (**Fig 4-5D**). For example, clusters 1 and 3 showed increased expression in the mortality group relative to other cohorts. In contrast, cluster 2 showed an increase in the survival group relative to mortality and cluster 4 showed a decrease in expression with infection, which went down further if the patient died. Cluster 5 had a strong healthy association and clusters 6 and 7 showed increased expression in the hospitalized, non-infected group. Again, we renamed the most interesting clusters according to their mortality expression directions and magnitude (C1: mMortality++, C3: mMortality+, C4: mMortality-; m signifies “metabolomics”).

While an association analysis tool like String-DB does not exist for metabolic data interpretation, metabolites can be grouped based on the similarity of their MS/MS spectra using a molecular networking approach(218). This analysis results in the formation of metabolite networks, ranging from 1 – 100 individual metabolites, that bear some structural similarity. To visualize our data more clearly, we overlaid the K-means cluster color onto the individual metabolites in these networks (**Fig 4-5E**). Taking a bird’s eye view of this data, we noted that nodes within a specific network were commonly assigned to the same expression cluster. In fact, >95% of all of our molecular networks had at least half of their nodes co-regulated (**Fig 4-5F**). We also noted that some clusters of similar expression profiles were often contained within the same networks, such as mMortality+ and mMortality++ (increased in infection/mortality) and clusters 6 and 7 (increased in hospitalization) (**Fig 4-5G**). Together these findings suggest that structurally related metabolites are often co-regulated, offering more support for their importance in the host response to infection and mortality.

Global Characterization of Metabolic Dysfunction in SaB Mortality Patients. As noted above, we detected a divergent regulation of IGFBPs in our mortality-associated protein modules. Specifically, we found that IGFBP1, 2, 4 and 7 were assigned to the pMortality+ cluster and thus had higher expression in mortality patients compared to controls and survival patients (**Fig 4-7A**). In contrast, IGFBP3, 5 and IGFALS were in pMortality--, demonstrating an opposite trend in expression (**Fig 4-7B**). IGFBPs function by binding to and stabilizing IGF-I and II in serum(240). The binding of IGFs to IGFBP1, 2, 4 or 7 results in the formation of binary complexes that extend the half-life of the IGF molecules from 2 - 30 minutes. However, if both IGFBP3 and IGFALS bind IGFs, forming a ternary complex, this stabilization is increased up to 24 hours. IGFBP5, can also form a ternary complex with IGFALS, albeit to a lesser extent than IGFBP3. Given the expression patterns above, we would expect the amount of IGFs in the blood to decrease as the constituents of the ternary complex decrease. Indeed, we noted a significant decrease in IGF-II with increasing disease severity and a similar trend with IGF-I, although the latter results did not attain statistical significance, likely due to the high number of missing values (**Fig 4-7C**). Further, a formal comparison of the correlations between IGFI and II with all the IGFBPs detected in our dataset revealed positive correlations of IGFs with IGFBP3, 5 and ALS, but mostly negative correlations with the rest of the IGFBPs (**Fig 4-7D**). Together, these data suggest that SaB mortality is associated with a decrease in the IGFBP ternary complex and an increase in binary complexes, ultimately resulting in lower levels of circulating IGF proteins.

Given the striking association of the IGF system with SaB mortality risk and the role of IGFs in host metabolism, we sought to further understand this metabolic

dysfunction. Thus, we mined the data for other molecular features related to general host metabolism. An obvious signature, derived from the GO analysis of our mortality-associated clusters, is the general decrease in apolipoproteins upon infection, which further decreases in SaB mortality patients (**Fig 4-7E**). A depletion of lipoproteins in response to infection is well known and has been proposed as a prognostic marker for severe sepsis(241, 242). Our results reinforce these findings further and establish a link between this phenomenon and SaB in addition to non-specific sepsis.

We also uncovered evidence for host metabolism dysfunction within our metabolomics dataset. The most prominent single feature we found was thyroxine (T4), a master regulator of host metabolism. T4 was assigned to cluster mMortality- and thus had a significant decrease in expression with infection, further extended in the mortality group (**Fig 4-7F**). In fact, T4 was our highest ranked, identified biomarker from our metabolomics screen. We also noted a similar trend in SERPINA7 (assigned to pMortality- cluster; **Fig 4-7F**), also known as thyroxine-binding globulin, which binds to and stabilizes T4 in circulation(243), further supporting the observed reduction in free T4 levels. Thyroid hormone dysfunction during non-specific sepsis, also known as euthyroid sick syndrome, is well characterized(244, 245); however, it has not been previously associated specifically with SaB infections nor mortality.

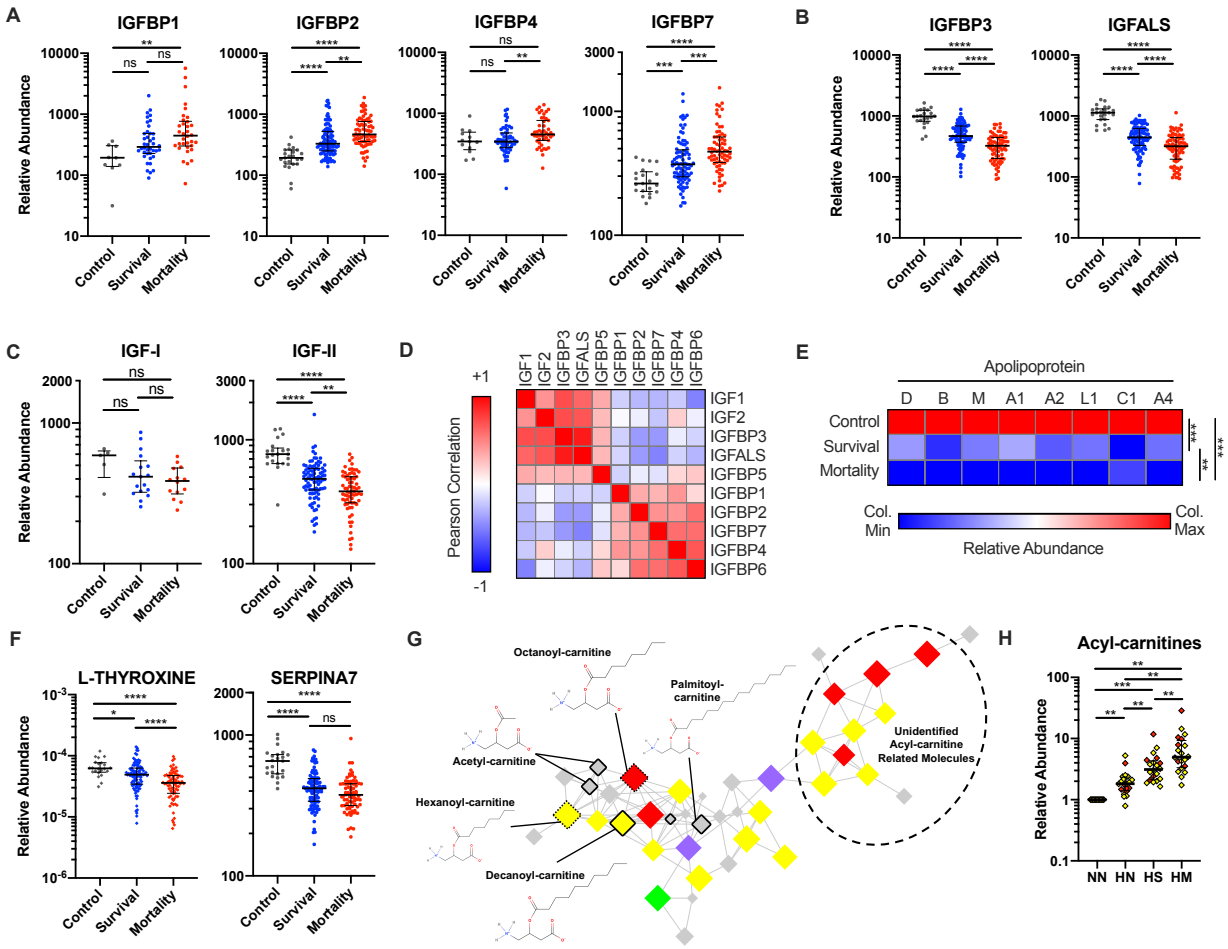


Figure 4-7 *Detection of Metabolic Dysfunction in SaB Mortality Patients*

(**A**) Relative protein abundance of IGFBP binary complex members. (**B**) Relative protein abundance of IGFBP tertiary complex members. (**C**) Relative protein abundance of IGF hormone molecules. (**D**) Correlation matrix of all IGFBP system family members. (**E**) Heatmap of relative apolipoprotein abundance. (**F**) Relative metabolite abundance of thyroxine signaling components. (**G**) Molecular network and (**H**) relative metabolite abundance of acyl-carnitines. In G and H, nodes and points are colored according to cluster designations in Fig 4D.

Another evident association of host metabolism with SaB mortality is the increase in decanoyl-carnitine abundance in mortality patients, which was our highest ranked, identified, metabolite biomarker with increasing expression (assigned to mMortality+ cluster). Decanoyl-carnitine is involved in fatty acid metabolism(246). Unlike T4, which was not grouped into a molecular network; rather, deconyl-carnitine was found to be part of larger network of related metabolites (**Fig 4-7G**). Observing this network as a whole, we found that the majority of nodes were assigned to mMortality+ (yellow) and mMortality++ (red) (**Fig 4-7G**), which increase in abundance with disease severity (**Fig 4-7H**). Most of these nodes were not identified in our initial GNPS analysis; however, we noted a large number of mass shifts of 28 Da, corresponding to two links in a fatty acid chain (Fig ie CH2-CH2). By following the mass shifts through the molecular network, we can assign identities to additional nodes such as hexanoyl-carntine and octanoyl-carnitine, which have stronger associations to mortality than the initially identified decanoyl-carnitine. Interestingly, there is a subset of this network that is moderately related to acyl-carnitines, which also demonstrates a strong association with SaB mortality and possess the 28 Da mass shifts suggestive of a fatty acid chain (**Fig 4-7G – circle**). Determining the molecular structures of these compounds and how they impact fatty acid metabolism would give us a deeper understanding of this metabolic dysfunction phenotype. Together, the above data indicate that the host metabolism dysfunction resulting from SaB may be predictive of patient outcomes.

Knowledge-based Analysis of Proteome Alterations Captures Underlying Cytokine Mortality Signatures. The above results describe a comprehensive assessment of the major molecular features associated with SaB mortality that are amenable to MS-based

analyses. However, major cytokine families, the focus of most infectious disease biomarker studies, were underrepresented in our dataset. These signaling molecule fall below the standard limit of detection in typical serum proteomic experiments(234), even in recent attempts at ultra-deep serum proteome coverage(236, 247), but have been shown to play major roles in disease, despite their low concentrations. Therefore, we designed a computational approach to infer the relative importance of major cytokine families from our proteomics data using functional protein association networks (**Fig 4-8A** – step 1). This analysis involves submitting a list of significantly altered proteins, along with lists of major cytokine families (eg. ILs, CXCLs, etc.), to the String-DB tool, resulting in a large network including both experimentally derived proteomic data and any known associations to cytokines. We can then determine which cytokines have the largest number of connections to proteins within our network, in specific expression clusters of interest (**Fig 4-8A** – step 2). By regenerating the networks using only the cytokines most associated with the clusters of interest, we can refine our networks to show the most likely underlying contributors to the observed protein expression (**Fig 4-8A** – step 3).

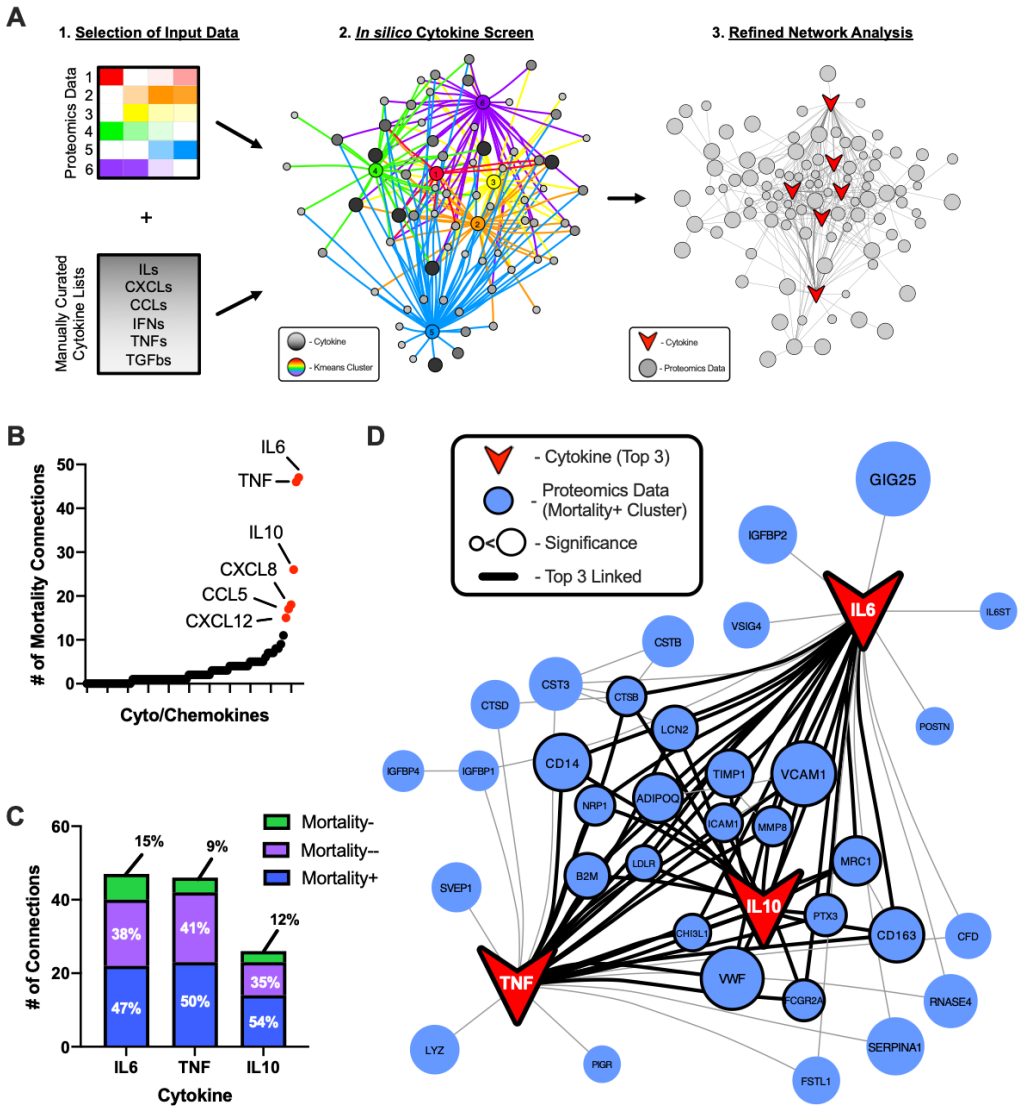


Figure 4-8 *Knowledge-based Analysis of Cytokines Predicts Major Contributors to Proteomic Alterations and Identifies Core of Modulated Proteins*

(**A**) Schematic for analysis of proteins that go undetected in proteomic analyses. (**B**) Inferred cytokines ranked by number of total mortality connections. (**C**) Connections of the top 3 ranked cytokine to each mortality-associated K means cluster. (**D**) Refined network of top 3 ranked cytokines and pMortality+ cluster proteins. In A, C, and D, heatmap, bars and circular nodes are colored according to cluster designations in Fig 4A.

In this experiment, we used complete lists of all the ILs, CXCLs, CCLs, IFNs, TNFs and TGFbs as our cytokine input lists. We then assessed which cytokines had the highest number of connections to the mortality-related proteomic profiles (ie. pMortality+, pMortality-, and pMortality--). We found that IL6, TNF, IL10, CXCL8, CCL5 and CXCL12 had the highest number of connections to our mortality networks (**Fig 4-8B**), suggesting they play a large role in the observed proteome. Notably, the top four inferred cytokines have demonstrated strong associations with SaB mortality in previous studies (TNF and IL10(208, 210); IL6 and CXCL8(214)), providing validation for this approach. This analysis further indicates that CCL5 and CXCL12 may also be linked to SaB mortality and merit further study. Focusing in on the top 3 cytokines predicted from this analysis (IL6, TNF and IL10) we found that the majority of the connections were to proteins in pMortality+ (**Fig 4-8C**). Regenerating a functional protein association network using only proteins from pMortality+ that are directly linked to a top 3 cytokine yielded a refined network (**Fig 4-8D**) that was much easier to interpret than the large, unwieldy networks initially generated. Similar networks can be generated with the proteins with reduced expression in mortality patients. Delving into the pMortality+ network, we found that ~50% of the proteins were connected to all of the top 3 cytokines (**Fig 4-8D** – black circles), suggesting they may be the most interesting proteins for further study. These include a number of proteins that contribute to the resolution of inflammation (eg. ADIPOQ, TIMP1, MRC1, CD163) and may represent actionable targets for new therapeutic interventions. Together, this analysis predicted the top cytokines that influence to the observed proteome landscape and enables researchers to start defining disease-associated pathways to biologically test in subsequent studies.

T4 and Adiponectin Signaling Influence SaB Outcomes in vivo. While the analyses above detail the biomarkers and pathways altered during SaB, it is unclear whether they are simply bystander markers or functionally contribute to disease outcomes. To address this gap, we utilized a mouse model of SaB to assess the influence of key features identified in the above data and determined their effects on bacterial burden and overall survival. We focused on two areas derived from the multi-omic analysis: thyroid hormones and adiponectin signaling.

Our multi-omic analysis captured a striking dysregulation of general host metabolism (**Fig 4-7**), including reduced levels of T4 in mortality patients. Exogenously administered T4 has been previously shown to be protective in mouse and rat models of polymicrobial sepsis(248); however, its contributions to SaB mortality remain unclear. Further, the impact of a hypothyroid state on SaB infections has not been tested. To address these questions, we designed an animal experiment to test whether pharmacologically altering T4 levels could affect survival in a SaB mouse model of infection (**Fig 4-9A**). We treated mice with either 1) a hypothyroid treatment (MMI and NaClO4) in their drinking water or 2) supplemental T4 through intraperitoneal (I.P.) injections to induce hyperthyroidism. Following these previously described treatment regimens(248, 249), we intravenously infected the mice with *S. aureus* and assessed survival. We found that hypothyroid mice had higher mortality rate than the control mice while the hyperthyroid group demonstrated four-time greater survival at 48 hours post-infection (p.i.) than the control group. It is worth noting that while the hypothyroid mice died more rapidly than control, control animals also succumbed to infection, leaving a marginal window to observe a significant difference between the groups. To interrogate

this finding further, we repeated the hypothyroid infection with a lower dose of S. aureus (50% original inoculum) and harvested organs for bacterial enumeration. Again, the hypothyroid mice had increased mortality than the control group (**Fig 4-9C**), and, consistently, the mice surviving at 48 hours p.i. had increased bacterial load in their hearts (**Fig 4-9D**) and kidneys (**Fig 4-9E**), indicating a defect in bacterial clearance.

Increased IL10 has emerged as a high confidence biomarker for the prediction of SaB patient mortality, which has been validated based on several studies from multiple groups, including our own. Through the cytokine-inference approach described above (**Fig 4-8**), we similarly predicted a significant contribution of IL10 to specific proteomic alterations. In particular, the IL10-linked, anti-inflammatory proteins may play a role in suppressing the overwhelming immune response observed in bacteremic patients. One of these anti-inflammatory proteins, adiponectin, had not previously been linked to SaB, but it is known to induce IL10 expression in human leukocytes(250). Given that IL10 is protective in a mouse SaB infection(251), we hypothesized that targeting adiponectin could also improve survival outcomes. To test this hypothesis, we treated mice with a commercially available small-molecule activator of the adiponectin receptor, AdipoRon, or vehicle control (**Fig 4-9A**), and utilized the same experimental scheme described above for the T4 experiments. Treatment with AdipoRon also markedly enhanced mouse survival (**Fig4-9F**) and induced significant decreases in organ CFUs (**Fig 4-9G-H**). Altogether, these *in vivo* studies demonstrate that stimulation of both the thyroid hormone system and the adiponectin receptor are protective in a mouse model of SaB.

