DISSERTATION

ENRICHMENT AND SEPARATION OF *MYCOBACTERIUM TUBERCULOSIS* EXTRACELLULAR VESICLES WITH A SIDE OF BIOSAFETY AND BIOSECURITY

Submitted by

Joan M. Ryan

Department of Microbiology, Immunology and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Fort Collins, Colorado

Spring 2022

Doctoral Committee:

Advisor: Karen M. Dobos
Co-Advisor: Nicole A. Kruh-Garcia

Robert P. Ellis
Dan Lark

ABSTRACT

ENRICHMENT AND SEPARATION OF *MYCOBACTERIUM TUBERCULOSIS* EXTRACELLULAR VESICLES WITH A SIDE OF BIOSAFETY AND BIOSECURITY

Mycobacterium tuberculosis (Mtb), the causative agent of tuberculosis (TB), is one of the world's deadliest pathogens. Recently, TB fell from the leading infectious cause of death to the second leading infectious killer due to COVID-19. Sadly, efforts to control and prevent TB have been negatively impacted by the global pandemic: 2020 saw the first year-over-year rise in cases since 2005. Mtb has been evolving with humans for several millennia. It has the ability to transition among various states of activation for long-term survival in a multitude of microenvironments. The high adaptability of Mtb has made it difficult to control and prevent TB disease, even though anti-tubercular treatments and a vaccine have been available for decades. New prevention and control strategies, including diagnostics, treatments, and vaccines, are desperately needed to combat this public health crisis. Understanding the role of extracellular vesicles in Mtb physiology and the host-pathogen interaction may unlock new avenues for combatting TB.

Extracellular vesicles (EVs) are released by virtually every living cell, serving numerous functions from biomolecule export to intercellular communication. Mtb generates EVs that are involved in response to environmental stress, virulence, and host immunomodulation. Mtb EVs have even shown promise as a highly antigenic vaccination component. Defining the role of Mtb EVs in both bacterial physiology and host-pathogen interactions is an attractive avenue for identifying new drug targets and interventions. Unfortunately, inconsistent and contradictory results in the literature have complicated a complete description of Mtb EV composition and functions. Highly reproducible enrichment techniques and a clear understanding of what biomolecules are Mtb EV associated versus co-enriched is required to fully investigate the biogenesis and roles of Mtb EVs.

This work explores Mtb EV enrichment and separation with a focus on technical and biological reproducibility. Common EV enrichment techniques including size-exclusion chromatography, ultracentrifugation, and density gradient separation are directly compared for quality and consistency of Mtb EV preparation. This comparison reveals that the most commonly used technique is also the most strongly influenced by biological variation, shedding light on a potential cause for irreproducibility seen in previous Mtb EV literature. Quality control markers for Mtb EV purity assessment are also defined.

For the first time, asymmetric flow field-flow fractionation (AF4) provides high resolution separation of Mtb EVs based on size. Proteomic exploration provides clarity of Mtb EV composition, with several proteins previously identified as EV associated cleanly separated from Mtb EV-containing AF4 fractions. Comparing AF4 with size-exclusion chromatography presents further insight into potential biases attributable to Mtb EV enrichment technique. The total proteomic variation between small and large Mtb EVs is negligible; however, differential detection of the lipoglycan lipoarabinomannan and changes in a subset of known secreted proteins suggest that Mtb EVs have potentially biologically relevant heterogeneity based on their size. This heterogeneity could be another contributing factor to inconsistencies in Mtb EV literature.

Finally, this dissertation includes a presentation of biosafety and biosecurity from the graduate student perspective. With a focus on protecting researchers, the community, and the environment, biosafety and biosecurity are essential for the advancement of life sciences. Entering this profession during COVID-19 exemplifies the dynamic nature of risk assessment and mitigation. The spotlight on these areas due to COVID-19 provides a renewed opportunity for engagement with scientists and the community to promote a culture of safe and secure scientific research.

ACKNOWLEDGEMENTS

I am astonished by the amount of support I have received throughout my graduate school journey. I would like to thank my advisor, Dr. Karen Dobos, for supporting my development as a scientist and for helping me pursue my passion: biosafety and biosecurity. My success would not have been possible without the support of my co-advisor and mentor, Dr. Nicole Kruh-Garcia. I simply cannot thank you enough for constantly believing in me, even when life hindered my progress. I also had the privilege of watching both of my advisors navigate significant professional changes. They have taught me the importance of prioritizing one's own wellbeing, even when the choice to do so seems impossible.

I would like to express my deepest appreciation to the members of my dissertation committee, Dr. Bob Ellis and Dr. Dan Lark, for providing valuable feedback throughout this process. In addition, I want to thank Dr. Angelo Izzo and Dr. Dean Crick, who were part of my preliminary examination committee.

Thank you to our collaborators, Dr. Kim Williams and Christine Plavchak, from Colorado School of Mines. I am grateful to Dr. Kitty Brown (ARC-BIO) for answering all of my proteomics questions and helping me wrangle my data. I also want to thank current and past members of the Dobos and Kruh-Garcia laboratories. I would not have succeeded without their technical guidance, friendship, and support. I give a special thanks to the students I had the privilege of teaching in the laboratory: I learned just as much (if not more) from you as you did from me.

The internship experiences I had with the Institutional Biosafety Committee and with the Biosafety Office have been invaluable. I want to thank Christy Johnson, Rebecca Moritz, Heather Blair, and Dr. Nikki Marlenee. I am exceptionally grateful for your patience, guidance, and mentorship.

Completion of my degree could not have happened without the unwavering love and support from my family. Thank you Mom, Dad, Grace, AliBeth, Jessica, Ted, Sheilagh, Teddy, Elizabeth, Mima, and Pipa for celebrating my successes and keeping me grounded in times of uncertainty. Finally, there are no words to express how grateful I am to my wife, Kelsey, without whom I would've lost myself in this process. You are the light of my life and my every day hero. I love you all of the loves.

DEDICATION

To my uncle, Christopher S. Coles

(June 10th, 1963 – October 5th, 2019)

Love,

Joanie Bear

TABLE OF CONTENTS

LIST OF TABLES

Main Text:

Appendix I:

Appendix II:

LIST OF FIGURES

Main Text:

Appendix I:

Appendix II:

Chapter 1 Literature Review and Overview of the Dissertation

1.1 History and Current State of Tuberculosis

As one of the oldest diseases known to affect humankind, Tuberculosis (TB) has been called by many other names over the course of history [1]. With the advent of polymerase chain reaction (PCR), it was proven that *Mycobacterium tuberculosis* (Mtb) DNA was present in tissues from Egyptian mummies with Pott's disease—confirming that TB was depicted in early Egyptian art [2]. Written descriptions from ancient Egypt are difficult to decipher, thus the earliest known word for tuberculosis was found in ancient Hebrew from the Old Testament: *schachepheth* [1,3]. Hippocrates used both phthisis and consumption to describe tuberculosis as fatal for many, particularly young adults in ancient Greece [1]. Scrofula, now known to have been extrapulmonary TB, plagued Europe as an affliction believed to be curable by the "royal touch," which was started by Clovis of France (481-511) and Edward the Confessor (1042-1066) [4]. This practice remained popular for centuries, fading in the early 1700s. As a disease of epidemic nature, the white plague and consumption remained part of literature and art as the world emerged from the Renaissance into the 18th and 19th centuries [5]. Johann Schonlein is credited for coining the term tuberculosis in 1834, though consumption would remain in common vernacular for decades to come [1,5].

Our understanding of TB as a pulmonary disease first evolved in the early 1800s with the invention of the stethoscope by Laennec who unearthed a plethora of knowledge regarding TB pathogenesis through patient interactions and autopsy [1]. In 1865 Jean-Antoine Villemin empirically demonstrated the transmissibility of TB as an infectious disease using rabbits [6]. Finally, on March 24th, 1882 Robert Koch presented "*Die Aetiologie der Tuberculose"* to the Berlin Physiological Society. At this meeting Koch introduced his famous postulates for demonstrating infectious disease etiology by describing the Mtb bacillus as the definitive causative agent of TB [1,7]. Koch went on to isolate tuberculin with the intention of using it as a type of treatment. Upon injecting himself and developing a

strong reaction to tuberculin, Koch concluded this material may better serve diagnostic purposes which was quickly adopted by veterinarians for cattle [1]. This delayed, type IV hypersensitivity reaction is the original basis for the purified protein derivative (PPD) test, which is a common TB surveillance tool today. Around the same time, Wilhelm Konrad von Röntgen discovered X-rays as an effective diagnostic method for TB, which is also still used as a screening tool [7].

At the end of the 19th century, tuberculosis was incredibly widespread and recognized by scientists, physicians, and politicians as a great concern. France began to identify soldiers with TB so they could be sent to health centers for treatment and education on prevention before returning to society [7]. These centers became the precursors for the public sanatorium, like the Adirondack Cottage Sanatorium opened in 1884 by Dr. Edward Trudeau in Saranac Lake, New York. Dr. Trudeau, who suffered from TB, became a prominent tuberculosis researcher and established the first US tuberculosis laboratory (now called the Trudeau Institute) in 1894 [8]. As sanitoriums emerged around the globe, so too did a push for the development of a tuberculosis vaccine. While several strategies for the development of a vaccine were being pursued in the early 1900s, it was ultimately Albert Calmette and Camille Guérin of the Pasteur Institute who generated the live-attenuated Bacille Calmette-Guérin (BCG) vaccine [1,7]. This vaccine, along with PPD testing, was used by United Nations Children's Fund (UNICEF) as part of a tuberculosis control program to combat the resurgence of TB following World War II in 1948. While it has been phased out as a routine vaccination in many areas since the 1970s, the BCG vaccine is still commonly used in many countries with high TB prevalence for children, healthcare professionals, and military members [7].

The development of clinical antibiotics and chemotherapies for tuberculosis started in the 20th century while BCG vaccination was emerging as a control strategy [7]. Streptomycin and *p*-aminosalicylic acid could be combined to treat TB patients, which was the most effective treatment regimen until isoniazid (INH) was shown to have antitubercular activity in the 1950s [7,9–11]. Several vital antitubercular drugs were developed relatively rapidly following INH, including pyrazinamide,

cycloserine, ethionamide, ethambutol, rifampicin, and others [10]. Current standard treatment of drug-sensitive infections still consists of four of these drugs: INH, rifampin, ethambutol, and pyrazinamide [11].

It has been almost 140 years since the *Mycobacterium tuberculosis* bacillus was first described, yet TB remains a public health crisis as the second leading infectious cause of death world-wide behind COVID-19 [12]. While there are diagnostic and surveillance methods and programs, treatments, and even a vaccine, this bacteria infects over 10 million people and claims the lives of 1.5 million individuals each year [12]. Although Mtb can infect anyone, it disproportionally impacts vulnerable groups like those in poverty, migrants, refugees, prisoners, and people with HIV. Drug resistance complicates the situation, with multidrug-resistant and extensively drug-resistant cases representing up to 4% of new cases annually [7,13]. Unfortunately the World Health Organization's (WHO) End TB Strategy has experienced a setback in the world of COVID-19, with TB deaths rising for the first time in 2020 in over a decade. Access to care has been severely impacted by the pandemic, and public health spending efforts for TB control have reverted to levels previously seen in 2016 [12].

1.2 Physiology of Tuberculosis Disease

Because *Mycobacterium tuberculosis* has been evolving with humans for thousands of years, it is extremely well adapted to the host environment. In fact this organism has no other environmental reservoir, resulting in a complex balance between being pathogenic and inconspicuous [14]. Mtb is a slow-growing facultative intracellular pathogen. Typically, the bacillus has a long doubling time of 16-22 hours, with the fastest reported doubling time of 14.7 hours in highly favorable culture conditions [15]. As a member of the Actinomycete family and the *Mycobacterium* genus, Mtb has a complex cell envelope comprised of lipids (including many that are unique to the genus), proteins, and carbohydrates [16,17]. This waxy, hydrophobic layer allows the bacilli to resist degradation by host enzymes and confers highly selective permeability keeping many toxic molecules, like most antibiotics, out of the cell

[17]. Additionally, the mycobacterial cell wall resists decolorization by acids used in many staining procedures, a characteristic known as acid-fastness [18].

Typically, a person becomes exposed to Mtb upon inhalation of droplets expelled when a pulmonary TB infected individual coughs [10,18]. Once inside the respiratory tract of a susceptible person, there are three main clinical outcomes that can occur. The first, and most likely outcome in a healthy individual, is the bacteria do not or cannot establish an infection because the immune system swiftly eradicates it. If infection is established, the second most likely outcome is latent tuberculosis infection (LTBI), whereby the body mounts an immune response to control or sequester the bacteria in the lung but cannot completely eliminate it. LTBI can be a life-long condition that goes undetected. The final outcome, which comprises only about 10% of cases, is the development of sub-clinical or active pulmonary TB disease. These states can be reached after initial infection or through activation of LTBI [7,19,20]. There are extrapulmonary manifestations of active TB, which can involve almost any organ system including the blood vessels, lymph nodes, bones, gastrointestinal tract, and central nervous system. In many of these cases, pulmonary based diagnostics are ineffective because the bacteria do not maintain residence in the lungs [21]. The varying spectrum of TB disease contributes to the difficulty in public health measures to eradicate it, with nearly a quarter of the world population harboring the bacteria, many of whom are completely asymptomatic [12].

Because Mtb must tolerate many types of microenvironments based on conditions in the host, there is vast heterogeneity in Mtb phenotype [20]. The bacterium is capable of entering and emerging from a non-replicative state within host tissues. Following inhalation, Mtb is translocated to the lower lung where it enters the alveoli. Here it encounters the dominant cell type that the bacteria infects: alveolar macrophages [22]. The macrophages internalize Mtb by receptor-mediated phagocytosis and attempt to eliminate the pathogen via lysosomal degradation. Once inside the phagosome, however, Mtb prevents phagosome acidification and actively blocks phago-lysosomal fusion, inhibiting progression of this pathogen elimination pathway [23]. The bacteria then break down the phagosome membrane to

release Mtb products, sometimes even whole bacilli, into the macrophage cytosol [24]. This translocation into the cytosol triggers a type I interferon (IFN) host response. This response causes recruitment additional phagocytic cells that can take up released bacteria, resulting in bacterial population expansion [22]. While residing in phagocytic cells, Mtb also employs several mechanisms to prolong the survival of the host cell by interfering with apoptosis and necrotic cell death to ensure sufficient time for bacterial replication [25].

After infecting alveolar macrophages Mtb accesses the lung parenchyma by either directly infecting alveolar epithelium or by inducing migration of the alveolar macrophage to the interstitial tissue [26]. Other early innate immune cells like neutrophils, inflammatory monocytes, and dendritic cells are recruited and become infected during bacterial expansion. Throughout this innate process, the onset of adaptive immunity is delayed, a hallmark of Mtb disease kinetics. Ultimately Mtb is transported by dendritic cells or inflammatory monocytes to pulmonary lymph nodes where T cell priming begins [27]. Initiation of the adaptive immune response leads to the recruitment of T cells and B cells, resulting in the hallmark granuloma formation. Onset of adaptive immunity typically arrests Mtb progression beyond the granuloma but does not eliminate the bacteria, resulting in latent infection where the person is asymptomatic and does not shed bacteria [22]. If the bacterial load in the granuloma becomes too great, either upon initial infection or even decades following persistence as LTBI, the infection can breach the granuloma barrier and bacteria disseminate. At this stage, Mtb exists extracellularly in necrotic lung cavities, which allows for transmission out of the host via aerosol [25,28]. What causes the change from latent to active infection remains an active area of investigation, with a variety of factors from genetics to the environment at play. Without treatment, approximately 50 - 70% of those with active TB succumb to the infection within ten years [12,26].

1.3 Diagnosis, Treatment, and Prevention of Tuberculosis

Tuberculosis is diagnosed using a combination of molecular detection methods, immunological assays, and direct observations via microscopy and bacterial culture [12,29]. Because the spectrum of

disease is broad, there is no one-size-fits-all technique to capture every state of TB infection [26,30]. Additionally, individual demographics and health status influence the effectiveness of several gold-standard methods for diagnosis [29,30]. **Table 1.1** highlights the most common TB diagnostic tools, including their advantages and limitations. Because many TB laden areas are under resourced, emerging techniques are limited in access based on the necessity for steady electricity, laboratory resources, and well-trained personnel [30]. Therefore, timely and effective diagnosis remains a core challenge and high priority in the End TB Strategy [12].

Table 1.1 Common diagnostic and screening tools for TB.

Diagnostic or Screening Tool	Advantages	Limitations
Chest X-ray	• Accessible • Non-invasive • Inexpensive (upfront costs offset by sending fewer samples for follow-up molecular test)	• Low specificity & sensitivity • Screening tool only • Limited to pulmonary TB
Sputum-smear microscopy	• Accessible • Rapid • Inexpensive	• Low sensitivity • Prone to human error • Limited to pulmonary TB in immunocompetent adults
Bacterial culture (conventional)	• High sensitivity & specificity • Drug sensitivity testing can occur simultaneously with drug diffusion discs	• Very slow (>28 days) • Requires high containment facility
Mycobacterial growth indicator tube (MGIT) culture	• High sensitivity & specificity • Faster than conventional culture • Simultaneous drug sensitivity testing	• Requires high containment facility • More expensive than conventional culture
GeneXpert MTB/RIF	• Rapid • Rifampin resistance tested simultaneously • High specificity	• Variable sensitivity, particularly for immunocompromised individuals and non-respiratory samples • Limited (but growing) accessibility
Line probe assay	• High specificity, moderate sensitivity • Detects MDR/XDR • Rapid	• Not a replacement for conventional phenotypic drug susceptibility testing • Significant infrastructure required
Lipoarabinomannan (LAM) antigen test	• Non-invasive • Rapid • Detects active TB in HIV+ individuals	• Not recommended for immunocompetent individuals • Variable sensitivity
Tuberculin skin test	• Can detect LTBI • Inexpensive	• Variable sensitivity and low specificity • Requires follow up visit for results
QuantiFERON-TB Gold	• Rapid • Can detect LTBI	• Expensive • Blood handling required

As previously mentioned, standard treatment for drug-susceptible active TB is a six-month course of four first-line drugs: isoniazid, rifampicin, ethambutol, and pyrazinamide. This standard course of oral medications costs approximately $40 USD per person, if supplied by the Global TB Drug Facility, and has an 85% success rate [12]. Those diagnosed with LTBI can be prophylactically treated with weekly or daily doses of different combinations of isoniazid, rifampicin, or rifapentine for three to nine months [12]. Although these treatments are fairly straightforward, patient access and compliance remain a challenge. Adverse reactions are common, with up to 75% of individuals who receive rifampicin reporting a range of problems from gastrointestinal distress to peripheral neuropathy and even drug-induced hepatitis [31]. WHO recommends directly observed therapy-short course (DOTS) to reduce noncompliance; however, that can be impossible for people living far from treatment locations [12].

Beyond access and compliance issues with antibiotic susceptible TB diagnosis and treatment, drug-resistance hinders TB control efforts with single, multi-, and extremely drug resistant cases continuing to rise [12]. Ideally, all individuals confirmed to have TB would receive further testing for drug resistance immediately. In reality these tests are not widely accessible and can take months, resulting in the person being on the wrong treatment course for an extended period of time [29]. This extends the likelihood of transmission of drug-resistant bacteria to others. Treatment for drug-resistant TB is over twenty-five times as expensive as the standard treatment and often includes injectable medications with even more unpleasant and debilitating side effects [12,32].

The live-inactivated BCG vaccine first used in humans one century ago in 1921 remains the only licensed vaccine for TB [1,7,12]. WHO recommends vaccination for all newborn infants in areas where TB incidence is high or if the child is going to be continually exposed to a person with TB that does not respond to treatment [12]. The BCG vaccine is effective for preventing severe forms of TB in children; however, the efficacy wanes significantly over time and there is no vaccine available to prevent TB disease in adults [33]. The BCG vaccine is not just one formulation either, with a variety of different *M. bovis* BCG seed strains used for manufacturing under distinct conditions across the world. Differences

based on the actual vaccine administered have been demonstrated to impact efficacy and innate immune activation [34]. Interestingly, numerous studies have demonstrated additional non-specific protection against other infections—including other respiratory infections, sepsis, and even malaria—in children [35,36]. Because BCG vaccination is extremely safe and does have beneficial effects in children, it remains the reference standard for other TB vaccine candidates [33]. Fortunately, many TB prevention candidates for a variety of demographics are currently in the pipeline as illustrated in **Figure 1.1**.

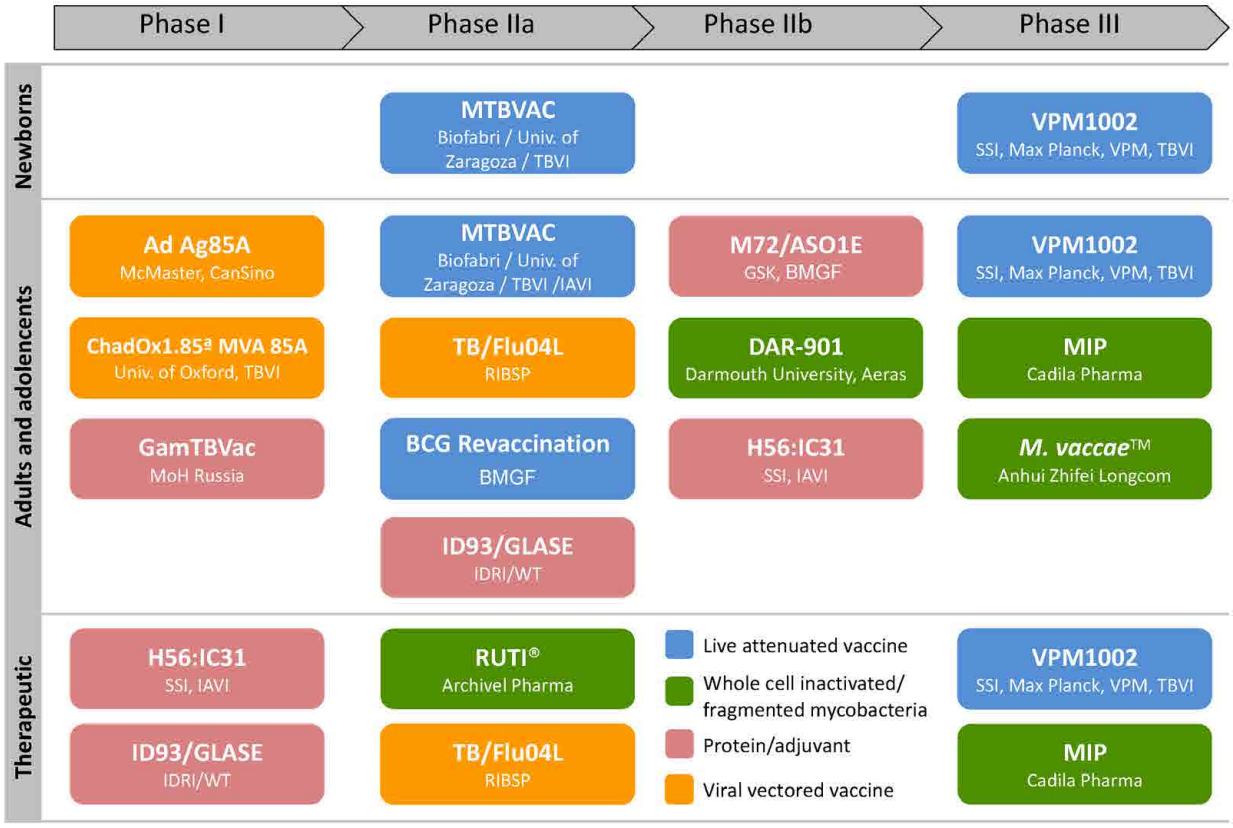


Figure 1.1 Current TB vaccine candidates in the pipeline. There are currently fourteen vaccine candidates across the phases for human clinical trials, colored according to vaccine strategy. This figure is adapted from [33]; the final version is free according to the Creative Commons CC-BY license.

Unfortunately, COVID-19 has had negative impacts on efforts to ensure diagnosis and treatment of TB. Not only has funding for TB control strategies declined, reduced access to both diagnosis and treatment have influenced an increase in TB deaths. This is expected to continue through 2022, with the full ramifications of COVID-19 related issues on TB control not yet realized. The WHO defines the immediate priority as restoration of "access to provisional and essential TB services such that the levels of

TB case detection and treatment can recover to at least 2019 levels, especially in the most badly-affected countries" [12]. Even before COVID-19 the world needed new strategies for TB detection, treatment, and prevention. With the incredible rate that COVID-19 vaccines were developed, it is even more apparent that wide-spread political commitment, funding, and research are essential for combatting the tuberculosis public health crisis.

1.4 Mtb Protein Secretion

Because the host-pathogen interaction during *Mycobacterium tuberculosis* infection is so complex, improving our understanding of how the bacterium interacts with its environment will help push forward research contributing to future diagnostic, treatment, and prevention strategies. One way that Mtb interacts with its environment is through protein secretion. The importance of the culture filtrate as a major source of antigens was first hinted at in 1932 in a search to understand what component of tuberculin causes the specific skin reaction observed in infected animals [37]. Since then the culture filtrate proteins (CFP) of Mtb have been studied for understanding of bacterial physiology, pathogen-host interactions, and potential vaccine components.

Complete genome sequencing of Mtb H37Rv in 1998 allowed for genome-based proteome analysis, which indicated that at least 257 proteins were potentially secreted [38,39]. These early predictions were confirmed and even expanded upon by numerous experiments. There are several active protein transport systems in Mtb, including the Sec-dependent general secretory pathway, the twin arginine translocation (TAT) pathway, and the type VII secretion systems (T7S) [40]. Each of these systems contribute to mycobacterial protein secretion.

The Sec secretion pathway is a highly conserved mechanism of protein export found in all classes of bacteria and across the three domains of life [41]. The initial genome analysis of Mtb H37Rv identified two different *secA* genes, *secA1* and *secA2*, as well as approximately 90 genes for preproteins with type I signal peptides (20 amino acid hydrophobic signal sequence at the N-terminus) that are expected to be transported via a Sec-dependent pathway [38]. These numbers have since been broadly expanded and the

Mtb Sec system is the primary route for protein export [42]. The Sec pathway translocates proteins in their unfolded state using a protein targeting component, motor protein, and translocase channel [41]. How the proteins get from the periplasm through the outer membrane following signal peptide cleavage is still not well understood for Mtb, but may involve the outer membrane protein CpnT (Rv3903c) [40]. Not surprisingly, the SecA1 protein Rv3240c and several other proteins that make up the system (SecD, E1, F, and Y) are essential for Mtb viability [40]. Interestingly, SecA2 is not essential for survival, but the mutants have an attenuated virulence in mice [43]. This accessory system is also implicated in the secretion of several proteins that do not contain the type I signal sequence such as SodA [43]. Due to the essentiality of SecA1 transport and residual export of SecA2 substrates even when the system is knocked out, there is not a complete list of Mtb SecA1 and SecA2 transported proteins [44].

Like the Sec system, the TAT pathway is conserved across all three domains of life and is essential for Mtb viability [40,41]. Proteins exported by the TAT pathway also have a specific N-terminal hydrophobic signal peptide that contains a pair of twin arginine residues (RR), but these proteins are already folded prior to transport [40]. The TAT system includes three integral membrane proteins, TatA, TatB, and TatC, with the larger TatC recognizing the twin arginine motif [45]. While there are around 90 predicted TAT secreted proteins, less than 20 have been confirmed experimentally [46]. The most well-known proteins confirmed to have the TAT signal sequence are those of the antigen 85 (Ag85) complex which is involved in biogenesis of cell wall components essential for maintaining structural integrity [40].