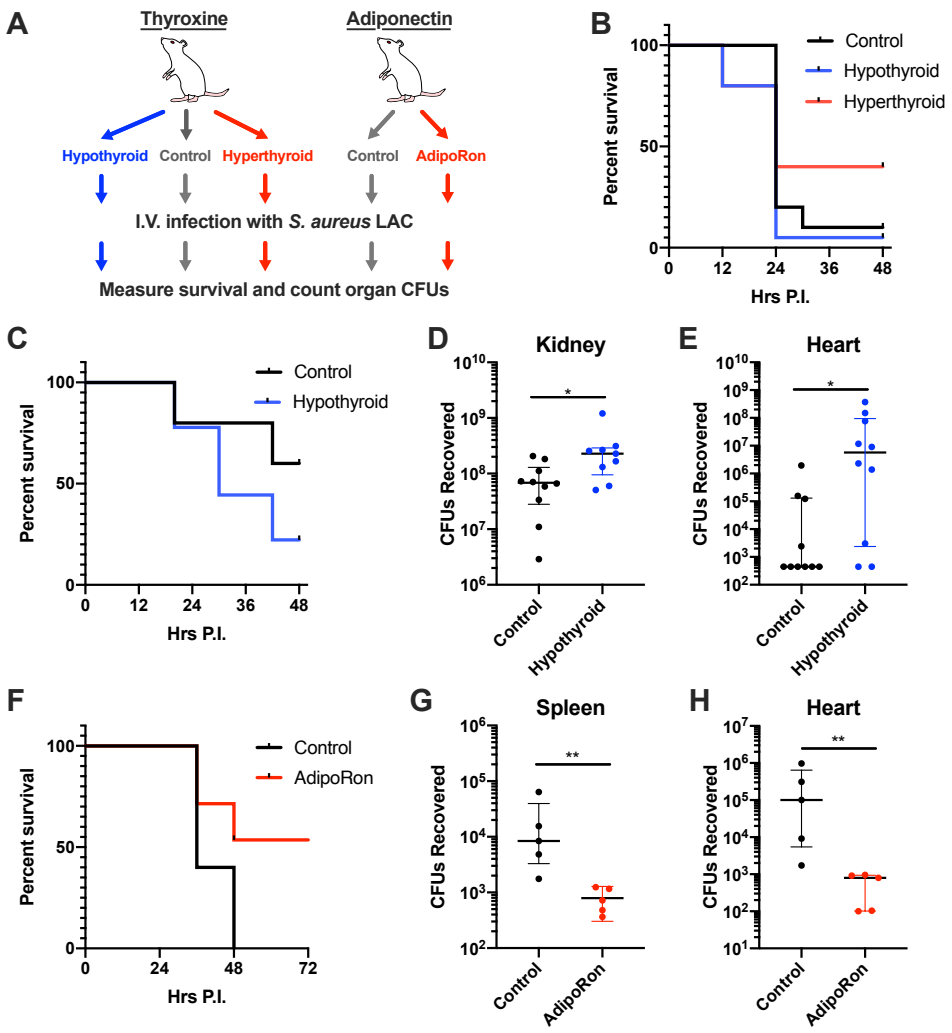


Figure 4-9 *Thyroid and Adiponectin Signaling Contribute to SaB Mortality in vivo*

(A) Schematic for treatment plan and mouse model of SaB. (B) Survival curve of mice given hyperthyroid, hypothyroid or control treatments then infected with 1x108 CFU S. aureus LAC. (C) Survival curve of mice given hypothyroid or control treatments then infected with 5x107 CFU S. aureus LAC. Bacterial CFUs recovered from the kidney (D) and heart (E) in hypothyroid or control mice 48 hours after infection with 5x107 CFU S. aureus LAC. (F) Survival curve of mice given AdipoRon or control treatments then infected with 5x107 CFU S. aureus LAC. Bacterial CFUs recovered from the spleen (D) and heart (E) in AdipoRon or control mice 48 hours after infection with 5x107 CFU S. aureus LAC.

Discussion

The traditional strategy for establishing prognostic biomarkers for infectious diseases has been based on analyses of a small subset of immunological response metrics such as cytokines. The current study sought to establish a new standard of infectious disease biomarker assessment by examining a much broader profile of the host response, eliminating the assumption that all clinically relevant responses to SaB are immunological. Beginning with a solid foundation of a large cohort with matched metadata in conjunction with a multi-omic approach, we defined numerous features whose predictive abilities equal or surpass that of previously published biomarkers(208-210, 214). In addition, the systems level of the analysis naturally led to the construction of multivariate models using combinations of biomarkers with predictive abilities that clearly outperform the current standard in the field. Importantly, this unbiased analysis uncovered features that would not be typically tested in the context of infection. Further, we associated all of the quantified features with all of the recorded metadata, thus laying the groundwork for subsequent follow-up studies in a variety of focus areas. While most biomarker studies may have stopped after simply defining predictive features, we expanded this study further through the application of additional computational analyses, an in-depth interrogation of disease-relevant host factor alterations, and validation of therapeutic relevance in an animal infection model.

Crude predictors of mortality in SaB can be based upon clinical, echocardiographic, and radiographic assessments of the patient(252, 253); however, they lack sufficient sensitivity and specificity to serve as reliable stratification methods upon which to individualize, de-escalate or stop therapy. Due to this ambiguity in

stratification, antimicrobial therapy selection and duration in SaB is largely based upon 'one size fits all' protocols whose foundations originate from the empiricism of clinical experience in the second half of the 20th century, when the antibiotic era was established(254). Clinical trials of SaB fail to discriminate the heterogenous patient types that make up a SaB clinical cohort and therefore have lacked sufficient granularity to enable clinically meaningful progress in an era of more potent antimicrobial therapy, particulary against MRSA(255). As a result, improvements in mortality in MRSA bacterema have not kept pace with progress in other fields of medicine over the past three decades, despite better drugs and faster diagnostics(255, 256).

Combination therapies offer an appealing approach to improve patient outcomes in SaB. For example, SaB patients with a higher mortality risk predicted by biomarker concentrations at the time of admission to the hospital may be candidates for daptomycin plus ceftaroline combination therapy. This combination therapy was recently to significantly reduce 30-day mortality compared to standard monotherapy (vancomycin or daptomycin) in a small prospective randomized trial, particularly in patients with elevated IL-10(257). The current challenge facing clinicians is that the cost of daptomycin plus ceftaroline is > 50 times that of vancomycin, posing considerable hospital pharmacy economic constraints. Attempts at developing less costly combinations have been wrought with drug toxicity(258-260). Therefore, utilizing the biomarkers uncovered herein to identify the 20 - 30% of patients with high SaB mortality risk on standard therapy would provide a compelling advance in the management of this disease.

In addition to defining standard protein and metabolite biomarkers for SaB mortality, this study utilized two advanced computational strategies to deepen the

analysis of proteomics data. First, a workflow was employed for the prediction and identification of PTMs from high-resolution proteomic datasets. Borrowing techniques used in the metabolomics field, molecular networking revealed that the paucity of identifications in serum datasets likely derives from the presence of highly modified peptides, including both serum glycoproteins and unexpected small PTMs. Using refined database searching techniques enabled us to match these modified features with high confidence. This approach resulted in the identification of our top predictive biomarkers, glycosylated peptides derived from fetuin A. Glycosylation has been used as biomarkers for various diseases, including cancer(261), Alzheimer's disease (262) and chronic inflammatory conditions(263). However, this is the first time that host glycosylation patterns have been linked to human SaB mortality. Specific interrogation of serum protein glycosylation patterns associated with mortality could provide a useful clinical tool in the future.

Intriguingly, our top unmodified biomarker was fetuin B and our top modified biomarker was glycosylation of fetuin A. Fetuins belong to the cystatin superfamily of proteins(264, 265) and can transport free fatty acids in the bloodstream(266). Both fetuin A and B are commonly studied in metabolic disorders such as obesity and diabetes, but fetuin A has also been shown to exert anti-inflammatory effects(267). Notably, fetuin A supplementation is protective in mouse models of systemic inflammation(266). Thus, defining the role of fetuin B in this process and whether glycosylation of either of these proteins impacts their activity could be of significant interest to the infectious and inflammatory disease communities. Regardless, both proteins can now be classified as high quality biomarkers of SaB patient mortality.

Another striking PTM finding uncovered in this dataset was a widespread increase in carbamylation of albumin and serum transferrin in mortality patients. Protein carbamylation is a non-enzymatic PTM(268) that is related to a number of pathological processes, including chronic kidney disease(269) and rheumatoid arthritis(270). In fact, multiple studies have proposed carbamylation of albumin as a prognostic factor for mortality in patients with kidney failure(269, 271). Of course, kidney disease and SaB are intimately related(232, 272), but the common signature of carbamylation suggests an underlying pathological process. Further, patients with rheumatoid arthritis, which can also be linked to *S. aureus* infections(273) and colonization(274), are reported to have higher levels of anti-carbamyl antibodies(275). Whether this modification is pathological or simply a marker of disease severity requires additional experimentation; nevertheless, it appears intimately linked with a variety of disease states, warranting further interrogation.

The second computational advancement in serum bioanalytics utilized in this study is the inferring of underlying cytokines signatures from serum proteomics data. A common drawback to unbiased proteomic approaches is when crucial proteins, known to be associated with a particular disease, go undetected in the MS data. This problem is exacerbated when analyzing serum samples due to the extreme dynamic range in protein concentrations. To mitigate this issue, we reasoned that relevant proteins could be predicted simply based on the observed proteomic alterations using a knowledge-based networking strategy. To demonstrate the application of this approach, we use our detected proteins to predict underlying cytokine signatures, which fall beneath the lower limit of detection in typical proteomic analyses(234, 236, 247). This analysis predicted

major alterations in TNF, IL6, IL10 and CXCL8 in mortality samples, all of which have exhibited strong associations to SaB human mortality in previous studies(208-210, 214), validating the method. This approach also enables researchers to link these major cytokine players to the observed proteomic alterations, facilitating the construction of new, readily testable hypotheses (such as the impact of adiponectin signaling in SaB). Additionally, we predict that CCL5 and CXCL12 may also play significant roles in the host response to SaB. Interestingly, receptors for both CCL5(276) and CXCL12(277) (CCR5 and CXCR4 respectively) are targets of *S. aureus* toxins that reduce host cell signal transduction and migration *in vitro*, further supporting their critical roles in host defense. Together, this method reinforced previous findings from the literature, further refined host response pathways and identified new potential players in disease.

In addition to defining high confidence biomarkers for the prediction of SaB mortality, we sought to gain a deeper understanding of the underlying biology leading to death. To do this, we parsed our multi-omics data into groups with similar expression profiles and performed a thorough interrogation of the mortality relevant clusters. Unexpectedly, the most striking findings from this were not related to the immune system, but rather a systemic dysfunction of host metabolism. While some of our findings have been previous described, such as the suppression of serum lipoproteins and T4 during severe infections, we also captured novel signatures of metabolic dysfunction, specifically linked to mortality. The most salient of these was the apparent shift from ternary to binary IGF-IGFBP complexes, resulting in lower circulating IGF levels, and the increase of acyl-carnitines and related molecular species in SaB mortality patients. The ultimate functional outcome of these perturbations is unclear; however, they may help uncover alternative

therapeutic avenues by which to stabilize patients while providing standard antimicrobial treatments.

Finally, we demonstrated that stimulation of both thyroid and adiponectin signaling pathways can enhance mouse survival in experimental SaB. Previous studies into thyroid signaling suggest that it's inhibition of macrophage migration inhibitory factor(248) or enhancement of intracellular bacterial killing(278) are responsible for its protective effects. In contrast, adiponectin is an adipokine (ie. produced primarily by adipocytes) that is mainly studied for its role in insulin resistance and diabetes(279). However, adiponectin also has anti-inflammatory, cardioprotective and vasoprotective effects(279) and adiponectin KO mice are more susceptible to polymicrobial sepsis(280). In line with these findings, our data indicate a protective role for adiponectin signaling in SaB infection models. Importantly, both T4(281) and AdipoRon(282) are orally bioavailable and T4 is already FDA-approved. If T4 or AdipoRon could offer similar protection in humans, they may be explored as adjunctive approaches to antibiotics for reducing SaB mortality.

Overall, our study aimed to set a new standard in the infectious disease biomarker field. We provide the best model for predicting SaB mortality reported to date and explore the use of computational approaches to enable a more complete analyses of our proteomic dataset. Conducting future studies to the same depth and rigor will likely uncover additional clinically useful findings and lead to a deeper understanding of mortality in infection. Ultimately, this study sets the groundwork for a multi-marker based tool for the rapid prediction of SaB patient mortality risk and patient stratification at the time of clinical presentation – a 'Rapid SaB Death Test'.

Chapter 4, in part, is a reprint of material submitted to *Cell*, 2020 (*manuscript in revision*), Jacob M. Wozniak, Joshua Olson, JR Caldera, Robert H. Mills, Marvic Carrillo-Terrazas, Chih-Ming Tsai, Fernando Vargas, Pieter C. Dorrestein, George Y. Liu, Victor Nizet, George Sakoulas, Warren Rose and David J. Gonzalez. The dissertation author was the primary author of this paper.

Chapter 5 – Future Directions

Summary

The above work represents foundational studies in a variety of research areas linked through the common theme of applying mass spectrometry-based approaches to better understand infectious diseases. The power in these techniques lies in their ability to identify and quantify thousands of molecular features in physiologically relevant compartments (ie. human serum, bacterial secretions, etc.) and the unbiased nature of data collection, which can include biologically relevant modified peptides and metabolites. Importantly, the biological significance of number of the identified features was validated through various *in vitro* and *in vivo* experiments. However, the systems level scale also enables to construction of numerous additional hypothesis that can be interrogated. Therefore, the Future Directions chapter of this thesis is presented in the form of a grant application to further investigate the findings above.

AIM 1: The human metabolome is a driver of SaB-induced death

1.1 Establish optimal timing and effective concentration of T4 SaB therapy

Rationale. Our preliminary data demonstrates that increased levels of T4 signaling is protective to the host during SaB infections, both in regard to mouse survival (Fig 21B-C) and organ CFU burden (Fig 21D-E). Other groups have reported T4 administration following infection to also be protective against other pathogens(248, 278). However, the pharmacokinetics of T4 were ill-defined in these studies. Therefore, to verify and further understand the contributions of T4 signaling in SaB, we aim to precisely define thyroid hormone pharmacokinetics during modulating treatments and SaB. This will allow us to properly understand the effects of maximal and minimal levels of T4 signaling in relation to SaB in subsequent experiments, which can subsequently be used to optimize drug administration regimens for favorable SaB outcomes.

Experiment 1.1.1 – Define thyroid hormone kinetics during thyroid modulating treatment. With the ultimate goal of establishing the effective concentration of T4, we will first need to define thyroid hormone kinetics during thyroid modulating treatment. Mice (n = 5) will be separated into hyperthyroid (T4+), hypothyroid (T4-) and control groups and treated as follows. T4+ mice will receive hyperthyroid treatment (I.P. injections of T4 once daily) for 1 week. T4- mice will receive hypothyroid treatment (MMI in drinking water) for 3 weeks. Control mice will receive standard drinking water and injections of 1X PBS at identical time points to the T4+ mice. T3 and T4 levels will be measured twice daily (12 hr apart) by taking facial bleeds and analyzing via ELISA assays. The minimum and maximum levels of T4 observed in this experiment will be hereafter referred to as $T4_{min}$ and $T4_{max}$, respectively.

Experiment 1.1.2 – Determine T4 effective concentration range. Once the optimal treatment strategies to generate $T4_{min}$ and $T4_{max}$ have been established, we will attempt to define the effective concentration of T4 in the blood. To do this, we will then select six concentrations representing an even distribution over the T4 range (ie. 0%, 20%, 40%, 60%, 80% and 100% of $T4_{max} – T4_{min}$). We will treat mice (n = 60; 10 mice each condition) with thyroid modulating therapy to achieve the desired T4 concentrations and infect mice with *S. aureus*. As in our preliminary studies, both survival and CFU burden experiments will be performed. For survival experiments, mice will be infected with a high dose of *S. aureus* ($1x10^8$ CFUs) and weight change and survival will be monitored every 12 hours for 5 days post infection (P.I.). For CFU burden experiments, mice (n = 30; 5 mice each condition) will be infected with a low dose of *S. aureus* ($1x10^7$ CFUs) and then euthanized 48 hrs P.I. for organ harvesting and CFU enumeration. The T4 levels that result in the best and worst outcomes (as determined by both survival and CFU burden) will be hereafter referred to as “optimal T4 SaB treatment”.

Experiment 1.1.3 – Assess the combinatorial effects of T4 plus standard SaB therapy. Once we have optimized T4 SaB treatment, we will explore the use of combinatorial therapy to more accurately replicate clinical settings. We will treat mice (n = 40; 5 mice each condition) with optimal T4 SaB treatment, infect mice as above for survival experiments and treat with common antibiotics used in the clinic. We will infect mice (n = 10 each) with both methicillin sensitive and methicillin resistant strains and treat with the appropriate antibiotic (flucloxacillin for MSSA and vancomycin for MRSA, n = 5 each). Untreated (n = 5 each strain) and monotherapy antibiotics (n = 5 each strain) will

be used as controls. Mice will be monitored and assessed for their response to SaB as in Experiment 1.1.2.

Anticipated results, potential pitfalls and alternative approaches. Based on previous studies, we expect the murine T4 levels range from 1 – 100 pg/L. Upon completion of this aim, we hope to have established a kinetic profile of thyroid hormone levels during thyroid modulating treatment. Utilizing these profiles, we also plan to identify the thyroid hormone levels that provide the best protection and worst outcomes. This information can be used in subsequent experiments investigating T4 protective effects and can be extrapolated and compared to human T4 levels to guide potential therapies in the clinic. It is possible that the above experiments do not deviate significantly from our preliminary models of T4 modulating treatment. If this is the case, it suggests that our current treatment therapy already represents the "optimal T4 SaB treatment".

1.2 Define the contribution of macrophages to T4 protective effects

Rationale. T4 signaling has been shown to be protective for multiple infections(248, 278). These studies linked the protection to macrophages, either through direct effects on cell function(278) or through blocking interaction with serum components(248). One study found minimal basal levels of THRA in macrophages, but that expression could be induced following exposure to pathogens(278). However, through the use of inhibitor studies, the authors attributed the effects of T4 to non-canonical signaling through ITGAV/B3 – PI3K – ERK1/2 pathway(283). Consistent with these findings, Human Protein Atlas (HPA)(284) data indicated basal expression of THRA and THRB is minimal in immune cells (**Fig 5-1A**), but, conflictingly, so are ITGAV and ITGB3 (**Fig 5-1A**). Another data resource (the DICE database(285)) confirmed low

immune cell expression of THRB (**Fig 5-1B**), ITGAV and ITGB3 (data not shown), but found appreciable levels of THRA (**Fig 5-1C**) in specific immune cells, such as T cells and NK cells. Interestingly, we found the expression of ITGAV/B3 (**Fig 5-1D**) to increase in THP1-derived macrophages following infection with *S. aureus*, similar to THRA in response to other pathogens(278). Therefore, it is clear that various immune cells express receptors for T4, either basally (T and NK cells) or induced (macrophages) which could contribute to the protective effects of T4 *in vivo*. With the wealth of data linking T4 protective signaling to macrophages, we will interrogate if macrophages play a similar role in the context of SaB infections *in vivo*. In parallel, we will characterize the impact of T4 on macrophage cell function and anti-*S. aureus* defense mechanisms *in vitro*.

Experiment 1.2.1 – Determine the role of macrophages in T4 SaB protective signaling. To investigate the contribution of macrophages to T4 SaB protective signaling, we will perform murine macrophage depletion/adoptive transfer experiments while simultaneously providing optimal T4 SaB therapy as described above. Mice (n = 30; 10 of each control, T4+ and T4- mice) will be depleted of macrophages using clodronate liposomes(286). Depletion of macrophages will be assessed by histological staining for the macrophage marker F4/80 in the liver, spleen and colon(286). Following optimal T4 SaB therapy and immune depletion, mice will be infected as above for survival and CFU burden experiments. In half of the depleted mice (n = 15; 5 of each control, T4+ and T4- mice), macrophages will be reconstituted through adoptive transfer experiments to control for the clodronate liposome-mediated depletion of other professional phagocytes(286). An additional control group of un-depleted mice (n = 15; 5 of each control, T4+ and T4- mice) will also be included.

Experiment 1.2.2 – Interrogate T4-treated macrophage activation. To interrogate effects of T4 on macrophage activation, three separate sources of cells will be utilized. For the first source, mice (n = 15; 5 of each control, T4+ and T4- mice) will be treated with optimal T4 SaB therapy and peripheral blood mononuclear cells (PBMCs) will be isolated from whole blood and differentiated into macrophages using described cytokines(287). For the second source, PBMCs will be isolated from untreated mice and differentiated into macrophages, but then treated *ex vivo* to mimic optimal T4 SaB therapy (n = 15; 5 of each control, T4+ and T4- mice) for 24 hours prior to experiments. The third source of macrophages will be differentiated from PBMCs isolated from human whole blood and treated *ex vivo* with T4 as above (n = 15; 5 of each control, T4+ and T4- mice). Macrophages will then be challenged with either IFN-γ (50 ng/ml) and LPS (10 ng/ml) to induced M1 polarization or with IL-4 (20 ng/ml) to induce M2a polarization(287). Transcript (Il12b and Ptgs2 for M1 and Klf4 and Irf4 for M2a), measured via quantitative polymerase chain reaction (qPCR), and cytokine markers (TNF and IL6 for M1 and IL10 and TGFb for M2a), measured via ELISA, will be used to assess activation(287).

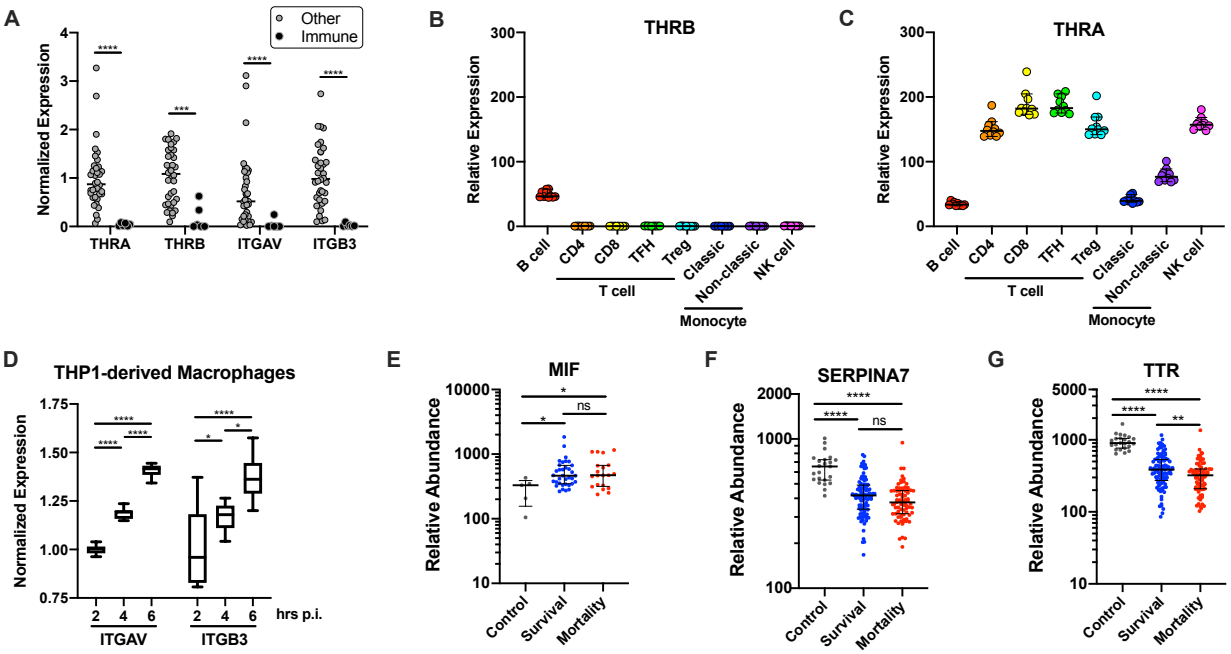


Figure 5-1 *Basal and Induced Expression of Thyroxine Interacting Proteins*

(**A**) Expression of known T4 receptors in various tissues (HPA). Expression of THRA (**B**) and THRB (**C**) in various immune subsets (DICE database). (**D**) Expression of ITGAV/B3 in THP1-derived macrophages upon exposure to *S. aureus*. Serum protein expression of (**E**) MIF, (**F**) SERPINA7, and (**G**) TTR in various human samples.

Experiment 1.2.3 – Interrogate T4-treated macrophage migration. To interrogate effects of T4 on macrophage migration, three cohorts of immune cells will be prepared for experiments as described in Experiment 1.2.2. However, instead of measuring activation markers, we will assess macrophage migration in transwell assays. Complement component 5a (C5a - 100 nM), to which both M1 and M2 macrophages are responsive, will be used as a chemoattractant, and migration quantified by lysing cells in both chambers and measuring relative lactate dehydrogenase (LDH) concentrations in the bottom chamber.

Experiment 1.2.4 – Interrogated the effects of T4 on macrophage S. aureus opsonization, phagocytosis and killing. To interrogate effects of T4 on macrophage interaction with *S. aureus*, again three cohorts of macrophages will be prepared for experiments as described above. However, these cells will now be infected with inoculums of *S. aureus* (MOI 10:1 bacteria:host cells) and phagocytosis and killing assays will be performed(288). Phagocytosis by macrophages will be determined by incubating *S. aureus* with macrophages for 0.5, 1 and 2 hrs then killing extracellular bacteria with antibiotics for 1 hr then immediately quantifying intracellular bacteria(289). To test for *S. aureus* killing, macrophage experiments will be carried out as above, but extending the time course to 4, 8, and 16 hrs P.I. In both experiments, host cell viability will be assessed using LDH released into the supernatant and *S. aureus* phagocytosis/viability will be assessed by serial dilution and CFU enumeration.

Anticipated results, potential pitfalls and alternative approaches. Upon completion of this aim, we anticipate having determined if macrophages are essential for the protective effects of T4 against SaB infections *in vivo*. We also hope to confirm and further

interrogate these cellular mechanisms of action using *in vitro* experiments. Once the contribution of macrophages has been verified, precise signaling mechanisms can be interrogated with conventional inhibitor and knockdown studies in cell lines and primary cells. For example, whether T4 SaB protective signaling is mediated by canonical receptors (ie. THRA and THRB) or non-canonical mechanisms (eg. ITGAV/B3 – PI3K – ERK1/2) can be determined. Once precise signaling mechanisms are pinpointed *in vitro*, we will validate their in vivo relevance using the SaB animal model described above. Since exposure to pathogens can induce T4 receptor expression, it is possible that additional cell types could mediate the protective effects of T4, such as other innate cells, adaptive immune cells or systemic components (eg. the serum or vasculature). If we are unable to assign the protective effects of T4 to macrophages in the experiments described above, we will begin to interrogate other immune responses using mice deficient neutrophils(290), NK cells(291) or adaptive immune cells (eg. Rag-I mice).

1.3. Define the serum contributions of the T4 protective mechanism

Rationale. In addition to the potential direct effects on macrophages as described above, T4 is known to interact with serum components such as thyroxine-binding globulin (SERPINA7), transthyretin (TTR) and macrophage inhibitory factor (MIF). Interestingly, we observed opposite trends in expression for these two proteins during SaB infection (**Fig 5-1 E-G**). T4-mediated antagonism of macrophage inhibitory factor (MIF) has been proposed to be beneficial for the host during polymicrobial sepsis(248). We hypothesize that, in addition to direct effects on immune cells, T4 imparts protective systemic effects during SaB, which can be mediated via serum components. The following experiments

are designed to interrogate if T4 protective signaling is also mediated through components in the serum, with a specific focus on MIF.

Experiment 1.3.1 – Interrogate the effects of T4 on S. aureus growth in serum. Serum has been shown to possess bacteriostatic effects on *S. aureus* growth in a dose dependent manner(292). To determine if T4 has any direct effects on serum that may contribute to the ability for *S. aureus* to survive *in vivo*, we will assess the ability of *S. aureus* to survive in serum under various conditions. First, mice will be treated with optimal T4 SaB therapy (n = 15; 5 of each control, T4+ and T4-) and serum will be collected for further experiments. Serum will be serially diluted to assess the dose-dependency of any observed effects. In tandem, mouse and human serum treated *ex vivo* to mimic optimal T4 SaB therapy (n = 15; 5 of each control, T4+ and T4-) will be interrogated for *S. aureus* bacteriostatic effects. *S. aureus* bacteriostatic effects will be assayed by incubating bacteria in serum at a concentration of 10^7 CFUs/ml for 2, 4, 6 and 8 hr. Following incubations, surviving *S. aureus* will be determined by serial dilution and CFU plating. An anti-MIF blocking antibody(293) (0.5 mg/mouse) and the MIF inhibitor, Z-590(294) (9 mg/kg), will be used to determine if any effects are related to MIF activity.

Experiment 1.3.2 – Interrogated the effects of T4 on S. aureus whole-blood survival. Whole blood contains both serum components and innate immune cells and therefore can be used to interrogate potential serum-mediated mechanisms. To determine in T4 has any effects on *S. aureus* whole-blood survival, whole blood will be prepared from three sources as the serum in Experiment 1.3.1 (ie. from mice receiving optimal T4 therapy and whole blood from untreated mice and humans spiked with T4 *ex vivo*). *S. aureus* survival will be assayed by incubating bacteria in the variously treated

whole blood at a concentration of 10^7 CFUs/ml for 2, 4 and 6 hours. The contributions of MIF will be interrogated as in Experiment 1.3.1.

Experiment 1.3.3 – Interrogated the effects of T4 on serum-enhanced, S. aureus opsonization, phagocytosis and killing. To verify and pinpoint the contributions of specific immune cells to T4 SaB protective effects, phagocytosis and killing assays will be performed as described in Experiment 1.2.3. However, this the current experiment, S. aureus will be opsonized in serum from various sources prior to co-culture with macrophages. Again, macrophages and serum will be prepared from three sources as described in Experiments 1.2.2 and 1.3.1, respectively. Macrophages and serum will be inter-mixed in all possible combinations (eg. control, T4+ and T4- macrophages each with control, T4+ and T4- serum). Phagocytosis and killing of *S. aureus*, and viability of host cells, will be assessed as in Experiment 1.2.3. The contributions of MIF will be interrogated as in Experiment 1.3.1.

Anticipated results, potential pitfalls and alternative approaches. Upon completion of this aim, we expect to have determined if T4 mediates protective effects through serum-mediated mechanisms. Guided by previous literature, we also anticipate confirming if any of the protective effects require MIF activity. However, it is possible that we will be unable to define significant contributions of T4 to serum-mediated mechanisms in the above experiments. T4 can also have general effects that promote cardiovascular health(295), which plays a large role in regulating sepsis-induced organ failure. Therefore, the cardiovascular system offers another avenue from which to explore T4 SaB protective signaling. If there is any residual protection of T4 that cannot be explained by immune cells (Subaim 1.2) or serum components (Subaim 1.3), we will turn our investigation to

the endovascular system.

Aim 2: The human proteome is a driver of SaB-induced death

2.1. Verify the protective contribution of adiponectin signaling in vivo

Rationale. Our preliminary data (**Fig 4-8C**) suggested anti-inflammatory proteins linked to multiple critical SaB cytokines may represent actionable targets for therapeutic interventions. We validated one of these targets, adiponectin, and demonstrated that increased adiponectin signaling is protective to the host during SaB infections, both in regard to mouse survival (**Fig 4-9F**) and organ CFU burden (**Fig 4-9 G-H**). However, our data is based on studies using a small molecule agonist (AdiopoRon(282)) of the adiponectin receptors (AdipoR1 and AdipoR2) and not adiponectin itself. To verify and further understand the contributions of adiponectin signaling, we will test if adiponectin itself can also impart protective effects during SaB infections. Additionally, while the adiponectin receptors are highly similar in primary sequence as well as crystal structures(296), they exhibited markedly different expression patterns across various tissues, with immune cells exhibiting high expression of ADIPOR1 and nearly undetectable expression of ADIPOR2(284) (**Fig 5-2 A-B**). Therefore, we will interrogate the receptor specificity of this signaling using knockout mice, which may also provide hints toward the cell types mediating these protective effects.

Experiment 2.1.1 – Confirm the protective contribution of adiponectin in SaB infections. While the small molecule AdipoRon is orally bioavailable, adiponectin protein needs to be administered intravenously (I.V.). Therefore, to assess the protective contributions of adiponectin in SaB infections, we will administer adiponectin [concentration] via I.V. 24 hours prior to infection with *S. aureus.* As in Experiment 1.1.2,

both survival and CFU burden experiments will be performed. For survival experiments, mice (n = 20; 10 control and adiponectin) will be infected with a high dose of *S. aureus* ($1x10^8$ CFUs) and weight change and survival will be monitored every 12 hours for 5 days post infection (P.I.). For CFU burden experiments, mice (n = 10; 5 control and 5 adiponectin) will be infected with a low dose of S. aureus ($1x10^7$ CFUs) and then euthanized 48 hrs P.I. for organ harvesting and CFU enumeration. Additionally, to determine the impact of adiponectin on central SaB perturbed cytokines (IL6, TGFB1, TNF, IL1b and IL10), we will probe for these cytokines in the CFU burden experiment mouse serum using ELISA assays.

Experiment 2.1.2 – Determine receptor specificity of protective adiponectin signaling. Adiponectin can signal through either AdipoR1 or AdipoR2. To determine which receptor is required for protective adiponectin signaling, we will utilize knockout mice of the AdipoR1 and AdipoR2 receptors. We will treat WT and knockout mice with adiponectin (I.V.) or AdipoRon (oral) 24 hours prior to infection with *S. aureus*. As above, both survival (n = 30; 10 control, adiponectin and AdipoRon) and CFU (n = 15; 5 control, adiponectin and AdipoRon) burden experiments will be performed (along with central cytokine interrogation).

Experiment 2.1.3 – Assess the combinatorial effects of adiponectin signaling plus standard SaB therapy. Once we have confirmed adiponectin protective monotherapy in SaB, we will explore the use of combinatorial therapy to more accurately replicate clinical settings. We will treat mice (n = 40; 5 mice each condition) with AdipoRon or adiponectin, infect mice as above for survival experiments and treat with common antibiotics used in the clinic. We will infect mice (n = 10 each) with both methicillin sensitive and methicillin

resistant strains and treat with the appropriate antibiotic (flucloxacillin for MSSA and vancomycin for MRSA, n = 5 each). Untreated (n = 5 each strain) and monotherapy antibiotics (n = 5 each strain) will be used as controls. Mice will be monitored and assessed for their response to SaB as in Experiment 2.1.1.

Anticipated results, potential pitfalls and alternative approaches. Upon completion of this aim, we anticipate having verified the SaB protective effects of the adiponectin protein itself as well as pinpointed the receptor specificity of these effects. If adiponectin and AdipoRon appear to have the same therapeutic value for murine SaB, we will continue to use AdipoRon due to its ease of administration and cost benefit. However, there is the possibility that adiponectin does not provide the same protective effects of AdipoRon. If this is the case, we will explore the use of higher doses of adiponectin to determine any potential therapeutic relevance. If we are still unable to assign any protection to adiponectin administration, the mechanisms of AdipoRon will be interrogated in subsequent assays and compared to adiponectin as a negative control. Since AdipoR1 and AdipoR2 exhibit distinct cell type enrichments, determining the receptor specificity will likely hint at the important cell populations (eg. immune vs other). However, it may be the case that signaling through both receptors is required for protection, or, alternatively, protection may be mediated through non-canonical mechanisms. Since we cannot rule out these possibilities at this time, Subaims 2.2 and 2.3 of the current proposal seek to define the immune and cardiovascular components of adiponectin SaB protective signaling, respectively. The completion of these aims will enable specific, protective, adiponectin (or AdipoRon if different) mechanisms to be further interrogated.

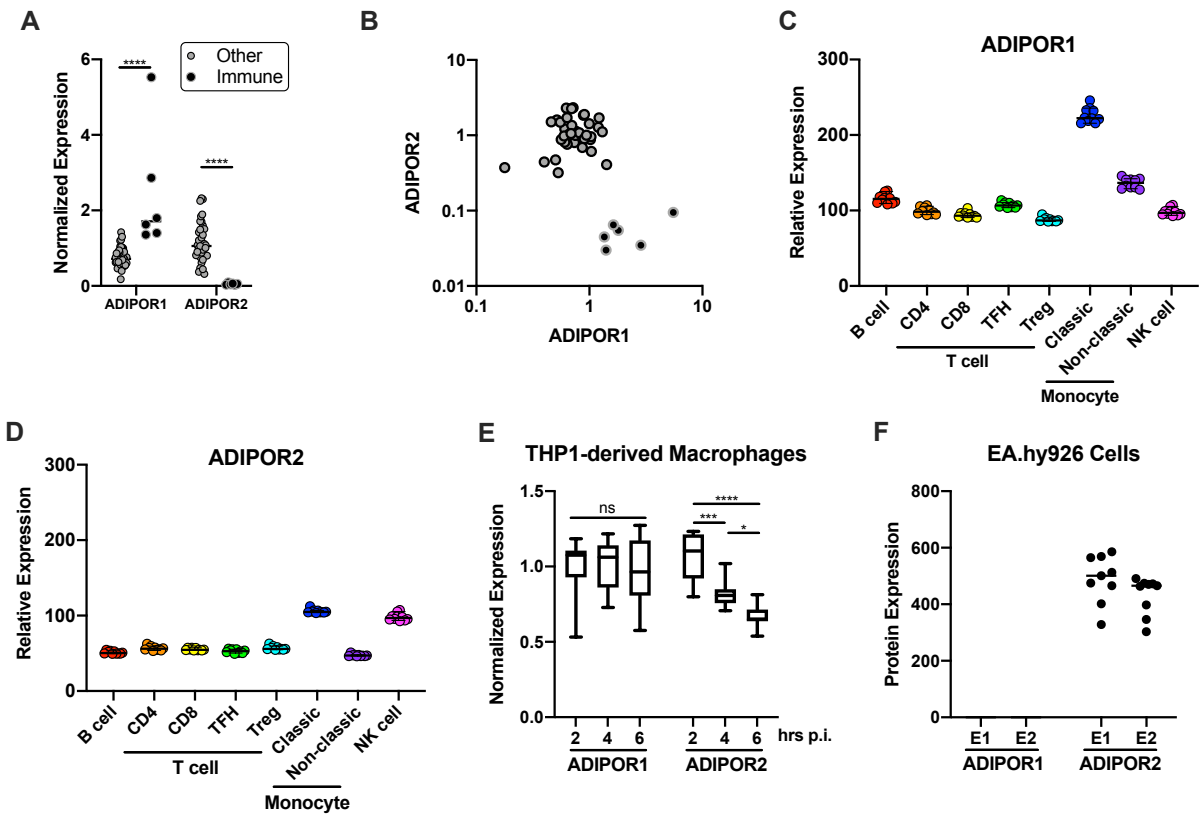


Figure 5-2 *Basal and Induced Adiponectin Receptor Expression*

(**A**) ADIPOR1 and ADIPOR2 expression in various tissues (HPA). (**B**) Correlation of ADIPOR1 and ADIPOR2 expression in various tissues (HPA). Expression of ADIPOR1 (**C**) and ADIPOR2 (**D**) in various immune subsets (DICE database). (**E**) Expression of ADIPOR1/2 in THP1-derived macrophages upon exposure to *S. aureus*. (**F**) Adiponectin receptor expression on EA.hy926 endothelial cell line from two different MS experiments (E1 and E2).