The T7S pathway was first defined in Mtb and has subsequently been observed throughout mycobacteria and the phylum *Actinobacteria* [45]. There are five T7S systems encoded in the Mtb genome. They are named ESX-1 through ESX-5 based on the discovery of ESX-1, which is responsible for secretion of EsxA and EsxB (also known as ESAT-6 and CFP-10 respectively) [45,47–49]. The ESX-2 and ESX-4 are known to be non-essential for virulence or *in vitro growth* of Mtb, but beyond this their role and their substrates remain obscure [45,50,51]. Although ESX-1 is not essential for growth in culture, this system has been intensely studied for more than a decade due to its role in virulence and

pathogenicity [40,45]. One of the key differences between *M. bovis* BCG and the pathogenic *M. bovis* and Mtb is in the region of difference 1 (RD1), which encodes for components of the ESX-1 system. The attenuated BCG strain lacks EsxA (ESAT-6), which contributes significantly to its lack of virulence [52,53]. ESX-1 is also largely responsible for Mtb intracellular survival, playing a crucial role in the lysis of the phagosomal membrane [24,46]. The ESX-3 system is one of the most conserved T7S, essential for viability due to its role in metal-ion homeostasis for iron and zinc [54]. Additionally, several of its substrates are involved in immune modulation with EsxH being required for inhibition of the phagosome maturation process [55]. Several (but not all) of the genes encoding the ESX-5 system are essential for Mtb viability; however, this system is only present in a subgroup of slow-growing mycobacteria, which happens to include most of the pathogenic species [46]. The ESX-5 system is most notably important for contribution to virulence, with the ESX-5 encoded PE/PPE proteins (Pro-Glu (PE) or Pro-Pro-Glu (PPE) N-terminal motif) highly immunogenic for T-cells [56]. Ultimately, the T7S of Mtb contribute to an abundance of secreted proteins that function in a myriad of essential functions for pathogenicity, yet there is not a good understanding of how these substrates cross the outer membrane [45].

Completion of the Mtb genome and technological advances in protein identification, particularly mass spectrometry, were instrumental in progressing the characterization of Mtb CFP, which began almost a century ago [37,40]. The impact of the environment and physiological state of the bacteria on CFP composition has been under investigation for several decades and these studies have advanced understanding of Mtb physiology both *in vitro* and *in vivo* [57,58]. Not only does growth stage and nutrient availability impact the profile of CFP, specific strains generate different CFP based on antibiotic exposure and resistance also changes the secretome [59–62]. **Table 1.2** highlights Mtb proteins from CFP (and other cellular fractions) that are of particular interest throughout this dissertation. Recently, the discovery of extracellular vesicle release by Mtb and other mycobacteria has presented a new area for characterization and understanding of bacterial physiology and infection through protein and other antigen secretion [63,64]. A visual summary of Mtb protein export pathways is provided in **Figure 1.2**.

Table 1.2 Mtb proteins of note in this dissertation. The Go Cellular Components codes are the following: Cyt = cytosol, PM = plasma membrane, Sec = secreted/extracellular, CS = cell surface, and CW = cell wall.

Protein	Molecular Weight	GO Cellular Component(s)	Mycobrowser Function(s)	Mycobrowser Functional Category
Ag85A/FbpA (Rv3804c)	36 kDa	CW, PM, Sec, CS	Secreted antigen 85-B, mycolytransferase	Lipid metabolism
Ag85B/FbpB (Rv1886c)	35 kDa	CW, PM, Sec	Secreted antigen 85-A, mycolytransferase	Lipid metabolism
Ag85C/FbpC (Rv0129c)	37 kDa	CW, Sec	Secreted antigen 85-C, mycolytransferase	Lipid metabolism
ArcA (Rv1001)	43 kDa	Cyt, PM	Probable arginine deiminase	Intermediary metabolism and respiration
BfrB (Rv3841)	20 kDa	Cyt, CW, PM, Sec	Bacterioferritin, iron storage	Intermediary metabolism and respiration
Cfp2/Mtb12 (Rv2376c)	17 kDa	CW, Sec	Low molecular weight antigen	Cell wall and cell processes
CpnT (Rv3903c)	88 kDa	Cyt, CW, Sec	Hypothetical alanine and proline rich protein	Conserved hypotheticals
Eis (Rv2416c)	44 kDa	Cyt, Sec	Acetylation, substrate unknown	Virulence, detoxification, adaptation
DnaK/HSP70 (Rv0350)	67 kDa	Cyt, PM, CS, CS, Sec	Probable chaperone protein, heat-shock response	Virulence, detoxification, adaptation
EsxA/ESAT-6 (Rv3875)	10 kDa	Cyt, CW, PM, Sec	6 kDa early secretory antigenic target	Cell wall and cell processes
EsxB/CFP-10 (Rv3874)	11 kDa	Cyt, CW, PM, Sec	10 kDa culture filtrate antigen	Cell wall and cell processes
EsxN (Rv1793)	10 kDa	PM, Sec	Putative ESAT-6 like protein	Cell wall and cell processes
GlcB (Rv1837c)	80 kDa	Cyt, PM, CS, CS, Sec	Malate synthase G, glyoxylate bypass	Intermediary metabolism and respiration
GlnA1 (Rv2220)	54 kDa	Cyt, CW, PM, Sec	Glutamine synthetase	Intermediary metabolism and respiration
GroEL2 (Rv0440)	57 kDa	Cyt, CW, PM, Sec	60 kDa chaperonin 2	Virulence, detoxification, adaptation
GroES (Rv3418c)	11 kDa	Cyt, PM, Sec	10 kDa chaperonin	Virulence, detoxification, adaptation
HspX (Rv2031c)	16 kDa	Cyt, CW, PM, Sec	Heat shock protein	Virulence, detoxification, adaptation
IniA (Rv0342)	70 kDa	CW, PM	Isoniazid inducible gene protein	Cell wall and cell processes
LpqH (Rv3763)	15 kDa	CW, PM, Sec, CS	19 kDa lipoprotein antigen precursor	Cell wall and cell processes
LprA (Rv1270c)	25 kDa	CW, PM, Sec	Possible lipoprotein	Cell wall and cell processes
LprG (Rv1411c)	25 kDa	Cyt, CW, PM, Sec	Conserved lipoprotein	Cell wall and cell processes
Mpt32/Apa (Rv1860)	33 kDa	PM, Sec	Alanine and proline rich secreted protein	Cell wall and cell processes
Mpt64 (Rv1980c)	25 kDa	CW, Sec	Immunogenic protein Mpt64	Cell wall and cell processes

Protein	Molecular Weight	GO Cellular Component(s)	Mycobrowser Function(s)	Mycobrowser Functional Category
PhoS1/PstS1 (Rv0934)	38 kDa	Cyt, CW, PM, Sec	Periplasmic phosphate-binding lipoprotein	Cell wall and cell processes
PPE41 (Rv2430c)	22 kda	CS, Sec	PPE family protein	PE/PPE
PstA1 (Rv0930)	33 kDa	PM	Probable phosphate-transport integral membrane ABC transporter	Cell wall and cell processes
Rv3371	49 kDa	PM	Possible triacylglycerol synthase	Lipid metabolism
SecA1 (Rv3240c)	106 kDa	Cyt, CW, PM	Probable preprotein translocase	Cell wall and cell processes
SecA2 (Rv1821)	89 kDa	Cyt, CW, PM	Possible preprotein translocase ATPase	Cell wall and cell processes
SecD (Rv2587c)	60 kDa	Cyt, CW, PM	Probable protein-export membrane protein	Cell wall and cell processes
SecE1 (Rv0638)	17 kDa	CW, PM	Probable preprotein translocase	Cell wall and cell processes
SecF (Rv2586c)	47 kDa	CW, PM	Probable protein-export membrane protein	Cell wall and cell processes
SecY (Rv0732)	48 kDa	PM, Sec	Probable preprotein translocase	Cell wall and cell processes
SodA/SodB (Rv3846)	23 kDa	Cyt, PM, Sec	Superoxide dismutase	Virulence, detoxification, adaptation
SodC (Rv0432)	24 kDa	CW, PM, Sec	Periplasmic superoxide dismutase	Virulence, detoxification, adaptation
TatA (Rv2094c)	9 kDa	PM	Sec-independent protein translocase	Cell wall and cell processes
TatB (Rv1224)	14 kDa	PM	Probable protein	Cell wall and cell processes
TatC (Rv2093c)	34 kDa	Cyt, PM	Sec-independent protein translocase transmembrane protein	Cell wall and cell processes
TlyA (Rv1694)	28 kDa	Cyt, CW, Sec	2'-O-methyltransferase, methylates 16S and 23S rRNA, also has contact-dependent hemolytic activity	Virulence, detoxification, adaptation
VirR (Rv0431)	17 kDa	CW	Putative tuberculin related protein	Cell wall and cell processes

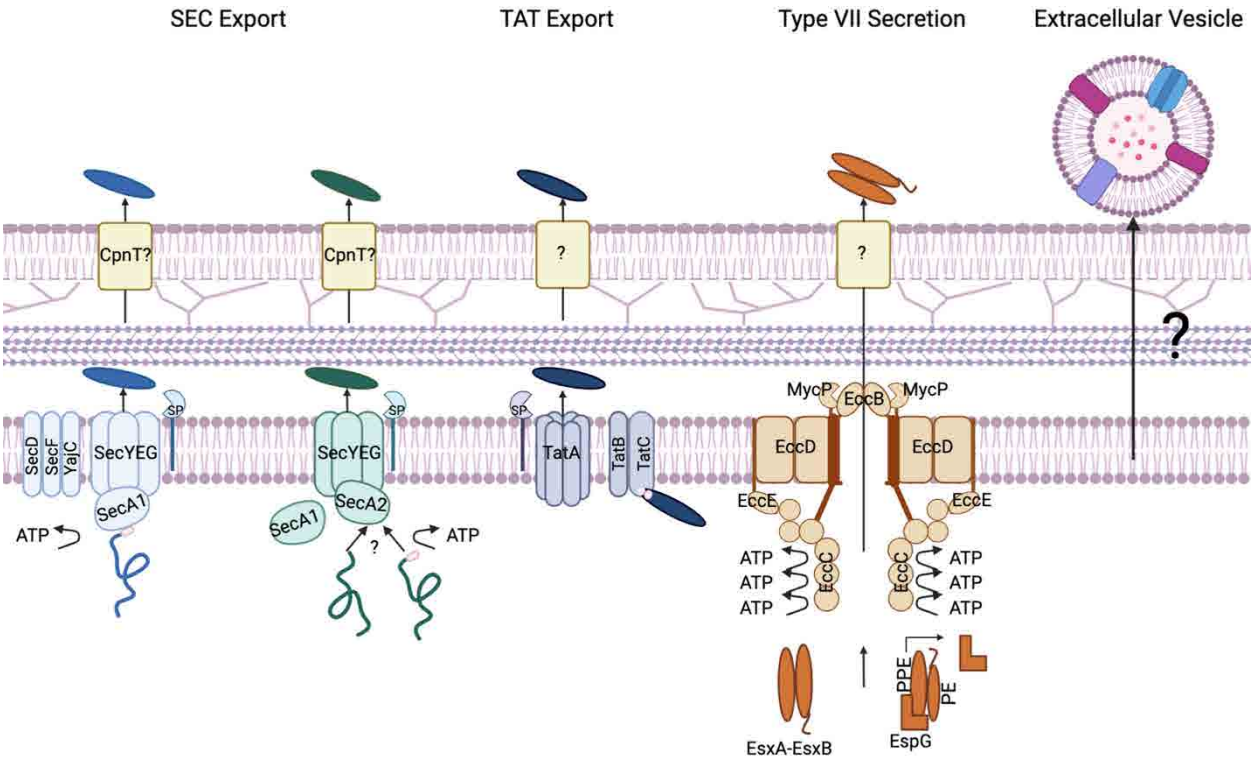


Figure 1.2 Mtb protein export. General depictions of both SEC transport systems, TAT export, type VII secretion, and extracellular vesicles. SP = signal peptidase, ATP = adenosine triphosphate, elements not drawn to scale. Created with BioRender.com

1.5 Mycobacterial Extracellular Vesicles

Extracellular vesicles (EVs) are secreted by nearly every cell and organism across all three domains of life [65]. In 1946, Chargaff and West reported the first observation of EVs—then referred to as pro-coagulant platelet-derived particles—which became known as "platelet dust" by Wolf in 1967 [66,67]. Anderson described matrix vesicles during bone calcification in 1969, and in the following decade several other mammalian cell types were noted to produce these structures [68]. In 1983 the first ultrastructural study demonstrating the release of vesicles by multi-vesicular bodies confirmed that production of some types of vesicles is a regulated phenomenon [69,70]. Fast forward to today, and we now know that mammals, plants, parasites, arthropods, fungi, archaea, and bacteria produce EVs.

EVs are membrane bound structures ranging from 20-1000 nm in diameter. While specific contents are influenced by the cell of origin and its physiologic state, EVs carry protein, lipids, carbohydrates, nucleic acids, and metabolites [65]. These vesicles have the capacity to transfer

biomolecules between cells, influencing physiological function in recipient cells [71]. The contents, size, and membrane composition are heterogeneous and change depending on the state of the cellular source and environmental conditions. Different classification systems have arisen in order to classify EVs based on things like cellular origin, size and morphology, biogenesis, and function [72].

Outer membrane vesicles (OMV), a subset of EVs specific to Gram-negative bacteria, have been a popular field of study since 1966 [73]. In contrast, the production of EVs by Gram-positive and mycobacteria has just recently gained attention. Because Gram-positive and mycobacteria differ in their surface composition, it was surprising to discover that vesicles can be released despite their complex cell wall structure and today the mechanisms of their EV biogenesis have yet to be completely elucidated [74]. The first report of extracellular vesicle production by mycobacteria was published in 2007, in which *M. ulcerans* biofilm formation and its role in transmission were described. Incidentally, scanning electron microscopy (SEM) imaging revealed vesicle association with the biofilm. This observation led to the first purification of mycobacterial EVs and immunoprecipitation of these EVs from infected mouse tail tissues [63]. Not long after this initial publication, production of EVs by many other environmental and pathogenic mycobacterial species, including *Mycobacterium tuberculosis*, was demonstrated [64]. Additional studies have confirmed EV production in other mycobacterial species, including the pathogenic *Mycobacterium avium* subspecies *paratuberculosis* (MAP) and *Mycobacterium avium* subspecies *hominissuis* (MAH) [75,76].

In their 2011 exploration of mycobacterial EVs, Prados-Rosales et al. began to describe the composition and role of Mtb EVs [64]. To show that production was occurring from live cells, C-acetate radiolabeled *Mycobacterium bovis* BCG were sub-cultured and given five days to replicate before vesicle isolation. The C-acetate label was detected only in vesicles from live cells and not in vesicle-like aggregates formed from heat killed cells. In addition to visualizing vesicles associated with the bacilli of Mtb H37Rv and BCG, they looked at a variety of mycobacteria and found that slow and fast growers, virulent, avirulent, pathogenic, and environmental species produce EVs that are roughly the same size

(20-300 nm) and exhibit a predominantly closed membrane morphology with both unilaminar and bilayered structures. Using 2D thin layer chromatography (TLC) and matrix-assisted laser desorption/ionization (MALDI) time of flight (TOF) mass spectrometry (MS), the group investigated the lipid composition of Mtb and BCG EVs. Polar lipids including phosphatidylinositol, phosphatidylinositol mannosides (PIMs), phosphatidylethanolamine (PE), and cardiolipin, accompanied by the lipoglycan Lipoarabinomannan (LAM), but not α-glucan, were found to make up the EV membrane. This points to the inner cytoplasmic membrane as the origin of mycobacterial EVs [64]. Contrastingly, a later study found that EVs from *M. avium* (MAH) contain lipids from the outer layers of the cell wall [75]. A comprehensive comparison of mycobacterial EV lipid composition across species, strains, and culture conditions would help clarify this contradiction.

Proteomic analysis of Mtb EVs consistently demonstrates lipoprotein enrichment, with LpqH and LprG strongly represented. Lipoproteins made up to 8% of the identified proteins in one study, while they comprise only 1-2% of proteins in the whole Mtb genome [77]. Functional categorization of EV proteins interrogated by mass spectrometry identify cell wall, membrane function, and intermediate metabolism and respiration as the predominant protein functions. Additionally, many identified proteins contribute to host-pathogen interactions, with several known toll-like receptor 2 (TLR-2) ligands consistently enriched [64,77]. Transmission electron microscopy (TEM) of immunogold antibody labeled EVs confirmed the presence of LpqH and LprG in the membrane of Mtb EVs, with DnaK in the lumen. LAM was also seen in these preparations in association with the EVs [64].

To date there has only been one mention of nucleic acid in Mtb EVs, and it was unpublished data indicating that DNA was detected in purified, intact Mtb EVs [64]. In MAH EVs, double stranded DNA has been reported in the lumen and on the surface with the majority present on the surface [75]. There was not enough DNA material to warrant sequencing based on the goals of this study. When comparing RNA content between EVs produced by infected macrophage cells and those from Mtb cell culture, very little

RNA was detected in the bacterial EVs [78]. Specific profiles of carbohydrates, metabolites, and nucleic acids from mycobacterial EVs have yet to be defined.

Defining the function of mycobacterial EVs remains an active area of investigation. The first reported role of virulence mediation was identified in *M. ulcerans* EVs. Lipid analysis revealed the presence of mycolactone in the vesicles. When applied to mouse bone marrow-derived macrophages (BMM) and COS cells (non-phagocytic fibroblast-like cell line from monkey kidney tissues), the vesicles were more cytotoxic than an equal amount of purified mycolactone [63]. *M. smegmatis* EVs carry TlyA, a protein previously shown to have hemolytic activity, which remains functional and active in the vesicles and is also produced by Mtb [79]. EVs allow for the transport of molecules in a protected manner, resulting in concentrated amounts of virulence factors or other important molecules, overcoming problems related to solubility or dilution through diffusion.

Mtb EVs also play a role in nutrient acquisition—specifically the uptake of iron, which is a scarce resource in the host. In normal iron conditions, Mtb EVs carry the siderophore carboxymycobactin which has a fairly low iron binding capacity. However, Mtb cultured in limited iron environments produced EVs with mycobactin, a very efficient iron binding molecule. The EVs produced in low iron can rescue an Mtb mutant incapable of making siderophores like mycobactin grown in low iron. The mycobactin rich EVs can even rescue wild type Mtb from severe iron deprivation [80]. The protein composition of MAH vesicles produced in minimal medium verses a metal mix medium (meant to mimic the phagosome) varies, with 52 proteins in common, 211 specific to the metal mix and 150 specific to the minimal medium. This further confirms that mycobacterial EVs are involved in response to nutrient availability in the environment [75].

Immunomodulation is a more complex function attributed to mycobacterial EVs. As previously mentioned, mycobacterial EVs from pathogenic species carry a variety of TLR-2 agonists including LpqH, LprG, LprA, PhoS1, and LAM among others [64]. TLRs recognize pathogen-associated molecular patterns (PAMPs) and initiate signal transduction pathways that regulate the expression of cytokines,

chemokines, and type I IFNs for innate and adaptive immune activation. TLR-2 signal transduction cascades in macrophages result in the induction of a CD4+ T cell response [81]. Prolonged TLR-2 activation, however, is immunosuppressive, leading to the production of immunosuppressive cytokines like IL-10 and ultimately inhibition of MHC-II antigen presentation [82]. The initial induction of TLR-2 starts CD4+ T cell activation, which promotes granuloma formation, but is not enough to eliminate the bacteria as the effects of prolonged TLR-2 signaling come in to play. Mtb TLR-2 agonists, including LpqH, LprG, LprA, LAM, and others, were shown to cause prolonged TLR-2 signaling [83–86]. This effect is also seen when cells are exposed to Mtb EVs both *in vitro* and *in vivo*, demonstrating that vesicles serve as a concentrated delivery system for these agonists providing some insight into how the prolonged signaling occurs [64].

Mtb EVs also influence immune cells through mechanisms other than TLR-2 signaling in macrophages. Direct exposure to Mtb EVs inhibits T cell activation, demonstrated by reduced IL-2 production and reduced T cell proliferation [87]. Mtb EVs induce partial anergy during primary stimulation of naïve T cells and they inhibit effector T cells in a transient manner. Temporary inhibition of effector T cells may facilitate local inhibition of Th1 responses at the site of infection without causing a systemic inhibition of T cell response [87]. In contrast to limiting antigen presentation in macrophages, exposure of mouse dendritic cells (DC) to Mtb EVs induced MHC-I, MHC-II, and CD86 expression, resulting in DC maturation and antigen presentation to Ag85B-specific CD4+ T cells [88]. Although it is clear that Mtb EVs have multiple functions, clarifying conflicting roles in the host-pathogen interaction remains an important avenue of inquiry.

Arguably the biggest mysteries remaining in mycobacterial EV biology are genetic regulation and specific mechanisms for biogenesis and release. As previously mentioned, early studies of Mtb and BCG EV lipid composition point to the cell membrane as the origin point for mycobacterial EVs [64]. In fact, it is specifically mentioned in a subsequent paper from the same group that no lipids from the outer membrane were detected in Mtb EVs, regardless of the concentration of iron in the media [80]. They

stated that they had no hypothesis for why or how that is, and that understanding vesicle synthesis may explain the absence. Contrastingly, outer membrane lipids were detected in EVs from MAH [75]. It is possible that individual species have different mechanisms of EV biogenesis. Another option is the production of different EV subpopulations produced by all species with various lipid compositions and certain conditions, like the culturing process or EV enrichment method, influence what is detected.

Mycobacterial EV production kinetics and composition vary based on the status of the cell of origin, indicating that biogenesis of EVs is a regulated process. Mycobacterial EV release kinetics during culture change based on the species, with *M. bovis* and Mtb having similar production patterns [76]. MAP is notably different with lower EV production, a lack of unilamellar vesicles, and slightly larger diameter. Under iron limitation, Mtb produces significantly more EVs (of the same size and morphology, but differing siderophore content) than in an iron-rich environment [80]. Similarly, MAH EVs vary in protein composition based on the metal nutrients present in culture [75]. Exposure to sub-inhibitory concentrations of INH triggers an increase in Mtb EV production [89]. Approximately 10% of the time when *M. smegmatis* divides, it happens asymmetrically resulting in very short (< 2 μm) and very long (> 7 μm) cells. This occurs regardless of the growth medium, indicating this is part of the normal population dynamics. Interestingly, these short cells contain more lipids and have a higher density of EVs on the surface compared to the normal and very long cells when imaged by SEM and TEM. The short cells are more sensitive to INH, rifampicin, H_2O_2, and acidification [90]. Perhaps the increase in EV production on these cells is related to their increased sensitivity to the environment.

The first experiments to reveal potential mechanisms of mycobacterial EV biogenesis involved a transposon mutant of Mtb H37Rv with a null allele insertion for *rv0431* that was successfully complemented back, resulting in overexpression [91]. This mutant was previously shown to grow normally in culture, but exhibited attenuation in mice [92]. Deficiency in this protein (now named VirR for “vesiculogenesis and immune response regulator”) results in similar protein released in CFP and normal cell wall integrity, but a significant increase in lipid release in CFP and EV production per cell

than the wild type (WT) [91]. The VirR mutant was not more susceptible to acidification, oxidative stress or nitrosative stress, and did not have changes in the other mycobacterial protein export pathways (SecA1, SecA2, TAT or T7S). Interestingly, there was more LpqH and SodC in the EVs of the mutant strain, which did go away upon complementation. Treatment of BMM with these EVs resulted in more TNF- α and IL-12 p40 production compared to treatment with WT EVs. Overall, this study indicates that VirR regulates EV generation and dampens the activation of macrophages, interfering with the host's ability to control Mtb proliferation, which is at least in part mediated by TLR-2. It is also likely that VirR is part of a higher-order complex with at least three likely binding partners (Rv1488, Rv0383c, and LpqH) [91].

Another gene seemingly involved in Mtb EV biogenesis is *rv3371*, which encodes for a triacylglycerol synthase (TGS) [93]. This gene plays a role in the deliberate metabolic slowdown of Mtb after establishing infection in the host [94]. Expression of *rv3371* in *M. smegmatis*, which does not naturally have it, increases the triacylglycerol levels, changes the surface from dry and rough to smooth and wet, increases the presence of bud-like structures on the surface, and results in shorter cells [93]. Together, these changes suggest that this TGS functions in the cell wall. Expression of *rv3371* in Mtb H37Rv increases during late log phase and through stationary phase. It also increases under hypoxia, nutrient starvation, and nitrosative and oxidative stress. All of these conditions are phases of slow or stopped growth. The authors found that iron deprivation caused a significant increase in *rv3371* transcription, but no increase was seen for any other TGS or the dormancy survival regulon (DosR). Although deletion of *rv3771* in H37Rv does not change growth rate or colony morphology, there is apparent cell wall alteration based on changes observed after staining cells that were grown in conditions known to increase *rv3771* expression. Interestingly, loss of *rv3771* leads to a decrease in EV secretion, which can be complemented back to normal levels. The decrease in EV secretion is also seen in low-iron conditions, and complementation results in a hypervesiculation phenotype in low-iron media [93]. This gene is essential for Mtb survival in the mouse. It is also required for growth arrest *in vitro*, but not in normal culture conditions. Based on this, and the fact that its transcripts have been found in the human

lung granuloma, the authors suggest that investigating the role of *rv3771* in *in vivo* persistence could determine if this gene is a good drug target.

The phosphate-specific transport (Pst) system activates the two-component SenX3-RegX3 system constitutively, resulting in an increase of ESX-5 protein secretion based on inorganic phosphate availability [95]. This signal transduction system plays a role in Mtb EV production. Deletion of *pstA1* (a transmembrane component of the Pst system) results in activation of SenX3-RegX3, no matter what the inorganic phosphate availability is. This mutant has a RegX3-dependent increase in *esx5* transcription and hypersecretion of two ESX-5 substrates associated with Mtb EVs: PPE41 and EsxN [95,96]. Further proteomic characterization showed that the *ΔpstA1* mutant secretes significantly more LpqH and PstS1, which is caused by a significant increase in EV production: up to fifteen times more vesicles than WT [97]. This increase in vesiculation was ultimately shown to be independent of ESX-5, but requires Reg3X. Interestingly, the *ΔpstA1* mutant's hypervesiculation is not dependent on VirR, but *ΔpstA1ΔvirR* produced 4 fold more EVs than *ΔpstA1* and 63 fold more EVs than *ΔvirR*, which suggests that the mechanisms driving EV production in theses strains synergize [97].

Most recently, investigation into the role of dynamin-like proteins (DLPs) for Mtb EV biogenesis was released as a pre-print (not peer reviewed) in 2020 [89]. DLPs are guanosine triphosphate hydrolase enzymes that mediate membrane fusion and fission in both prokaryotes and eukaryotes. A soluble DLP, IniA, was shown to deform membranes, contribute to membrane fission and maintenance of plasma membrane integrity, and contribute to drug-resistance based on studies with *M. smegmatis* [98]. As a member of the LytR-CpsA-Psr protein family, VirR is also likely involved in cell envelope integrity since other members of the family include enzymes that transfer glycopolymers from membrane-linked precursors to cell envelope proteins or peptidoglycan [99]. Examination of both *ΔvirR* and iron-limited Mtb revealed changes in the cell envelope, with loss of VirR causing increased thickness and iron limitation causing thinning [89]. In both strains, the iniBAC operon and [Fe-S] assembly genes were induced as compared to WT Mtb. Since the metabolic stress from VirR inactivation and iron limitation is

not identical, the upregulated genes in both conditions likely converge on a shared function. Indeed deletion of *iniA* results in decreased vesiculation, but protein secretion concentrations are maintained as compared to the WT [89]. The impacts on vesiculation from *ΔiniA* could not be rescued by *iniA* or *iniC* individually, but could be rescued by a plasmid with the entire WT *iniBAC* operon. Additionally, EVs from *ΔiniA* cannot rescue cells from iron-limitation even though the strain makes enough mycobactin/carboxymycobactin in normal media for itself [89]. Also of note, the *ΔiniA* strain replicated similarly to WT during macrophage infection, but had seemingly decreased EV production based on reduction of EVs containing bacterial components in the cell culture. These experiments further confirm a link between cell envelope alterations and Mtb EV release, demonstrate convergence of several factors influencing Mtb EV release, and provide a potential mechanism for targeting vesicle production *in vivo* [89].