2.2 Define the immune-mediated contributions of AdipoRon protective signaling

Rationale. As mentioned above, AdipoR1 is highly expressed on innate immune cells and, as such, is poised to mediate adiponectin mediated protection against infection. Adiponectin has been demonstrated to affect a variety of innate immune cells(297), including macrophages, neutrophils, dendritic cells and NK cells, which are critical in the host response to *S. aureus(298, 299)*. The DICE database(285) confirmed the high expression of ADIPOR1 in immune cells, specifically in macrophages (**Fig 5-2C**), and indicated macrophages and NK cells expressed appreciable levels of ADIPOR2 (**Fig 5-2D**). Further, we detected a time-dependent decrease in ADIPOR2 expression in macrophages upon exposure to *S. aureus*, while ADIPOR1 expression remained stable (**Fig 5-2E**). However, there is currently a gap in knowledge as to what, if any, innate immune cells mediate adiponectin protection in SaB. We hypothesize that adiponectin could signal through any of the above cell types, so we will interrogate if any play a role in the context of SaB infections *in vivo* using depletion and adoptive transfer experiments and *in vitro* using purified cell populations.

*Experiment 2.2.1 – Determine the role of various innate immune cell subsets in AdipoRon SaB protective signaling***.** Similar to Experiment 1.2.1, we will deplete mice of immune cells while simultaneously treating mice with protective SaB therapy to investigate the contribution of various innate immune cell subsets to adiponectin SaB protective signaling. We will focus our initial analyses on the major types of innate immune cells as they have been previously shown to respond to adiponectin and are likely most important for single-challenge mouse models of SaB infection. The following innate immune cell subsets will be depleted in this assay as previously described: macrophages

and dendritic cells (clodronate liposomes(286)), neutrophils (anti-Lys6G antibody(290)), NK cells (anti-NK1.1 antibody(291)). Depletion/adoptive transfer will be confirmed for each cell type as follows (macrophage/dendritic cells - F4/80+ cells in spleen(286); neutrophils - whole blood absolute neutrophil counts; NK cells - NKp46+ cells in spleen(300)). Since both macrophages and dendritic cells are depleted via clondronate liposomes, the specific contributions of each cell type will be verified by adoptive- transfer experiments. As in Experiment 2.1.1, both survival and CFU burden experiments will be performed (n = 40 each for survival and CFU burden; 5 untreated control and 5 AdipoRon for each depletion (3 total) and an un-depleted control). To verify cell type contributions, each depletion will be repeated, and adoptive transfer experiments will be performed with the respective depleted cells (n = 40 for each survival and CFU burden; 5 for each cell type (4 total) for both untreated control and AdipoRon). For any cell type showing significant differences in the first round of experiments, this experiment will be repeated three independent times (n = 5 each condition).

Experiment 2.2.2 – Interrogate AdipoRon-treated innate immune cell activation. To further interrogate adiponectin cellular mechanisms, we will assess the effects that adiponectin has on innate immune cell activation *in vitro*. Similar to Experiment 1.2.2, we will purify innate immune cells from three distinct sources (ie. from mice receiving optimal adiponectin therapy and from untreated mice and humans spiked with adiponectin *ex vivo*), challenge them *in vitro* with their respective activating ligands and measure specific cell markers to assess activation. Primary neutrophils will be isolated from whole blood and dendritic cells(301) and NK cells(302) isolated from PBMCs. Macrophages will be differentiated from PBMCs as previously described (Experiment 1.2.2). Cell activating

ligands and markers are previously described (macrophages(286), neutrophils(303), dendritic cells(304), NK cells(305)). Individual experiments will be conducted on cells isolated from n = 5 samples.

Experiment 2.2.3 – Interrogate AdipoRon-treated innate immune cell migration. To further interrogate adiponectin cellular mechanisms, we will assess the effects that adiponectin has on innate immune cell migration *in vitro* using transwell assays. Cells will be isolated/differentiated as in Experiment 2.2.2. Following differentiation, cells will be assessed for migratory potential toward ligands for each cell type (macrophages: C5a(306), neutrophils: fMLP(307), dendritic cells: CCL19(308), and NK cells: CXCL9(309)). Migration quantified by lysing cells in both chambers and measuring relative LDH concentrations in the bottom chamber. Individual experiments will be conducted on cells isolated from n = 5 samples.

Experiment 2.2.4 – Interrogate AdipoRon-treated innate immune cell interaction with S. aureus. To further interrogate adiponectin cellular mechanisms, we will assess the effects that adiponectin has on innate immune cell interactions with *S. aureus* using *in vitro* killing assays. Cells will be isolated/differentiated as in Experiment 2.2.2. Both macrophages(289) and dendritic cells(310) can kill *S. aureus* through phagocytosis and intracellular anti-microbial activity, which will be assessed with phagocytosis and killing assays as described in Experiment 1.2.4. Neutrophils will be assessed for their ability to kill S. aureus through both intracellular and extracellular mechanisms in a single assay(311). Finally, while NK cells do not directly kill *S. aureus,* they can enhance anti-staphylococcal activity of macrophages in co-culture experiments(312), which will be performed as previously described. Again, an MOI of 10:1 bacteria:host cells will be used.

Individual experiments will be conducted on cells isolated from n = 5 samples. Host cell viability will be assessed using LDH released into the supernatant and *S. aureus* viability will be assessed by serial dilution and CFU enumeration.

Anticipated results, potential pitfalls and alternative approaches. Upon completion of this aim, we hope to construct an innate immunological profile of the protective effects of adiponectin in response to SaB. We aim to pinpoint specific innate immune subsets that are required for adiponectin's protective effects and delineate the cellular mechanisms that mediate this protection. By unbiasedly assessing four major innate immune subsets, we will be able to determine if multiple cellular mechanisms exist. Once precise signaling mechanisms are pinpointed *in vitro*, we will validate their *in vivo* relevance using the SaB animal model described above. However, it is possible that adiponectin mediates SaB protection by other routes, such as adaptive immune cells or systemic components (eg. serum, vasculature, etc.). If we are unable to assign AdipoRon protection to any innate immune cells, we will explore the possibility of adaptive immune system contributions (eg. using RagI mice). Further, we plan to interrogate any protective effects mediated through the vasculature in Subaim 2.3 and present the interrogation of serum components as an alternative approach in that Subaim.

2.3. Define the systemic contributions of AdipoRon protective effects

Rationale. S. aureus is a leading cause of infective endocarditis(207, 313) and SaB causes severe damage to the vasculature(207, 210), both of which can contribute to multi-organ dysfunction, failure and, ultimately, death(314). While it is likely that adiponectin can improve the host response to SaB by signaling through AdipoR1 on innate immune cells, adiponectin is also noted for its cardio and vaso-protective

activity(279, 315), and we have detected AdipoR2 protein expression in endothelial cells in multiple experiments (**Fig 5-2F**). In fact, the decreased adiponectin observed in obese individuals is linked to cardiovascular risk and exogenous adiponectin has therapeutic value in rat models of macrovascular disease(315). The mechanisms of adiponectin-mediated protective effects vary depending on the target cell type and tissue, but evidence of action on endothelial cells and cardiomyocytes is strong(279, 280, 315). This includes preliminary data from our demonstrating that AdipoRon treatment reduced bacterial CFU burden in hearts of infected mice (**Fig 4-9H**). Therefore, adiponectin sits in a prime position to act on two primary sites of SaB-induced host damage, the heart and vasculature, which may underlie its therapeutic value. To interrogate these cellular mechanisms, we will monitor cardiovascular health during SaB infections in mice treated with AdipoRon and investigate cell-specific signaling events using phospho-proteomic approaches.

Experiment 2.3.1 – Monitor hemostasis during AdipoRon SaB treatment. We will assess the effects that adiponectin signaling has on hemostasis by monitoring blood pressure, heart rate and coagulation status during a SaB infection. Mice (n = 20; 10 control and 10 AdipoRon) will be infected with a low dose of *S. aureus* for 48 hr as in CFU burden experiments (Experiment 2.1.1). During the infection, blood pressure, heart rate and temperature will be measured every 6 hr using tail cuff systems(316). Following the infection time course, whole blood will be harvested and the coagulation of the mice will be assessed using thrombelastography(317, 318).

Experiment 2.3.2 – Interrogate cardiac health during AdipoRon SaB treatment. We will assess the effects that adiponectin has on cardiac health during a SaB infection using

classical histological techniques. Hearts will be harvested from the same animals used in Experiment 2.3.1 and snap frozen in liquid nitrogen. Hearts will then be section and stained for both fibrosis (Sirius Red/Fast Green Collagen Stain) and immune cell infiltration (hematoxylin and eosin stain followed by counting immune cell nuclei).

Experiment 2.3.3 – Measure endovascular permeability during AdipoRon SaB treatment. We will assess the effects that adiponectin has on endovascular permeability during an SaB infection using Evans blue dye(319). Mice (n = 20; 10 control and 10 AdipoRon) will be infected with a low dose of *S. aureus* as above for 48 hr. Thirty minutes prior to euthanization, mice will be injected with Evans blue dye (1% in 0.9% saline – 2 mg/kg tail vein). Following euthanization, mice will be perfused with saline, and organs (heart, liver, lung, kidney, and spleen) will be harvested, weighed and prepared for permeability quantification as previously described. Final comparisons will be between ug Evan's blue dye/ug total organ protein for each isolation.

Experiment 2.3.4 – Profile the cell-specific phospho-proteomic signaling during AdipoRon SaB treatment. To support the above results and identify proteins involved in Adiponectin-promoted endothelial barrier stability and heart function, we will comparative profile both endothelial cell and cardiomyocyte protein expression and phosphorylation dynamics during *S. aureus* model infections. To provide enough starting material for the phospho-proteomic workflow, we will use human umbilical vein endothelial cells (HUVECs(320)) as a source of endothelial cells and human pluripotent stem cell-derived cardiomyocytes (hPSC-CMs(321)) as a source of cardiomyocytes. These cells will be cultured *in vitro*, treated with AdipoRon or a vehicle control and infected with *S. aureus* at an MOI of 1:1 bacteria:human cells. Infections will be allowed to persist for 2, 4, 8 and 16

hr P.I. and, at each time point, cells will be washed, harvested and processed for phospho-proteomic analysis as previously described(322). Differentially expressed proteins will be identified via ANOVA ($p < 0.05$), clustered using the k means algorithm and interrogated for functional group using gene ontology(122) and STRING-db(123).

Anticipated results, potential pitfalls and alternative approaches. Upon completion of this aim, we hope to determine if adiponectin signaling protects the heart and/or vasculature during SaB infections. By merging classical interrogation of cardiovascular health with state-of-the-art phospho-proteomic analyses, we anticipate building a multi-scale model of AdipoRon SaB treatment, from overt clinical features (eg. blood pressure) to molecular signaling pathways (ie. the phospho-proteomic analysis). Once this information is obtained, knockout and inhibitor studies can be planned to directly interrogate the prioritized pathways. While it is likely that the effects of adiponectin signaling on the vascular have some effects in SaB, it is possible that they may not significantly influence infection outcomes. If this is the case, we will profile the serum component, as in subaim 1.3.3, for any contributions to AdiopoRon-mediated protective effects.

Aim 3: The Rapid Index of SaB Mortality Kinetics (RISK) test

3.1. Construct and validate the RISK-24 test

Rationale. In our preliminary experiments, we used MS-based approaches to identify high-confidence biomarkers for the prediction of SaB patient mortality. These markers can be used in combination to enhance predictive abilities (**Fig 5-3A**). While some hospital labs may be able to rapidly conduct MS-based workflows, immune-based assays, such as enzyme-linked immunosorbent assays (ELISAs), are routinely used in a

clinical setting and offer an attractive approach to reduce operator training and analysis times to clinically tractable quantitative values. However, while patient serum is routinely used as a source of biomarkers, preparation of serum from whole blood can be variable depending on the hospital, lab or individual technician involved. The use of whole blood as a source of biomarkers can remove the variation induced by serum preparation. The goal of this subaim is to verify top SaB biomarkers, from our dataset, as well as literature reported cytokines, using ELISAs, and construct a multiplexed assay that could be conducted in-clinic within a 24 hr period.

Experiment 3.1.1 – Verify top biomarkers via immuno-assays. We hypothesize that the top biomarkers captured in our proteomics data are amenable to rapid analysis via immuno-assays. To test this hypothesis, we have selected 10 biomarkers (5 cytokines from the literature and 5 proteins from our preliminary data) for verification via commercially available ELISA kits. We will interrogate these markers in a cohort of 100 SaB patients (50 survival and 50 mortality) as well as 30 control samples (15 non-hospitalized, non-infected, 15 hospitalized, non-infected). Biomarker concentrations will be assessed for their predictive abilities using Mann-Whitney U, ROC curves and logistic regression analyses and profiled against all available metadata to control for confounding variables (eg. age, gender, etc.).

Experiment 3.1.2 – Define biomarker linearity and LOD in serum and whole blood. We hypothesize that some of the biomarkers verified in serum will also be measurable in whole blood. To test this hypothesis, we will separate 20 human blood samples (obtained from de-identified individuals at the SD Blood Bank) into two aliquots. One aliquot will be retained for biomarker analysis in whole blood while serum will be prepared from the

second aliquot. Absolute concentrations and coefficients of variation will be determined for each marker in both whole blood and serum. Following the determination of average human levels of each biomarker, the samples will be diluted and re-analyzed to determine the linearity and LOD in both whole blood and serum.

Experiment 3.1.3 – Construct and validate a multiplexed, immuno-assay-based RISK-24 test. Once biomarkers have been confirmed, we will exhaustively compare all combinations of up to 5 biomarkers using multivariate logistic regression. The optimal combination, as determined by ROC AUC, will be used to construct a multiplexed, immuno-assay-based test. This test, termed the RISK-24 test, will have an approximate completion time of 3-4 hrs. Together, with the additional time requirements of clinical laboratories (eg. sample preparation, transport and batching), this test could reasonably occur, and clinicians could make an informed intervention, within a 24 hr period (or earlier depending on sample priorities). We will first verify this test on our optimization cohort used above. Then, we will ship these tests to independent infectious disease clinicians (established by our co-I Warren Rose; see LOC) for use by diverse personnel groups in a multi-center prospective study. Ultimate functionality of the test will be determined by Mann-Whitney U tests, ROC curves, logistic regression and % of tests completed within 24 hrs on a per center and total study basis.

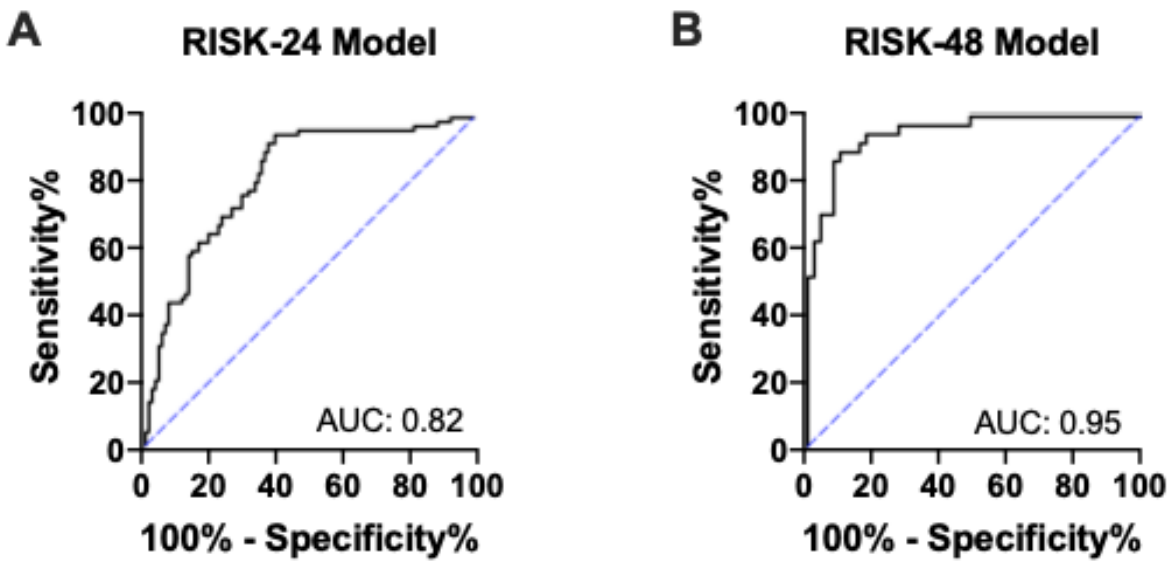


Figure 5-3 *Predictive Model Performance of RISK-24 and RISK-48 Tests*

(**A**) RISK24 predictive model based on the expression of seven proteins. (**B**) RISK48 predictive model based on the expression of four modified proteins and two metabolites.

Anticipated results, potential pitfalls and alternative approaches. Upon completion of this aim, we anticipate the construction of a high confidence, multiplexed, immuno-assay that can predict SaB patient mortality within a 24 hr period. This time prediction is based on extensive literature of clinical laboratory testing. Potential pitfalls include being unable to validate our biomarkers using immuno-assays or the multiplexed tests. If any biomarker doesn't validate at any stage, the biomarker will be discarded, and the next top biomarker will be included.

3.2. Construct and validate an MS-based, RISK-48 test:

Rationale. While our efforts in Subaim 3.1 seek to provide the field with a clinically applicable, immuno-assay-based test, our preliminary data demonstrated that the predictive strength of multi-marker assays can be increased by including PTMs or yet-to-be-identified metabolomic features (**Fig 5-3B**). Measuring these types of markers is more amenable via MS, which can be targeted to quantify specific peptides (with and without modifications) or simply m/z ratios (even if the identity of the feature is unknown). Provided the infrastructure for targeted proteomics and metabolomics analyses is present, clinical labs should be able to quantify any reasonably abundant serum ion within this time frame. Therefore, this aim will establish MS-based tests for unconventional biomarkers not accessible via immuno-assays that could be conducted in-clinic within 48 hr.

Experiment 3.2.1 – Verify top biomarkers via targeted MS methods. Based on our preliminary results, specific protein PTMs and unidentified metabolites can be combined with total protein abundances to enhance SaB mortality predictions. To verify these markers, we will establish targeted MS-based assays (ie. parallel reaction monitoring -

PRM) for their quantification. For PTM quantification (5 markers), synthetic, heavy peptides (both the modified and unmodified forms) will be purchased and used as standards. For unidentified metabolite quantification (20 markers), m/z values will be targeted and compared to standard MS markers. We will interrogate these markers in the same cohort as described in experiment 3.1.1. Biomarker concentrations will be assessed for their predictive abilities as described in experiment 3.1.3.

Experiment 3.2.2 – Define biomarker linearity and LOD in serum and whole blood. We hypothesize that some of the biomarkers verified in serum will also be measurable in whole blood. To test this hypothesis, we will separate 20 human blood samples (obtained from de-identified individuals at the San Diego Blood Bank) into two aliquots. One aliquot will be retained for biomarker analysis in whole blood while serum will be prepared from the second aliquot. Absolute concentrations and coefficients of variation will be determined for each marker in both whole blood and serum. Following the determination of average human levels of each biomarker, the samples will be diluted and re-analyzed to determine the linearity and LOD in both whole blood and serum.

Experiment 3.2.3 – Construct and validate a multiplexed, MS-based RISK-48 test. Once biomarkers have been rigorously confirmed in our targeted cohort, we will exhaustively compare all combinations of up to 10 biomarkers using multivariate logistic regression. The optimal combination, as determined by ROC AUC, will be used to construct a multiplexed, MS-based test. This test, termed the RISK-48 test, will have an approximate completion time of 36 hrs. Together, with the additional time requirements of clinical laboratories (eg. sample preparation, transport and batching), this test could reasonably occur, and clinicians could make an informed intervention, within a 48 hr

period (or earlier depending on sample priorities). We will first verify this test on our optimization cohort used above. Then, we will ship these tests to infectious disease clinicians and assess them as described in experiment 3.1.3.

Anticipated results, potential pitfalls and alternative approaches. Upon completion of this aim, we anticipate the construction of a high confidence, multiplexed, MS-based test that can predict SaB patient mortality within 48 hrs. Potential pitfalls include being unable to validate our biomarkers using targeted MS approaches or in the multi-target approach. If any biomarker doesn't validate at any stage, the biomarker will be discarded, and the next top biomarker will be included. Another hurdle is that the in-clinic turnaround time may be longer than 48 hrs. If this is the case, we will explore the use of rapid proteomic methods (eg. RapidFire MS) for sample preparation, gradient reductions and additional computational approaches to improve the turnaround time of the assay.

References

1. Dunn WB. Mass spectrometry in systems biology an introduction. Methods Enzymol. 2011;500:15-35.

2. Finehout EJ, Lee KH. An introduction to mass spectrometry applications in biological research. Biochem Mol Biol Educ. 2004;32(2):93-100.

3. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J Proteome Res. 2003;2(1):43-50.

4. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003;422(6928):198-207.

5. Bensimon A, Heck AJ, Aebersold R. Mass spectrometry-based proteomics and network biology. Annu Rev Biochem. 2012;81:379-405.

6. Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, Neumann T, Johnstone R, Mohammed AK, Hamon C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem. 2003;75(8):1895-904.

7. Ting L, Rad R, Gygi SP, Haas W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. Nat Methods. 2011;8(11):937-40.

8. McAlister GC, Huttlin EL, Haas W, Ting L, Jedrychowski MP, Rogers JC, Kuhn K, Pike I, Grothe RA, Blethrow JD, Gygi SP. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. Anal Chem. 2012;84(17):7469-78.

9. McAlister GC, Nusinow DP, Jedrychowski MP, Wuhr M, Huttlin EL, Erickson BK, Rad R, Haas W, Gygi SP. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. Anal Chem. 2014;86(14):7150-8.

10. Devabhaktuni A, Lin S, Zhang L, Swaminathan K, Gonzalez CG, Olsson N, Pearlman SM, Rawson K, Elias JE. TagGraph reveals vast protein modification landscapes from large tandem mass spectrometry datasets. Nat Biotechnol. 2019;37(4):469-79.

11. Chick JM, Kolippakkam D, Nusinow DP, Zhai B, Rad R, Huttlin EL, Gygi SP. A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. Nat Biotechnol. 2015;33(7):743-9.

12. Hunter T. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. Cell. 1995;80(2):225-36.

13. Cohen P. The origins of protein phosphorylation. Nat Cell Biol. 2002;4(5):E127-30.

14. Kouzarides T. Acetylation: a regulatory modification to rival phosphorylation? EMBO J. 2000;19(6):1176-9.

15. Verdin E, Ott M. 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. Nat Rev Mol Cell Biol. 2015;16(4):258-64.

16. Hicke L. Protein regulation by monoubiquitin. Nat Rev Mol Cell Biol. 2001;2(3):195-201.

17. Hicke L, Dunn R. Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. Annu Rev Cell Dev Biol. 2003;19:141-72.

18. Dennis JW, Granovsky M, Warren CE. Protein glycosylation in development and disease. Bioessays. 1999;21(5):412-21.

19. Lis H, Sharon N. Protein glycosylation. Structural and functional aspects. Eur J Biochem. 1993;218(1):1-27.

20. Shiloh Y. ATM and related protein kinases: safeguarding genome integrity. Nat Rev Cancer. 2003;3(3):155-68.

21. Sebolt-Leopold JS, Herrera R. Targeting the mitogen-activated protein kinase cascade to treat cancer. Nat Rev Cancer. 2004;4(12):937-47.

22. Cohen P. The role of protein phosphorylation in human health and disease. The Sir Hans Krebs Medal Lecture. Eur J Biochem. 2001;268(19):5001-10.

23. Lapek JD, Jr., Lewinski MK, Wozniak JM, Guatelli J, Gonzalez DJ. Quantitative Temporal Viromics of an Inducible HIV-1 Model Yields Insight to Global Host Targets and Phospho-Dynamics Associated with Vpr. Mol Cell Proteomics. 2017.

24. Schmutz C, Ahrne E, Kasper CA, Tschon T, Sorg I, Dreier RF, Schmidt A, Arrieumerlou C. Systems-level overview of host protein phosphorylation during Shigella flexneri infection revealed by phosphoproteomics. Mol Cell Proteomics. 2013;12(10):2952-68.

25. Mustelin T, Vang T, Bottini N. Protein tyrosine phosphatases and the immune response. Nat Rev Immunol. 2005;5(1):43-57.

26. Macek B, Mann M, Olsen JV. Global and site-specific quantitative phosphoproteomics: principles and applications. Annu Rev Pharmacol Toxicol. 2009;49:199-221.

27. Riley NM, Coon JJ. Phosphoproteomics in the Age of Rapid and Deep Proteome Profiling. Anal Chem. 2016;88(1):74-94.

28. Schwartz D, Gygi SP. An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. Nat Biotechnol. 2005;23(11):1391-8.

29. Chou MF, Schwartz D. Biological sequence motif discovery using motif-x. Curr Protoc Bioinformatics. 2011;Chapter 13:Unit 13 5-24.

30. Wagih O, Sugiyama N, Ishihama Y, Beltrao P. Uncovering Phosphorylation-Based Specificities through Functional Interaction Networks. Mol Cell Proteomics. 2016;15(1):236-45.

31. He Z, Yang C, Guo G, Li N, Yu W. Motif-All: discovering all phosphorylation motifs. BMC Bioinformatics. 2011;12 Suppl 1:S22.

32. Ritz A, Shakhnarovich G, Salomon AR, Raphael BJ. Discovery of phosphorylation motif mixtures in phosphoproteomics data. Bioinformatics. 2009;25(1):14-21.

33. Zaman SB, Hussain MA, Nye R, Mehta V, Mamun KT, Hossain N. A Review on Antibiotic Resistance: Alarm Bells are Ringing. Cureus. 2017;9(6):e1403.

34. Pechere JC, Hughes D, Kardas P, Cornaglia G. Non-compliance with antibiotic therapy for acute community infections: a global survey. Int J Antimicrob Agents. 2007;29(3):245-53.

35. Chambers HF, Deleo FR. Waves of resistance: Staphylococcus aureus in the antibiotic era. Nat Rev Microbiol. 2009;7(9):629-41.

36. Esposito S, Noviello S, Leone S. Epidemiology and microbiology of skin and soft tissue infections. Curr Opin Infect Dis. 2016;29(2):109-15.

37. Magill SS, Edwards JR, Bamberg W, Beldavs ZG, Dumyati G, Kainer MA, Lynfield R, Maloney M, McAllister-Hollod L, Nadle J, Ray SM, Thompson DL, Wilson LE, Fridkin SK, Emerging Infections Program Healthcare-Associated I, Antimicrobial Use Prevalence Survey T. Multistate point-prevalence survey of health care-associated infections. N Engl J Med. 2014;370(13):1198-208.

38. Infectious Diseases Society of A, Spellberg B, Blaser M, Guidos RJ, Boucher HW, Bradley JS, Eisenstein BI, Gerding D, Lynfield R, Reller LB, Rex J, Schwartz D, Septimus E, Tenover FC, Gilbert DN. Combating antimicrobial resistance: policy recommendations to save lives. Clin Infect Dis. 2011;52 Suppl 5:S397-428.

39. David MZ, Daum RS. Community-associated methicillin-resistant Staphylococcus aureus: epidemiology and clinical consequences of an emerging epidemic. Clin Microbiol Rev. 2010;23(3):616-87.

40. Wang R, Braughton KR, Kretschmer D, Bach TH, Queck SY, Li M, Kennedy AD, Dorward DW, Klebanoff SJ, Peschel A, DeLeo FR, Otto M. Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. Nat Med. 2007;13(12):1510-4.

41. Perez-Llarena FJ, Bou G. Proteomics As a Tool for Studying Bacterial Virulence and Antimicrobial Resistance. Front Microbiol. 2016;7:410.

42. Bakour S, Sankar SA, Rathored J, Biagini P, Raoult D, Fournier PE. Identification of virulence factors and antibiotic resistance markers using bacterial genomics. Future Microbiol. 2016;11(3):455-66.

43. Lapek JD, Jr., Lewinski MK, Wozniak JM, Guatelli J, Gonzalez DJ. Quantitative Temporal Viromics of an Inducible HIV-1 Model Yields Insight to Global Host Targets and Phospho-Dynamics Associated with Protein Vpr. Mol Cell Proteomics. 2017;16(8):1447-61.

44. Nizet V. Stopping superbugs, maintaining the microbiota. Sci Transl Med. 2015;7(295):295ed8.

45. Eriksen NH, Espersen F, Rosdahl VT, Jensen K. Carriage of Staphylococcus aureus among 104 healthy persons during a 19-month period. Epidemiol Infect. 1995;115(1):51-60.

46. Kluytmans J, van Belkum A, Verbrugh H. Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. Clin Microbiol Rev. 1997;10(3):505-20.

47. Tong SY, Davis JS, Eichenberger E, Holland TL, Fowler VG, Jr. Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. Clin Microbiol Rev. 2015;28(3):603-61.

48. CDC.gov. General Information About MRSA in the Community 2016 [updated June 16th, 2017. Available from: https://www.cdc.gov/mrsa/community/index.html.

49. Klevens RM, Morrison MA, Nadle J, Petit S, Gershman K, Ray S, Harrison LH, Lynfield R, Dumyati G, Townes JM, Craig AS, Zell ER, Fosheim GE, McDougal LK, Carey RB, Fridkin SK, Active Bacterial Core surveillance MI. Invasive methicillin-resistant Staphylococcus aureus infections in the United States. JAMA. 2007;298(15):1763-71.

50. Cadena J, Richardson AM, Frei CR. Risk factors for methicillin-resistant Staphylococcus aureus skin and soft tissue infection in MRSA-colonized patients discharged from a Veterans Affairs hospital. Epidemiol Infect. 2016;144(3):647-51.

51. D'Lima NG, Ma J, Winkler L, Chu Q, Loh KH, Corpuz EO, Budnik BA, Lykke-Andersen J, Saghatelian A, Slavoff SA. A human microprotein that interacts with the mRNA decapping complex. Nat Chem Biol. 2017;13(2):174-80.

52. Slavoff SA, Mitchell AJ, Schwaid AG, Cabili MN, Ma J, Levin JZ, Karger AD, Budnik BA, Rinn JL, Saghatelian A. Peptidomic discovery of short open reading frame-encoded peptides in human cells. Nat Chem Biol. 2013;9(1):59-64.

53. Thoendel M, Kavanaugh JS, Flack CE, Horswill AR. Peptide signaling in the staphylococci. Chem Rev. 2011;111(1):117-51.

54. Periasamy S, Joo HS, Duong AC, Bach TH, Tan VY, Chatterjee SS, Cheung GY, Otto M. How Staphylococcus aureus biofilms develop their characteristic structure. Proc Natl Acad Sci U S A. 2012;109(4):1281-6.

55. Cheung GY, Joo HS, Chatterjee SS, Otto M. Phenol-soluble modulins--critical determinants of staphylococcal virulence. FEMS Microbiol Rev. 2014;38(4):698-719.

56. Frith MC, Forrest AR, Nourbakhsh E, Pang KC, Kai C, Kawai J, Carninci P, Hayashizaki Y, Bailey TL, Grimmond SM. The abundance of short proteins in the mammalian proteome. PLoS Genet. 2006;2(4):e52.

57. Andrews SJ, Rothnagel JA. Emerging evidence for functional peptides encoded by short open reading frames. Nat Rev Genet. 2014;15(3):193-204.

58. Shvedova AA, Tyurina JY, Kawai K, Tyurin VA, Kommineni C, Castranova V, Fabisiak JP, Kagan VE. Selective peroxidation and externalization of phosphatidylserine in normal human epidermal keratinocytes during oxidative stress induced by cumene hydroperoxide. J Invest Dermatol. 2002;118(6):1008-18.

59. Ponec M, Weerheim A, Kempenaar J, Mommaas AM, Nugteren DH. Lipid-Composition of Cultured Human Keratinocytes in Relation to Their Differentiation. J Lipid Res. 1988;29(7):949-61.

60. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic acids research. 2001;29(9):e45.

61. Haas W, Faherty BK, Gerber SA, Elias JE, Beausoleil SA, Bakalarski CE, Li X, Villen J, Gygi SP. Optimization and use of peptide mass measurement accuracy in shotgun proteomics. Mol Cell Proteomics. 2006;5(7):1326-37.

62. Tolonen AC, Haas W. Quantitative proteomics using reductive dimethylation for stable isotope labeling. J Vis Exp. 2014(89).

63. Wang Y, Yang F, Gritsenko MA, Wang Y, Clauss T, Liu T, Shen Y, Monroe ME, Lopez-Ferrer D, Reno T, Moore RJ, Klemke RL, Camp DG, 2nd, Smith RD. Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. Proteomics. 2011;11(10):2019-26.

64. Eng JK, McCormack AL, Yates JR. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom. 1994;5(11):976-89.

65. Charif D, Lobry, J. R. . SeqinR 1.0-2: A Contributed Package to the R Project for Statistical Computing Devoted to Biological Sequences Retrieval and Analysis. Structural Approaches to Sequence Evolution2007. p. 207-32

66. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods. 2007;4(3):207-14.

67. Elias JE, Haas W, Faherty BK, Gygi SP. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. Nature Methods. 2005;2(9):667-75.

68. Lapek JD, Jr., Greninger P, Morris R, Amzallag A, Pruteanu-Malinici I, Benes CH, Haas W. Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. Nat Biotechnol. 2017.

69. Gonzalez DJ, Corriden R, Akong-Moore K, Olson J, Dorrestein PC, Nizet V. N-Terminal ArgD Peptides from the Classical Staphylococcus aureus Agr System Have Cytotoxic and Proinflammatory Activities. Chemistry & Biology. 2014;21(11):1457-62.

70. Payne SH, Bonissone S, Wu S, Brown RN, Ivankov DN, Frishman D, Pasa-Tolic L, Smith RD, Pevzner PA. Unexpected diversity of signal peptides in prokaryotes. MBio. 2012;3(6).

71. Martoglio B. Intramembrane proteolysis and post-targeting functions of signal peptides. Biochem Soc Trans. 2003;31(Pt 6):1243-7.

72. Bufe B, Schumann T, Kappl R, Bogeski I, Kummerow C, Podgorska M, Smola S, Hoth M, Zufall F. Recognition of bacterial signal peptides by mammalian formyl peptide receptors: a new mechanism for sensing pathogens. J Biol Chem. 2015;290(12):7369-87.

73. Joo HS, Cheung GY, Otto M. Antimicrobial activity of community-associated methicillin-resistant Staphylococcus aureus is caused by phenol-soluble modulin derivatives. J Biol Chem. 2011;286(11):8933-40.

74. Gonzalez DJ, Okumura CY, Hollands A, Kersten R, Akong-Moore K, Pence MA, Malone CL, Derieux J, Moore BS, Horswill AR, Dixon JE, Dorrestein PC, Nizet V. Novel phenol-soluble modulin derivatives in community-associated methicillin-resistant Staphylococcus aureus identified through imaging mass spectrometry. J Biol Chem. 2012;287(17):13889-98.

75. Zhu W, Smith JW, Huang CM. Mass spectrometry-based label-free quantitative proteomics. J Biomed Biotechnol. 2010;2010:840518.

76. Silva MT. Neutrophils and macrophages work in concert as inducers and effectors of adaptive immunity against extracellular and intracellular microbial pathogens. J Leukoc Biol. 2010;87(5):805-13.

77. Nauseef WM. How human neutrophils kill and degrade microbes: an integrated view. Immunol Rev. 2007;219:88-102.

78. Kretschmer D, Gleske AK, Rautenberg M, Wang R, Koberle M, Bohn E, Schoneberg T, Rabiet MJ, Boulay F, Klebanoff SJ, van Kessel KA, van Strijp JA, Otto M, Peschel A. Human formyl peptide receptor 2 senses highly pathogenic Staphylococcus aureus. Cell Host Microbe. 2010;7(6):463-73.

79. Le Y, Li B, Gong W, Shen W, Hu J, Dunlop NM, Oppenheim JJ, Wang JM. Novel pathophysiological role of classical chemotactic peptide receptors and their communications with chemokine receptors. Immunol Rev. 2000;177:185-94.

80. Nestle FO, Di Meglio P, Qin JZ, Nickoloff BJ. Skin immune sentinels in health and disease. Nat Rev Immunol. 2009;9(10):679-91.

81. Rasigade JP, Trouillet-Assant S, Ferry T, Diep BA, Sapin A, Lhoste Y, Ranfaing J, Badiou C, Benito Y, Bes M, Couzon F, Tigaud S, Lina G, Etienne J, Vandenesch F, Laurent F. PSMs of hypervirulent Staphylococcus aureus act as intracellular toxins that kill infected osteoblasts. PLoS One. 2013;8(5):e63176.

82. Surewaard BG, Nijland R, Spaan AN, Kruijtzer JA, de Haas CJ, van Strijp JA. Inactivation of staphylococcal phenol soluble modulins by serum lipoprotein particles. PLoS Pathog. 2012;8(3):e1002606.

83. Corrigan RM, Foster TJ. An improved tetracycline-inducible expression vector for Staphylococcus aureus. Plasmid. 2009;61(2):126-9.

84. Zhu C, Bai Y, Liu Q, Li D, Hong J, Yang Z, Cui L, Hua X, Yuan C. Depolymerization of cytokeratin intermediate filaments facilitates intracellular infection of HeLa cells by Bartonella henselae. J Infect Dis. 2013;207(9):1397-405.

85. Labudova M, Tomaskova J, Skultety L, Pastorek J, Pastorekova S. The nucleoprotein of lymphocytic choriomeningitis virus facilitates spread of persistent infection through stabilization of the keratin network. J Virol. 2009;83(16):7842-9.

86. Chen L, Tai PC, Briggs MS, Gierasch LM. Protein translocation into Escherichia coli membrane vesicles is inhibited by functional synthetic signal peptides. J Biol Chem. 1987;262(4):1427-9.

87. Da F, Joo HS, Cheung GYC, Villaruz AE, Rohde H, Luo X, Otto M. Phenol-Soluble Modulin Toxins of Staphylococcus haemolyticus. Front Cell Infect Microbiol. 2017;7:206.

88. Chatterjee SS, Joo HS, Duong AC, Dieringer TD, Tan VY, Song Y, Fischer ER, Cheung GY, Li M, Otto M. Essential Staphylococcus aureus toxin export system. Nat Med. 2013;19(3):364-7.

89. Steverding D. The history of Chagas disease. Parasit Vectors. 2014;7:317.

90. Basile L, Jansa JM, Carlier Y, Salamanca DD, Angheben A, Bartoloni A, Seixas J, Van Gool T, Canavate C, Flores-Chavez M, Jackson Y, Chiodini PL, Albajar-Vinas P,

Working Group on Chagas D. Chagas disease in European countries: the challenge of a surveillance system. Euro Surveill. 2011;16(37).