Understanding mycobacterial EV composition, functions, and secretion mechanisms is important for several potential applications including drug development, biomarker-based diagnostics, and vaccine production. To date, there are no completely null Mtb mutants for vesiculation, which strongly indicates this function as essential for bacterial viability [100]. Some of the most successful medications for tuberculosis target cell-envelope related processes, including INH and ethambutol (among others) [101]. Since EV release undoubtedly involves cell-envelope alterations—and given the ties between vesiculation and environmental stress like what the bacteria encounters in the host—defining various vesicle biogenesis mechanisms may provide new targets for antitubercular drugs.

Because the composition of mycobacterial EVs vary based on the bacterial species and includes highly immunogenic biomolecules in pathogenic species, use as a diagnostic serology test for evidence of infection and even discrimination between infection by closely related species is an active area of investigation. In 2013, Ziegenbalg et al. discovered that serum from TB positive individuals reacts strongly and significantly to Mtb EVs [102]. The IgG responses in particular were significantly higher in sputum smear positive versus smear negative cases and absent in healthy, BCG vaccinated, tuberculin

skin test positive serum samples. Paratuberculosis, caused by MAP, impacts up to 50% of European and North American bovine herds [103]. There is a vaccine, but it does not provide complete protection and can cause interference with tests for bovine tuberculosis caused by *M. bovis* so there is a need to easily distinguish among bovine TB, paratuberculosis infection, and paratuberculosis vaccination. Mycobacterial EVs may provide the answer since sera from *M. bovis* infected cows recognized LpqH and other proteins from Mtb EVs while sera from healthy, MAP infected, and MAP vaccinated cows did not [76]. Interestingly, the EVs from *M. bovis* and MAP did not demonstrate good diagnostic serology capacity for any of these groups.

Mtb EVs were demonstrated to be strong stimulators of the inflammatory immune response in mice, with increased and rapid granuloma formation in the lungs of mice administered H37Rv EVs through the trachea [64]. Subsequently, subcutaneous administration of mycobacterial EVs alone and in combination with BCG-vaccination indicated that these vesicles may serve as effective vaccine components [104]. Both BCG and Mtb EV immunization elicit a strong humoral and cellular response to membrane and cell wall components of Mtb. Vaccination using Mtb EVs without adjuvant protected mice against H37Rv infection through a very strong Th1 mediated response. The Mtb EV vaccinated mice were protected as well as the regular BCG immunization group; however, BCG immunization did not generate as strong of a Th1 activation. Protection was not seen in mice vaccinated with BCG derived EVs. Additionally, Mtb EV administration following BCG immunization boosted the immunity afforded by BCG-vaccination alone, resulting in a more diverse antibody response. Unfortunately, there were challenges with reproducibility within the study whereby only two of the three studies of Mtb EVs alone as a vaccine component demonstrated protection [104]. Based on this study, use of Mtb EVs as part of a vaccine requires a better understanding of the source of variability. Additionally, if consistent preparation of Mtb EVs is not achievable, then understanding what parts of the EVs mediate the protective immune response may allow for the generation of artificial EVs carrying those components. In order to move from

preclinical to clinical proof of concept studies, the active component characterization is of paramount importance as well as developing a straightforward manufacturing and quality control process [17].

1.6 Extracellular Vesicles in the Context of Mtb Infection

As previously mentioned, essentially every cell generates extracellular vesicles. This makes many (if not all) biofluids a complex mixture of vesicles from hundreds of cell types. Additionally, many biofluids are home to natural flora and pathogens whose vesicles also contribute to the global population for that fluid. This concept holds true for culture media from *in vitro* experiments where multiple organisms are present. Because of this, the previous section reviews only publications involving EVs generated by mycobacterial culture for at least one experiment, and is limited to the results from those specific experiments. Literature focusing on host-derived vesicles and/or EVs that are likely of mixed host and bacterial origin will now be summarized.

EVs as biomarkers of TB disease has emerged as a potential tool to help move away from sputum-based diagnostics to more accessible biofluids like blood. This is vital because sputum-based diagnostics are difficult to use for children and HIV+ individuals, and are limited to active TB cases. In 2014, Kruh-Garcia et al. identified 33 unique Mtb peptides in EVs from the serum of active TB patients by multiple reaction monitoring mass spectrometry (MRM-MS) [105]. Further optimization of these methods allowed for greater than 90% accuracy in distinguishing TB positive versus TB negative people [106]. Peptides from Cfp2, Mpt32, Mpt64, and BfrB were the most consistently identified. Most recently, this assay was able to identify 95% of individuals with LTBI considered at high risk for progression to active disease [107]. A peptide from GlnA1 was detected in the majority of the serum EV samples. These studies almost certainly represent assays for human origin EVs; it is highly unlikely that Mtb EVs travel beyond the lung environment and circulate in the blood. This is especially true when the infection is dormant, as with LTBI. In addition to monitoring Mtb proteins in EVs as an avenue for biomarkers of infection, the host EV proteome changes significantly following infection [108,109]. Simultaneously

evaluating host and Mtb proteins in serum EVs may allow for a multi-faceted approach to defining TB infection status.

While Mtb EV biogenesis is the biggest unknown from the bacterial physiology perspective, the fate of Mtb EVs *in vivo* remains one of the most challenging questions from the host-pathogen interaction lens. Several studies have visually demonstrated through electron microscopy that mycobacteria release vesicles while intracellular [40,64,75]. Additionally, there is nothing to refute mycobacterial EV release *in vivo* while the bacteria is extracellular. Athman et al. (2015) provided strong evidence for the presence of Mtb EVs in Mtb-infected macrophage culture supernatants. Using dual immunogold labeling, they visualized MHC-II and CD-9 on vesicles. A general anti-Mtb gold-labeled antibody was seen primarily on vesicles without those markers (very occasionally, a mix of host and Mtb markers were seen on the same vesicles) [110]. After attempting to separate the vesicle populations and evaluating their impacts on human embryonic kidney cells and mouse macrophages, they conclude that "the TLR-2 agonist and proinflammatory activities attributed to EVs released from *M. tuberculosis*-infected cells is derived from *M. tuberculosis* [EVs] and not from host cell-derived [EVs]," [110].

In contrast with Athman et al. (2015), there is evidence that host EVs, rather than Mtb EVs, are the dominant EV type containing Mtb proteins enriched from infected cell cultures. Smith et al. (2015) found that inhibiting ubiquitination in Mtb-infected macrophages significantly reduced Mtb protein levels in subsequently released EVs [111]. This is important because ubiquitin-interacting motifs are found on subunits of the endosomal sorting complex required for transport (ESCRT) machinery [112]. ESCRTs are involved with human exosome biogenesis [113]. In fact, when inhibiting ESCRT, the production of EVs by uninfected macrophages was reduced by over 80% [111]. Mtb proteins were strongly detected after enrichment of ubiquinated proteins from Mtb-infected macrophage EVs by immunoprecipitation. The authors found that ubiquination of GroES and HspX was required for those proteins to be detected in Mtb-infected macrophage EVs [111]. While mycobacteria have a process called pupylation that is somewhat similar to ubiquination for protein degradation, the eukaryotic ubiquitin tag is completely

different from the pupylation tag [114]. Therefore, the most plausible conclusion for GroES and HspX presence in the Mtb-infected macrophage EVs is through the host ESCRT pathway [111].

Another study with Mtb-infected mice demonstrated similar conclusions using Rab27a, a protein involved in exosome biogenesis. Macrophages from mice deficient in Rab27a released 80% fewer EVs than macrophages from WT mice following Mtb infection [115]. The Rab27a deficient Mtb-infected macrophages released less EVs that also had fewer Mtb proteins. These EVs failed to elicit pro-inflammatory responses in macrophages. Combined with the decrease in Mtb proteins in EVs following ubiquitination inhibition [111], this supports the hypothesis that the dominant Mtb protein-containing EVs from infected cell cultures are actually host derived. Interestingly, the levels of LpqH present in EVs from the Rab27a deficient and WT macrophages after Mtb infection remained the same, which might support the presence of Mtb EVs [115].

Authors from the ubiquination and Rab27a studies suggest that the conclusions reached by Athman et al. (2015) were premature [116]. Schorey et al. (2021) state that "this conclusion is based on the detection of LAM-positive vesicles lacking the exosomal markers CD9 and CD63; however recent data for exosomes and other [human]EVs indicate a significantly greater heterogeneity in vesicle composition than previously appreciated, with classic exosome markers such as CD9 and CD63 being present on only a subset of exosomes (Kowal et al., 2016)" [116]. There are a few problems with this statement. First, the TEM immunogold labeling was performed with polyclonal anti-Mtb, which "detects LAM and LM, although its specificity is not limited to these lipoglycans"[110]. It is therefore impossible to determine if the antibody is staining non-LAM/LM Mtb biomolecules, although LAM and LM are expected to be the dominantly stained antigens. Second, and more importantly, the antibodies targeting host proteins in the TEM portion were against CD-9 and MHC-II, not CD-63 [110]. Ironically, the paper cited by Schorey et al. (2021) to support their claim that the Athman et al. (2015) conclusions were premature demonstrated MHC-II as a non-exosome specific EV marker, at least from DCs *in-vitro* [117]. Without a clear answer to the fate of Mtb EVs in mixed culture or the origin of Mtb proteins in EVs from

infected cell cultures, the following discussion assumes that EVs from cell cultures infected with mycobacteria or cultures exposed to mycobacterial EVs via CFP contain a mix of host and bacterial vesicles. Therefore, functions observed in these studies cannot be specifically ascribed to EVs of either origin based on the information provided.

Extracellular vesicles from mycobacteria-infected cells contribute to the innate immune response by promoting the recruitment and activation of immune cells. In a TLR-2 dependent manner, EVs from *M. avium*-infected macrophages stimulate a proinflammatory response in mouse BMMs [118]. Similarly, EVs from Mtb- or BCG-infected macrophages induce bone marrow-derived DC maturation and activation [119]. MAP-infected macrophage EVs carry mycobacterial proteins ESAT-6, Mpt63, SodA, Mpt51, and Ag85. Treating naïve macrophages with these EVs increases their rate of phagocytosis and secretion of proinflammatory cytokines, but does not induce apoptosis and necrosis like exposure to the actual MAP bacteria [120]. This latter fact supports potential use of these MAP-infected macrophage EVs in vaccines or therapeutics. Along these lines, intranasal injection of mice with EVs from BCG- or Mtb-infected macrophage cultures induced proinflammatory cytokine production and resulted in the recruitment of neutrophils and macrophages to the lungs [121].

EVs from Mtb-infected macrophages have also been shown to activate endothelial cells [122]. This results in greater expression of chemokine receptors, cell adhesion molecules, and release of chemokines. The same endothelial cell responses are observed with exposure to EVs from the serum of Mtb-infected mice. Additionally, BMMs can migrate through endothelial cells that have been exposed to EVs from infected cells, but not when exposed EVs from uninfected cells [122].

Finally, mycobacterial RNA is present in EVs of Mtb-infected macrophages, which depends on functional SecA2 in the bacteria [71]. These EVs induce a type I IFN response in macrophages based on the RIG-I/MAVS-dependent foreign RNA detection pathway. Activation of this pathway is required for the EVs released by Mtb-infected macrophages to induce restriction of Mtb replication in neighboring macrophages [71]. Finally, treatment of Mtb-infected macrophages with EVs from Mtb-infected

neutrophils increases superoxide anion production and autophagy in the macrophage, leading to increased bacterial killing [123].Together, these studies demonstrate that EVs from a variety of mycobacterium-infected innate immune cells incite further innate immune activation both *in vitro* and *in vivo*.

Adaptive immune activation can also be facilitated by EVs from mycobacterium-infected cells. EVs from BCG- and Mtb-infected macrophages can present peptide-MHC-II complexes to T cells [124]. However, this process is more effective in the presence of other antigen presenting cells because they are more efficient at stimulating memory T cells than naïve T cells [124]. Mice treated with EVs from BCG-infected macrophages produce CD4+ and CD8+ T cells in the spleen, lung, and mediastinal lymph nodes [119]. Theses T cells produce IFN- γ upon *ex vivo* stimulation with BCG [119]. These findings were reproduced using EVs from macrophages treated with Mtb CFP rather than mycobacterial infection [125]. Interestingly the *ex vivo* stimulation of splenic cells from Mtb CFP vaccinated mice with EVs from CFP-treated macrophages resulted in a more robust cytokine response, suggesting that enrichment of the molecules in the vesicles provides a more efficient antigen presentation to the splenic cells [125]. Also, EVs from CFP treated macrophages were shown to protect mice when administered intranasally, at a level equivalent to BCG vaccination, following low-dose aerosol Mtb infection [126].

In contrast to immune activation, EVs from mycobacterial infected cells also demonstrate immune suppression. Macrophage MHC-II and CD-64 expression in response to IFN-γ is partially inhibited following treatment with EVs from Mtb-infected macrophages, which is dependent on TLR2 [127]. Additionally, EVs from Mtb-infected macrophages inhibit CD4+ T cell activation similarly to, but more strongly than, LAM alone. This resulted in reduced IL-2 production and T cell proliferation [87]. Without the ability to efficiently and completely separate mycobacterial EVs from eukaryotic EVs and/or the ability to block EV production by Mtb or eukaryotic cells, determining which EVs are causing which responses throughout the course of infection is incredibly challenging. Attempts to separate the two vesicle populations suggest a density gradient may work to achieve this [110], however, there were aspects of the experimental design that were problematic. First, the vesicles were loaded at the top of the

gradient while flotation gradients by upward displacement are more appropriate for separation [128,129]. Additionally the reported density for the bacterial EVs, which likely have a higher lipid to protein ratio, is higher than the human vesicles in the study and still in the reported range for EVs of human origin [110]. It is impossible to confirm complete separation of human and Mtb EVs based on the data provided. Further proteomic analysis of the density gradient-enriched populations, as suggested by Schorey et al. (2021), would help answer this question [116]. In summary, the functions specifically attributed to mycobacterial EVs from mycobacterial culture verses EVs generated during *in vitro* or *in vivo* infection is provided in **Figure 1.3**.

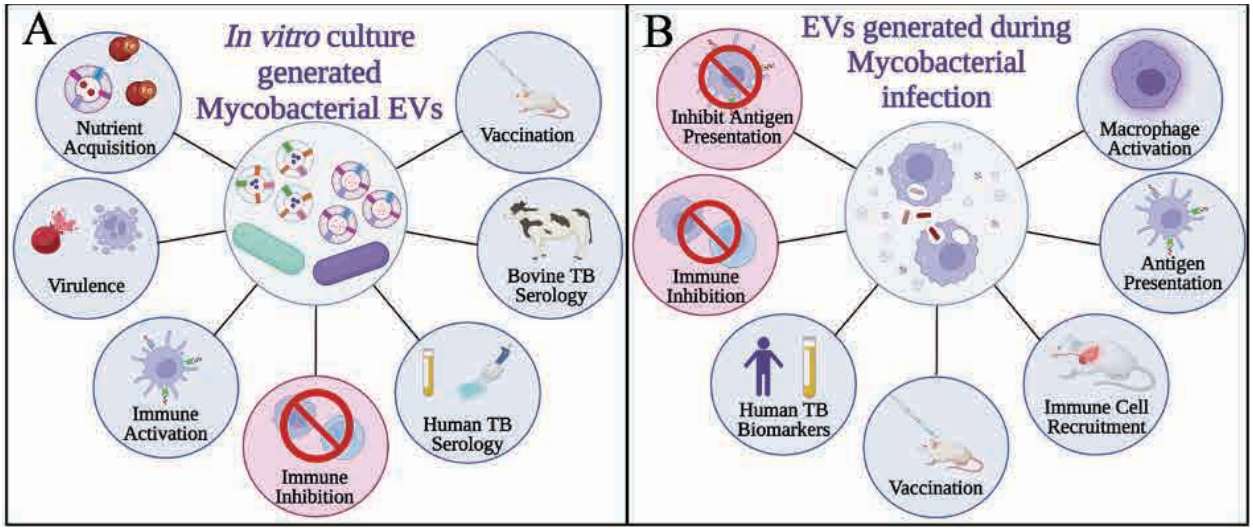


Figure 1.3 Summary of mycobacterial EVs. General functions and applications of EVs (A) generated strictly by mycobacteria and (B) those generated by mycobacterium-infected cells. Created with BioRender.com.

1.7 Overview of the Dissertation

While progress has been made to uncover the composition and functions of Mtb extracellular vesicles, there is still a plethora of outstanding questions. Many things need to be addressed in order to fully assess the role of Mtb EVs in the host-pathogen interaction and their potential for development of surveillance, prevention, and treatment components. Current literature demonstrates conflicting information, with reports of immune stimulation and immune suppression by Mtb EVs. It is possible that both facets are true, especially given the complicated immunomodulation of the host by the Mtb bacilli as

a whole, but confirmation of these previous studies is warranted. Several groups have stated that a significant hinderance to further elucidating the role Mtb EVs during infection and evaluating their use in things like vaccines is variation across and within experiments in terms of composition and host response [17,104,110,130–132].

The first attempt to use Mtb EVs as a vaccine showed that in two out of three replications of five mice per immunization group, Mtb EV vaccination elicited a protective response similar to BCG. In the third iteration this finding was not reproduced. The authors cite subtle changes in culture conditions influencing vesicle production as the most likely cause [104]. Additionally, the protein content of EVs from the same Mtb strain analyzed by very similar mass spectrometry methods showed 48 proteins in 2011, but over five times as many, 287, in 2015 [64,77]. Work in Diaz's dissertation detected 40 proteins from the same strain using similar proteomic methods in 2017 [109]. Given that Mtb EV composition varies based on the environment [80], it is important to understand what contents are present under standard culture conditions versus various states of stress. Without this, it becomes impossible to determine what components are facilitating which immune reactions, and what aspects of Mtb EVs can be exploited to fight this pathogen. At this point, pinpointing exactly what components are EV associated is difficult based on variation in separation methods.

In 2014 the Minimal Information for Studies of Extracellular Vesicles (MISEV) was published to help combat variations in rigor and reproducibility for EV studies [128]. While this publication and the subsequent update in 2018 focus heavily on eukaryotic EVs, many of the challenges noted and the principals discussed apply to EVs of all origins [129]. Variation in EV composition based on the recovery or enrichment protocol has been well documented for eukaryotic EVs and extends to mycobacterial EVs as well [133–137]. In fact, technical reproducibility issues within density gradient separation are physically visible with *M. smegmatis* EVs, and the difference in protein, particle, and RNA recovery between enrichment techniques is quite significant [137]. This technical inconsistency in combination

with variations in the culture conditions for Mtb EV generation may be influencing the data collected and conclusions drawn regarding Mtb EVs.

We hypothesize that enrichment technique is significantly impacting the biological characteristics observed in *Mycobacterium tuberculosis* extracellular vesicles. Additionally, we propose that Mtb EVs are heterogeneous in protein composition based on their size and it is possible that the enrichment techniques used are biased toward or against certain subsets of the population. This dissertation explores these hypotheses through the following aims. First, the published Mtb EV isolation techniques are optimized and evaluated for reproducibility when performed on the secretome of Mtb cultivated under the same conditions. In the second aim, the potential for biological variation in Mtb EV subpopulations grouped by size is explored. In summation, we provide insight into the influence of technical versus biological reproducibility in Mtb EV enrichment, which is essential for further investigation of Mtb EVs as tools in the fight against tuberculosis.

1.8 References

[1] Daniel TM. The history of tuberculosis. Respiratory Medicine 2006;100:1862–70. https://doi.org/10.1016/J.RMED.2006.08.006.

[2] Nerlich AG, Haas CJ, Zink A, Szeimies U, Hagedorn HG. Molecular evidence for tuberculosis in an ancient Egyptian mummy. The Lancet 1997;350:1404. https://doi.org/10.1016/S0140-6736(05)65185-9.

[3] Daniel VS, Daniel TM. Old Testament Biblical References to Tuberculosis. Clinical Infectious Diseases 1999;29:1557–8. https://doi.org/10.1086/313562.

[4] Grzybowski S, Allen EA. History and importance of scrofula. The Lancet 1995;346:1472–4. https://doi.org/10.1016/S0140-6736(95)92478-7.

[5] Chalke HD. The impact of tuberculosis on history, literature, and art. Medical History 1962;6:301. https://doi.org/10.1017/S0025727300027642.

[6] Daniel TM. Jean-Antoine Villemin and the infectious nature of tuberculosis. International Journal of Tuberculosis and Lung Disease 2015;19:267–8. https://doi.org/10.5588/IJTLD.06.0636.

[7] Martini M, Besozzi G, Barberis I. The never-ending story of the fight against tuberculosis: from Koch's bacillus to global control programs. Journal of Preventive Medicine and Hygiene 2018;59:E241–E241. https://doi.org/10.15167/2421-4248/JPMH2018.59.3.1051.

[8] Shampo MA, Kyle RA, Steensma DP. Edward L. Trudeau—Founder of a Sanatorium for Treatment of Tuberculosis. Mayo Clinic Proceedings 2010;85:e48. https://doi.org/10.4065/MCP.2010.0379.

[9] Fox HH. The chemical approach to the control of tuberculosis. Science 1952;116:129–34. https://doi.org/10.1126/SCIENCE.116.3006.129/ASSET/EF2F9E3C-8E38-4688-8896-61B0240C9961/ASSETS/SCIENCE.116.3006.129.FP.PNG.

[10] Chakraborty S, Rhee KY. Tuberculosis Drug Development: History and Evolution of the Mechanism-Based Paradigm. Cold Spring Harbor Perspectives in Medicine 2015;5:a021147. https://doi.org/10.1101/CSHPERSPECT.A021147.

[11] Vilchèze C, Jacobs WR. The Mechanism of Isoniazid Killing: Clarity Through the Scope of Genetics. Http://DxDoiOrg/101146/AnnurevMicro61111606122346 2007;61:35–50. https://doi.org/10.1146/ANNUREV.MICRO.61.111606.122346.

[12] World Health Organization. Global Tuberculosis Report 2021.

[13] Kurz SG, Furin JJ, Bark CM. Drug-Resistant Tuberculosis: Challenges and Progress. Infectious Disease Clinics of North America 2016;30:509–22. https://doi.org/10.1016/J.IDC.2016.02.010.

[14] Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nature Genetics 2013 45:10 2013;45:1176–82. https://doi.org/10.1038/ng.2744.

[15] James BW, Williams A, Marsh PD. The physiology and pathogenicity of Mycobacterium tuberculosis grown under controlled conditions in a defined medium. Journal of Applied Microbiology 2000;88:669–77. https://doi.org/10.1046/J.1365-2672.2000.01020.X.

[16] Brennan PJ. Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. Tuberculosis 2003;83:91–7. https://doi.org/10.1016/S1472-9792(02)00089-6.

[17] Sarmiento ME, Alvarez N, Chin KL, Bigi F, Tirado Y, García MA, et al. Tuberculosis vaccine candidates based on mycobacterial cell envelope components. Tuberculosis 2019;115:26–41. https://doi.org/10.1016/J.TUBE.2019.01.003.

[18] Talbot EA, Raffa BJ. Mycobacterium tuberculosis. Molecular Medical Microbiology: Second Edition 2015;3:1637–53. https://doi.org/10.1016/B978-0-12-397169-2.00092-5.

[19] Salgame P, Geadas C, Collins L, Jones-López E, Ellner JJ. Latent tuberculosis infection - Revisiting and revising concepts. Tuberculosis 2015;95:373–84. https://doi.org/10.1016/j.tube.2015.04.003.

[20] Dhar N, McKinney J, Manina G. Phenotypic Heterogeneity in Mycobacterium tuberculosis . Microbiology Spectrum 2016;4. https://doi.org/10.1128/MICROBIOLSPEC.TBTB2-0021-2016/ASSET/6CF8B71D-D868-4B02-837E-1FDE82518442/ASSETS/GRAPHIC/TBTB2-0021-2016-FIG5.GIF.

[21] Robertson BD, Schurr E, Cirillo JD, Moule MG. Article 65 Citation: Moule MG and Cirillo JD (2020) Mycobacterium tuberculosis Dissemination Plays a Critical Role in. Frontiers in Cellular and Infection Microbiology 2020;10:65. https://doi.org/10.3389/fcimb.2020.00065.

[22] Ernst JD. The immunological life cycle of tuberculosis. Nature Reviews Immunology 2012 12:8 2012;12:581–91. https://doi.org/10.1038/nri3259.

[23] Vergne I, Chua J, Lee HH, Lucas M, Belisle J, Deretic V. Mechanism of phagolysosome biogenesis block by viable Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences of the United States of America 2005;102:4033–8. https://doi.org/10.1073/PNAS.0409716102.

[24] van der Wel N, Hava D, Houben D, Fluitsma D, van Zon M, Pierson J, et al. M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. Cell 2007;129:1287–98. https://doi.org/10.1016/J.CELL.2007.05.059.

[25] Lam A, Prabhu R, Gross CM, Riesenberg LA, Singh V, Aggarwal S. Role of apoptosis and autophagy in tuberculosis. American Journal of Physiology - Lung Cellular and Molecular Physiology 2017;313:L218. https://doi.org/10.1152/AJPLUNG.00162.2017.

[26] Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. Nature Reviews Disease Primers 2016 2:1 2016;2:1–23. https://doi.org/10.1038/nrdp.2016.76.

[27] Wolf AJ, Desvignes L, Linas B, Banaiee N, Tamura T, Takatsu K, et al. Initiation of the adaptive immune response to Mycobacterium tuberculosis depends on antigen production in the local lymph node, not the lungs. The Journal of Experimental Medicine 2008;205:105. https://doi.org/10.1084/JEM.20071367.

[28] Ryan GJ, Hoff DR, Driver ER, Voskuil MI, Gonzalez-Juarrero M, Basaraba RJ, et al. Multiple M. tuberculosis Phenotypes in Mouse and Guinea Pig Lung Tissue Revealed by a Dual-Staining Approach. PLOS ONE 2010;5:e11108. https://doi.org/10.1371/JOURNAL.PONE.0011108.

[29] World Health Organization. WHO consolidated guidelines on tuberculosis rapid diagnostics for tuberculosis detection; Module 3: Diagnosis. 2020:137.

[30] MacGregor-Fairlie M, Wilkinson S, Besra GS, Oppenheimer PG. Tuberculosis diagnostics: overcoming ancient challenges with modern solutions. Emerging Topics in Life Sciences 2020;4:435–48. https://doi.org/10.1042/ETLS20200335.

[31] Boeree MJ, Heinrich N, Aarnoutse R, Diacon AH, Dawson R, Rehal S, et al. High-dose rifampicin, moxifloxacin, and SQ109 for treating tuberculosis: a multi-arm, multi-stage randomised controlled trial. The Lancet Infectious Diseases 2017;17:39–49. https://doi.org/10.1016/S1473-3099(16)30274-2.

[32] World Health Organization. WHO consolidated guidelines on tuberculosis. Module 4, Treatment : drug-resistant tuberculosis treatment. 2020:98.

[33] Martin C, Aguilo N, Marinova D, Gonzalo-Asensio J. Update on TB Vaccine Pipeline. Applied Sciences 2020, Vol 10, Page 2632 2020;10:2632. https://doi.org/10.3390/APP10072632.

[34] Angelidou A, Conti MG, Diray-Arce J, Benn CS, Shann F, Netea MG, et al. Licensed Bacille Calmette-Guérin (BCG) formulations differ markedly in bacterial viability, RNA content and innate immune activation. Vaccine 2020;38:2229–40. https://doi.org/10.1016/J.VACCINE.2019.11.060.

[35] de Castro MJ, Pardo-Seco J, Martinón-Torres F. Nonspecific (Heterologous) Protection of Neonatal BCG Vaccination Against Hospitalization Due to Respiratory Infection and Sepsis. Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America 2015;60:1611–9. https://doi.org/10.1093/CID/CIV144.

[36] Berendsen MLT, van Gijzel SWL, Smits J, de Mast Q, Aaby P, Benn CS, et al. BCG vaccination is associated with reduced malaria prevalence in children under the age of 5 years in sub-Saharan Africa. BMJ Global Health 2019;4. https://doi.org/10.1136/BMJGH-2019-001862.

[37] Siebert FB, Munday B. The Chemical Composition of the Active Principle of Tuberculin | XV. Precipitated Purified Tuberculin Protein Suitable for the Preparation of a Standard Tuberculin1,2 | American Review of Tuberculosis. American Review of Tuberculosis 1932. https://www.atsjournals.org/doi/abs/10.1164/art.1932.25.6.724?journalCode=art (accessed December 20, 2021).

[38] Cole S, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 1998;393:537–44.

[39] Jungblut PR, Schaible UE, Mollenkopf HJ, Zimny-Arndt U, Raupach B, Mattow J, et al. Comparative proteome analysis of Mycobacterium tuberculosis and Mycobacterium bovis BCG strains: towards functional genomics of microbial pathogens. Molecular Microbiology 1999;33:1103–17. https://doi.org/10.1046/J.1365-2958.1999.01549.X.

[40] Majlessi L, Prados-Rosales R, Casadevall A, Brosch R. Release of mycobacterial antigens. Immunological Reviews 2015;264:25–45. https://doi.org/10.1111/IMR.12251.

[41] Green ER, Mecsas J. Bacterial Secretion Systems – An overview. Microbiology Spectrum 2016;4. https://doi.org/10.1128/MICROBIOLSPEC.VMBF-0012-2015.

[42] Stamm CE, Pasko BL, Chaisavaneeyakorn S, Franco LH, Nair VR, Weigele BA, et al. Screening Mycobacterium tuberculosis Secreted Proteins Identifies Mpt64 as a Eukaryotic Membrane-Binding Bacterial Effector. MSphere 2019;4. https://doi.org/10.1128/MSPHERE.00354-19/SUPPL_FILE/MSPHERE.00354-19-ST002.XLSX.

[43] Braunstein M, Espinosa BJ, Chan J, Belisle JT, Jacobs WR. SecA2 functions in the secretion of superoxide dismutase A and in the virulence of Mycobacterium tuberculosis. Molecular Microbiology 2003;48:453–64. https://doi.org/10.1046/J.1365-2958.2003.03438.X.

[44] Miller BK, Zulauf KE, Braunstein M. The Sec Pathways and Exportomes of Mycobacterium tuberculosis . Microbiology Spectrum 2017;5.

https://doi.org/10.1128/MICROBIOLSPEC.TBTB2-0013-2016/ASSET/E819EB28-B71C-411E-A2AD-E5E957C83242/ASSETS/GRAPHIC/TBTB2-0013-2016-FIG4.GIF.

[45] van Winden VJC, Houben ENG, Braunstein M. Protein Export into and across the Atypical Diderm Cell Envelope of Mycobacteria. Microbiology Spectrum 2019;7. https://doi.org/10.1128/MICROBIOLSPEC.GPP3-0043-2018/ASSET/F51404BC-5167-4FF0-8DCC-1F766DC8C8A7/ASSETS/GRAPHIC/GPP3-0043-2018-FIG2.GIF.

[46] Ligon LS, Hayden JD, Braunstein M. The ins and outs of Mycobacterium tuberculosis protein export. Tuberculosis 2012;92:121–32. https://doi.org/10.1016/J.TUBE.2011.11.005.

[47] Sorensen AL, Nagai S, Houen G, Andersen P, Andersen AB. Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis. Infection and Immunity 1995;63:1710. https://doi.org/10.1128/iai.63.5.1710-1717.1995.

[48] Berthet FX, Rasmussen PB, Rosenkrands I, Andersen P, Gicquel B. A Mycobacterium tuberculosis operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). Microbiology (Reading, England) 1998;144 (Pt 11):3195–203. https://doi.org/10.1099/00221287-144-11-3195.

[49] Stanley SA, Raghavan S, Hwang WW, Cox JS. Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system. Proceedings of the National Academy of Sciences of the United States of America 2003;100:13001–6. https://doi.org/10.1073/PNAS.2235593100.

[50] Sassetti M, Boyd DH, Rubin EJ, Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. Molecular Microbiology 2003;48:77–84. https://doi.org/10.1046/J.1365-2958.2003.03425.X.

[51] Sassetti CM, Rubin EJ. Genetic requirements for mycobacterial survival during infection. Proceedings of the National Academy of Sciences 2003;100:12989–94. https://doi.org/10.1073/PNAS.2134250100.

[52] Mahairas GG, Sabo PJ, Hickey MJ, Singh DC, Stover CK. Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent M. bovis. Journal of Bacteriology 1996;178:1274–82. https://doi.org/10.1128/JB.178.5.1274-1282.1996.

[53] Harboe M, Oettinger T, Wiker HG, Rosenkrands I, Andersen P. Evidence for occurrence of the ESAT-6 protein in Mycobacterium tuberculosis and virulent Mycobacterium bovis and for its absence in Mycobacterium bovis BCG. Infection and Immunity 1996;64:16–22. https://doi.org/10.1128/IAI.64.1.16-22.1996.

[54] Serafini A, Boldrin F, Palù G, Manganelli R. Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: essentiality and rescue by iron and zinc. Journal of Bacteriology 2009;191:6340–4. https://doi.org/10.1128/JB.00756-09.

[55] Portal-Celhay C, Tufariello JM, Srivastava S, Zahra A, Klevorn T, Grace PS, et al. Mycobacterium tuberculosis EsxH inhibits ESCRT-dependent CD4 + T-cell activation. Nature Microbiology 2016;2. https://doi.org/10.1038/NMICROBIOL.2016.232.

[56] Sayes F, Sun L, di Luca M, Simeone R, Degaiffier N, Fiette L, et al. Strong immunogenicity and cross-reactivity of Mycobacterium tuberculosis ESX-5 type VII secretion: encoded PE-PPE proteins predicts vaccine potential. Cell Host & Microbe 2012;11:352–63. https://doi.org/10.1016/J.CHOM.2012.03.003.

[57] Andersen P, Askgaard D, Ljungqvist L, Bennedsen J, Heron I. Proteins released from Mycobacterium tuberculosis during growth. Infection and Immunity 1991;59:1905. https://doi.org/10.1128/iai.59.6.1905-1910.1991.

[58] Sonnenberg MG, Belisle JT. Definition of Mycobacterium tuberculosis culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. Infection and Immunity 1997;65:4515–24. https://doi.org/10.1128/IAI.65.11.4515-4524.1997.

[59] Albrethsen J, Agner J, Piersma SR, Højrup P, Pham T v., Weldingh K, et al. Proteomic Profiling of Mycobacterium tuberculosis Identifies Nutrient-starvation-responsive Toxin–antitoxin Systems. Molecular & Cellular Proteomics : MCP 2013;12:1180. https://doi.org/10.1074/MCP.M112.018846.

[60] Mehaffy C, Hess A, Prenni JE, Mathema B, Kreiswirth B, Dobos KM. Descriptive proteomic analysis shows protein variability between closely related clinical isolates of Mycobacterium tuberculosis. Proteomics 2010;10:1966–84. https://doi.org/10.1002/PMIC.200900836.

[61] Kruh-Garcia NA, Murray M, Prucha JG, Dobos KM. Antigen 85 variation across lineages of Mycobacterium tuberculosis-implications for vaccine and biomarker success. Journal of Proteomics 2014;97:141–50. https://doi.org/10.1016/J.JPROT.2013.07.005.

[62] Nieto R LM, Mehaffy C, Dobos KM. Comparing isogenic strains of Beijing genotype Mycobacterium tuberculosis after acquisition of Isoniazid resistance: A proteomics approach. Proteomics 2016;16:1376–80. https://doi.org/10.1002/PMIC.201500403.

[63] Marsollier L, Brodin P, Jackson M, Korduláková J, Tafelmeyer P, Carbonnelle E, et al. Impact of Mycobacterium ulcerans biofilm on transmissibility to ecological niches and Buruli ulcer pathogenesis. PLoS Pathogens 2007;3:0582–94. https://doi.org/10.1371/journal.ppat.0030062.

[64] Prados-Rosales R, Baena A, Martinez LR, Luque-Garcia J, Kalscheuer R, Veeraraghavan U, et al. Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. Journal of Clinical Investigation 2011;121:1471–83. https://doi.org/10.1172/JCI44261.

[65] Gill S, Catchpole R, Forterre P. Extracellular membrane vesicles in the three domains of life and beyond. FEMS Microbiology Reviews 2019;042:273–303. https://doi.org/10.1093/femsre/fuy042.

[66] Chargaff E, West R. The Biological Significance of the Thromboplastic Protein of Blood. Journal of Biological Chemistry 1946.

[67] Wolf P. The Nature and Significance of Platelet Products in Human Plasma. British Journal of Haematology 1967;13:269–88. https://doi.org/10.1111/j.1365-2141.1967.tb08741.x.

[68] Anderson HC. Vesicles associated with calcification in the matrix of epiphyseal cartilage. The Journal of Cell Biology 1969;41:59–72. https://doi.org/10.1083/jcb.41.1.59.

[69] Pan BT, Johnstone RM. Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: Selective externalization of the receptor. Cell 1983;33:967–78. https://doi.org/10.1016/0092-8674(83)90040-5.

[70] Harding C, Heuser J, Stahl P. Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. The Journal of Cell Biology 1983;97:329–39. https://doi.org/10.1083/jcb.97.2.329.

[71] Cheng Y, Schorey JS. Extracellular vesicles deliver Mycobacterium RNA to promote host immunity and bacterial killing. EMBO Reports 2019;20:e46613. https://doi.org/10.15252/EMBR.201846613.

[72] Gould SJ, Raposo G. As we wait: coping with an imperfect nomenclature for extracellular vesicles. Https://DoiOrg/103402/JevV2i020389 2013;2. https://doi.org/10.3402/JEV.V2I0.20389.

[73] Knox KW, Vesk M, Work E. Relation between excreted lipopolysaccharide complexes and surface structures of a lysine-limited culture of Escherichia coli. Journal of Bacteriology 1966;92:1206–17. https://doi.org/10.1128/jb.92.4.1206-1217.1966.

[74] Brown L, Wolf JM, Prados-Rosales R, Casadevall A. Through the wall: Extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. Nature Reviews Microbiology 2015;13:620–30. https://doi.org/10.1038/nrmicro3480.

[75] Chiplunkar SS, Silva CA, Bermudez LE, Danelishvili L. Characterization of membrane vesicles released by Mycobacterium avium in response to environment mimicking the macrophage phagosome. Future Microbiology 2019;14:293–313. https://doi.org/10.2217/FMB-2018-0249/ASSET/IMAGES/LARGE/FIGURE6.JPEG.

[76] Palacios A, Sampedro L, Sevilla IA, Molina E, Gil D, Azkargorta M, et al. Mycobacterium tuberculosis extracellular vesicle-associated lipoprotein LpqH as a potential biomarker to distinguish paratuberculosis infection or vaccination from tuberculosis infection. BMC Veterinary Research 2019;15:188. https://doi.org/10.1186/s12917-019-1941-6.

[77] Lee J, Kim SH, Choi DS, Lee JS, Kim DK, Go G, et al. Proteomic analysis of extracellular vesicles derived from Mycobacterium tuberculosis. Proteomics 2015;15:3331–7. https://doi.org/10.1002/pmic.201500037.

[78] Singh PP, Li L, Schorey JS. Exosomal RNA from Mycobacterium tuberculosis-Infected Cells Is Functional in Recipient Macrophages. Traffic 2015;16:555–71. https://doi.org/10.1111/TRA.12278.

[79] Kumar S, Mittal E, Deore S, Kumar A, Rahman A, Krishnasastry M v. Mycobacterial tlyA gene product is localized to the cell-wall without signal sequence. Frontiers in Cellular and Infection Microbiology 2015;5:60. https://doi.org/10.3389/FCIMB.2015.00060.

[80] Prados-Rosales R, Weinrick BC, Piqué DG, Jacobs WR, Casadevall A, Rodriguez GM. Role for mycobacterium tuberculosis membrane vesicles in iron acquisition. Journal of Bacteriology 2014;196:1250–6. https://doi.org/10.1128/JB.01090-13.

[81] Oliveira-Nascimento L, Massari P, Wetzler LM. The role of TLR2 in infection and immunity. Frontiers in Immunology 2012;3:79. https://doi.org/10.3389/FIMMU.2012.00079/BIBTEX.

[82] Harding C v., Boom WH. Regulation of antigen presentation by Mycobacterium tuberculosis: a role for Toll-like receptors. Nature Reviews Microbiology 2010 8:4 2010;8:296–307. https://doi.org/10.1038/nrmicro2321.

[83] Tapping RI, Tobias PS. Mycobacterial lipoarabinomannan mediates physical interactions between TLR1 and TLR2 to induce signaling. Journal of Endotoxin Research 2003;9:264–8. https://doi.org/10.1179/096805103225001477.

[84] Pecora ND, Gehring AJ, Canaday DH, Boom WH, Harding C v. Mycobacterium tuberculosis LprA Is a Lipoprotein Agonist of TLR2 That Regulates Innate Immunity and APC Function. The Journal of Immunology 2006;177:422–9. https://doi.org/10.4049/JIMMUNOL.177.1.422.

[85] Gehring AJ, Dobos KM, Belisle JT, Harding C v., Boom WH. Mycobacterium tuberculosis LprG (Rv1411c): a novel TLR-2 ligand that inhibits human macrophage class II MHC antigen processing. Journal of Immunology (Baltimore, Md : 1950) 2004;173:2660–8. https://doi.org/10.4049/JIMMUNOL.173.4.2660.

[86] Noss EH, Pai RK, Sellati TJ, Radolf JD, Belisle J, Golenbock DT, et al. Toll-Like Receptor 2-Dependent Inhibition of Macrophage Class II MHC Expression and Antigen Processing by 19-kDa Lipoprotein of Mycobacterium tuberculosis. The Journal of Immunology 2001;167:910–8. https://doi.org/10.4049/JIMMUNOL.167.2.910.

[87] Athman JJ, Sande OJ, Groft SG, Reba SM, Nagy N, Wearsch PA, et al. Mycobacterium tuberculosis Membrane Vesicles Inhibit T Cell Activation . The Journal of Immunology 2017;198:2028–37. https://doi.org/10.4049/jimmunol.1601199.

[88] Jurkoshek KS, Wang Y, Athman JJ, Barton MR, Wearsch PA. Interspecies Communication between Pathogens and Immune Cells via Bacterial Membrane Vesicles. Frontiers in Cell and Developmental Biology 2016;4:125. https://doi.org/10.3389/fcell.2016.00125.

[89] Gupta S, Palacios A, Khataokar A, Weinrick B, Lavín JL, Sampedro L, et al. Dynamin-like proteins are essential for vesicle biogenesis in Mycobacterium tuberculosis. BioRxiv 2020:2020.01.14.906362. https://doi.org/10.1101/2020.01.14.906362.

[90] Vijay S, Nair RR, Sharan D, Jakkala K, Mukkayyan N, Swaminath S, et al. Mycobacterial cultures contain cell size and density specific sub-populations of cells with significant differential susceptibility to antibiotics, oxidative and nitrite stress. Frontiers in Microbiology 2017;8:463. https://doi.org/10.3389/FMICB.2017.00463/BIBTEX.

[91] Rath P, Huang C, Wang T, Wang T, Li H, Prados-Rosales R, et al. Genetic regulation of vesiculogenesis and immunomodulation in Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences of the United States of America 2013;110. https://doi.org/10.1073/pnas.1320118110.

[92] Beaulieu AM, Rath P, Imhof M, Siddall ME, Roberts J, Schnappinger D, et al. Genome-Wide Screen for Mycobacterium tuberculosis Genes That Regulate Host Immunity. PLOS ONE 2010;5:e15120. https://doi.org/10.1371/JOURNAL.PONE.0015120.

[93] Rastogi S, Singh AK, Chandra G, Kushwaha P, Pant G, Singh K, et al. The diacylglycerol acyltransferase Rv3371 of Mycobacterium tuberculosis is required for growth arrest and involved in stress-induced cell wall alterations. Tuberculosis 2017;104:8–19. https://doi.org/10.1016/J.TUBE.2017.02.001.

[94] Baek SH, Li AH, Sassetti CM. Metabolic regulation of mycobacterial growth and antibiotic sensitivity. PLOS Biology 2011;9. https://doi.org/10.1371/JOURNAL.PBIO.1001065.

[95] Elliott SR, Tischler AD. Phosphate starvation: a novel signal that triggers ESX-5 secretion in Mycobacterium tuberculosis. Molecular Microbiology 2016;100:510–26. https://doi.org/10.1111/MMI.13332.

[96] Tischler AD, Leistikow RL, Kirksey MA, Voskuil MI, McKinney JD. Mycobacterium tuberculosis requires phosphate-responsive gene regulation to resist host immunity. Infection and Immunity 2013;81:317–28. https://doi.org/10.1128/IAI.01136-12.

[97] White DW, Elliott SR, Odean E, Bemis LT, Tischler AD. Mycobacterium tuberculosis Pst/SenX3-RegX3 regulates membrane vesicle production independently of ESX-5 activity. MBio 2018;9. https://doi.org/10.1128/MBIO.00778-18/SUPPL_FILE/MBO003183934ST2.DOCX.

[98] Wang M, Guo X, Yang X, Zhang B, Ren J, Liu A, et al. Mycobacterial dynamin-like protein IniA mediates membrane fission. Nature Communications 2019 10:1 2019;10:1–13. https://doi.org/10.1038/s41467-019-11860-z.

[99] Kawai Y, Marles-Wright J, Cleverley RM, Emmins R, Ishikawa S, Kuwano M, et al. A widespread family of bacterial cell wall assembly proteins. The EMBO Journal 2011;30:4931–41. https://doi.org/10.1038/EMBOJ.2011.358.

[100] Coelho C, Casadevall A. Answers to naysayers regarding microbial extracellular vesicles. Biochemical Society Transactions 2019;47:1005–12. https://doi.org/10.1042/BST20180252.

[101] Jackson M, McNeil MR, Brennan PJ. Progress in targeting cell envelope biogenesis in Mycobacterium tuberculosis. Future Microbiology 2013;8:855–75. https://doi.org/10.2217/FMB.13.52.

[102] Ziegenbalg A, Prados-Rosales R, Jenny-Avital ER, Kim RS, Casadevall A, Achkar JM. Immunogenicity of mycobacterial vesicles in humans: Identification of a new tuberculosis antibody biomarker. Tuberculosis 2013;93:448–55. https://doi.org/10.1016/j.tube.2013.03.001.

[103] McAloon CG, Roche S, Ritter C, Barkema HW, Whyte P, More SJ, et al. A review of paratuberculosis in dairy herds — Part 2: On-farm control. The Veterinary Journal 2019;246:54–8. https://doi.org/10.1016/J.TVJL.2019.01.009.

[104] Prados-Rosales R, Carreño LJ, Batista-Gonzalez A, Baena A, Venkataswamy MM, Xu J, et al. Mycobacterial membrane vesicles administered systemically in mice induce a protective immune response to surface compartments of mycobacterium tuberculosis. MBio 2014;5. https://doi.org/10.1128/mBio.01921-14.

[105] Kruh-Garcia NA, Wolfe LM, Chaisson LH, Worodria WO, Nahid P, Schorey JS, et al. Detection of Mycobacterium tuberculosis peptides in the exosomes of patients with active and latent M. tuberculosis infection using MRM-MS. PloS One 2014;9. https://doi.org/10.1371/JOURNAL.PONE.0103811.

[106] Mehaffy C, Dobos KM, Nahid P, Kruh-Garcia NA. Second generation multiple reaction monitoring assays for enhanced detection of ultra-low abundance Mycobacterium tuberculosis peptides in human serum. Clinical Proteomics 2017;14. https://doi.org/10.1186/s12014-017-9156-y.

[107] Mehaffy C, Kruh-Garcia NA, Graham B, Jarlsberg LG, Willyerd CE, Borisov A, et al. Identification of Mycobacterium tuberculosis Peptides in Serum Extracellular Vesicles from Persons with Latent Tuberculosis. Journal of Clinical Microbiology 2020;58. https://doi.org/10.1128/JCM.00393-20.

[108] Diaz G, Wolfe LM, Kruh-Garcia NA, Dobos KM. Changes in the Membrane-Associated Proteins of Exosomes Released from Human Macrophages after Mycobacterium tuberculosis Infection. Scientific Reports 2016;6. https://doi.org/10.1038/SREP37975.

[109] Diaz G. Exosomes: A Potential Novel Source of Biomarkers for Tuberculosis. 2017. Doctoral dissertation, Colorado State University, Fort Collins, CO. Mountain Scholar.

[110] Athman JJ, Wang Y, McDonald DJ, Boom WH, Harding C v., Wearsch PA. Bacterial Membrane Vesicles Mediate the Release of Mycobacterium tuberculosis Lipoglycans and Lipoproteins from Infected Macrophages . The Journal of Immunology 2015;195:1044–53. https://doi.org/10.4049/jimmunol.1402894.

[111] Smith VL, Jackson L, Schorey JS. Ubiquitination as a Mechanism To Transport Soluble Mycobacterial and Eukaryotic Proteins to Exosomes. Journal of Immunology (Baltimore, Md : 1950) 2015;195:2722–30. https://doi.org/10.4049/JIMMUNOL.1403186.

[112] Katzmann DJ, Babst M, Emr SD. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. Cell 2001;106:145–55. https://doi.org/10.1016/S0092-8674(01)00434-2.

[113] Vietri M, Radulovic M, Stenmark H. The many functions of ESCRTs. Nature Reviews Molecular Cell Biology 2020;21:25–42. https://doi.org/10.1038/S41580-019-0177-4.

[114] Akhter Y, Thakur S. Targets of ubiquitin like system in mycobacteria and related actinobacterial species. Microbiological Research 2017;204:9–29. https://doi.org/10.1016/J.MICRES.2017.07.002.

[115] Smith VL, Cheng Y, Bryant BR, Schorey JS. Exosomes function in antigen presentation during an in vivo Mycobacterium tuberculosis infection. Scientific Reports 2017 7:1 2017;7:1–12. https://doi.org/10.1038/srep43578.

[116] Schorey JS, Cheng Y, Mcmanus WR. Bacteria-and host-derived extracellular vesicles-two sides of the same coin? Journal of Cell Science 2021;134. https://doi.org/10.1242/jcs.256628.

[117] Kowal J, Arras G, Colombo M, Jouve M, Morath JP, Primdal-Bengtson B, et al. Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. Proceedings of the National Academy of Sciences of the United States of America 2016;113:E968–77. https://doi.org/10.1073/PNAS.1521230113/-/DCSUPPLEMENTAL.

[118] Bhatnagar S, Schorey JS. Exosomes Released from Infected Macrophages Contain Mycobacterium avium Glycopeptidolipids and Are Proinflammatory. Journal of Biological Chemistry 2007;282:25779–89. https://doi.org/10.1074/JBC.M702277200.

[119] Giri PK, Schorey JS, Bishai W, Hopkins J. Exosomes Derived from M. Bovis BCG Infected Macrophages Activate Antigen-Specific CD4+ and CD8+ T Cells In Vitro and In Vivo. PLOS ONE 2008;3:e2461. https://doi.org/10.1371/JOURNAL.PONE.0002461.

[120] Wang J, Yao Y, Xiong J, Wu J, Tang X, Li G. Evaluation of the Inflammatory Response in Macrophages Stimulated with Exosomes Secreted by Mycobacterium avium-Infected Macrophages. BioMed Research International 2015. https://doi.org/10.1155/2015/658421.

[121] Bhatnagar S, Shinagawa K, Castellino FJ, Schorey JS. Exosomes released from macrophages infected with intracellular pathogens stimulate a proinflammatory response in vitro and in vivo. Blood 2007;110:3234–44. https://doi.org/10.1182/BLOOD-2007-03-079152.

[122] Li L, Cheng Y, Emrich S, Schorey J. Activation of endothelial cells by extracellular vesicles derived from Mycobacterium tuberculosis infected macrophages or mice. PLoS ONE 2018;13:e0198337. https://doi.org/10.1371/journal.pone.0198337.

[123] Alvarez-Jiménez VD, Leyva-Paredes K, García-Martínez M, Vázquez-Flores L, García-Paredes VG, Campillo-Navarro M, et al. Extracellular vesicles released from Mycobacterium tuberculosis-Infected neutrophils promote macrophage autophagy and decrease intracellular mycobacterial survival. Frontiers in Immunology 2018;9:272. https://doi.org/10.3389/FIMMU.2018.00272/BIBTEX.

[124] Ramachandra L, Qu Y, Wang Y, Lewis CJ, Cobb BA, Takatsu K, et al. Mycobacterium tuberculosis synergizes with ATP to induce release of microvesicles and exosomes containing major histocompatibility complex class II molecules capable of antigen presentation. Infection and

Immunity 2010;78:5116–25. https://doi.org/10.1128/IAI.01089-09/SUPPL_FILE/SUPPLEMENTAL_FILE_1D__MTB_ATP_.ZIP.

[125] Giri PK, Kruh NA, Dobos KM, Schorey JS. Proteomic analysis identifies highly antigenic proteins in exosomes from M. tuberculosis-infected and culture filtrate protein-treated macrophages. PROTEOMICS 2010;10:3190–202. https://doi.org/10.1002/PMIC.200900840.

[126] Cheng Y, Schorey JS. Exosomes carrying mycobacterial antigens can protect mice against Mycobacterium tuberculosis infection. European Journal of Immunology 2013;43:3279–90. https://doi.org/10.1002/EJI.201343727.

[127] Singh PP, LeMaire C, Tan JC, Zeng E, Schorey JS. Exosomes Released from M.tuberculosis Infected Cells Can Suppress IFN-γ Mediated Activation of Naïve Macrophages. PLOS ONE 2011;6:e18564. https://doi.org/10.1371/JOURNAL.PONE.0018564.

[128] Lötvall J, Hill AF, Hochberg F, Buzás EI, Vizio D di, Gardiner C, et al. Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles. Https://DoiOrg/103402/Jev.v326913 2014;3. https://doi.org/10.3402/JEV.V3.26913.

[129] Théry C, Witwer KW, Aikawa E, Alcaraz MJ, Anderson JD, Andriantsitohaina R, et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. Journal of Extracellular Vesicles 2018;7. https://doi.org/10.1080/20013078.2018.1535750.

[130] Rodriguez GM, Prados-Rosales R. Functions and importance of mycobacterial extracellular vesicles. Applied Microbiology and Biotechnology 2016;100:3887. https://doi.org/10.1007/S00253-016-7484-X.

[131] Gupta S, Rodriguez GM. Mycobacterial extracellular vesicles and host pathogen interactions. Pathogens and Disease 2018;76. https://doi.org/10.1093/FEMSPD/FTY031.

[132] Wang J, Wang Y, Tang L, Garcia RC. Extracellular vesicles in mycobacterial infections: Their potential as molecule transfer vectors. Frontiers in Immunology 2019;10:1929. https://doi.org/10.3389/FIMMU.2019.01929/BIBTEX.

[133] Zonneveld MI, Brisson AR, van Herwijnen MJC, Tan S, van de Lest CHA, Redegeld FA, et al. Recovery of extracellular vesicles from human breast milk is influenced by sample collection and vesicle isolation procedures. Journal of Extracellular Vesicles 2014;3. https://doi.org/10.3402/JEV.V3.24215.

[134] van Deun J, Mestdagh P, Sormunen R, Cocquyt V, Vermaelen K, Vandesompele J, et al. The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling. Journal of Extracellular Vesicles 2014;3. https://doi.org/10.3402/JEV.V3.24858.