91. Manne-Goehler J, Umeh CA, Montgomery SP, Wirtz VJ. Estimating the Burden of Chagas Disease in the United States. PLoS Negl Trop Dis. 2016;10(11):e0005033.

92. Bern C, Montgomery SP. An estimate of the burden of Chagas disease in the United States. Clin Infect Dis. 2009;49(5):e52-4.

93. Bonney KM, Luthringer DJ, Kim SA, Garg NJ, Engman DM. Pathology and Pathogenesis of Chagas Heart Disease. Annu Rev Pathol. 2019;14:421-47.

94. WHO. Neglected tropical diseases 2019 [Available from: https://www.who.int/neglected_diseases/diseases/en/.

95. Lee BY, Bacon KM, Bottazzi ME, Hotez PJ. Global economic burden of Chagas disease: a computational simulation model. Lancet Infect Dis. 2013;13(4):342-8.

96. Perez-Molina JA, Molina I. Chagas disease. Lancet. 2018;391(10115):82-94.

97. Benziger CP, do Carmo GAL, Ribeiro ALP. Chagas Cardiomyopathy: Clinical Presentation and Management in the Americas. Cardiol Clin. 2017;35(1):31-47.

98. Davila DF, Rossell O, de Bellabarba GA. Pathogenesis of chronic chagas heart disease: parasite persistence and autoimmune responses versus cardiac remodelling and neurohormonal activation. Int J Parasitol. 2002;32(1):107-9.

99. Bonney KM, Engman DM. Autoimmune pathogenesis of Chagas heart disease: looking back, looking ahead. Am J Pathol. 2015;185(6):1537-47.

100. Sanchez-Valdez FJ, Padilla A, Wang W, Orr D, Tarleton RL. Spontaneous dormancy protects Trypanosoma cruzi during extended drug exposure. Elife. 2018;7.

101. Esper L, Talvani A, Pimentel P, Teixeira MM, Machado FS. Molecular mechanisms of myocarditis caused by Trypanosoma cruzi. Curr Opin Infect Dis. 2015;28(3):246-52.

102. Chandrasekar B, Melby PC, Troyer DA, Colston JT, Freeman GL. Temporal expression of pro-inflammatory cytokines and inducible nitric oxide synthase in experimental acute Chagasic cardiomyopathy. Am J Pathol. 1998;152(4):925-34.

103. Cardillo F, Voltarelli JC, Reed SG, Silva JS. Regulation of Trypanosoma cruzi infection in mice by gamma interferon and interleukin 10: role of NK cells. Infect Immun. 1996;64(1):128-34.

104. Araujo-Jorge TC, Waghabi MC, Soeiro Mde N, Keramidas M, Bailly S, Feige JJ. Pivotal role for TGF-beta in infectious heart disease: The case of Trypanosoma cruzi infection and consequent Chagasic myocardiopathy. Cytokine Growth Factor Rev. 2008;19(5-6):405-13.

105. Rath PC, Aggarwal BB. TNF-induced signaling in apoptosis. J Clin Immunol. 1999;19(6):350-64.

106. Calvet CM, Oliveira FO, Jr., Araujo-Jorge TC, Pereira MC. Regulation of extracellular matrix expression and distribution in Trypanosoma cruzi-infected cardiomyocytes. Int J Med Microbiol. 2009;299(4):301-12.

107. Udoko AN, Johnson CA, Dykan A, Rachakonda G, Villalta F, Mandape SN, Lima MF, Pratap S, Nde PN. Early Regulation of Profibrotic Genes in Primary Human Cardiac Myocytes by Trypanosoma cruzi. PLoS Negl Trop Dis. 2016;10(1):e0003747.

108. Suman S, Rachakonda G, Mandape SN, Sakhare SS, Villalta F, Pratap S, Lima MF, Nde PN. Phospho-proteomic analysis of primary human colon epithelial cells during the early Trypanosoma cruzi infection phase. PLoS Negl Trop Dis. 2018;12(9):e0006792.

109. Silva TAF, L.F.D.C.; Pereira, M.C.D.S.; Calvet, C.M. Differential Role of TGF-β in Extracellular Matrix Regulation during Trypanosoma cruzi—Host Cell Interaction. Preprints. 2019;2019050171.

110. Si H, Liu D. Phytochemical genistein in the regulation of vascular function: new insights. Curr Med Chem. 2007;14(24):2581-9.

111. Qin W, Du N, Zhang L, Wu X, Hu Y, Li X, Shen N, Li Y, Yang B, Xu C, Fang Z, Lu Y, Zhang Y, Du Z. Genistein alleviates pressure overload-induced cardiac dysfunction and interstitial fibrosis in mice. Br J Pharmacol. 2015;172(23):5559-72.

112. Bennett BL, Sasaki DT, Murray BW, O'Leary EC, Sakata ST, Xu W, Leisten JC, Motiwala A, Pierce S, Satoh Y, Bhagwat SS, Manning AM, Anderson DW. SP600125, an anthrapyrazolone inhibitor of Jun N-terminal kinase. Proc Natl Acad Sci U S A. 2001;98(24):13681-6.

113. Wu W, Shan J, Bonne G, Worman HJ, Muchir A. Pharmacological inhibition of c-Jun N-terminal kinase signaling prevents cardiomyopathy caused by mutation in LMNA gene. Biochim Biophys Acta. 2010;1802(7-8):632-8.

114. Garg NJ, Soman KV, Zago MP, Koo SJ, Spratt H, Stafford S, Blell ZN, Gupta S, Nunez Burgos J, Barrientos N, Brasier AR, Wiktorowicz JE. Changes in Proteome Profile of Peripheral Blood Mononuclear Cells in Chronic Chagas Disease. PLoS Negl Trop Dis. 2016;10(2):e0004490.

115. Baez AL, Reynoso MN, Lo Presti MS, Bazan PC, Strauss M, Miler N, Pons P, Rivarola HW, Paglini-Oliva P. Mitochondrial dysfunction in skeletal muscle during experimental Chagas disease. Exp Mol Pathol. 2015;98(3):467-75.

116. Lopez M, Tanowitz HB, Garg NJ. Pathogenesis of Chronic Chagas Disease: Macrophages, Mitochondria, and Oxidative Stress. Curr Clin Microbiol Rep. 2018;5(1):45-54.

117. Calvet CM, Choi JY, Thomas D, Suzuki B, Hirata K, Lostracco-Johnson S, de Mesquita LB, Nogueira A, Meuser-Batista M, Silva TA, Siqueira-Neto JL, Roush WR, de Souza Pereira MC, McKerrow JH, Podust LM. 4-aminopyridyl-based lead compounds targeting CYP51 prevent spontaneous parasite relapse in a chronic model and improve cardiac pathology in an acute model of Trypanosoma cruzi infection. PLoS Negl Trop Dis. 2017;11(12):e0006132.

118. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676-82.

119. Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, Villen J, Haas W, Sowa ME, Gygi SP. A tissue-specific atlas of mouse protein phosphorylation and expression. Cell. 2010;143(7):1174-89.

120. Xiao Y, Hsiao TH, Suresh U, Chen HI, Wu X, Wolf SE, Chen Y. A novel significance score for gene selection and ranking. Bioinformatics. 2014;30(6):801-7.

121. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37(1):1-13.

122. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57.

123. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-D13.

124. Wozniak JM, Gonzalez DJ. PTMphinder: an R package for PTM site localization and motif extraction from proteomic datasets. PeerJ. 2019;7:e7046.

125. Xue Y, Liu Z, Cao J, Ma Q, Gao X, Wang Q, Jin C, Zhou Y, Wen L, Ren J. GPS 2.1: enhanced prediction of kinase-specific phosphorylation sites with an algorithm of motif length selection. Protein Eng Des Sel. 2011;24(3):255-60.

126. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. 2015;43(Database issue):D512-20.

127. Brito BOF, Ribeiro ALP. Electrocardiogram in Chagas disease. Rev Soc Bras Med Trop. 2018;51(5):570-7.

128. Ferreira RR, Abreu RDS, Vilar-Pereira G, Degrave W, Meuser-Batista M, Ferreira NVC, da Cruz Moreira O, da Silva Gomes NL, Mello de Souza E, Ramos IP, Bailly S, Feige JJ, Lannes-Vieira J, de Araujo-Jorge TC, Waghabi MC. TGF-beta inhibitor therapy

decreases fibrosis and stimulates cardiac improvement in a pre-clinical study of chronic Chagas' heart disease. PLoS Negl Trop Dis. 2019;13(7):e0007602.

129. Rossi MA. The pattern of myocardial fibrosis in chronic Chagas' heart disease. Int J Cardiol. 1991;30(3):335-40.

130. Kuzmanov U, Guo H, Buchsbaum D, Cosme J, Abbasi C, Isserlin R, Sharma P, Gramolini AO, Emili A. Global phosphoproteomic profiling reveals perturbed signaling in a mouse model of dilated cardiomyopathy. Proc Natl Acad Sci U S A. 2016;113(44):12592-7.

131. Fruh K, Yang Y. Antigen presentation by MHC class I and its regulation by interferon gamma. Curr Opin Immunol. 1999;11(1):76-81.

132. Watts C. The endosome-lysosome pathway and information generation in the immune system. Biochim Biophys Acta. 2012;1824(1):14-21.

133. Paiva CN, Figueiredo RT, Kroll-Palhares K, Silva AA, Silverio JC, Gibaldi D, Pyrrho Ados S, Benjamim CF, Lannes-Vieira J, Bozza MT. CCL2/MCP-1 controls parasite burden, cell infiltration, and mononuclear activation during acute Trypanosoma cruzi infection. J Leukoc Biol. 2009;86(5):1239-46.

134. Machado FS, Koyama NS, Carregaro V, Ferreira BR, Milanezi CM, Teixeira MM, Rossi MA, Silva JS. CCR5 plays a critical role in the development of myocarditis and host protection in mice infected with Trypanosoma cruzi. J Infect Dis. 2005;191(4):627-36.

135. Huang H, Calderon TM, Berman JW, Braunstein VL, Weiss LM, Wittner M, Tanowitz HB. Infection of endothelial cells with Trypanosoma cruzi activates NF-kappaB and induces vascular adhesion molecule expression. Infect Immun. 1999;67(10):5434-40.

136. Kulkarni MM, Varikuti S, Terrazas C, Kimble JL, Satoskar AR, McGwire BS. Signal transducer and activator of transcription 1 (STAT-1) plays a critical role in control of Trypanosoma cruzi infection. Immunology. 2015;145(2):225-31.

137. Goncalves VM, Matteucci KC, Buzzo CL, Miollo BH, Ferrante D, Torrecilhas AC, Rodrigues MM, Alvarez JM, Bortoluci KR. NLRP3 controls Trypanosoma cruzi infection through a caspase-1-dependent IL-1R-independent NO production. PLoS Negl Trop Dis. 2013;7(10):e2469.

138. Reis DD, Jones EM, Tostes S, Lopes ER, Chapadeiro E, Gazzinelli G, Colley DG, McCurley TL. Expression of major histocompatibility complex antigens and adhesion molecules in hearts of patients with chronic Chagas' disease. Am J Trop Med Hyg. 1993;49(2):192-200.

139. Deng X, Sabino EC, Cunha-Neto E, Ribeiro AL, Ianni B, Mady C, Busch MP, Seielstad M, International RCSGftNREDS-IC. Genome wide association study (GWAS)

of Chagas cardiomyopathy in Trypanosoma cruzi seropositive subjects. PLoS One. 2013;8(11):e79629.

140. Castillo C, Carrillo I, Libisch G, Juiz N, Schijman A, Robello C, Kemmerling U. Host-parasite interaction: changes in human placental gene expression induced by Trypanosoma cruzi. Parasit Vectors. 2018;11(1):479.

141. Soares MB, de Lima RS, Rocha LL, Vasconcelos JF, Rogatto SR, dos Santos RR, Iacobas S, Goldenberg RC, Iacobas DA, Tanowitz HB, de Carvalho AC, Spray DC. Gene expression changes associated with myocarditis and fibrosis in hearts of mice with chronic chagasic cardiomyopathy. J Infect Dis. 2010;202(3):416-26.

142. Burke MA, Chang S, Wakimoto H, Gorham JM, Conner DA, Christodoulou DC, Parfenov MG, DePalma SR, Eminaga S, Konno T, Seidman JG, Seidman CE. Molecular profiling of dilated cardiomyopathy that progresses to heart failure. JCI Insight. 2016;1(6).

143. Freshney NW, Rawlinson L, Guesdon F, Jones E, Cowley S, Hsuan J, Saklatvala J. Interleukin-1 activates a novel protein kinase cascade that results in the phosphorylation of Hsp27. Cell. 1994;78(6):1039-49.

144. Park Y, Shon SK, Kim A, Kim KI, Yang Y, Cho DH, Lee MS, Lim JS. SOCS1 induced by NDRG2 expression negatively regulates STAT3 activation in breast cancer cells. Biochem Biophys Res Commun. 2007;363(2):361-7.

145. Liu X, Niu T, Liu X, Hou W, Zhang J, Yao L. Microarray profiling of HepG2 cells ectopically expressing NDRG2. Gene. 2012;503(1):48-55.

146. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res. 2006;34(Database issue):D668-72.

147. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 2018;46(D1):D1074-D82.

148. Mukherjee S, Machado FS, Huang H, Oz HS, Jelicks LA, Prado CM, Koba W, Fine EJ, Zhao D, Factor SM, Collado JE, Weiss LM, Tanowitz HB, Ashton AW. Aspirin treatment of mice infected with Trypanosoma cruzi and implications for the pathogenesis of Chagas disease. PLoS One. 2011;6(2):e16959.

149. Carvalho de Freitas R, Lonien SCH, Malvezi AD, Silveira GF, Wowk PF, da Silva RV, Yamauchi LM, Yamada-Ogatta SF, Rizzo LV, Bordignon J, Pinge-Filho P. Trypanosoma cruzi: Inhibition of infection of human monocytes by aspirin. Exp Parasitol. 2017;182:26-33.

150. Wen JJ, Garg NJ. Proteome expression and carbonylation changes during Trypanosoma cruzi infection and Chagas disease in rats. Mol Cell Proteomics. 2012;11(4):M111 010918.

151. Wen JJ, Zago MP, Nunez S, Gupta S, Burgos FN, Garg NJ. Serum proteomic signature of human chagasic patients for the identification of novel potential protein biomarkers of disease. Mol Cell Proteomics. 2012;11(8):435-52.

152. Zago MP, Wiktorowicz JE, Spratt H, Koo SJ, Barrientos N, Nunez Burgos A, Nunez Burgos J, Iniguez F, Botelli V, Leon de la Fuente R, Garg NJ. Potential Utility of Protein Targets of Cysteine-S-Nitrosylation in Identifying Clinical Disease Status in Human Chagas Disease. Front Microbiol. 2018;9:3320.

153. Lucena ACR, Amorim JC, de Paula Lima CV, Batista M, Krieger MA, de Godoy LMF, Marchini FK. Quantitative phosphoproteome and proteome analyses emphasize the influence of phosphorylation events during the nutritional stress of Trypanosoma cruzi: the initial moments of in vitro metacyclogenesis. Cell Stress Chaperones. 2019;24(5):927-36.

154. Queiroz RM, Charneau S, Mandacaru SC, Schwammle V, Lima BD, Roepstorff P, Ricart CA. Quantitative proteomic and phosphoproteomic analysis of Trypanosoma cruzi amastigogenesis. Mol Cell Proteomics. 2014;13(12):3457-72.

155. de Godoy LM, Marchini FK, Pavoni DP, Rampazzo Rde C, Probst CM, Goldenberg S, Krieger MA. Quantitative proteomics of Trypanosoma cruzi during metacyclogenesis. Proteomics. 2012;12(17):2694-703.

156. Astrom A, Tavakkol A, Pettersson U, Cromie M, Elder JT, Voorhees JJ. Molecular cloning of two human cellular retinoic acid-binding proteins (CRABP). Retinoic acid-induced expression of CRABP-II but not CRABP-I in adult human skin in vivo and in skin fibroblasts in vitro. J Biol Chem. 1991;266(26):17662-6.

157. Palace VP, Khaper N, Qin Q, Singal PK. Antioxidant potentials of vitamin A and carotenoids and their relevance to heart disease. Free Radic Biol Med. 1999;26(5-6):746-61.

158. Huang Z, Liu Y, Qi G, Brand D, Zheng SG. Role of Vitamin A in the Immune System. J Clin Med. 2018;7(9).

159. Wyss R, Bucheli F. Determination of endogenous levels of 13-cis-retinoic acid (isotretinoin), all-trans-retinoic acid (tretinoin) and their 4-oxo metabolites in human and animal plasma by high-performance liquid chromatography with automated column switching and ultraviolet detection. J Chromatogr B Biomed Sci Appl. 1997;700(1-2):31-47.

160. Reigada C, Valera-Vera EA, Saye M, Errasti AE, Avila CC, Miranda MR, Pereira CA. Trypanocidal Effect of Isotretinoin through the Inhibition of Polyamine and Amino Acid Transporters in Trypanosoma cruzi. PLoS Negl Trop Dis. 2017;11(3):e0005472.

161. Providello MV, Carneiro ZA, Portapilla GB, do Vale GT, Camargo RS, Tirapelli CR, de Albuquerque S. Benefits of Ascorbic Acid in Association with Low-Dose Benznidazole in Treatment of Chagas Disease. Antimicrob Agents Chemother. 2018;62(9).

162. Marques RE, Guabiraba R, Russo RC, Teixeira MM. Targeting CCL5 in inflammation. Expert Opin Ther Targets. 2013;17(12):1439-60.

163. Ge B, Li J, Wei Z, Sun T, Song Y, Khan NU. Functional expression of CCL8 and its interaction with chemokine receptor CCR3. BMC Immunol. 2017;18(1):54.

164. Vestal DJ, Jeyaratnam JA. The guanylate-binding proteins: emerging insights into the biochemical properties and functions of this family of large interferon-induced guanosine triphosphatase. J Interferon Cytokine Res. 2011;31(1):89-97.

165. Kim BH, Chee JD, Bradfield CJ, Park ES, Kumar P, MacMicking JD. Interferon-induced guanylate-binding proteins in inflammasome activation and host defense. Nat Immunol. 2016;17(5):481-9.

166. Virreira Winter S, Niedelman W, Jensen KD, Rosowski EE, Julien L, Spooner E, Caradonna K, Burleigh BA, Saeij JP, Ploegh HL, Frickel EM. Determinants of GBP recruitment to Toxoplasma gondii vacuoles and the parasitic factors that control it. PLoS One. 2011;6(9):e24434.

167. Degrandi D, Kravets E, Konermann C, Beuter-Gunia C, Klumpers V, Lahme S, Wischmann E, Mausberg AK, Beer-Hammer S, Pfeffer K. Murine guanylate binding protein 2 (mGBP2) controls Toxoplasma gondii replication. Proc Natl Acad Sci U S A. 2013;110(1):294-9.

168. Lentini G, Dos Santos Pacheco N, Burleigh BA. Targeting host mitochondria: A role for the Trypanosoma cruzi amastigote flagellum. Cell Microbiol. 2018;20(2).

169. Wen JJ, Garg NJ. Mitochondrial complex III defects contribute to inefficient respiration and ATP synthesis in the myocardium of Trypanosoma cruzi-infected mice. Antioxid Redox Signal. 2010;12(1):27-37.

170. Wen JJ, Yachelini PC, Sembaj A, Manzur RE, Garg NJ. Increased oxidative stress is correlated with mitochondrial dysfunction in chagasic patients. Free Radic Biol Med. 2006;41(2):270-6.

171. Gazos-Lopes F, Martin JL, Dumoulin PC, Burleigh BA. Host triacylglycerols shape the lipidome of intracellular trypanosomes and modulate their growth. PLoS Pathog. 2017;13(12):e1006800.