[135] Kalra H, Adda CG, Liem M, Ang CS, Mechler A, Simpson RJ, et al. Comparative proteomics evaluation of plasma exosome isolation techniques and assessment of the stability of exosomes in normal human blood plasma. Proteomics 2013;13:3354–64. https://doi.org/10.1002/PMIC.201300282.

[136] Cvjetkovic A, Lötvall J, Lässer C. The influence of rotor type and centrifugation time on the yield and purity of extracellular vesicles. Https://DoiOrg/103402/Jev.v323111 2014;3. https://doi.org/10.3402/JEV.V3.23111.

[137] Dauros Singorenko P, Chang V, Whitcombe A, Simonov D, Hong J, Phillips A, et al. Isolation of membrane vesicles from prokaryotes: a technical and biological comparison reveals heterogeneity. Journal of Extracellular Vesicles 2017;6. https://doi.org/10.1080/20013078.2017.1324731.

Chapter 2 Comparison of Mtb EV Enrichment Methods

2.1 Introduction

Experimental variation in mycobacterial extracellular vesicle (EV) literature has complicated a thorough understanding of the contents and functions of Mtb EVs. Given the impact that enrichment technique and culture parameters have on EV preparation and the influence of the environment (including media components) on Mtb EVs [1], it is important to determine if the separation methods are influencing the findings in Mtb EV experiments. Unfortunately, lack of parameter reporting and significant inconsistency in the methods used—even within the same research groups—make it impossible to evaluate whether challenges with reproducibility are due to technical or biological variation. Because the full extent of individual parameters on EV separation is far from elucidated, the Minimal Information for Studies of Extracellular Vesicles (MISEV) guidelines strongly encourage accurate reporting of every experimental parameter [2]. In fact, in order to facilitate evaluating the rigor in method reporting, the EV-TRACK database was formed where publications can be scored based on their adherence to MISEV reporting guidelines [3].

The impact of EV separation technique on the composition and function of the resulting preparation has been documented for a variety of systems, from human biofluids and eukaryotic cell culture to helminth and bacterial cell culture [4–7]. Parameters within purification methods also significantly impact the resulting EV material. For example, in centrifugation-based methods the type of rotor, fixed angle or swinging bucket, and the amount of time used significantly influences the protein to RNA ratio in the EV pellet [8]. This study also demonstrated that more time for ultracentrifugation is not always better, with an increase in protein aggregates in the pellet but not actual EVs after longer centrifugation times [8]. The wide array of variables within EV enrichment techniques reinforces the importance of thoroughly reporting all experimental details.

Regardless of their cell of origin, EVs are released into a complex matrix of various biomolecules with a wide range of sizes and structures. The non-EV components may have functional consequences that are additive, deleterious, or synergistic, but not attributable to the EVs themselves. The growth phase at which *in vitro* produced EVs are collected from a culture is of utmost importance. Mid-log phase is preferred to avoid lysis; lipid membranes in solution can spontaneously form unilamellar and bilayered vesicles, which can occur with membranes from cell lysis, generating artificial EV-like structures [9,10]. Other factors shown to impact EV yield, integrity, and content beyond the general separation technique include storage time and temperature, media components, biofluid collection steps, and even equipment and operator dependent technical variability [5,6,11–14].

In this study, we reduce sources of biological variation and focus on the technical aspects of EV enrichment methods. There is a plethora of techniques for EV enrichment and separation, each with its own set of benefits and drawbacks [15]; these are summarized as they pertain to mycobacterial EV preparations in **Table 2.1**. Density gradient ultracentrifugation was one of the first methods for separating EVs, and is the primary method used for mycobacterial EV enrichment [1,16–26]. A handful of publications have employed alternative workflows including ultracentrifugation alone [27–30] or size-exclusion chromatography (SEC) [22,31–33]. The requisite level of separation of the EVs from the matrix varies based on the downstream application, with the most stringent purity required for defining specific vesicular contents or functions [2]. Based on numerous comparisons of density gradient ultracentrifugation and size exclusion chromatography for EV preparations [4,6,22,34], we hypothesize that the Mtb EV enrichment technique significantly impacts the quality and yield of the EV preparation, and SEC methods provide more reproducible Mtb EV preparations than ultracentrifugation based methods. We evaluated four Mtb EV separation methods on three different culture filtrate protein (CFP) preparations from Mtb and performed them in three technical replicates. The quantity and quality of the Mtb EVs was then evaluated and compared for technical repeatability and biological variation.

Table 2.1 Advantages and disadvantages of mycobacterial EV preparation methods. There are three methods used in the literature thus far for separation and enrichment of mycobacterial EVs.

Technique	Advantages	Disadvantages
Ultracentrifugation	• Straightforward process • Inexpensive after centrifuge purchase • Widely used	• Instrument dependent • Low separation efficiency • Vesicle aggregation/potential damage or lysis • Published methods often lack details about instrumentation, rotors, and conditions
Density Gradient Ultracentrifugation	• High separation efficiency • Inexpensive after centrifuge purchase • Widely used • Additional dimension of separation by density vs size only	• Time consuming • Instrument dependent • Complex process • Low yield • Vesicle aggregation • Gradient medium removal required prior to analysis • Published methods often lack details about instrumentation, rotors, and conditions
Size Exclusion Chromatography	• High separation efficiency • Scalable • Straightforward process • Automation capable • Preserves EV integrity • Gentle	• Reagent and consumables cost • Dilute output • Can be challenging to define desired fractions based on overlap of EV markers and contaminants

2.2 Materials and Methods

2.2.1 Media Preparation

2.2.1.1 7H11 + 10% oleic albumin dextrose catalase (OADC) medium

Agar plates were prepared by stirring 21 g of 7H11 dehydrated agar (Middlebrook) into 900 mL of water. Once the dehydrated powder was thoroughly mixed, 5 mL of glycerol was added. The mixture was autoclaved at 121 °C for 30 min with slow exhaust. The sterile medium was cooled in a 55 °C using a water bath, prior to the addition of 100 mL sterile OADC solution in a biosafety cabinet (BSC). Approximately 140 mL liquid agar per plate was poured into seven sterile plates (15 x 150 mm) and allowed to cool. Once the agar solidified, plates were transferred to 37 °C for 12 h to ensure sterility, and then stored at 4 °C until used.

2.2.1.2 Glycerol alanine salts (GAS) medium

Liquid medium was prepared by adding the listed ingredients, in order, per liter of water while stirring: 0.3 g Bacto™ Casitone (Life Technologies), 0.05 g ferric ammonium citrate, 4.0 g potassium phosphate dibasic anhydrous, 2.0 g citric acid, 1.0 g L-alanine, 1.2 g magnesium chloride, 0.6 g potassium sulfate, and 2.0 g ammonium chloride. Once fully dispersed, 1.8 mL 10 M sodium hydroxide was added followed by 10.0 mL glycerol. The pH was measured and adjusted to 6.6 with the addition of 10 M sodium hydroxide as needed. The medium was then autoclaved at 121 °C for 1 h with slow exhaust. Medium was stored at room temperature until used.

2.2.1.3 Nutrient agar medium

Agar plates were prepared by stirring 23 g dehydrated nutrient agar (Difco) into 1 L of water. Once completely in solution, the media was autoclaved at 121 °C for 30 min for slow exhaust then cooled to 55 °C using a water bath. Approximately 50 mL liquid agar per plate was poured into twenty sterile plates (15 x 100 mm) and allowed to cool. Solidified plates were transferred to 37 °C for 12 h to ensure sterility, and then stored at 4 °C until used.

2.2.2 Mtb Culture and Culture Filtrate Protein Generation

All procedures involving live, virulent *Mycobacterium tuberculosis* were performed in a biosafety level 3 laboratory (BSL-3) at Colorado State University. These cultivation and CFP generation methods were adapted from [33,35].

2.2.2.1 Bacterial culture

M. tuberculosis strain H37Rv was plated using 200 µL of frozen glycerol stock on 7H11 + 10% OADC agar plates (15 x 150 mm). After four weeks of growth at 37 °C, the bacterial lawn was collected using a sterile cell scraper with 25 cm handle and 3 cm blade (Falcon). The cells from one plate were transferred using the cell scraper to a Fernbach flask containing 900 mL GAS medium, plugged with a sterile roll cotton wrapped in cheese cloth and covered with aluminum foil. The culture was incubated for two weeks at 37 °C on an orbital platform shaker at 65 rpm. Bacteria from one flask were swirled and allowed to

settle for 5 min to create a loose pellet of cells in the center of the flask. Approximately 1/20th of the amassed cells was transferred with a pipette to twenty 1 L roller bottles with 400 mL GAS medium. Roller bottles were incubated at 37 °C on a roller apparatus rotating at 1.65 rpm for two weeks. At every manipulation, the culture was checked for contamination visually and by plating one loopful of culture on a nutrient agar plate with a 48 h incubation at 37 °C.

2.2.2.2 CFP harvest and concentration

CFP was separated from bacterial cells by 0.2 µm filtration with a VacuCap 90 vacuum filtration device (Pall Corporation). Sterile CFP was removed from BSL-3 for further processing at BSL-2. A 2 L stirred-cell ultrafiltration unit was assembled with a 150 mm diameter 5 kDa MWCO filter (Millipore Sigma). The filter was first prepared by equilibration in 90% isopropyl alcohol for 10 min followed by 30 min in water. Eight liters of CFP combined from 20 roller bottles was concentrated (40-fold) to 200 mL by stirred-cell ultrafiltration, using a pressure reservoir connected to compressed nitrogen kept at 4 °C. Buffer exchange (20-fold) was performed with 4 L of 10 mM ammonium bicarbonate. Once reduced to approximately 100 mL, the CFP was recovered from the stirred-cell and filtered with a 0.2 µm PES 250 mL unit. Total protein was determined using the bicinchoninic acid assay (BCA) (Pierce™ Thermo Scientific) in a 96-well format following the manufacturer recommendations with standards diluted in 10 mM ammonium bicarbonate. CFP was then qualified by silver stain and Western blots (see **2.2.4.4 Western blot**) for the presence of DnaK (Rv0350), PstS1 (Rv0934), GroES (Rv3418c), Ag85 complex (Rv3804c, Rv1866c, Rv0129c), and SodA (Rv3846). GroEL2 (Rv0440), a cell lysis marker, was used as a negative control Western blot. If any of the expected markers were absent or GroEL2 was present in moderate intensity, the CFP was discarded. CFP was stored at 4 °C prior to use.

2.2.2.3 CFP ultrafiltration

Ultrafiltration was performed using a 100 kDa MWCO Centricon Plus -70 centrifugal filter (Millipore Sigma). The filter was prepared by adding 25 mL phosphate buffered saline (1XPBS) (0.144 g/L potassium dihydrogen phosphate, 9.0 g/L sodium chloride, 0.795 g/L disodium phosphate, pH 7.3-7.5) to

the top of the unit followed by centrifugation for 10 min at 2,800 x g, 4 °C. Residual 1X PBS was discarded from the unit. CFP was added to the top of the filter unit up to 60 mL at a time and centrifuged in 10 min increments at 2,800 x g at 4 °C. The first 50 mL of 100 kDa CFP flow through (100FT) was saved at 4 °C for downstream analysis. Once the CFP was reduced to the volume of the filter, 60 mL 1X PBS was added to the unit and concentrated for a total of 5 washes. One milliliter of 1X PBS was added to each filter column directly followed by 3 mL 1X PBS in the top of the unit. The filter was inverted into the collection cup for a recovery spin at 57 x g for 5 min at 4 °C. An additional 1 mL of 1X PBS was added to each filter and allowed to sit while the retentate was transferred out of the recovery cup. A second recovery spin was performed and that retentate pooled with the first for a total of approximately 7 mL retentate. The 100 kDa retentate (100R) was passed through a 0.8 µm/0.2 µm dual PES filter (Pall Corporation) using a 10 mL syringe. Total protein for the 100R and 100FT was determined by BCA using a 96-well plate format following the manufacturer recommendations with standards diluted in 1X PBS. The filtered 100 kDa CFP retentate was stored at 4 °C prior to use.

2.2.3 Mtb EV Enrichment

Following protocol optimization, all enrichment techniques were performed in triplicate on three different batches of 100R for method comparison. Details regarding protocol optimization are included in **Appendix I**: Chapter 2 Supplementary Information. Total protein for each enriched sample was determined by microBCA (Pierce™ Thermo Scientific) using a 96-well plate format following the manufacturer recommendations with standards diluted in 1X PBS. All EV preparations were stored at 4 °C when not in use.

2.2.3.1 Capto™Core size exclusion chromatography

Capto™Core 700 (GE Healthcare Life Sciences) resin was packed into Poly-Prep® Columns (Bio-Rad) by adding 0.5 mL Capto™Core slurry followed by 5 mL 1X PBS to remove the resin storage solution. The 1X PBS wash of the slurry was performed by gravity-flow. The column was then capped at the bottom and 3 mg 100R in 5 mL 1X PBS was gently added to the top of the column. The slurry was

allowed to settle for 5 min before the cap was removed and the flow through recovered in a clean tube. After collection the column was re-capped and the flow through added for a second pass over the resin. The resin was allowed to settle before the cap was removed and the final flow through recovered in a clean tube. The resin was washed twice with 0.5 mL 1X PBS to maximize recovery, which was collected with the final flow through. Sample concentration was performed using a 4 mL 3 kDa MWCO Amicon Ultra centrifugal filter (Millipore Sigma). The filter was prepared by adding 4 mL 1X PBS to the top of the unit followed by centrifugation for 5 min at 2,800 x g, 4 °C. Residual 1X PBS was discarded from the unit. Capto™Core enriched EVs (CC-EVs) were added to the unit and reduced to approximately 200 µL. The concentrated CC-EVs were transferred to a new tube and the filter rinsed twice with 200 µL 1X PBS. The wash material was pooled with the concentrated CC-EVs and the final volume recorded. For each batch of 100R, the three technical replicates were normalized by volume with the addition of 1X PBS.

2.2.3.2 qEV size exclusion chromatography

qEVoriginal 35 nm columns were used with an automatic fraction collector (AFC) (Izon) following the manufacturer recommendations. After equilibrating to room temperature, the qEV column was installed on the AFC and flushed with 60 mL 1X PBS. The AFC was programmed to collect 20 x 500 µL fractions. Three mg of 100R brought to 0.5 mL with 1X PBS was applied to the top of the column. Once the sample had run into the column, 15 mL 1X PBS was added to the top of the column. The AFC collected fractions after the column void volume had flowed through. The first three fractions collected contain the qEV enriched EVs (qEV-EVs) and were pooled together. After collection was complete, the qEV column was cleaned with 0.5 mL of 0.5 M NaOH followed by 60 mL 1X PBS. Once all three replicates had been run, the column was stored in 0.05% NaN_3 at 4 °C. Each column can be used a total of 5 times. One column was used per 100R batch to avoid lot-to-lot cross contamination.

2.2.3.3 Ultracentrifugation

Centrifuge rotor buckets for a SW 32.1 Ti swinging bucket rotor (Beckman Colter) were cooled to 4 °C. Thin-wall polypropylene tubes 16 x 96 mm (Beckman Coulter) were rinsed with 70% ethanol and

allowed to air dry. Tubes were filled with 16 mL 1X PBS followed by the addition of 3 mg 100R and then weighed. 1X PBS was added as needed to ensure all tubes were within 20 mg of each other for proper balancing. The rotor buckets and O-rings were carefully inspected for deformities, then the O-rings were lightly and evenly coated with silicone vacuum grease. Spinkote lubricant was applied to cap grooves in the bucket tops with a lint-free cotton swab, and the bucket caps were matched with the numbered buckets. Balanced tubes were loaded into the centrifuge buckets which were installed in the SW 32.1 Ti rotor inside of an L7-80 Ultracentrifuge (Beckman Coulter). Samples were centrifuged at 100,000 x g for 1 h at 4 °C. Buckets were retrieved from the rotor and tubes carefully removed from the buckets. Supernatant was decanted from the invisible pellets by pouring smoothly and swiftly in one motion away from the expected pellet location. The pellets were each resuspended in 500 µL 1X PBS and transferred to a new tube.

2.2.3.4 OptiPrep™ density gradient separation

The same ultracentrifuge, rotor, and tubes described in **2.2.3.3 Ultracentrifugation** were used for density gradient ultracentrifugation. Ultracentrifuge concentrated EVS (UC-EVs) were transferred to a new, dry polypropylene tube 16 x 96 mm (Beckman Coulter) and mixed with 1.5 mL OptiPrep™ density gradient medium (Millipore Sigma) to achieve a 45% iodixanol solution. Subsequently, 1.5 mL layers of 40 – 5% OptiPrep in 1X PBS in decreasing increments of 5% were gently overlaid on the sample followed by 2 mL 1X PBS. The tubes were weighed and prepared in ultracentrifuge buckets as described in **2.2.3.3 Ultracentrifugation**. Samples were centrifuged at 100,000 x g for 16 h at 4 °C. Buckets were retrieved from the rotor and tubes carefully removed from the buckets. Fractions were collected with a 1 mL pipette starting from the top of each gradient and transferred to new tubes. Fractions 8-12 were pooled, as they contain the EVs. Buffer exchange to 1X PBS was performed with a 15 mL 100 kDa MWCO Amicon Ultra centrifugal filter (Millipore Sigma). The filter was prepared by adding 15 mL 1X PBS to the top of the unit followed by centrifugation for 10 min at 2,800 x g, 4 °C. Residual 1X PBS was discarded from the unit. Eleven milliliters of 1X PBS were added to the filter followed by the 4 mL pool of density

gradient enriched EVs (DG-EVs) and reduced to approximately 250 µL. The filter was filled with 1X PBS and the sample was reduced again. This process was repeated a total of five times to ensure removal of residual iodixanol. The concentrated DG-EVs were transferred to a new tube and the filter rinsed twice with 200 µL 1X PBS. The wash material was pooled with the concentrated DG-EVs and the final volume recorded.

2.2.4 Mtb EV Analysis

2.2.4.1 Nanoparticle tracking analysis (NTA)

The concentration and size distribution of the enriched EVs were evaluated by NTA using the NanoSight NS300 (Malvern Panalytical) with an automatic syringe pump and software version 3.4. EV samples were diluted in 1X PBS such that the final concentration was in the range of 10^7-10^9 particles per mL, per instrument recommendations. The screen gain was set to 13, the camera level 13, and the syringe pump 10. The focus was adjusted as needed for each sample. Three 30 second videos were captured for each sample. Analysis settings included a screen gain of 10 and detection level 5 and reported an average of the three videos. Final sample particle concentration was calculated based on the NTA read out and the volume of sample diluted for measurement.

2.2.4.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Samples normalized by either total protein or NTA particle count were dried using a SpeedVac, and then resuspended in 10 µL of Laemmli buffer (5% 2-mercaptoethanol, 0.01% bromophenol blue, 10% glycerol, 2% SDS, 63 nM Tris-HCL) and heated for 5 min at 100 °C. After a brief centrifugation to collect any condensation, samples were loaded on a NuPAGE Novex 4-12% Bis-Tris Gel (Life Technologies) in 1X NuPAGE MES SDS Running Buffer (Life Technologies). At least one lane contained 3-5 µL Precision Plus Protein™ Dual Color Standard (Bio-Rad) to serve as a molecular weight reference. Resolution was accomplished using constant voltage of 200 for 35 min.

2.2.4.3 Silver stain

All incubation steps occur rocking and at room temperature. Following SDS-PAGE resolution, the gel was removed from the cassette and incubated for 1 h in a solution of 40% methanol and 10% glacial acetic acid in water then decanted. Next the gel was covered with a solution of 5% methanol and 7% glacial acetic acid in water for 5 min then decanted. The proteins were then fixed by 2% glutaraldehyde in water for 5 min and decanted, followed by three water washes for 10 min each. Oxidation was performed by incubation in 0.17 mM dithiothreitol (DTT) for 5 minutes then decanted. Next, 6 mM silver nitrate was added for 5 min followed by three brief water washes. Approximately one-third the volume of developer solution (6 g Na_2CO_3 and 3-6 drops 37% formaldehyde in 200 mL water) was added to rinse any residual silver nitrate and discarded before the remaining two-thirds of the developer was allowed to incubate on the gel. Once the desired color saturation was reached, approximately 20 mL of 50% citric acid was used to stop the reaction. After 20 min the gel was transferred to water before imaging.

2.2.4.4 Western blot

Following SDS-PAGE resolution, the gel was removed from the cassette and transferred to 0.2 μm nitrocellulose (Bio-Rad) for 1 h at 50 V in transfer buffer (25 mM Tris base and 195 mM glycine, in 20% methanol). The membrane was removed and blocked with 2% bovine serum albumin in TBS-T (1.21 g/L Tris, 8.77 g/L NaCl, 0.5 mL/L 100% Tween 80, pH 7.4) for 1 h at room temperature or overnight at 4 °C. The blocking solution was decanted, and the membrane rinsed with TBS-T followed by the addition of primary antibody diluted in TBS-T. Primary antibody was incubated for 1 h at room temperature or overnight at 4 °C then decanted and the membrane rinsed three times with TBS-T. Alkaline-phosphatase conjugated secondary antibody was diluted in TBS (1.21 g/L Tris, 8.77 g/L NaCl, pH 7.4) and added to the membrane at room temperature for 30 min. After removal of the secondary antibody, the membrane was rinsed for 5 min in TBS then exposed to the developer reagent (NBT/BCIP ready to use tablet, Roche, dissolved in 10 mL water). The reaction was stopped by the addition of water then the membrane was rinsed and allowed to air dry before imaging.

2.2.4.5 Transmission electron microscopy (TEM)

Vesicle samples were fixed by adding an equal volume of 4% TEM grade paraformaldehyde, final concentration of 2%, then stored at 4 °C for a minimum of 24 h. Formvar/carbon coated grids (Electron Microscopy Science) were glow discharged for 30 seconds before 10 µL of fixed EV sample was dropped onto the grid and allowed to sit for 5 min. Excess liquid was blotted away with filter paper. The grid was washed by floating on a 50 µL drop of ultrapure water for 30 seconds. Excess water was blotted away then the grid was stained by floating on a 50 µL drop of 2% uranyl acetate for 2 min. Excess stain was blotted away and the grid air dried. Grids were imaged on a JOEL JEM-2100F transmission electron microscope at 200 kV with spot size 5 and alpha 2.

2.2.5 Data Analysis

2.2.5.1 Western blot image analysis

Western blot images were converted to grayscale for analysis in ImageJ (version 1.53) [36]. The largest band of interest on the blot was used for setting the region of interest (ROI) size for analysis. Measurement for each band was performed using the mean gray value of the ROI. Background values were also generated by selecting three random areas as the ROI. All values were inverted by subtracting the pixel density recorded by Image J from 255, the maximum pixel value. The background was then subtracted to obtain the final intensity value for relative quantification of band intensity.

2.2.5.2 Statistical analysis

RStudio (2021.09.0 Build 351) using R for statistical computing (version 4.1.1) was used to generate graphs and plots. All coding for statistical analyses performed in RStudio is provided in Appendix 1 (section **6.4 R Code**). Prior to statistical tests for differences among means, samples were evaluated for normality by histogram, QQ Plot, and Shapiro-Wilk's test (Appendix I **Figure AI.1)**. If normality was met, equal variance was determined by Bartlett's test. If normality and equal variance assumptions were met, ANOVA with post-hoc Tukey or Student's t-test was used. If either assumption was violated (p-

value <0.05), the nonparametric Kruskal-Wallis test followed by pairwise Wilcoxon signed-rank test was performed.

2.3 Results and Discussion

2.3.1 EV Morphology and Size

Visualization of formaldehyde-fixed EVs from each method demonstrates closed-membrane vesicles of the anticipated size range (**Figure 2.1)**. Images reveal that regardless of the method, all of the samples include EVs of slightly differing sizes and morphologies. Some of the variation in vesicle shape is artifactual and can be attributed to the drying process for TEM grid preparation. NTA supports the TEM data and **Figure 2.2** demonstrates that each method results in EVs of similar size (majority within 95-110 nm range). ANOVA demonstrated that NTA sizes show no significant difference across the means.

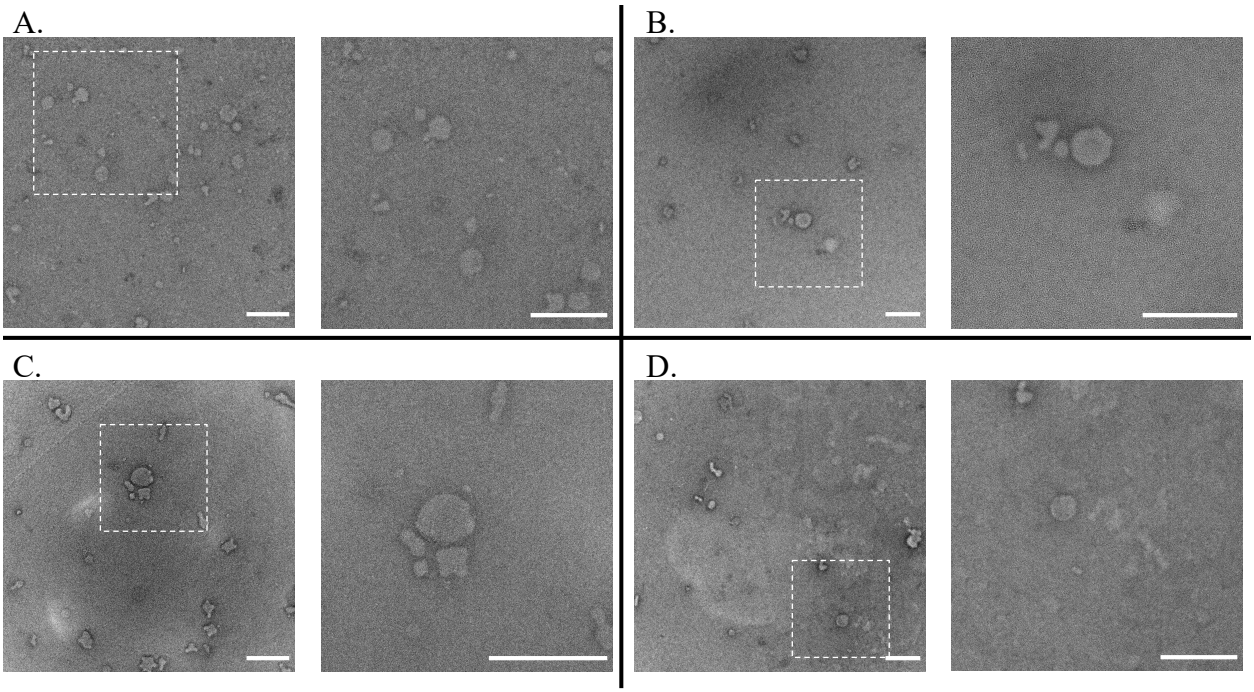


Figure 2.1 Representative TEM images by enrichment method. (A) Capto™Core, (B) Density Gradient, (C) qEV, and (D) Ultracentrifugation. The right image for each method is the zoomed in view of the dotted-line white box. Scale bars represent 200 nm.

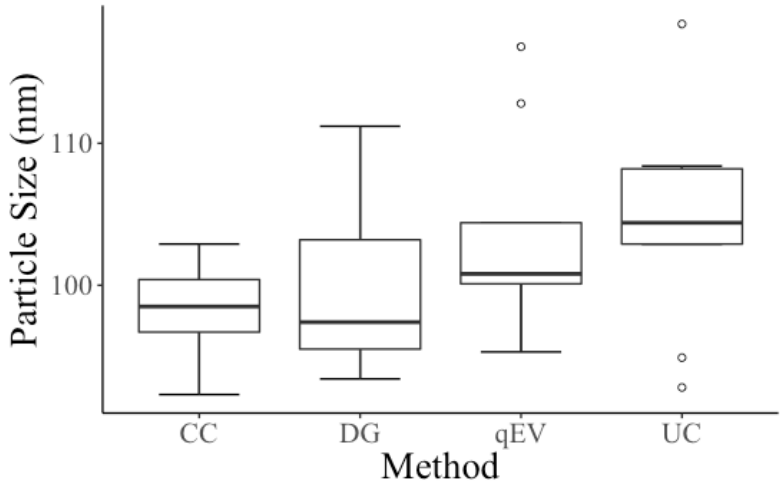


Figure 2.2 Box plots of mean particle size by NTA across methods. ANOVA testing result was p = 0.1 which means failure to reject the null hypothesis that there is a significant difference among the means.