172. Meyers DE, Basha HI, Koenig MK. Mitochondrial cardiomyopathy: pathophysiology, diagnosis, and management. Tex Heart Inst J. 2013;40(4):385-94.

173. Sussman MA. Developing hearts need their SPEG. Circulation. 2009;119(2):213-4.

174. Bernau K, Torr EE, Evans MD, Aoki JK, Ngam CR, Sandbo N. Tensin 1 Is Essential for Myofibroblast Differentiation and Extracellular Matrix Formation. Am J Respir Cell Mol Biol. 2017;56(4):465-76.

175. Knezevic T, Myers VD, Gordon J, Tilley DG, Sharp TE, 3rd, Wang J, Khalili K, Cheung JY, Feldman AM. BAG3: a new player in the heart failure paradigm. Heart Fail Rev. 2015;20(4):423-34.

176. Kioka N, Ueda K, Amachi T. Vinexin, CAP/ponsin, ArgBP2: a novel adaptor protein family regulating cytoskeletal organization and signal transduction. Cell Struct Funct. 2002;27(1):1-7.

177. Granados-Riveron JT, Ghosh TK, Pope M, Bu'Lock F, Thornborough C, Eason J, Kirk EP, Fatkin D, Feneley MP, Harvey RP, Armour JA, David Brook J. Alpha-cardiac myosin heavy chain (MYH6) mutations affecting myofibril formation are associated with congenital heart defects. Hum Mol Genet. 2010;19(20):4007-16.

178. Carrier L, Mearini G, Stathopoulou K, Cuello F. Cardiac myosin-binding protein C (MYBPC3) in cardiac pathophysiology. Gene. 2015;573(2):188-97.

179. Horsthemke M, Nutter LMJ, Bachg AC, Skryabin BV, Honnert U, Zobel T, Bogdan S, Stoll M, Seidl MD, Muller FU, Ravens U, Unger A, Linke WA, van Gorp PRR, de Vries AAF, Bahler M, Hanley PJ. A novel isoform of myosin 18A (Myo18Agamma) is an essential sarcomeric protein in mouse heart. J Biol Chem. 2019;294(18):7202-18.

180. Sheikh F, Lyon RC, Chen J. Functions of myosin light chain-2 (MYL2) in cardiac muscle and disease. Gene. 2015;569(1):14-20.

181. Siegert R, Perrot A, Keller S, Behlke J, Michalewska-Wludarczyk A, Wycisk A, Tendera M, Morano I, Ozcelik C. A myomesin mutation associated with hypertrophic cardiomyopathy deteriorates dimerisation properties. Biochem Biophys Res Commun. 2011;405(3):473-9.

182. Osio A, Tan L, Chen SN, Lombardi R, Nagueh SF, Shete S, Roberts R, Willerson JT, Marian AJ. Myozenin 2 is a novel gene for human hypertrophic cardiomyopathy. Circ Res. 2007;100(6):766-8.

183. Ferreira RR, de Souza EM, de Oliveira FL, Ferrao PM, Gomes LH, Mendonca-Lima L, Meuser-Batista M, Bailly S, Feige JJ, de Araujo-Jorge TC, Waghabi MC. Proteins involved on TGF-beta pathway are up-regulated during the acute phase of experimental Chagas disease. Immunobiology. 2016;221(5):587-94.

184. Seidel P, Sun Q, Costa L, Lardinois D, Tamm M, Roth M. The MNK-1/eIF4E pathway as a new therapeutic pathway to target inflammation and remodelling in asthma. Cell Signal. 2016;28(10):1555-62.

185. Israel A. The IKK complex, a central regulator of NF-kappaB activation. Cold Spring Harb Perspect Biol. 2010;2(3):a000158.

186. Karin M, Delhase M. The I kappa B kinase (IKK) and NF-kappa B: key elements of proinflammatory signalling. Semin Immunol. 2000;12(1):85-98.

187. Waskiewicz AJ, Flynn A, Proud CG, Cooper JA. Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. EMBO J. 1997;16(8):1909-20.

188. Waskiewicz AJ, Johnson JC, Penn B, Mahalingam M, Kimball SR, Cooper JA. Phosphorylation of the cap-binding protein eukaryotic translation initiation factor 4E by protein kinase Mnk1 in vivo. Mol Cell Biol. 1999;19(3):1871-80.

189. Nihira NT, Yoshida K. Engagement of DYRK2 in proper control for cell division. Cell Cycle. 2015;14(6):802-7.

190. Taira N, Nihira K, Yamaguchi T, Miki Y, Yoshida K. DYRK2 is targeted to the nucleus and controls p53 via Ser46 phosphorylation in the apoptotic response to DNA damage. Mol Cell. 2007;25(5):725-38.

191. Slepak TI, Salay LD, Lemmon VP, Bixby JL. Dyrk kinases regulate phosphorylation of doublecortin, cytoskeletal organization, and neuronal morphology. Cytoskeleton (Hoboken). 2012;69(7):514-27.

192. Blanco S, Sanz-Garcia M, Santos CR, Lazo PA. Modulation of interleukin-1 transcriptional response by the interaction between VRK2 and the JIP1 scaffold protein. PLoS One. 2008;3(2):e1660.

193. Blanco S, Santos C, Lazo PA. Vaccinia-related kinase 2 modulates the stress response to hypoxia mediated by TAK1. Mol Cell Biol. 2007;27(20):7273-83.

194. Eng GWL, Edison, Virshup DM. Site-specific phosphorylation of casein kinase 1 delta (CK1delta) regulates its activity towards the circadian regulator PER2. PLoS One. 2017;12(5):e0177834.

195. Lin A, Frost J, Deng T, Smeal T, al-Alawi N, Kikkawa U, Hunter T, Brenner D, Karin M. Casein kinase II is a negative regulator of c-Jun DNA binding and AP-1 activity. Cell. 1992;70(5):777-89.

196. Singh NN, Ramji DP. Protein kinase CK2, an important regulator of the inflammatory response? J Mol Med (Berl). 2008;86(8):887-97.

197. Glushkova OV, Parfenyuk SB, Novoselova TV, Khrenov MO, Lunin SM, Novoselova EG. The Role of p38 and CK2 Protein Kinases in the Response of RAW 264.7 Macrophages to Lipopolysaccharide. Biochemistry (Mosc). 2018;83(6):746-54.

198. Moussazadeh M, Greene JJ. Redox regulation of casein kinase II autophosphorylation and its effect on Jun-DNA binding. Cell Mol Biol (Noisy-le-grand). 2000;46(7):1265-75.

199. Bussel J, Arnold DM, Grossbard E, Mayer J, Trelinski J, Homenda W, Hellmann A, Windyga J, Sivcheva L, Khalafallah AA, Zaja F, Cooper N, Markovtsov V, Zayed H, Duliege AM. Fostamatinib for the treatment of adult persistent and chronic immune thrombocytopenia: Results of two phase 3, randomized, placebo-controlled trials. Am J Hematol. 2018;93(7):921-30.

200. Smailhodzic D, van Asten F, Blom AM, Mohlin FC, den Hollander AI, van de Ven JP, van Huet RA, Groenewoud JM, Tian Y, Berendschot TT, Lechanteur YT, Fauser S, de Bruijn C, Daha MR, van der Wilt GJ, Hoyng CB, Klevering BJ. Zinc supplementation inhibits complement activation in age-related macular degeneration. PLoS One. 2014;9(11):e112682.

201. Schmaldienst S, Horl WH. Bacterial infections during immunosuppression - immunosuppressive agents interfere not only with immune response, but also with polymorphonuclear cell function. Nephrol Dial Transplant. 1996;11(7):1243-5.

202. Maywald M, Wessels I, Rink L. Zinc Signals and Immunity. Int J Mol Sci. 2017;18(10).

203. Wang FD, Chen YY, Chen TL, Liu CY. Risk factors and mortality in patients with nosocomial Staphylococcus aureus bacteremia. Am J Infect Control. 2008;36(2):118-22.

204. van Hal SJ, Jensen SO, Vaska VL, Espedido BA, Paterson DL, Gosbell IB. Predictors of mortality in Staphylococcus aureus Bacteremia. Clin Microbiol Rev. 2012;25(2):362-86.

205. Kern WV. Management of Staphylococcus aureus bacteremia and endocarditis: progresses and challenges. Curr Opin Infect Dis. 2010;23(4):346-58.

206. Rasmussen RV, Fowler VG, Jr., Skov R, Bruun NE. Future challenges and treatment of Staphylococcus aureus bacteremia with emphasis on MRSA. Future Microbiol. 2011;6(1):43-56.

207. Holland TL, Arnold C, Fowler VG, Jr. Clinical management of Staphylococcus aureus bacteremia: a review. JAMA. 2014;312(13):1330-41.

208. Rose WE, Eickhoff JC, Shukla SK, Pantrangi M, Rooijakkers S, Cosgrove SE, Nizet V, Sakoulas G. Elevated serum interleukin-10 at time of hospital admission is predictive of mortality in patients with Staphylococcus aureus bacteremia. J Infect Dis. 2012;206(10):1604-11.

209. Minejima E, Bensman J, She RC, Mack WJ, Tuan Tran M, Ny P, Lou M, Yamaki J, Nieberg P, Ho J, Wong-Beringer A. A Dysregulated Balance of Proinflammatory and Anti-Inflammatory Host Cytokine Response Early During Therapy Predicts Persistence and Mortality in Staphylococcus aureus Bacteremia. Crit Care Med. 2016;44(4):671-9.

210. Rose WE, Shukla SK, Berti AD, Hayney MS, Henriquez KM, Ranzoni A, Cooper MA, Proctor RA, Nizet V, Sakoulas G. Increased Endovascular Staphylococcus aureus

Inoculum Is the Link Between Elevated Serum Interleukin 10 Concentrations and Mortality in Patients With Bacteremia. Clin Infect Dis. 2017;64(10):1406-12.

211. Ferrer R, Martin-Loeches I, Phillips G, Osborn TM, Townsend S, Dellinger RP, Artigas A, Schorr C, Levy MM. Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. Crit Care Med. 2014;42(8):1749-55.

212. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, Sevransky JE, Sprung CL, Douglas IS, Jaeschke R, Osborn TM, Nunnally ME, Townsend SR, Reinhart K, Kleinpell RM, Angus DC, Deutschman CS, Machado FR, Rubenfeld GD, Webb SA, Beale RJ, Vincent JL, Moreno R, Surviving Sepsis Campaign Guidelines Committee including the Pediatric S. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Crit Care Med. 2013;41(2):580-637.

213. Ersoy SC, Heithoff DM, Barnes Lt, Tripp GK, House JK, Marth JD, Smith JW, Mahan MJ. Correcting a Fundamental Flaw in the Paradigm for Antimicrobial Susceptibility Testing. EBioMedicine. 2017;20:173-81.

214. Guimaraes AO, Cao Y, Hong K, Mayba O, Peck MC, Gutierrez J, Ruffin F, Carrasco-Triguero M, Dinoso JB, Clemenzi-Allen A, Koss CA, Maskarinec SA, Chambers HF, Fowler VG, Baruch A, Rosenberger CM. A Prognostic Model of Persistent Bacteremia and Mortality in Complicated Staphylococcus aureus Bloodstream Infection. Clin Infect Dis. 2019;68(9):1502-11.

215. Mikkelsen ME, Miltiades AN, Gaieski DF, Goyal M, Fuchs BD, Shah CV, Bellamy SL, Christie JD. Serum lactate is associated with mortality in severe sepsis independent of organ failure and shock. Crit Care Med. 2009;37(5):1670-7.

216. Ralto KM, Parikh SM. Mitochondria in Acute Kidney Injury. Semin Nephrol. 2016;36(1):8-16.

217. Singer M. The role of mitochondrial dysfunction in sepsis-induced multi-organ failure. Virulence. 2014;5(1):66-72.

218. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapono CA, Luzzatto-Knaan T, Porto C, Bouslimani A, Melnik AV, Meehan MJ, Liu WT, Crusemann M, Boudreau PD, Esquenazi E, Sandoval-Calderon M, Kersten RD, Pace LA, Quinn RA, Duncan KR, Hsu CC, Floros DJ, Gavilan RG, Kleigrewe K, Northen T, Dutton RJ, Parrot D, Carlson EE, Aigle B, Michelsen CF, Jelsbak L, Sohlenkamp C, Pevzner P, Edlund A, McLean J, Piel J, Murphy BT, Gerwick L, Liaw CC, Yang YL, Humpf HU, Maansson M, Keyzers RA, Sims AC, Johnson AR, Sidebottom AM, Sedio BE, Klitgaard A, Larson CB, P CAB, Torres-Mendoza D, Gonzalez DJ, Silva DB, Marques LM, Demarque DP, Pociute E, O'Neill EC, Briand E, Helfrich EJN, Granatosky EA, Glukhov E, Ryffel F, Houson H, Mohimani H, Kharbush JJ, Zeng Y, Vorholt JA, Kurita KL, Charusanti P, McPhail KL, Nielsen KF, Vuong L, Elfeki M, Traxler MF, Engene N, Koyama N, Vining OB, Baric R, Silva RR, Mascuch SJ, Tomasi S, Jenkins S, Macherla

V, Hoffman T, Agarwal V, Williams PG, Dai J, Neupane R, Gurr J, Rodriguez AMC, Lamsa A, Zhang C, Dorrestein K, Duggan BM, Almaliti J, Allard PM, Phapale P, Nothias LF, Alexandrov T, Litaudon M, Wolfender JL, Kyle JE, Metz TO, Peryea T, Nguyen DT, VanLeer D, Shinn P, Jadhav A, Muller R, Waters KM, Shi W, Liu X, Zhang L, Knight R, Jensen PR, Palsson BO, Pogliano K, Linington RG, Gutierrez M, Lopes NP, Gerwick WH, Moore BS, Dorrestein PC, Bandeira N. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat Biotechnol. 2016;34(8):828-37.

219. Scheubert K, Hufsky F, Petras D, Wang M, Nothias LF, Duhrkop K, Bandeira N, Dorrestein PC, Bocker S. Significance estimation for large scale metabolomics annotations by spectral matching. Nat Commun. 2017;8(1):1494.

220. Pluskal T, Castillo S, Villar-Briones A, Oresic M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics. 2010;11:395.

221. Olivon F, Grelier G, Roussi F, Litaudon M, Touboul D. MZmine 2 Data-Preprocessing To Enhance Molecular Networking Reliability. Anal Chem. 2017;89(15):7836-40.

222. Bern M, Kil YJ, Becker C. Byonic: advanced peptide and protein identification software. Curr Protoc Bioinformatics. 2012;Chapter 13:Unit13 20.

223. WHO WHO-. C-reactive protein concentrations as a marker of inflammation or infection for interpreting biomarkers of micronutrient status 2014 [Available from: https://www.who.int/nutrition/publications/micronutrients/indicators_c-reactive_protein/en/.

224. Povoa P, Coelho L, Almeida E, Fernandes A, Mealha R, Moreira P, Sabino H. C-reactive protein as a marker of infection in critically ill patients. Clin Microbiol Infect. 2005;11(2):101-8.

225. Park JH, Kim DH, Jang HR, Kim MJ, Jung SH, Lee JE, Huh W, Kim YG, Kim DJ, Oh HY. Clinical relevance of procalcitonin and C-reactive protein as infection markers in renal impairment: a cross-sectional study. Crit Care. 2014;18(6):640.

226. Liu A, Bui T, Van Nguyen H, Ong B, Shen Q, Kamalasena D. Serum C-reactive protein as a biomarker for early detection of bacterial infection in the older patient. Age Ageing. 2010;39(5):559-65.

227. Neumann U, Genze N, Heider D. EFS: an ensemble feature selection tool implemented as R-package and web-application. BioData Min. 2017;10:21.

228. He Z, Yu W. Stable feature selection for biomarker discovery. Comput Biol Chem. 2010;34(4):215-25.

229. Williams FM. Biomarkers: in combination they may do better. Arthritis Res Ther. 2009;11(5):130.

230. Ensor JE. Biomarker validation: common data analysis concerns. Oncologist. 2014;19(8):886-91.

231. Levey AS, Perrone RD, Madias NE. Serum creatinine and renal function. Annu Rev Med. 1988;39:465-90.

232. Nielsen LH, Jensen-Fangel S, Benfield T, Skov R, Jespersen B, Larsen AR, Ostergaard L, Stovring H, Schonheyder HC, Sogaard OS. Risk and prognosis of Staphylococcus aureus bacteremia among individuals with and without end-stage renal disease: a Danish, population-based cohort study. BMC Infect Dis. 2015;15:6.

233. Vandecasteele SJ, Boelaert JR, De Vriese AS. Staphylococcus aureus infections in hemodialysis: what a nephrologist should know. Clin J Am Soc Nephrol. 2009;4(8):1388-400.

234. Geyer PE, Holdt LM, Teupser D, Mann M. Revisiting biomarker discovery by plasma proteomics. Mol Syst Biol. 2017;13(9):942.

235. Chandramouli K, Qian PY. Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. Hum Genomics Proteomics. 2009;2009.

236. Dey KK, Wang H, Niu M, Bai B, Wang X, Li Y, Cho JH, Tan H, Mishra A, High AA, Chen PC, Wu Z, Beach TG, Peng J. Deep undepleted human serum proteome profiling toward biomarker discovery for Alzheimer's disease. Clin Proteomics. 2019;16:16.

237. Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, Stephens R, Baseler MW, Lane HC, Lempicki RA. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007;8(9):R183.

238. Soderquist B, Sundqvist KG, Vikerfors T. Adhesion molecules (E-selectin, intercellular adhesion molecule-1 (ICAM-1) and vascular cell adhesion molecule-1 (VCAM-1)) in sera from patients with Staphylococcus aureus bacteraemia with or without endocarditis. Clin Exp Immunol. 1999;118(3):408-11.

239. van Leeuwen HJ, Heezius EC, Dallinga GM, van Strijp JA, Verhoef J, van Kessel KP. Lipoprotein metabolism in patients with severe sepsis. Crit Care Med. 2003;31(5):1359-66.

240. Baxter RC. IGF binding proteins in cancer: mechanistic and clinical insights. Nat Rev Cancer. 2014;14(5):329-41.

241. Christoffersen C, Nielsen LB. Apolipoprotein M--a new biomarker in sepsis. Crit Care. 2012;16(3):126.

242. Sharma NK, Ferreira BL, Tashima AK, Brunialti MKC, Torquato RJS, Bafi A, Assuncao M, Azevedo LCP, Salomao R. Lipid metabolism impairment in patients with sepsis secondary to hospital acquired pneumonia, a proteomic analysis. Clin Proteomics. 2019;16:29.

243. Schussler GC. The thyroxine-binding proteins. Thyroid. 2000;10(2):141-9.

244. Plikat K, Langgartner J, Buettner R, Bollheimer LC, Woenckhaus U, Scholmerich J, Wrede CE. Frequency and outcome of patients with nonthyroidal illness syndrome in a medical intensive care unit. Metabolism. 2007;56(2):239-44.

245. Bello G, Pennisi MA, Montini L, Silva S, Maviglia R, Cavallaro F, Bianchi A, De Marinis L, Antonelli M. Nonthyroidal illness syndrome and prolonged mechanical ventilation in patients admitted to the ICU. Chest. 2009;135(6):1448-54.

246. McCoin CS, Knotts TA, Adams SH. Acylcarnitines--old actors auditioning for new roles in metabolic physiology. Nat Rev Endocrinol. 2015;11(10):617-25.

247. Keshishian H, Burgess MW, Gillette MA, Mertins P, Clauser KR, Mani DR, Kuhn EW, Farrell LA, Gerszten RE, Carr SA. Multiplexed, Quantitative Workflow for Sensitive Biomarker Discovery in Plasma Yields Novel Candidates for Early Myocardial Injury. Mol Cell Proteomics. 2015;14(9):2375-93.