2.3.2 Comparing EV Recovery Across Enrichment Methods

Protein recovery by method illustrated in **Figure 2.3** is significantly higher for CC-EVs than all other methods. UC-EVs have significantly a higher protein recovery than qEV and DG, while qEV has a significantly higher recovery than DG. The standard deviations within biological replicates increase as protein recovery increases.

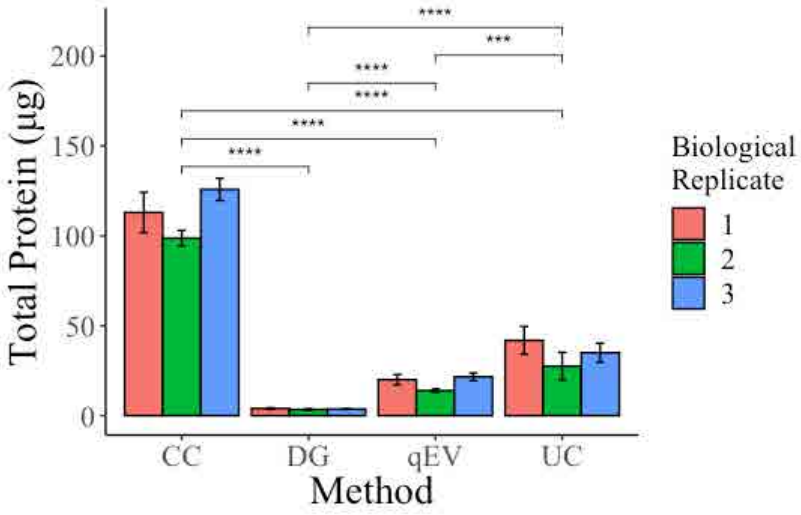


Figure 2.3 Bar chart of protein recovery. The total protein recovered is charted by method and colored by biological replicate. The error bars represent the standard deviation of the technical replicates within the biological replicate. Significance from pairwise Wilcoxon signed-rank test indicated as p<0.001***, and p<0.0001****.

Consistent with total protein recovery, CC also has the highest particle recovery, but the statistical significance between CC and qEV disappears (**Figure 2.4**). There are significantly more particles recovered by qEV than DG, and by UC than DG, but no difference between qEV and UC. Therefore, SEC results in significantly more particle recovery than density gradient ultracentrifugation.

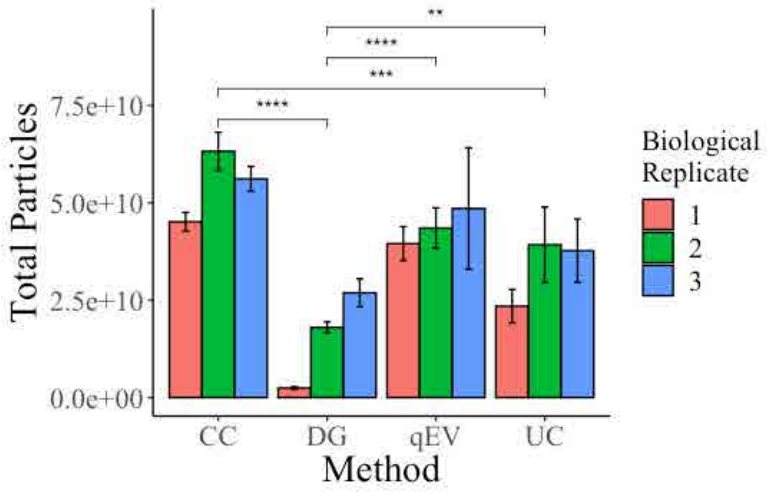


Figure 2.4 Bar chart of total particles. The total particles recovered is charted by method and colored by biological replicate. The error bars represent the standard deviation of the technical replicates within the biological replicate. Significance from Student's t-test indicated as p<0.01**, p<0.001***, and p<0.0001****.

2.3.3 Evaluating Repeatability Within Enrichment Methods

Establishing technical repeatability is paramount for determining if the discrepancies noted within and between published studies are more likely due to biological or technical variation. In **Figure 2.5,** it is evident that regardless of biological replicate, DG and CC produce technical replicates that are similar to each other in total protein and particle yield. The average range for DG total protein is 0.26 µg and total particles is 1.94E9. The average range for CC total protein is 12.93 µg and total vesicles is 6.33E9. qEV-EV technical replicates cluster fairly closely for biological replicates 1 and 2, but there is significant spread in biological replicate 3, primarily due to variation in particle yield. The average range for qEV total protein is narrow at 3.78 µg while the total particle average range is wide at 2.14E10. UC-EV technical replicates are the least tightly grouped when considering all three biological replicates, with an

average protein range of 16.06 µg and an average particle range of 1.63E10. Tables reporting the numerical values are located in Appendix I (**Table AI6.4** and **Table AI6.5**).

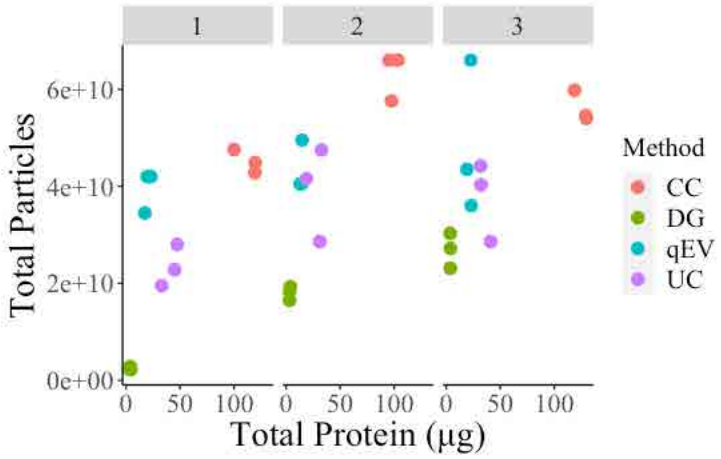


Figure 2.5 Scatter plots of particles by protein for each method. The total particles are plotted by total protein for each biological replicate (1, 2, 3) and color coded by method.

Absolute EV quantitation and preparation evaluation is difficult to achieve through any single metric. To combat this, EV preparations are often evaluated based on a ratio of two measurements, such as particles per µg of protein. A technically consistent method is one that produces similar particles per unit protein independent of biological replicate. **Figure 2.6** demonstrates that although DG-EVs have technical repeatability in the protein and particle measurements, the particle to protein ratio varies greatly across biological replicates. The standard deviation of this ratio within the technical replicates increases over biological replicates. CC-EVs demonstrate the most consistent protein to particle ratio with the smallest standard deviations. qEV-EVs and UC-EVs both have slight fluctuation in the protein to particle ratio and the technical replicate standard deviations.

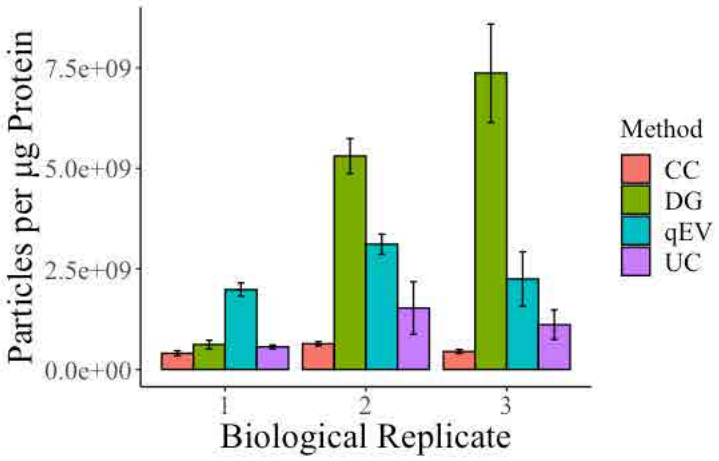


Figure 2.6 Bar chart of particle to protein ratio. The number of particles per µg of protein is charted by method and colored by biological replicate. The error bars represent the standard deviation of the technical replicates within the biological replicate.

Coefficient of variation (CV) is the measure of the relative dispersion of data points around the mean, calculated as the ratio of the standard deviation to the mean. This ratio is often used for assessing technical variability, and is presented for particle to protein ratio (**Table 2.2**). The CV is compared among technical replicates as well as across the method as a whole because each method was performed using the same amount of 100R for technical replicates across the biological replicates. This gives valuable insight into the difference between technical variation and biological variation on the output. Density gradient ultracentrifugation has low CV among the technical replicates, but a very high CV when looking across the biological replicates. This indicates that biological variation is linked to technical repeatability for this method. Considering CV, the size exclusion chromatography methods perform more consistently across the biological replicates than ultracentrifugation-based methods for particle to protein ratio. This means that the preparation is consistent in the reduction of soluble proteins compared to vesicle recovery. Ultimately, this data supports the hypothesis that SEC based methods provide more technically reproducible methods than UC based methods.

Table 2.2 Particle to protein ratio coefficients of variation.

Method	Biological Replicate	Technical Replicate Average Particles per µg	Technical Replicate Standard Deviation	Technical Replicate CoV (%)	Average Technical Replicate CoV (%)	Method Average Particles per µg	Method Standard Deviation	Method CoV (%)
CC	1	4.03E+08	6.33E+07	15.7	11.55	4.97E+08	1.19E+08	24.0
CC	2	6.40E+08	5.15E+07	8.0				
CC	3	4.48E+08	4.88E+07	10.9				
qEV	1	1.98E+09	1.62E+08	8.2	15.39	2.45E+09	6.32E+08	25.8
qEV	2	3.12E+09	2.47E+08	7.9				
qEV	3	2.25E+09	6.77E+08	30.1				
UC	1	5.62E+08	4.86E+07	8.7	28.18	1.07E+09	5.65E+08	52.9
UC	2	1.53E+09	6.56E+08	42.9				
UC	3	1.11E+09	3.68E+08	33.0				
DG	1	6.21E+08	1.06E+08	17.1	13.94	4.43E+09	3.07E+09	69.1
DG	2	5.31E+09	4.32E+08	8.1				
DG	3	7.37E+09	1.22E+09	16.6				

2.3.4 Visualizing Protein Variation Across Enrichment Methods

CC-EVs have the highest protein and particle recovery (**Figure 2.3** and **Figure 2.4**) but the lowest particle to protein ratio (**Figure 2.6**), which is significantly lower than all three of the other methods. This indicates that although particle recovery is high, soluble protein reduction is poor. UC-EVs also had a low particle to protein ratio across the three biological replicates. In **Figure 2.7**, the more efficient reduction in soluble protein in qEV-EVs and DG-EVs is visible by silver stain (A). Interestingly, GroES (C) appears in UC-EVs, but not CC-EVs. This non-EV associated contaminate is still removed by CC, likely because it is small enough to enter the resin beads and remain bound. The multi-modal characteristics of CC is evident in the silver stain (A) where excess soluble protein appears to be larger in size when comparing the profiles of CC-EVs and UC-EVs. This suggests that the CC resin is more efficient at removing smaller soluble proteins while UC-EVs contain soluble proteins across the size range. Silver stain and Western blot images for biological replicates 1 and 3 are available in Appendix I (**Figure AI.7** and **Figure AI.8**). Note: due to recovery, biological replicate 1 DG-EVs were not included for silver stain and Western blots.

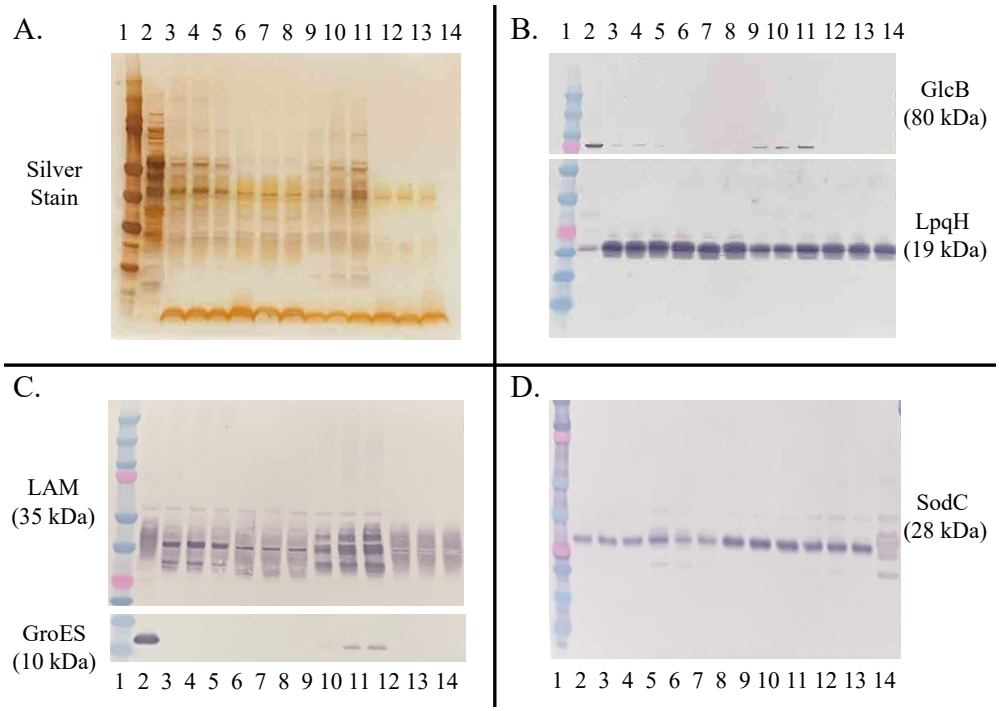


Figure 2.7 Silver stain and Western blots of biological replicate 2. (A) Silver stain and (B-D) Western blots for each technical replicate of biological replicate 2. All are loaded in the following format: 1 = ladder, 2 = 5 µg 100R, 3-5 = 1E9 CC-EV 2.1-2.3, 6-8 = 1E9 qEV-EV 2.1-2.3, 9-11 = 1E9 UC 2.1-2.3, and 12-14 = 1E9 DG 2.1-2.3.

Western blot analysis demonstrates that UC-EVs have the most remaining soluble contaminants (GlcB and GroES). CC-EVs remain slightly positive for GlcB, likely due to its size. Neither of these are visible for qEV-EVs or DG-EVs. Intensity analysis of the Western blots (**Figure 2.8**) confirms that all the methods demonstrate LpqH enrichment (C) when compared to the 100R material. Reduction in GlcB (A) and GroES (B) varies slightly based on the biological replicate, particularly for UC-EVs compared to 100R. While the LAM banding patterns vary by method (**Figure 2.7 C**), the overall intensity does not demonstrate many significant differences (**Figure 2.8 D**). Although SEC methods have higher intensity of LpqH (C) and SodC (E) than ultracentrifugation methods, it is variable in rising to statistically significant differences. Because Western blot intensity is relative quantitation, further proteomic investigation into the abundance of these markers will increase confidence in these findings.

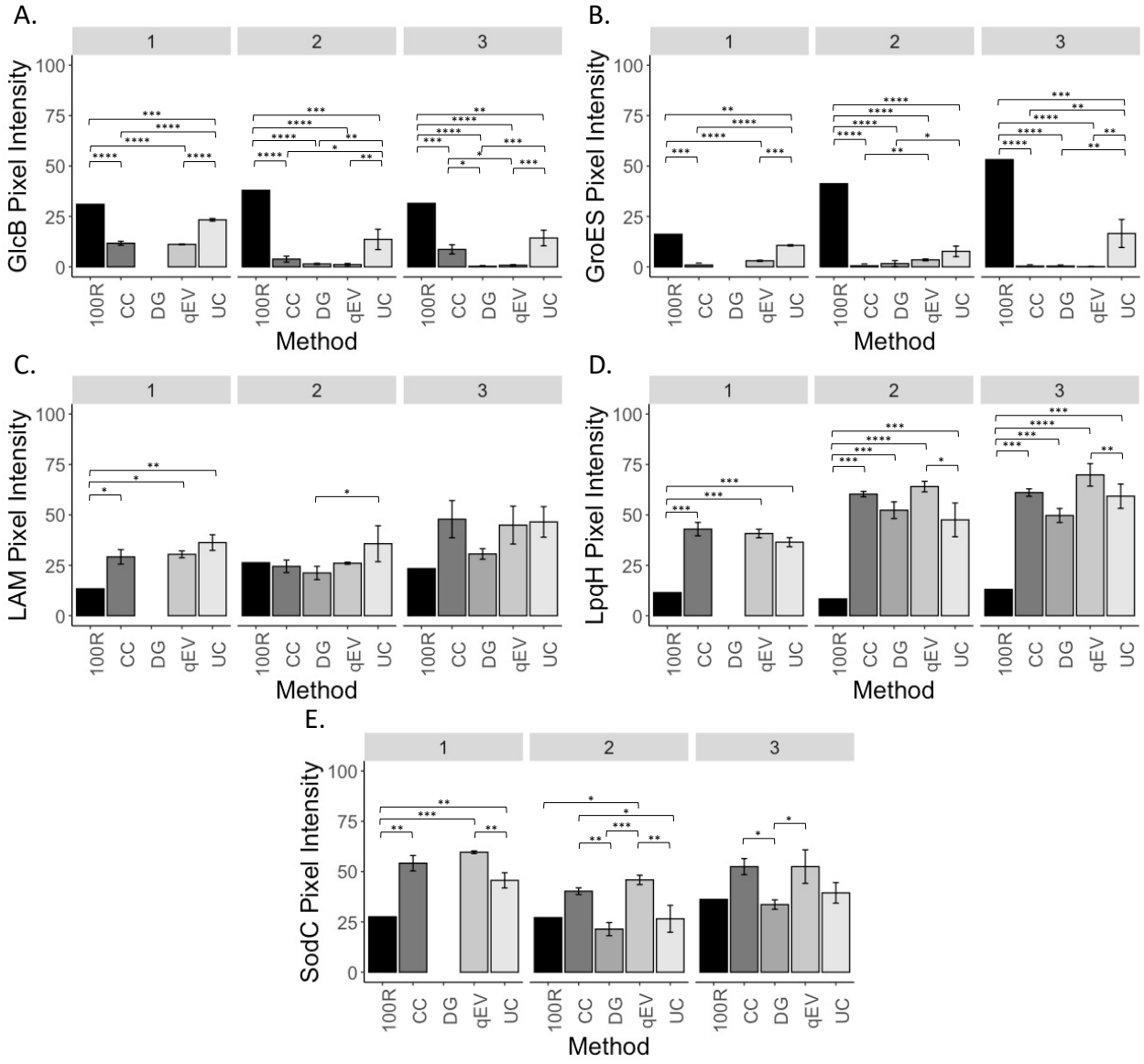


Figure 2.8 Western blot intensity comparisons. Average pixel intensity for each technical replicate is shown for Western blots against (A) GlcB, (B) GroES, (C) LAM, (D) LpqH, and (E) SodC. Each biological replicate had one Western blot per antibody so a comparison among pixel intensity was performed using the Student's t-test. Significance is indicated as $p<0.05$*, $p<0.01$**, $p<0.001$***, and $p<0.0001$****.

2.3.5 Conclusions

Ultimately, TEM and NTA agree that vesicle size does not vary across the methods. However, purification method was a source of total protein and particle yield variation, independent of biological variation. Based on particle to protein ratio, SEC methods are more technically reproducible than

ultracentrifugation-based methods. Because DG is the primary method used in the literature, the inconsistency in technical repeatability is likely influencing the findings within and between publications. The separation of EVs from soluble material varies, with qEV-EVs and DG-EVs having the least amount of contaminants like GlcB and GroES. Although CC-EVs had the highest particle recovery, the low particle to protein ratio and presence of GlcB indicates this method is not as effective at separating soluble proteins. Considering all of the analyses summarized in **Table 2.3**, qEV-EVs represent the highest yield with the best separation. Ultimately the choice of EV enrichment technique depends on the down-stream application. If purity is of lower concern but high yield is very important, CC might be the best choice for that study. In cases where purity is important, qEV is going to provide the highest vesicle recovery. Further investigation into the proteomic differences across these methods would be the next step for further validating the level of separation achieved and determining the impact of each method on the proteins represented in the vesicle preparations.

Table 2.3 Mtb EV enrichment method comparison summary.

	CC	**DG**	**qEV**	**UC**
Size	No difference	No difference	No difference	No difference
Total protein	++++	+	++	+++
Total particles	++++	+	+++	++
GroEL/GlcB	-/+	-/-	-/-	+/+
Vesicle to protein ratio	High reproducibility	Impacted by biological variation	High reproducibility	Impacted by biological variation
Yield/Purity	High/Moderate	Low/High	Moderate/High	Moderate/Low

2.4 References

[1] Prados-Rosales R, Weinrick BC, Piqué DG, Jacobs WR, Casadevall A, Rodriguez GM. Role for mycobacterium tuberculosis membrane vesicles in iron acquisition. Journal of Bacteriology 2014;196:1250–6. https://doi.org/10.1128/JB.01090-13.

[2] Théry C, Witwer KW, Aikawa E, Alcaraz MJ, Anderson JD, Andriantsitohaina R, et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. Journal of Extracellular Vesicles 2018;7. https://doi.org/10.1080/20013078.2018.1535750.

[3] van Deun J, Mestdagh P, Agostinis P, Akay Ö, Anand S, Anckaert J, et al. EV-TRACK: transparent reporting and centralizing knowledge in extracellular vesicle research. Nature Methods 2017 14:3 2017;14:228–32. https://doi.org/10.1038/nmeth.4185.

[4] Brennan K, Martin K, FitzGerald SP, O'Sullivan J, Wu Y, Blanco A, et al. A comparison of methods for the isolation and separation of extracellular vesicles from protein and lipid particles in human serum. Scientific Reports 2020 10:1 2020;10:1–13. https://doi.org/10.1038/s41598-020-57497-7.

[5] van Deun J, Mestdagh P, Sormunen R, Cocquyt V, Vermaelen K, Vandesompele J, et al. The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling. Journal of Extracellular Vesicles 2014;3. https://doi.org/10.3402/JEV.V3.24858.

[6] Davis CN, Phillips H, Tomes JJ, Swain MT, Wilkinson TJ, Brophy PM, et al. The importance of extracellular vesicle purification for downstream analysis: A comparison of differential centrifugation and size exclusion chromatography for helminth pathogens. PLOS Neglected Tropical Diseases 2019;13:e0007191. https://doi.org/10.1371/JOURNAL.PNTD.0007191.

[7] Dauros Singorenko P, Chang V, Whitcombe A, Simonov D, Hong J, Phillips A, et al. Isolation of membrane vesicles from prokaryotes: a technical and biological comparison reveals heterogeneity. Journal of Extracellular Vesicles 2017;6:1324731. https://doi.org/10.1080/20013078.2017.1324731.

[8] Cvjetkovic A, Lötvall J, Lässer C. The influence of rotor type and centrifugation time on the yield and purity of extracellular vesicles. Https://DoiOrg/103402/Jev.v323111 2014;3. https://doi.org/10.3402/JEV.V3.23111.

[9] Walker SA, Kennedy MT, Zasadzinski JA. Encapsulation of bilayer vesicles by self-assembly. Nature 1997;387:61–4. https://doi.org/10.1038/387061A0.

[10] Edwards DA, Schneck F, Zhang I, Davis AMJ, Chen H, Langer R. Spontaneous vesicle formation at lipid bilayer membranes. Biophysical Journal 1996;71:1208. https://doi.org/10.1016/S0006-3495(96)79334-7.

[11] Deville S, Berckmans P, van Hoof R, Lambrichts I, Salvati A, Nelissen I. Comparison of extracellular vesicle isolation and storage methods using high-sensitivity flow cytometry. PLOS ONE 2021;16:e0245835. https://doi.org/10.1371/JOURNAL.PONE.0245835.

[12] Torres Crigna A, Fricke F, Nitschke K, Worst T, Erb U, Karremann M, et al. Inter-Laboratory Comparison of Extracellular Vesicle Isolation Based on Ultracentrifugation. Transfusion Medicine and Hemotherapy : Offizielles Organ Der Deutschen Gesellschaft Fur Transfusionsmedizin Und Immunhamatologie 2021;48:48–59. https://doi.org/10.1159/000508712.

[13] Zonneveld MI, Brisson AR, van Herwijnen MJC, Tan S, van de Lest CHA, Redegeld FA, et al. Recovery of extracellular vesicles from human breast milk is influenced by sample collection and vesicle isolation procedures. Journal of Extracellular Vesicles 2014;3. https://doi.org/10.3402/JEV.V3.24215.

[14] Kalra H, Adda CG, Liem M, Ang CS, Mechler A, Simpson RJ, et al. Comparative proteomics evaluation of plasma exosome isolation techniques and assessment of the stability of exosomes in normal human blood plasma. Proteomics 2013;13:3354–64. https://doi.org/10.1002/PMIC.201300282.

[15] Liangsupree T, Multia E, Riekkola ML. Modern isolation and separation techniques for extracellular vesicles. Journal of Chromatography A 2021;1636:461773. https://doi.org/10.1016/J.CHROMA.2020.461773.

[16] Prados-Rosales R, Baena A, Martinez LR, Luque-Garcia J, Kalscheuer R, Veeraraghavan U, et al. Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. Journal of Clinical Investigation 2011;121:1471–83. https://doi.org/10.1172/JCI44261.

[17] Rath P, Huang C, Wang T, Wang T, Li H, Prados-Rosales R, et al. Genetic regulation of vesiculogenesis and immunomodulation in Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences of the United States of America 2013;110. https://doi.org/10.1073/pnas.1320118110.

[18] Prados-Rosales R, Carreño LJ, Batista-Gonzalez A, Baena A, Venkataswamy MM, Xu J, et al. Mycobacterial membrane vesicles administered systemically in mice induce a protective immune response to surface compartments of mycobacterium tuberculosis. MBio 2014;5. https://doi.org/10.1128/mBio.01921-14.

[19] Prados-Rosales R, Brown L, Casadevall A, Montalvo-Quirós S, Luque-Garcia JL. Isolation and identification of membrane vesicle-associated proteins in Gram-positive bacteria and mycobacteria. MethodsX 2014;1:e124–9. https://doi.org/10.1016/j.mex.2014.08.001.

[20] Lee J, Kim SH, Choi DS, Lee JS, Kim DK, Go G, et al. Proteomic analysis of extracellular vesicles derived from Mycobacterium tuberculosis. Proteomics 2015;15:3331–7. https://doi.org/10.1002/pmic.201500037.

[21] Athman JJ, Wang Y, McDonald DJ, Boom WH, Harding C v., Wearsch PA. Bacterial Membrane Vesicles Mediate the Release of Mycobacterium tuberculosis Lipoglycans and Lipoproteins from Infected Macrophages . The Journal of Immunology 2015;195:1044–53. https://doi.org/10.4049/jimmunol.1402894.

[22] Dauros Singorenko P, Chang V, Whitcombe A, Simonov D, Hong J, Phillips A, et al. Isolation of membrane vesicles from prokaryotes: a technical and biological comparison reveals heterogeneity. Journal of Extracellular Vesicles 2017;6. https://doi.org/10.1080/20013078.2017.1324731.

[23] Palacios A, Sampedro L, Sevilla IA, Molina E, Gil D, Azkargorta M, et al. Mycobacterium tuberculosis extracellular vesicle-associated lipoprotein LpqH as a potential biomarker to distinguish paratuberculosis infection or vaccination from tuberculosis infection. BMC Veterinary Research 2019;15:188. https://doi.org/10.1186/s12917-019-1941-6.

[24] Chiplunkar SS, Silva CA, Bermudez LE, Danelishvili L. Characterization of membrane vesicles released by Mycobacterium avium in response to environment mimicking the macrophage

phagosome. Future Microbiology 2019;14:293–313. https://doi.org/10.2217/FMB-2018-0249/ASSET/IMAGES/LARGE/FIGURE6.JPEG.

[25] Gupta S, Rodriguez GM. Isolation and Characterization of Extracellular Vesicles Produced by Iron-limited Mycobacteria. JoVE (Journal of Visualized Experiments) 2019;2019:e60359. https://doi.org/10.3791/60359.

[26] Gupta S, Palacios A, Khataokar A, Weinrick B, Lavín JL, Sampedro L, et al. Dynamin-like proteins are essential for vesicle biogenesis in Mycobacterium tuberculosis. BioRxiv 2020:2020.01.14.906362. https://doi.org/10.1101/2020.01.14.906362.

[27] Marsollier L, Brodin P, Jackson M, Korduláková J, Tafelmeyer P, Carbonnelle E, et al. Impact of Mycobacterium ulcerans biofilm on transmissibility to ecological niches and Buruli ulcer pathogenesis. PLoS Pathogens 2007;3:0582–94. https://doi.org/10.1371/journal.ppat.0030062.

[28] Ziegenbalg A, Prados-Rosales R, Jenny-Avital ER, Kim RS, Casadevall A, Achkar JM. Immunogenicity of mycobacterial vesicles in humans: Identification of a new tuberculosis antibody biomarker. Tuberculosis 2013;93:448–55. https://doi.org/10.1016/j.tube.2013.03.001.

[29] White DW, Elliott SR, Odean E, Bemis LT, Tischler AD. Mycobacterium tuberculosis Pst/SenX3-RegX3 regulates membrane vesicle production independently of ESX-5 activity. MBio 2018;9. https://doi.org/10.1128/MBIO.00778-18/SUPPL_FILE/MBO003183934ST2.DOCX.

[30] Kumar S, Mittal E, Deore S, Kumar A, Rahman A, Krishnasastry M v. Mycobacterial tlyA gene product is localized to the cell-wall without signal sequence. Frontiers in Cellular and Infection Microbiology 2015;5:60. https://doi.org/10.3389/FCIMB.2015.00060.

[31] Jurkoshek KS, Wang Y, Athman JJ, Barton MR, Wearsch PA. Interspecies Communication between Pathogens and Immune Cells via Bacterial Membrane Vesicles. Frontiers in Cell and Developmental Biology 2016;4:125. https://doi.org/10.3389/fcell.2016.00125.

[32] Athman JJ, Sande OJ, Groft SG, Reba SM, Nagy N, Wearsch PA, et al. Mycobacterium tuberculosis Membrane Vesicles Inhibit T Cell Activation . The Journal of Immunology 2017;198:2028–37. https://doi.org/10.4049/jimmunol.1601199.

[33] Lucas M, Ryan JM, Watkins J, Early K, Kruh-Garcia NA, Mehaffy C, et al. Extraction and Separation of Mycobacterial Proteins. Methods in Molecular Biology 2021;2314:77–107. https://doi.org/10.1007/978-1-0716-1460-0_3.

[34] Benedikter BJ, Bouwman FG, Vajen T, Heinzmann ACA, Grauls G, Mariman EC, et al. Ultrafiltration combined with size exclusion chromatography efficiently isolates extracellular vesicles from cell culture media for compositional and functional studies. Scientific Reports 2017 7:1 2017;7:1–13. https://doi.org/10.1038/s41598-017-15717-7.

[35] Wallace E, Hendrickson D, Tolli N, Mehaffy C, Peña M, Nick JA, et al. Culturing Mycobacteria. Methods in Molecular Biology 2021;2314:1–58. https://doi.org/10.1007/978-1-0716-1460-0_1.

[36] Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 2012;9:671–5. https://doi.org/10.1038/NMETH.2089.

Chapter 3 Asymmetric Flow Field-Flow Fractionation for Mtb EV Separation

3.1 Introduction

Heterogeneity in extracellular vesicle (EV) populations is well documented in eukaryotic EV biology; it is primarily attributed to the existence of multiple biogenesis pathways, resulting in apoptotic bodies, microvesicles, and exosomes [1]. Originally, this diversity was thought to be represented due to variation in size depending on the mechanism of origin [2,3]. As separation and evaluation techniques evolved, so too did the understanding that size alone does not distinguish one EV population from another [1,3]. However, defining the functional characteristics of different EV populations remains challenging for a number of reasons, including the lack of subtype specific markers. Still, separation of populations based on size does provide insight into the functional differences among EV subpopulations and highlights the necessity to consider various classification systems for these different EV types [4].

Bacterial EV biogenesis remains elusive for Gram positive and mycobacterial species. Although distinct biogenesis pathways have yet to be described, there are multiple independent mechanisms reported to influence EV release in a variety of species including *Mycobacterium tuberculosis* [5–7]. Additionally, biophysical EV heterogeneity within a bacterial culture has been established through atomic force microscopy [8]. The presence of multiple subpopulations of Mtb EVs may help explain the demonstrated heterogeneity in composition and functions across publications, especially if certain enrichment techniques bias toward or against certain EV subsets. Given the broad range of 50-300 nm for Mtb EVs [9], variation based on size could be a significant factor impacting the various separation techniques. As with eukaryotic EV biology, assessing Mtb EV heterogeneity based on size may provide vital information for understanding the potential for multiple biogenesis regulation mechanisms and functional consequences.

Asymmetric flow field-flow fractionation (AF4) is an emerging technique for EV separation and characterization, especially when coupled with orthogonal sizing techniques like multi-angle light

scattering (MALS) or dynamic light scattering (DLS) [10]. AF4 has been instrumental in pushing the understanding eukaryotic EV heterogeneity forward, including the discovery of human exomeres [4]. One advantage to AF4 over size exclusion chromatography is the ability to adjust parameters to change separation efficiency based on multiple flow rates rather than be reliant on a static resin. In fact, AF4 has been used to separate biological content as large as whole cells to as small as EVs, viruses, and even ribosomes [11]. **Figure 3.1** gives an overview of how analytes are separated by AF4.

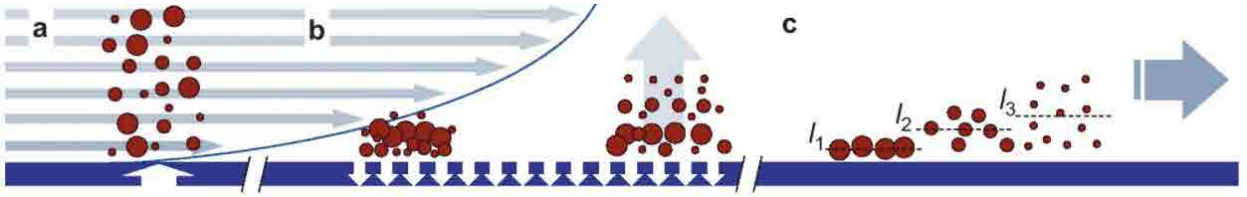


Figure 3.1 Asymmetric flow field-flow fractionation. A) The analytes are injected into the AF4 channel where the eluent moves by laminar parabolic flow with layers of increasing speed to the center. The bottom channel wall is covered by a permeable membrane over a ceramic frit while the top is solid. B) The analytes are focused by crossflow resulting in accumulation against the channel wall, then analytes diffuse against the crossflow depending on their diffusion coefficients. C) Analytes elute smallest to largest based on size, with the smallest components reaching the highest velocity based on diffusion away from the channel wall. This figure is adapted from [12]; the final version is free according to the Creative Commons CC-BY license.

AF4 was used to obtain high-resolution separation of Mtb EVs to explore the hypothesis that Mtb EVs are heterogeneous in protein composition based on their size. This heterogeneity is suspected to contribute to variation seen in the literature across different Mtb EV separation techniques. In addition, AF4 allows a clearer understanding of what proteins are truly vesicle associated versus co-enriched as an artifact of the separation techniques.

3.2 Materials and Methods

Several of the following sections appear in Chapter 2 and are copied verbatim for ease of reference.

*3.2.1 Media Preparation (equivalent to section **2.2.1** Media Preparation)*

3.2.1.1 7H11 + 10% oleic albumin dextrose catalase (OADC) medium:

Agar plates were prepared by stirring 21 g of 7H11 dehydrated agar (Middlebrook) into 900 mL of water. Once the dehydrated powder was thoroughly mixed, 5 mL of glycerol was added. The mixture was

autoclaved at 121 °C for 30 min with slow exhaust. The sterile medium was cooled in a 55 °C using a water bath, prior to the addition of 100 mL sterile OADC solution in a biosafety cabinet (BSC). Approximately 140 mL liquid agar per plate was poured into seven sterile plates (15 x 150 mm) and allowed to cool. Once the agar solidified, plates were transferred to 37 °C for 12 h to ensure sterility, and then stored at 4 °C until used.

3.2.1.2 Glycerol alanine salts (GAS) medium

Liquid medium was prepared by adding the listed ingredients, in order, per liter of water while stirring: 0.3 g Bacto™ Casitone (Life Technologies), 0.05 g ferric ammonium citrate, 4.0 g potassium phosphate dibasic anhydrous, 2.0 g citric acid, 1.0 g L-alanine, 1.2 g magnesium chloride, 0.6 g potassium sulfate, and 2.0 g ammonium chloride. Once fully dispersed, 1.8 mL 10 M sodium hydroxide was added followed by 10.0 mL glycerol. The pH was measured and adjusted to 6.6 with the addition of 10 M sodium hydroxide as needed. The medium was then autoclaved at 121 °C for 1 h with slow exhaust. Medium was stored at room temperature until used.

3.2.1.3 Nutrient agar medium

Agar plates were prepared by stirring 23 g dehydrated nutrient agar (Difco) into 1 L of water. Once completely in solution, the media was autoclaved at 121 °C for 30 min for slow exhaust then cooled to 55 °C using a water bath. Approximately 50 mL liquid agar per plate was poured into twenty sterile plates (15 x 100 mm) and allowed to cool. Solidified plates were transferred to 37 °C for 12 h to ensure sterility, and then stored at 4 °C until used.

*3.2.2 Mtb Culture and Culture Filtrate Protein Generation (equivalent to section **2.2.2** Mtb Culture and Culture Filtrate Protein Generation)*

All procedures involving live, virulent *Mycobacterium tuberculosis* were performed in a biosafety level 3 laboratory (BSL-3) at Colorado State University. These cultivation and CFP generation methods were adapted from [13, 14].

3.2.2.1 Bacterial culture

M. tuberculosis strain H37Rv was plated using 200 µL of frozen glycerol stock on 7H11 + 10% OADC agar plates (15 x 150 mm). After four weeks of growth at 37 °C, the bacterial lawn was collected using a sterile cell scraper with 25 cm handle and 3 cm blade (Falcon). The cells from one plate were transferred using the cell scraper to a Fernbach flask containing 900 mL GAS medium, plugged with a sterile roll cotton wrapped in cheese cloth and covered with aluminum foil. The culture was incubated for two weeks at 37 °C on an orbital platform shaker at 65 rpm. Bacteria from one flask were swirled and allowed to settle for 5 min to create a loose pellet of cells in the center of the flask. Approximately $1/20^{th}$ of the amassed cells was transferred with a pipette to twenty 1 L roller bottles with 400 mL GAS medium. Roller bottles were incubated at 37 °C on a roller apparatus rotating at 1.65 rpm for two weeks. At every manipulation, the culture was checked for contamination visually and by plating one loopful of culture on a nutrient agar plate with a 48 h incubation at 37 °C.

3.2.2.2 CFP harvest and concentration

CFP was separated from bacterial cells by 0.2 µm filtration with a VacuCap 90 vacuum filtration device (Pall Corporation). Sterile CFP was removed from BSL-3 for further processing at BSL-2. A 2 L stirred-cell ultrafiltration unit was assembled with a 150 mm diameter 5 kDa MWCO filter (Millipore Sigma). The filter was first prepared by equilibration in 90% isopropyl alcohol for 10 min followed by 30 min in water. Eight liters of CFP combined from 20 roller bottles was concentrated (40-fold) to 200 mL by stirred-cell ultrafiltration, using a pressure reservoir connected to compressed nitrogen kept at 4 °C. Buffer exchange (20-fold) was performed with 4 L of 10 mM ammonium bicarbonate. Once reduced to approximately 100 mL, the CFP was recovered from the stirred-cell and filtered with a 0.2 µm PES 250 mL unit. Total protein was determined using the bicinchoninic acid assay (BCA) (Pierce™ Thermo Scientific) in a 96-well format following the manufacturer recommendations with standards diluted in 10 mM ammonium bicarbonate. CFP was then qualified by silver stain and Western blots (see **3.2.4.4** Western blot) for the presence of DnaK (Rv0350), PstS1 (Rv0934), GroES (Rv3418c), Ag85 complex

(Rv3804c, Rv1866c, Rv0129c), and SodA (Rv3846). GroEL2 (Rv0440), a cell lysis marker, was used as a negative control Western blot. If any of the expected markers were absent or GroEL2 was present in moderate intensity, the CFP was discarded. CFP was stored at 4 °C prior to use.

3.2.2.3 CFP ultrafiltration

Ultrafiltration was performed using a 100 kDa MWCO Centricon Plus -70 centrifugal filter (Millipore Sigma). The filter was prepared by adding 25 mL phosphate buffered saline (1XPBS) (0.144 g/L potassium dihydrogen phosphate, 9.0 g/L sodium chloride, 0.795 g/L disodium phosphate, pH 7.3-7.5) to the top of the unit followed by centrifugation for 10 min at 2,800 x g, 4 °C. Residual 1X PBS was discarded from the unit. CFP was added to the top of the filter unit up to 60 mL at a time and centrifuged in 10 min increments at 2,800 x g at 4 °C. The first 50 mL of 100 kDa CFP flow through (100FT) was saved at 4 °C for downstream analysis. Once the CFP was reduced to the volume of the filter, 60 mL 1X PBS was added to the unit and concentrated for a total of 5 washes. One milliliter of 1X PBS was added to each filter column directly followed by 3 mL 1X PBS in the top of the unit. The filter was inverted into the collection cup for a recovery spin at 57 x g for 5 min at 4 °C. An additional 1 mL of 1X PBS was added to each filter and allowed to sit while the retentate was transferred out of the recovery cup. A second recovery spin was performed and that retentate pooled with the first for a total of approximately 7 mL retentate. The 100 kDa retentate (100R) was passed through a 0.8 µm/0.2 µm dual PES filter (Pall Corporation) using a 10 mL syringe. Total protein for the 100R and 100FT was determined by BCA using a 96-well plate format following the manufacturer recommendations with standards diluted in 1X PBS. The filtered 100 kDa CFP retentate was stored at 4 °C prior to use.

*3.2.3 Mtb EV Enrichment (equivalent to section **2.2.3.2** qEV size exclusion chromatography)*

3.2.3.2 qEV size exclusion chromatography

qEVoriginal 35 nm columns were used with an automatic fraction collector (AFC) (Izon) following the manufacturer recommendations. After equilibrating to room temperature, the qEV column was installed on the AFC and flushed with 60 mL 1X PBS. The AFC was programmed to collect 20 x 500 µL fractions.

Three mg of 100R brought to 0.5 mL with 1X PBS was applied to the top of the column. Once the sample had run into the column, 15 mL 1X PBS was added to the top of the column. The AFC collected fractions after the column void volume had flowed through. The first three fractions collected contain the qEV enriched EVs (qEV-EVs) and were pooled together. After collection was complete, the qEV column was cleaned with 0.5 mL of 0.5 M NaOH followed by 60 mL 1X PBS. Once all three replicates had been run, the column was stored in 0.05% NaN_3 at 4 °C. Each column can be used a total of 5 times. One column was used per 100R batch to avoid lot-to-lot cross contamination.

3.2.4 Asymmetric Flow Field-Flow Fractionation (AF4)

This experiment was conducted in collaboration with Kim R. Williams' Lab at Colorado School of Mines. The Williams Lab performed the AF4 separation. Specific methodological details will be described in a future publication of this work.

3.2.4.1 Sample preparation

Three different 100R samples were used for this experiment. Based on AF4 volume limits, the desired protein concentration was 5 mg/mL. Each sample was further concentrated using a 4 mL 100 kDa MWCO Amicon Ultra centrifugal filter (Millipore Sigma). The filter was prepared by adding 4 mL 1X PBS to the top of the unit followed by centrifugation for 10 min at 2,800 x g, 4 °C. Residual 1X PBS was discarded from the unit. Ten milligrams of 100R was added to the unit and reduced to approximately 2 mL. The 100R was transferred to a new tube and the filter rinsed three times with 200 µL 1X PBS. The wash material was pooled with the 100R. Protein concentration was determined by diluting the material 1:10 in 1X PBS for BCA. Each sample was brought to 5.56 mg/mL with 1X PBS and stored at 4 °C.

3.2.4.2 AF4

Seven milligrams of three 100R concentrated samples at 5.56 mg/mL were given to the Williams Lab for AF4 fractionation. A total of 5.5 mg of sample were fractionated by multiple injections on the AF4 and resulting fractions were pooled by sample. Five fractions were collected per sample that were anticipated to correspond with F1: < 50 nm, F2: 50 – 70 nm, F3: 70 – 100 nm, F4: 100 – 130 nm, and F5: >130 nm.

The approximate final volumes per fraction were F1: 60 mL, F2: 10 mL, F3: 10 mL, F4: 10 mL, and F5: 70 mL. All fractions were stored at 4 °C.

3.2.4.3 Fraction concentration and protein measurement

Fractions were concentrated using 15 mL 3 kDa Amicon Ultra centrifugal filters (Millipore Sigma). The filter was prepared by adding 10 mL 1X PBS to the top of the unit followed by centrifugation for 10 min at 2,800 x g, 4 °C. Residual 1X PBS was discarded from the unit. Up to 10 mL of sample was added at a time to the filter unit and concentration was performed at 2,800 x g, 4 °C until the final volume was less than 1 mL. The retentate was transferred to a new tube and the filter rinsed with 200 µL 1X PBS. The wash material was pooled with the retentate and the final volume for each sample fraction was recorded. Protein concentration was determined for F1 diluting the material 1:10 in 1X PBS for BCA. The remaining four fractions were diluted 1:6 in 1X PBS for microBCA.

*3.2.5 Mtb EV and AF4 Fraction Analysis (equivalent to section **2.2.4** Mtb EV Analysis)*

3.2.4.1 Nanoparticle tracking analysis (NTA)

The concentration and size distribution of the enriched EVs were evaluated by NTA using the NanoSight NS300 (Malvern Panalytical) with an automatic syringe pump and software version 3.4. EV samples were diluted in 1X PBS such that the final concentration was in the range of 10^7-10^9 particles per mL, per instrument recommendations. The screen gain was set to 13, the camera level 13, and the syringe pump 10. The focus was adjusted as needed for each sample. Three 30 second videos were captured for each sample. Analysis settings included a screen gain of 10 and detection level 5 and reported an average of the three videos. Final sample particle concentration was calculated based on the NTA read out and the volume of sample diluted for measurement.

3.2.4.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Samples normalized by either total protein or NTA particle count were dried using a SpeedVac, and then resuspended in 10 µL of Laemmli buffer (5% 2-mercaptoethanol, 0.01% bromophenol blue, 10%

glycerol, 2% SDS, 63 nM Tris-HCL) and heated for 5 min at 100 °C. After a brief centrifugation to collect any condensation, samples were loaded on a NuPAGE Novex 4-12% Bis-Tris Gel (Life Technologies) in 1X NuPAGE MES SDS Running Buffer (Life Technologies). At least one lane contained 3-5 µL Precision Plus Protein™ Dual Color Standard (Bio-Rad) to serve as a molecular weight reference. Resolution was accomplished using constant voltage of 200 for 35 min.

3.2.4.3 Silver stain

All incubation steps occur rocking and at room temperature. Following SDS-PAGE resolution, the gel was removed from the cassette and incubated for 1 h in a solution of 40% methanol and 10% glacial acetic acid in water then decanted. Next the gel was covered with a solution of 5% methanol and 7% glacial acetic acid in water for 5 min then decanted. The proteins were then fixed by 2% glutaraldehyde in water for 5 min and decanted, followed by three water washes for 10 min each. Oxidation was performed by incubation in 0.17 mM dithiothreitol (DTT) for 5 minutes then decanted. Next, 6 mM silver nitrate was added for 5 min followed by three brief water washes. Approximately one-third the volume of developer solution (6 g Na_2CO_3 and 3-6 drops 37% formaldehyde in 200 mL water) was added to rinse any residual silver nitrate and discarded before the remaining two-thirds of the developer was allowed to incubate on the gel. Once the desired color saturation was reached, approximately 20 mL of 50% citric acid was used to stop the reaction. After 20 min the gel was transferred to water before imaging.

3.2.4.4 Western blot

Following SDS-PAGE resolution, the gel was removed from the cassette and transferred to 0.2 µm nitrocellulose (Bio-Rad) for 1 h at 50 V in transfer buffer (25 mM Tris base and 195 mM glycine, in 20% methanol). The membrane was removed and blocked with 2% bovine serum albumin in TBS-T (1.21 g/L Tris, 8.77 g/L NaCl, 0.5 mL/L 100% Tween 80, pH 7.4) for 1 h at room temperature or overnight at 4 °C. The blocking solution was decanted, and the membrane rinsed with TBS-T followed by the addition of primary antibody diluted in TBS-T. Primary antibody was incubated for 1 h at room temperature or overnight at 4 °C then decanted and the membrane rinsed three times with TBS-T. Alkaline-phosphatase

conjugated secondary antibody was diluted in TBS (1.21 g/L Tris, 8.77 g/L NaCl, pH 7.4) and added to the membrane at room temperature for 30 min. After removal of the secondary antibody, the membrane was rinsed for 5 min in TBS then exposed to the developer reagent (NBT/BCIP ready to use tablet, Roche, dissolved in 10 mL water). The reaction was stopped by the addition of water then the membrane was rinsed and allowed to air dry before imaging.

3.2.4.5 Transmission electron microscopy (TEM)

Vesicle samples were fixed by adding an equal volume of 4% TEM grade paraformaldehyde, final concentration of 2%, then stored at 4 °C for a minimum of 24 h. Formvar/carbon coated grids (Electron Microscopy Science) were glow discharged for 30 seconds before 10 µL of fixed EV sample was dropped onto the grid and allowed to sit for 5 min. Excess liquid was blotted away with filter paper. The grid was washed by floating on a 50 µL drop of ultrapure water for 30 seconds. Excess water was blotted away then the grid was stained by floating on a 50 µL drop of 2% uranyl acetate for 2 min. Excess stain was blotted away and the grid air dried. Grids were imaged on a JOEL JEM-2100F transmission electron microscope at 200 kV with spot size 5 and alpha 2.

3.2.6 Shotgun Proteomics

3.2.6.1 Trypsin digestion for peptides

Samples normalized by either total protein or NTA particle count were resolved by SDS-PAGE as described in section **3.2.4.2** Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) with the total run time at 200 V reduced to 10 min. The gel was removed from the cassette and washed in water for 5 min. SimplyBlue SafeStain (Invitrogen) was used to visualize the protein samples by incubating the gel in 50 mL stain at room temperature for a minimum of 1 h. After decanting the stain, the background was de-stained in water for a minimum of 1 h at room temperature. Individual sample lanes were excised from the gel and cut into approximately 1 mm^3 pieces then transferred to a de-plasticized 0.65 mL centrifuge tube. Two-hundred microliters of de-stain solution (60% acetonitrile (ACN) in 0.2 M ammonium bicarbonate) was added and the samples placed at 37 °C for a minimum of 1 h after which the

liquid was removed. This process was repeated at least once. In the event the gel pieces remained blue, they were rehydrated with 50 µL 0.2 M ammonium bicarbonate briefly, followed by an additional round of de-staining. Once the stain was removed, the gel pieces were dried with a SpeedVac and stored at 4 °C prior to trypsin digestion. Sequencing grade trypsin (Roche) was dissolved in 0.2 M ammonium bicarbonate at 12 µL solution per 1 µg trypsin. Trypsin was added directly to the gel pieces at a ratio of 50:1 (sample: trypsin) approximating 1E9 vesicles as equivalent to 5 µg protein. After the trypsin solution was absorbed by the gel pieces for five minutes, 150 µL of 0.2 M ammonium bicarbonate was used to completely cover the gel pieces and the samples were placed at 37 °C for 16 h. The next day the tryptic peptides were extracted by adding 100 µL of 60% ACN, 0.1% trifluoroacetic acid in water and incubated for 1 h at 37 °C. The solution was transferred to a new de-plasticized tube and the extraction repeated, pooling both supernatants prior to vacuum drying the extracted peptide solution. Dried peptides were suspended in Solvent A (3% ACN and 0.1% formic acid in water) and centrifuged at 13,000 x g for 5 min at 4 °C to pellet larger debris. The supernatant was carefully transferred into a new de-plasticized tube and quantitated using a NanoDrop UV-Vis spectrophotometer A205 scopes method. The samples were brought to 1 µg/µL with Solvent A then stored at -20 °C.

3.2.6.2 Liquid chromatography tandem mass spectrometry (LC-MS/MS)

Reverse phase chromatography was performed using water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). A total of 1 µg of peptides were purified and concentrated using an on-line enrichment column (Waters Symmetry Trap C18 100 Å, 5 µm, 180 µm ID x 20mm column). Subsequent chromatographic separation was performed on a reverse phase nanospray column (Waters, Peptide BEH C18; 1.7 µm, 75 µm ID x 150 mm column, 45 °C) using a 90 min gradient: 5 - 30% B over 85 minutes followed by 30 - 45% B over 5 minutes at a flow rate of 350 nL/min. Peptides were eluted directly into the mass spectrometer (Orbitrap Velos Pro, Thermo Scientific) equipped with a Nanospray Flex ion source (Thermo Scientific) and spectra were collected over a *m*/*z* range of 400–2000, positive mode ionization. Ions with charge state +2 or +3 were accepted for MS/MS using a dynamic exclusion limit of

2 MS/MS spectra of a given *m*/*z* value for 30 seconds (exclusion duration of 90 seconds). The instrument was operated in FT mode for MS detection (resolution of 60,000) and ion trap mode for MS/MS detection with a normalized collision energy set to 35%. Compound lists of the resulting spectra were generated using Xcalibur 3.0 software (Thermo Scientific) with a S/N threshold of 1.5 and 1 scan/group. Samples were randomized and each sample was injected three separate times. To monitor instrument performance, a blank injection followed by bovine serum albumin and another blank injection were run after every seven samples.

3.2.6.3 Database searching

Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by ProteoWizard (MSConvert; version 3.0). Raw data files were converted to mzXML and submitted to Sorcerer2 (Sage-N Research; version 5.1.1). All MS/MS analysis was performed using SEQUEST (Thermo Fisher Scientific, San Jose CA, USA; version 1.0). SEQUEST was set up to search the uniprot-proteome UP000001584 Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) database (May 24, 2021 update with 3,997 entries) with automatic decoy generation assuming the enzymatic digestion with trypsin (after Arg or Lys). The search was performed with a fragment ion mass tolerance of 1.00 Da and a parent ion tolerance of 20 PPM. Carbamidomethyl of cysteine was specified as a fixed modification and as a variable modification. Oxidation of methionine was also specified as a variable modification.

3.2.6.4 Individual injection protein identification

Scaffold (Proteome Software Inc., Portland, OR; version 5.0.1) was used to validate MS/MS based on peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability to achieve a false discovery rate (FDR) of less than 0.1% by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two peptides. Protein probabilities were assigned by the Protein Prophet algorithm [15]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins were annotated with

gene ontology (GO) terms from the National Center for Biotechnology Information (NCBI) (downloaded Nov 22, 2021)[16].