248. Al-Abed Y, Metz CN, Cheng KF, Aljabari B, VanPatten S, Blau S, Lee H, Ochani M, Pavlov VA, Coleman T, Meurice N, Tracey KJ, Miller EJ. Thyroxine is a potential endogenous antagonist of macrophage migration inhibitory factor (MIF) activity. Proc Natl Acad Sci U S A. 2011;108(20):8224-7.

249. Tsourdi E, Rijntjes E, Kohrle J, Hofbauer LC, Rauner M. Hyperthyroidism and Hypothyroidism in Male Mice and Their Effects on Bone Mass, Bone Turnover, and the Wnt Inhibitors Sclerostin and Dickkopf-1. Endocrinology. 2015;156(10):3517-27.

250. Wolf AM, Wolf D, Rumpold H, Enrich B, Tilg H. Adiponectin induces the anti-inflammatory cytokines IL-10 and IL-1RA in human leukocytes. Biochem Biophys Res Commun. 2004;323(2):630-5.

251. Leech JM, Lacey KA, Mulcahy ME, Medina E, McLoughlin RM. IL-10 Plays Opposing Roles during Staphylococcus aureus Systemic and Localized Infections. J Immunol. 2017;198(6):2352-65.

252. Hawkins C, Huang J, Jin N, Noskin GA, Zembower TR, Bolon M. Persistent Staphylococcus aureus bacteremia: an analysis of risk factors and outcomes. Arch Intern Med. 2007;167(17):1861-7.

253. Pastagia M, Kleinman LC, Lacerda de la Cruz EG, Jenkins SG. Predicting risk for death from MRSA bacteremia. Emerg Infect Dis. 2012;18(7):1072-80.

254. Liu C, Bayer A, Cosgrove SE, Daum RS, Fridkin SK, Gorwitz RJ, Kaplan SL, Karchmer AW, Levine DP, Murray BE, M JR, Talan DA, Chambers HF, Infectious Diseases Society of A. Clinical practice guidelines by the infectious diseases society of america for the treatment of methicillin-resistant Staphylococcus aureus infections in adults and children. Clin Infect Dis. 2011;52(3):e18-55.

255. Fowler VG, Jr., Boucher HW, Corey GR, Abrutyn E, Karchmer AW, Rupp ME, Levine DP, Chambers HF, Tally FP, Vigliani GA, Cabell CH, Link AS, DeMeyer I, Filler SG, Zervos M, Cook P, Parsonnet J, Bernstein JM, Price CS, Forrest GN, Fatkenheuer G, Gareca M, Rehm SJ, Brodt HR, Tice A, Cosgrove SE, Endocarditis Sa, Bacteremia Study G. Daptomycin versus standard therapy for bacteremia and endocarditis caused by Staphylococcus aureus. N Engl J Med. 2006;355(7):653-65.

256. Rehm SJ, Boucher H, Levine D, Campion M, Eisenstein BI, Vigliani GA, Corey GR, Abrutyn E. Daptomycin versus vancomycin plus gentamicin for treatment of bacteraemia and endocarditis due to Staphylococcus aureus: subset analysis of patients infected with methicillin-resistant isolates. J Antimicrob Chemother. 2008;62(6):1413-21.

257. Geriak M, Haddad F, Rizvi K, Rose W, Kullar R, LaPlante K, Yu M, Vasina L, Ouellette K, Zervos M, Nizet V, Sakoulas G. Clinical Data on Daptomycin plus Ceftaroline versus Standard of Care Monotherapy in the Treatment of Methicillin-Resistant Staphylococcus aureus Bacteremia. Antimicrob Agents Chemother. 2019;63(5).

258. Gomes DM, Smotherman C, Birch A, Dupree L, Della Vecchia BJ, Kraemer DF, Jankowski CA. Comparison of acute kidney injury during treatment with vancomycin in combination with piperacillin-tazobactam or cefepime. Pharmacotherapy. 2014;34(7):662-9.

259. Burgess LD, Drew RH. Comparison of the incidence of vancomycin-induced nephrotoxicity in hospitalized patients with and without concomitant piperacillin-tazobactam. Pharmacotherapy. 2014;34(7):670-6.

260. Tong SYC, Lye DC, Yahav D, Sud A, Robinson JO, Nelson J, Archuleta S, Roberts MA, Cass A, Paterson DL, Foo H, Paul M, Guy SD, Tramontana AR, Walls GB, McBride S, Bak N, Ghosh N, Rogers BA, Ralph AP, Davies J, Ferguson PE, Dotel R, McKew GL, Gray TJ, Holmes NE, Smith S, Warner MS, Kalimuddin S, Young BE, Runnegar N, Andresen DN, Anagnostou NA, Johnson SA, Chatfield MD, Cheng AC, Fowler VG, Jr., Howden BP, Meagher N, Price DJ, van Hal SJ, O'Sullivan MVN, Davis JS, Australasian Society for Infectious Diseases Clinical Research N. Effect of Vancomycin or Daptomycin With vs Without an Antistaphylococcal beta-Lactam on Mortality, Bacteremia, Relapse, or Treatment Failure in Patients With MRSA Bacteremia: A Randomized Clinical Trial. JAMA. 2020;323(6):527-37.

261. Silsirivanit A. Glycosylation markers in cancer. Adv Clin Chem. 2019;89:189-213.

262. Regan P, McClean PL, Smyth T, Doherty M. Early Stage Glycosylation Biomarkers in Alzheimer's Disease. Medicines (Basel). 2019;6(3).

263. Gornik O, Lauc G. Glycosylation of serum proteins in inflammatory diseases. Dis Markers. 2008;25(4-5):267-78.

264. Olivier E, Soury E, Ruminy P, Husson A, Parmentier F, Daveau M, Salier JP. Fetuin-B, a second member of the fetuin family in mammals. Biochem J. 2000;350 Pt 2:589-97.

265. Dabrowska AM, Tarach JS, Wojtysiak-Duma B, Duma D. Fetuin-A (AHSG) and its usefulness in clinical practice. Review of the literature. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2015;159(3):352-9.

266. Cayatte AJ, Kumbla L, Subbiah MT. Marked acceleration of exogenous fatty acid incorporation into cellular triglycerides by fetuin. J Biol Chem. 1990;265(10):5883-8.

267. Wang H, Sama AE. Anti-inflammatory role of fetuin-A in injury and infection. Curr Mol Med. 2012;12(5):625-33.

268. Jaisson S, Pietrement C, Gillery P. Protein Carbamylation: Chemistry, Pathophysiological Involvement, and Biomarkers. Adv Clin Chem. 2018;84:1-38.

269. Berg AH, Drechsler C, Wenger J, Buccafusca R, Hod T, Kalim S, Ramma W, Parikh SM, Steen H, Friedman DJ, Danziger J, Wanner C, Thadhani R, Karumanchi SA. Carbamylation of serum albumin as a risk factor for mortality in patients with kidney failure. Sci Transl Med. 2013;5(175):175ra29.

270. Pruijn GJ. Citrullination and carbamylation in the pathophysiology of rheumatoid arthritis. Front Immunol. 2015;6:192.

271. Kalim S, Tamez H, Wenger J, Ankers E, Trottier CA, Deferio JJ, Berg AH, Karumanchi SA, Thadhani RI. Carbamylation of serum albumin and erythropoietin resistance in end stage kidney disease. Clin J Am Soc Nephrol. 2013;8(11):1927-34.

272. Alobaidi R, Basu RK, Goldstein SL, Bagshaw SM. Sepsis-associated acute kidney injury. Semin Nephrol. 2015;35(1):2-11.

273. Joost I, Kaasch A, Pausch C, Peyerl-Hoffmann G, Schneider C, Voll RE, Seifert H, Kern WV, Rieg S. Staphylococcus aureus bacteremia in patients with rheumatoid arthritis - Data from the prospective INSTINCT cohort. J Infect. 2017;74(6):575-84.

274. Goodman SM, Nocon AA, Selemon NA, Shopsin B, Fulmer Y, Decker ME, Grond SE, Donlin LT, Figgie MP, Sculco TP, Russell LA, Henry ME, Bass AR, Miller AO, Sculco PK. Increased Staphylococcus aureus Nasal Carriage Rates in Rheumatoid Arthritis Patients on Biologic Therapy. J Arthroplasty. 2019;34(5):954-8.

275. Shi J, van de Stadt LA, Levarht EW, Huizinga TW, Hamann D, van Schaardenburg D, Toes RE, Trouw LA. Anti-carbamylated protein (anti-CarP) antibodies precede the onset of rheumatoid arthritis. Ann Rheum Dis. 2014;73(4):780-3.

276. Alonzo F, 3rd, Kozhaya L, Rawlings SA, Reyes-Robles T, DuMont AL, Myszka DG, Landau NR, Unutmaz D, Torres VJ. CCR5 is a receptor for Staphylococcus aureus leukotoxin ED. Nature. 2013;493(7430):51-5.

277. Walenkamp AM, Boer IG, Bestebroer J, Rozeveld D, Timmer-Bosscha H, Hemrika W, van Strijp JA, de Haas CJ. Staphylococcal superantigen-like 10 inhibits CXCL12-induced human tumor cell migration. Neoplasia. 2009;11(4):333-44.

278. Chen Y, Sjolinder M, Wang X, Altenbacher G, Hagner M, Berglund P, Gao Y, Lu T, Jonsson AB, Sjolinder H. Thyroid hormone enhances nitric oxide-mediated bacterial clearance and promotes survival after meningococcal infection. PLoS One. 2012;7(7):e41445.

279. Achari AE, Jain SK. Adiponectin, a Therapeutic Target for Obesity, Diabetes, and Endothelial Dysfunction. Int J Mol Sci. 2017;18(6).

280. Teoh H, Quan A, Bang KW, Wang G, Lovren F, Vu V, Haitsma JJ, Szmitko PE, Al-Omran M, Wang CH, Gupta M, Peterson MD, Zhang H, Chan L, Freedman J, Sweeney G, Verma S. Adiponectin deficiency promotes endothelial activation and profoundly exacerbates sepsis-related mortality. Am J Physiol Endocrinol Metab. 2008;295(3):E658-64.

281. Tanguay M, Girard J, Scarsi C, Mautone G, Larouche R. Pharmacokinetics and Comparative Bioavailability of a Levothyroxine Sodium Oral Solution and Soft Capsule. Clin Pharmacol Drug Dev. 2019;8(4):521-8.

282. Okada-Iwabu M, Yamauchi T, Iwabu M, Honma T, Hamagami K, Matsuda K, Yamaguchi M, Tanabe H, Kimura-Someya T, Shirouzu M, Ogata H, Tokuyama K, Ueki K, Nagano T, Tanaka A, Yokoyama S, Kadowaki T. A small-molecule AdipoR agonist for type 2 diabetes and short life in obesity. Nature. 2013;503(7477):493-9.

283. Cayrol F, Diaz Flaque MC, Fernando T, Yang SN, Sterle HA, Bolontrade M, Amoros M, Isse B, Farias RN, Ahn H, Tian YF, Tabbo F, Singh A, Inghirami G, Cerchietti L, Cremaschi GA. Integrin alphavbeta3 acting as membrane receptor for thyroid hormones mediates angiogenesis in malignant T cells. Blood. 2015;125(5):841-51.

284. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Ponten F. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.

285. Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, Ha B, Altay G, Greenbaum JA, McVicker G, Seumois G, Rao A, Kronenberg M, Peters B, Vijayanand P. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. Cell. 2018;175(6):1701-15 e16.

286. Kozicky LK, Sly LM. Depletion and Reconstitution of Macrophages in Mice. Methods Mol Biol. 2019;1960:101-12.

287. Murray PJ, Wynn TA. Protective and pathogenic functions of macrophage subsets. Nat Rev Immunol. 2011;11(11):723-37.

288. Nahm MH, Briles DE, Yu X. Development of a multi-specificity opsonophagocytic killing assay. Vaccine. 2000;18(24):2768-71.

289. Lacoma A, Cano V, Moranta D, Regueiro V, Dominguez-Villanueva D, Laabei M, Gonzalez-Nicolau M, Ausina V, Prat C, Bengoechea JA. Investigating intracellular persistence of Staphylococcus aureus within a murine alveolar macrophage cell line. Virulence. 2017;8(8):1761-75.

290. Stackowicz J, Jonsson F, Reber LL. Mouse Models and Tools for the in vivo Study of Neutrophils. Front Immunol. 2019;10:3130.

291. Waggoner SN, Daniels KA, Welsh RM. Therapeutic depletion of natural killer cells controls persistent infection. J Virol. 2014;88(4):1953-60.

292. Ehrenkranz NJ, Elliott DF, Zarco R. Serum Bacteriostasis of Staphylococcus aureus. Infect Immun. 1971;3(5):664-70.

293. Kerschbaumer RJ, Rieger M, Volkel D, Le Roy D, Roger T, Garbaraviciene J, Boehncke WH, Mullberg J, Hoet RM, Wood CR, Antoine G, Thiele M, Savidis-Dacho H, Dockal M, Ehrlich H, Calandra T, Scheiflinger F. Neutralization of macrophage migration inhibitory factor (MIF) by fully human antibodies correlates with their specificity for the beta-sheet structure of MIF. J Biol Chem. 2012;287(10):7446-55.

294. Zhang Z, Zhang R, Li L, Zhu L, Gao S, Lu Q, Gu Y, Zhang Y, Yang H, Hou T, Zhen X, Zheng LT. Macrophage migration inhibitory factor (MIF) inhibitor, Z-590 suppresses cartilage destruction in adjuvant-induced arthritis via inhibition of macrophage inflammatory activation. Immunopharmacol Immunotoxicol. 2018;40(2):149-57.

295. Udovcic M, Pena RH, Patham B, Tabatabai L, Kansara A. Hypothyroidism and the Heart. Methodist Debakey Cardiovasc J. 2017;13(2):55-9.

296. Tanabe H, Fujii Y, Okada-Iwabu M, Iwabu M, Nakamura Y, Hosaka T, Motoyama K, Ikeda M, Wakiyama M, Terada T, Ohsawa N, Hato M, Ogasawara S, Hino T, Murata T, Iwata S, Hirata K, Kawano Y, Yamamoto M, Kimura-Someya T, Shirouzu M, Yamauchi T, Kadowaki T, Yokoyama S. Crystal structures of the human adiponectin receptors. Nature. 2015;520(7547):312-6.

297. Luo Y, Liu M. Adiponectin: a versatile player of innate immunity. J Mol Cell Biol. 2016;8(2):120-8.

298. Fournier B, Philpott DJ. Recognition of Staphylococcus aureus by the innate immune system. Clin Microbiol Rev. 2005;18(3):521-40.

299. Small CL, McCormick S, Gill N, Kugathasan K, Santosuosso M, Donaldson N, Heinrichs DE, Ashkar A, Xing Z. NK cells play a critical protective role in host defense against acute extracellular Staphylococcus aureus bacterial infection in the lung. J Immunol. 2008;180(8):5558-68.

300. Waggoner SN, Cornberg M, Selin LK, Welsh RM. Natural killer cells act as rheostats modulating antiviral T cells. Nature. 2011;481(7381):394-8.

301. Nair S, Archer GE, Tedder TF. Isolation and generation of human dendritic cells. Curr Protoc Immunol. 2012;Chapter 7:Unit7 32.

302. Meinhardt K, Kroeger I, Abendroth A, Muller S, Mackensen A, Ullrich E. Influence of NK cell magnetic bead isolation methods on phenotype and function of murine NK cells. J Immunol Methods. 2012;378(1-2):1-10.

303. Kuhns DB, Priel DAL, Chu J, Zarember KA. Isolation and Functional Analysis of Human Neutrophils. Curr Protoc Immunol. 2015;111:7 23 1-7 16.

304. Ghoneum M, Seto Y, Agrawal S. Activation of human monocyte-derived dendritic cells in vitro by Thymax, a gross thymic extract. Anticancer Res. 2009;29(11):4367-71.

305. Tognarelli S, Jacobs B, Staiger N, Ullrich E. Flow Cytometry-based Assay for the Monitoring of NK Cell Functions. J Vis Exp. 2016(116).

306. Taylor L, Brodermann MH, McCaffary D, Iqbal AJ, Greaves DR. Netrin-1 Reduces Monocyte and Macrophage Chemotaxis towards the Complement Component C5a. PLoS One. 2016;11(8):e0160685.

307. Nuzzi PA, Lokuta MA, Huttenlocher A. Analysis of neutrophil chemotaxis. Methods Mol Biol. 2007;370:23-36.

308. Tiberio L, Del Prete A, Schioppa T, Sozio F, Bosisio D, Sozzani S. Chemokine and chemotactic signals in dendritic cell migration. Cell Mol Immunol. 2018;15(4):346-52.

309. Gustafsson K, Ingelsten M, Bergqvist L, Nystrom J, Andersson B, Karlsson-Parra A. Recruitment and activation of natural killer cells in vitro by a human dendritic cell vaccine. Cancer Res. 2008;68(14):5965-71.

310. Nagl M, Kacani L, Mullauer B, Lemberger EM, Stoiber H, Sprinzl GM, Schennach H, Dierich MP. Phagocytosis and killing of bacteria by professional phagocytes and dendritic cells. Clin Diagn Lab Immunol. 2002;9(6):1165-8.

311. Hampton MB, Vissers MC, Winterbourn CC. A single assay for measuring the rates of phagocytosis and bacterial killing by neutrophils. J Leukoc Biol. 1994;55(2):147-52.

312. Haller D, Serrant P, Granato D, Schiffrin EJ, Blum S. Activation of human NK cells by staphylococci and lactobacilli requires cell contact-dependent costimulation by autologous monocytes. Clin Diagn Lab Immunol. 2002;9(3):649-57.

313. Murray RJ. Staphylococcus aureus infective endocarditis: diagnosis and management guidelines. Intern Med J. 2005;35 Suppl 2:S25-44.

314. Matsuda N, Hattori Y. Vascular biology in sepsis: pathophysiological and therapeutic significance of vascular dysfunction. J Smooth Muscle Res. 2007;43(4):117-37.

315. Goldstein BJ, Scalia RG, Ma XL. Protective vascular and myocardial effects of adiponectin. Nat Clin Pract Cardiovasc Med. 2009;6(1):27-35.

316. Zhao X, Ho D, Gao S, Hong C, Vatner DE, Vatner SF. Arterial Pressure Monitoring in Mice. Curr Protoc Mouse Biol. 2011;1:105-22.

317. da Luz LT, Nascimento B, Rizoli S. Thrombelastography (TEG(R)): practical considerations on its clinical use in trauma resuscitation. Scand J Trauma Resusc Emerg Med. 2013;21:29.

318. Reikvam H, Steien E, Hauge B, Liseth K, Hagen KG, Storkson R, Hervig T. Thrombelastography. Transfus Apher Sci. 2009;40(2):119-23.

319. Chao CH, Chen HR, Chuang YC, Yeh TM. Macrophage Migration Inhibitory Factor-Induced Autophagy Contributes to Thrombin-Triggered Endothelial Hyperpermeability in Sepsis. Shock. 2018;50(1):103-11.

320. Lidington EA, Moyes DL, McCormack AM, Rose ML. A comparison of primary endothelial cells and endothelial cell lines for studies of immune interactions. Transpl Immunol. 1999;7(4):239-46.

321. Park M, Yoon YS. Cardiac Regeneration with Human Pluripotent Stem Cell-Derived Cardiomyocytes. Korean Circ J. 2018;48(11):974-88.

322. Lin Y, Wozniak JM, Grimsey NJ, Girada S, Patwardhan A, Molinar-Inglis O, Smith TH, Lapek JD, Gonzalez DJ, Trejo J. Phosphoproteomic analysis of protease-activated receptor-1 biased signaling reveals unique modulators of endothelial barrier function. Proc Natl Acad Sci U S A. 2020;117(9):5039-48.