3.2.6.5 Multidimensional protein identification technology (MudPIT) protein identification

Following evaluation of individual LC-MS/MS injections, triplicate injections were combined into one sample using the MudPIT function in Scaffold (Proteome Software Inc., Portland, OR; version 5.0.1). Peptide identifications were accepted if they could be established at greater than 96.0% probability to achieve and FDR less than 0.1% by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two peptides. Protein probabilities were assigned by the Protein Prophet algorithm [15]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins were annotated with GO terms from NCBI (downloaded Nov 22, 2021)[16].

3.2.7 Data Analysis

3.2.7.1 Statistics and visualization

RStudio (2021.09.0 Build 351) using R for statistical computing (version 4.1.1) was used to generate graphs and plots. All coding for statistical analyses performed in RStudio will be provided in a supplementary file, uploaded with this dissertation. Prior to statistical tests for differences among means, samples were evaluated for normality by histogram, quantile-quantile (QQ) Plot, and Shapiro-Wilk's test. If normality was met, equal variance was determined by Bartlett's test. If normality and equal variance assumptions were met, ANOVA with post-hoc Tukey or Student's t-test was used. If either assumption was violated (p-value <0.05), the nonparametric Kruskal-Wallis test followed by pairwise Wilcoxon signed-rank test was performed. Total ion chromatograms were viewed using Xcaliber 3.0 software (Thermo Fisher Scientific). Additional statistical analysis was performed in Scaffold including ANOVA testing across different sample groups using the normalized spectral abundance factor (NSAF) to account

for peptides and spectral counts shared across proteins [17]. The Benjamini-Hochberg multiple tests correction was applied to control for type I error as recommended by the software.

3.2.7.2 Comparison of proteins with signal peptides

The uniprot-proteome UP000001584 Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) database (May 24, 2021 update with 3,997 entries) was uploaded to SignalP v. 6.0 [18]. The output was sorted based on predicted signal peptides including signal peptide (SP), the lipobox motif (LIPO), the twin-arginine translocase (TAT) motif, and no signal peptide. All proteins with a predicted signal sequence were searched in the MudPIT LC-MS/MS data.

3.3 Results and Discussion

3.3.1 Transmission Electron Microscopy

Visualization of each AF4 fraction by TEM (**Figure 3.2**) demonstrates closed-morphology vesicles in F2-F5, with no in-tact Mtb EVs in F1. Fractions 2-5 increase in vesicular diameter, as expected based on the AF4 technique. The qEV image demonstrates vesicles corresponding to F3-F4 in size indicating that qEV enrichment captures the middle range of vesicles. Variation in vesicle morphology is due to dehydration and drying that occurs during the preparation process. Vesicle concentration cannot be analyzed with this TEM preparation method because it relies on vesicles adhering to the grid, which will not capture all the vesicles in the sample and can vary based on a multitude of factors. These images verify AF4 as an effective technique for Mtb EV separation based on size.

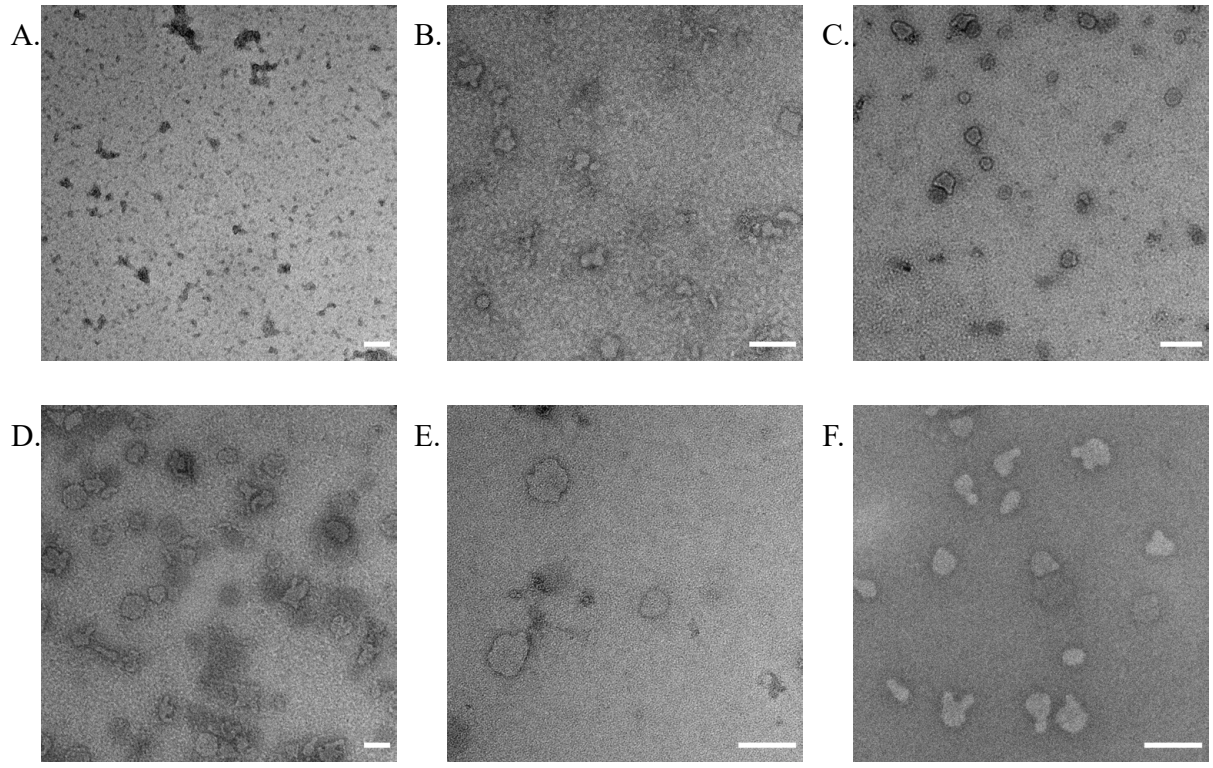


Figure 3.2 TEM visualization of AF4 and qEV fractions. (A) F1, (B) F2, (C) F3, (D) F4, (E) F5, and (F) qEV were pooled by sample and fixed for TEM. Scale bars represent 100 nm.

3.3.2 AF4 Protein and Vesicle Recovery

The protein and vesicle data associated with each fraction and qEV enrichment is reported in **Table 3.1**. The protein concentration and total protein recovered decreases over time during the AF4 sample elution, which is expected as the small soluble proteins elute first. The overall protein recovery varied up to 12% among the samples. The total protein recovered for sample 3 was higher than the injection load, however, the bulk of this was eluted in F1. The percentage of protein recovered by fraction also varies slightly across the samples. This indicates biological variation between each sample, which is expected. Individual AF4 injections (n = 6) within each sample were consistent in MALS intensity supporting technical reproducibility within the technique (data not shown).

Table 3.1 AF4 protein and vesicle recovery. Protein concentration and vesicle counts for AF4 fractions as well as qEV of each sample is reported.

Sample	Fraction	Particles/µL	Mean (nm)	Mode (nm)	Total Particles	Protein (µg/µL)	Total Protein (µg)	Total AF4 Protein Recovered (µg)	Percentage of Total AF4 Protein	Total AF4 Particles Recovered	Percentage of Total AF4 Particles	Particle:Protein (# Particles/ µg)
1	1	7.50E+06	101	100	6.90E+09	5.320	4894.40		98.06			
1	2	9.70E+06	76	86	9.22E+09	0.033	31.23		0.63		5.32	2.95E+08
1	3	5.40E+07	81	78	5.02E+10	0.036	33.05	4991.05	0.66	1.73E+11	28.98	1.52E+09
1	4	5.90E+07	95	95	5.31E+10	0.021	18.50		0.37		30.64	2.87E+09
1	5	5.60E+07	105	101	6.08E+10	0.013	13.88		0.28		35.06	4.38E+09
1	qEV	1.40E+08	96	91	2.10E+11	0.018	26.55					7.91E+09
2	1	7.00E+06	103	88	7.14E+09	3.985	4064.70		97.40			
2	2	2.80E+06	78	65	2.37E+09	0.039	32.68		0.78		1.89	7.24E+07
2	3	4.25E+07	73	71	4.93E+10	0.031	35.48	4173.12	0.85	1.25E+11	39.42	1.39E+09
2	4	3.80E+07	84	85	3.42E+10	0.024	21.71		0.52		27.35	1.58E+09
2	5	4.00E+07	102	95	3.92E+10	0.019	18.56		0.44		31.34	2.11E+09
2	qEV	5.80E+07	111	93	8.70E+10	0.023	34.79					2.50E+09
3	1	3.00E+07	166	134	2.96E+10	5.562	5478.57		98.43			
3	2	6.00E+06	106	105	4.92E+09	0.042	34.28		0.62		1.54	1.44E+08
3	3	5.30E+07	88	76	4.88E+10	0.033	29.95	5566.08	0.54	3.19E+11	15.27	1.63E+09
3	4	1.20E+08	88	82	1.03E+11	0.016	13.66		0.25		32.32	7.56E+09
3	5	1.90E+08	105	103	1.62E+11	0.011	9.62		0.17		50.87	1.69E+10
3	qEV	4.90E+07	118	107	7.35E+10	0.013	19.62					3.75E+09

As with TEM, the overall sizes estimated by NTA indicate an appropriate increase across the fractions (except for sample 3 F2 which was abnormally high). The number of particles by NTA increases from F2-F5 for all samples. **Figure 3.3** shows the NTA profiles for each sample including qEV. The high particle numbers associated with F1 are likely due to protein aggregates and complexes, rather than vesicles, based on TEM and the larger error bars and multimodal peak distribution. F2 also demonstrates wider peak distribution, likely due to residual larger soluble proteins. Several of the limitations of NTA are illustrated in these data. First, NTA cannot distinguish between vesicles and other aggregates of protein complexes of similar size, which is likely contributing to the reports for all three F1. Additionally, the sizes reported can only be considered estimates, not exact [19]. The protein and vesicle recovery for qEV samples cannot be directly compared to the fractions because less 100R was loaded. The reported NTA profile, however, can be compared and demonstrates similar sizing to the EVs recovered in F2-F5. The peak is slightly broader which is expected since AF4 fractions represent only a portion while qEV is all of the recovered EVs from that method. In combination with TEM, these data indicate successful fraction by AF4, with no vesicles in F1 and little soluble protein in the higher fractions.

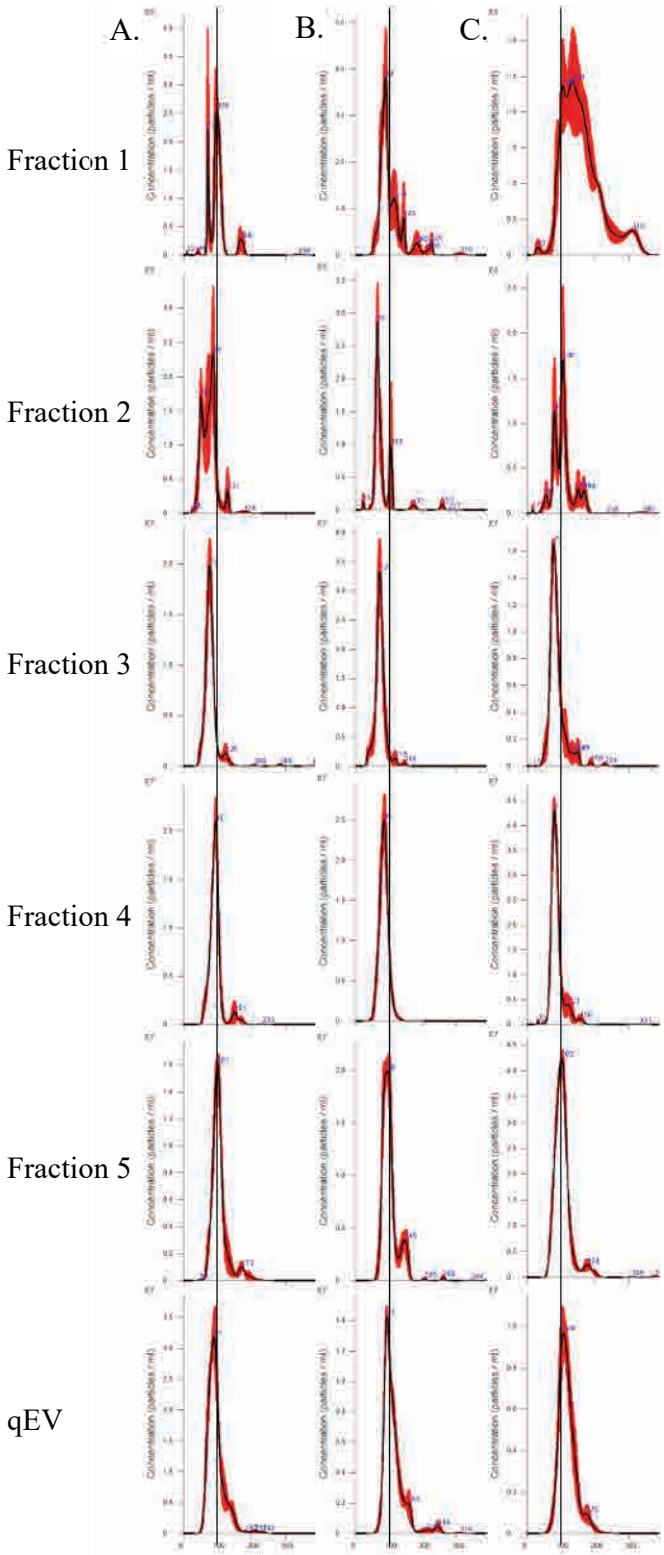


Figure 3.3 AF4 fraction NTA profiles. The NTA profile for (A) sample 1, (B) sample 2, and (C) sample 3 increase from F1 on the top to F5 then qEV on the bottom. Red error bars indicate +/- 1 standard error of the mean. A vertical line is drawn for each sample set at 100 nm to emphasize the shift in EV size during elution.

3.3.3 Protein Visualization by Silver Stain and Western Blot

The total protein profile is visualized for sample 1 and sample 3 in **Figure 3.4**; sample 2 is not included due to low yield resulting in prioritization for LC-MS/MS. Normalization for the silver stained gels and Western blots was determined by considering the protein and particle concentrations as established in Appendix I **Figure AI.1**; in this case the quantities are cut in half based on total yield. The overall protein profiles fade as fraction number increases, matching the particle to protein ratio in **Table 3.1**. This again suggests the presence of non-vesicle associated proteins in the early fractions, with F1 having a similar banding pattern to the 100R, and F2 staining more strongly than the remaining fractions. The qEV material has banding patterns similar to the AF4 fractions where vesicles are present by TEM (F2-F5).

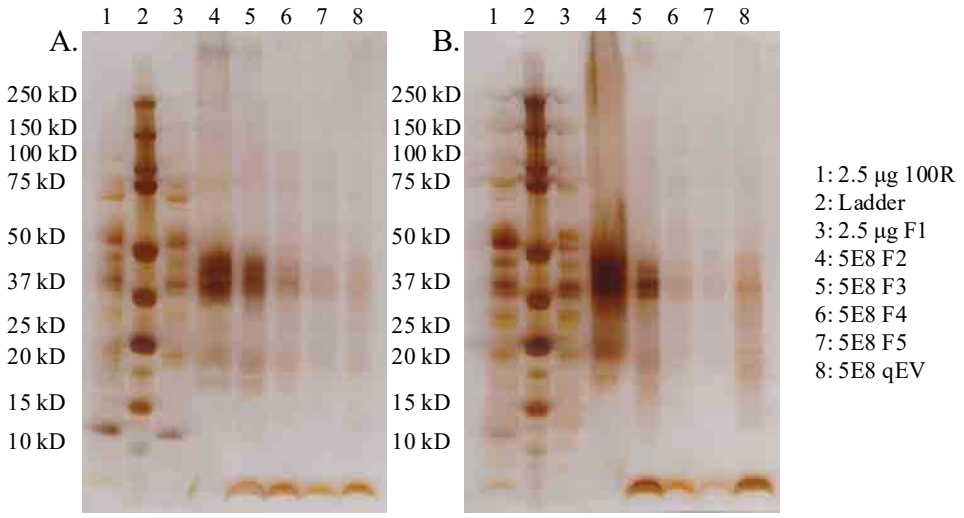


Figure 3.4 Silver stain of samples 1 & 3. Sample 1 (A) and Sample 3 (B) were resolved by SDS-PAGE and silver stained. The layout for each gel is identical (right).

AF4 fractions were probed with specific antibodies based on sample availability. Because sample 1 had enough material in F2 specifically for multiple full-length gels, it was used to demonstrate the total Western blot stain for the majority of the markers in **Figure 3.5**. Samples 2 and 3 were cut for each marker and were not probed for SodC in **Figure 3.6**.

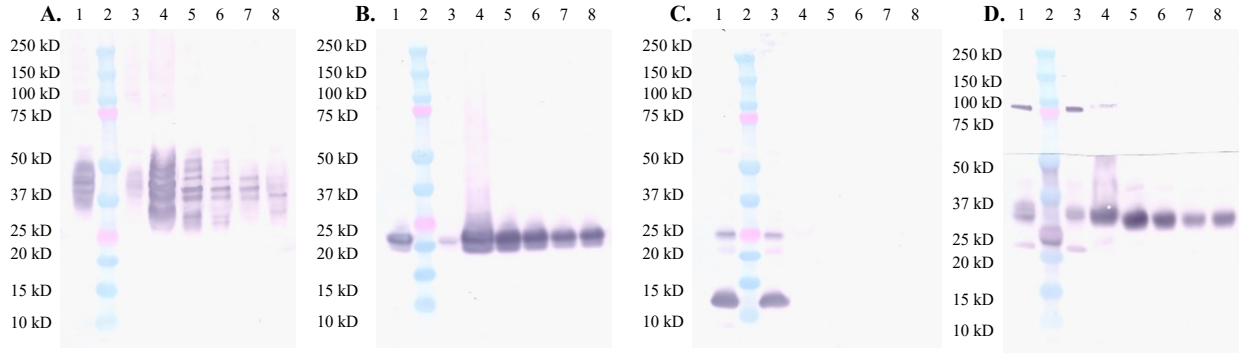


Figure 3.5 Western blots of sample 1. Sample 1 was interrogated by Western blot for antibodies against (A) LAM (30-45 kD), (B) LpqH (19 kD), (C) GroES (10 kD), (D, top) GlcB (80 kD) and (D, bottom) SodC (28 kD). The layout for each blot is identical (bottom).

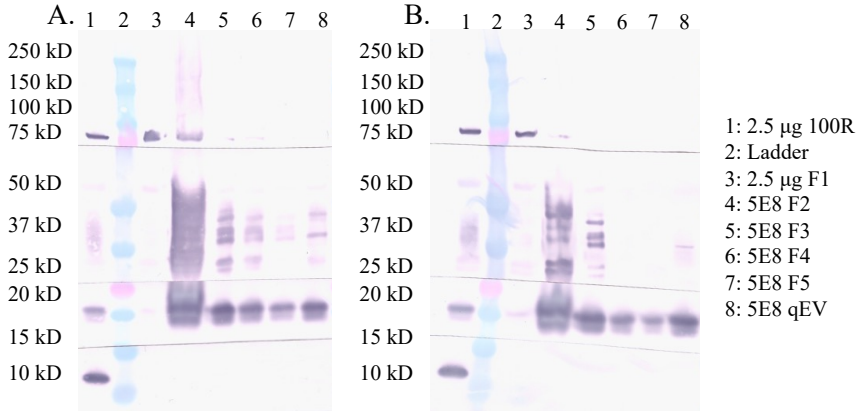


Figure 3.6 Western blots of samples 2 & 3. Samples (A) 2 & (B) 3 were interrogated by Western blot for antibodies against GlcB, LAM, LpqH, and GroES (top to bottom). The layout for each blot is identical (right).

For all three samples, LpqH is the strongest in F2 and wanes in intensity by F5, with very little in F1 (**Figure 3.5 B, Figure 3.6**). This matches the protein to particle ratio observed in **Table 3.1** and confirms Mtb EV presence in F2-F5. SodC has been shown to be enriched in Mtb EVs [20], and follows the same pattern as LpqH in sample 1 (**Figure 3.5 D**). LAM follows a similar pattern, but presents with greater sample to sample variation. The lack of LAM in F4 and F5 of sample 3 (**Figure 3.6 B**) indicates the presence of LAM negative Mtb EVs in these fractions. Given the sharp decrease across the fractions for the other two samples, it is possible there are LAM negative Mtb EVs; however, this is surprising given that the larger vesicles have more surface area for the LAM to be integrated. GroES, a negative marker for Mtb EVs, indicates a relatively pure population of vesicles in F2-F5 in sample 1 (**Figure 3.5**

C) as it is strong in the 100R and F1 but not detectable in the remaining fractions. Interestingly, it does not appear in F1 for samples 2 or 3 (**Figure 3.6**) indicating sample variation during the separation or fraction concentration process, given the strong intensity for all three 100R samples. GlcB is also variable by sample, with slight positives in F3 and F4 for sample 2 but no continuation after F2 for sample 1 and sample 3. Taken together these data support AF4 for separation of Mtb EVs and suggest proteomic and lipoglycan variation across the EVs that correlates with their size.

3.3.4 Validation of Normalization Strategy by Tryptic Digestion

Samples subjected to SDS-PAGE were normalized by either protein concentration or NTA particle count. Because 100R and F1 contain high amounts of soluble protein, NTA is not a very reliable way to measure EVs due to protein aggregation artifacts. Normalizing F2-F5 by protein for SDS-PAGE would result in poor resolution based on very high lipid presence. Samples were therefore normalized by either protein or particle count assuming that 5 µg protein is approximately 1E9 vesicles as discussed in Appendix I **Figure AI.2**. The validity of this approach was confirmed by peptide recovery following trypsin digest and gel extraction as measured by NanoDrop (Appendix II **Table AII7.1** and **Figure 3.7**). The data was not normally distributed with a Shapiro-Wilk test p-value of 0.04445 (Appendix II **Figure AII7.1**). Comparison of means was therefore performed using the non-parametric Kruskal-Wallis test. Samples digested based on protein have roughly the same peptide recovery as those digested based on particle count, with no significant difference among the means. Because the individual fraction with the lowest particle count, 1.5E9, resulted in similar recovery to the other starting amounts, it is likely there was increased trypsin autodigestion in this sample. There is variation by sample, with sample 3 having a larger spread and significantly higher mean than samples 1 and 2. However, overall the peptide recovery is very consistent and indicates that the normalization strategy is appropriate for comparing across fractions. Additionally, for LC-MS/MS 1 µg of peptides normalized to 1 µg/µL was injected to facilitate direct comparison across all the samples.

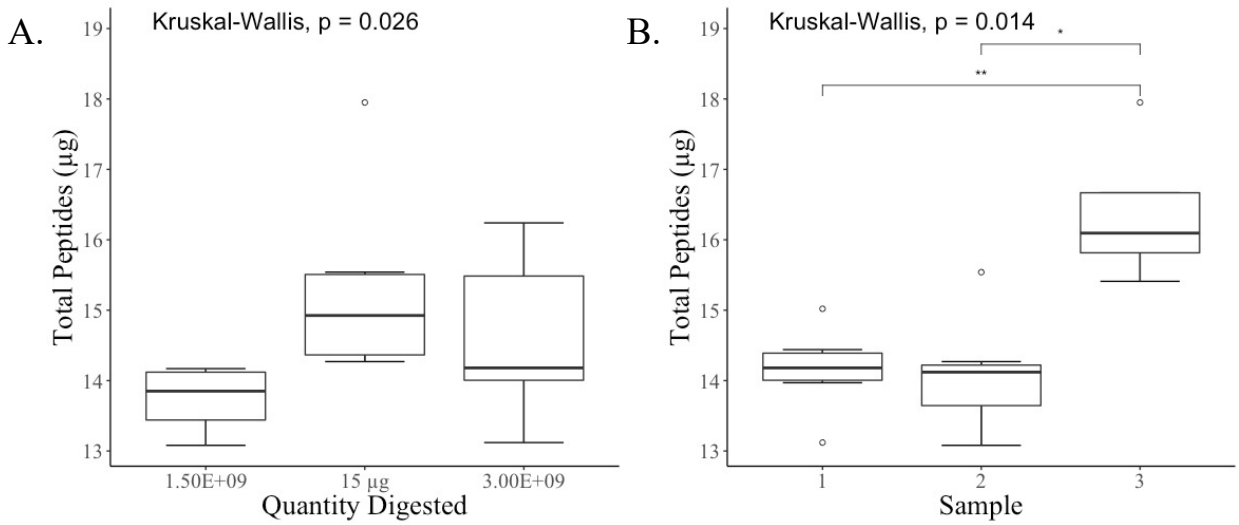


Figure 3.7 Plots of peptide recovery. Comparison of peptide recovery by (A) the quantity digested and (B) sample. Significance from pairwise Wilcoxon signed-rank test indicated as $p < 0.05$* and $p < 0.01$**.

Because the peptides from each sample were injected in triplicate, the individual injections were evaluated for quality and consistency (Appendix II section **7.2** Assessment of Individual LC-MS/MS Injections). LC-MS/MS instrument performance was also examined and determined to be consistent over time (Appendix II **Figure AII7.2**). Across the individual injections, there were 779 proteins identified including 14 decoys for a 1.8% decoy FDR from 395,494 spectra with a 0.1% FDR. In any of the statistical analyses with individual injections considered, if an outlier was found or the coefficient of variation (CV) among the three injections for any statistic was higher than 10%, the total ion chromatogram (TIC) for that individual injection was manually compared to the other TICs for that sample (Appendix II **Figure AII7.6**). Manual validation for all of those samples revealed a similar intensity and pattern by TIC, therefore all injections were included in MudPIT analyses.

3.3.5 Proteomic Biological Variation

Proteomic analysis indicates that there is variation among the three CFP batches used for AF4. Because this is a biological system, some variation is expected; however, too much inconsistency across samples can complicate data interpretation. Sample 3 had significantly more peptide recovery than samples 1 and 2, as illustrated in **Figure 3.7**. This may indicate technical differences in digestion or

extraction efficiency, contamination by keratin, an increase in trypsin autodigestion, or a biological increase in post-translational modifications (and not detectable during MS analysis). There is very little statistically significant difference across the samples when considering individual LC-MS/MS injection statistics (Appendix II **Figure AII7.4**). Differences between samples 1 and 3 as well as 2 and 3 meet the threshold of significance when comparing the number of proteins identified, with sample 1 having more than the other two. Differences between samples 1 and 3 meet the threshold of significance when comparing the percentage of spectra identified, again with sample 1 being higher than sample 3. Given that there is no difference in the number of spectra or the number of spectra identified across the samples, the variation in peptide recovery did not impact ionization by sample and likely reflects differences in the proteins present.

Following MudPIT combination of triplicate injections, less variation is observed across the samples. There were 811 proteins with at least 2 peptides with 18 decoys identified for a 2.3% decoy FDR from 387,026 spectra with 0.06% decoy FDR. Of the combined injection LC-MS/MS statistics (Appendix II **Table AII7.3 LC-MS/MS statistics for combined injections.**), all the data are normally distributed (Appendix II **Figure AII7.7**). The Bartlett's test for equal variance was performed and the p-values were not statistically significant. Therefore, ANOVA was used to compare means across samples for all data. **Figure 3.8** shows that there is no significant difference across means by sample for the combined injections data.

Over 50% (410) of the 811 proteins identified are present in all three samples **(Figure 3.9 A**). Sample 1 has 717 proteins (88.41%) identified, sample 2 has 545 (67.20%), and sample 3 has 485 (59.80%). Only 38 proteins (4.68%) are significantly different across the samples by NSAF (**Figure 3.9 B**). Although the samples do vary a bit by presence/absence of specific proteins, the majority of that variation includes proteins of low abundance that do not meet the level of significance.

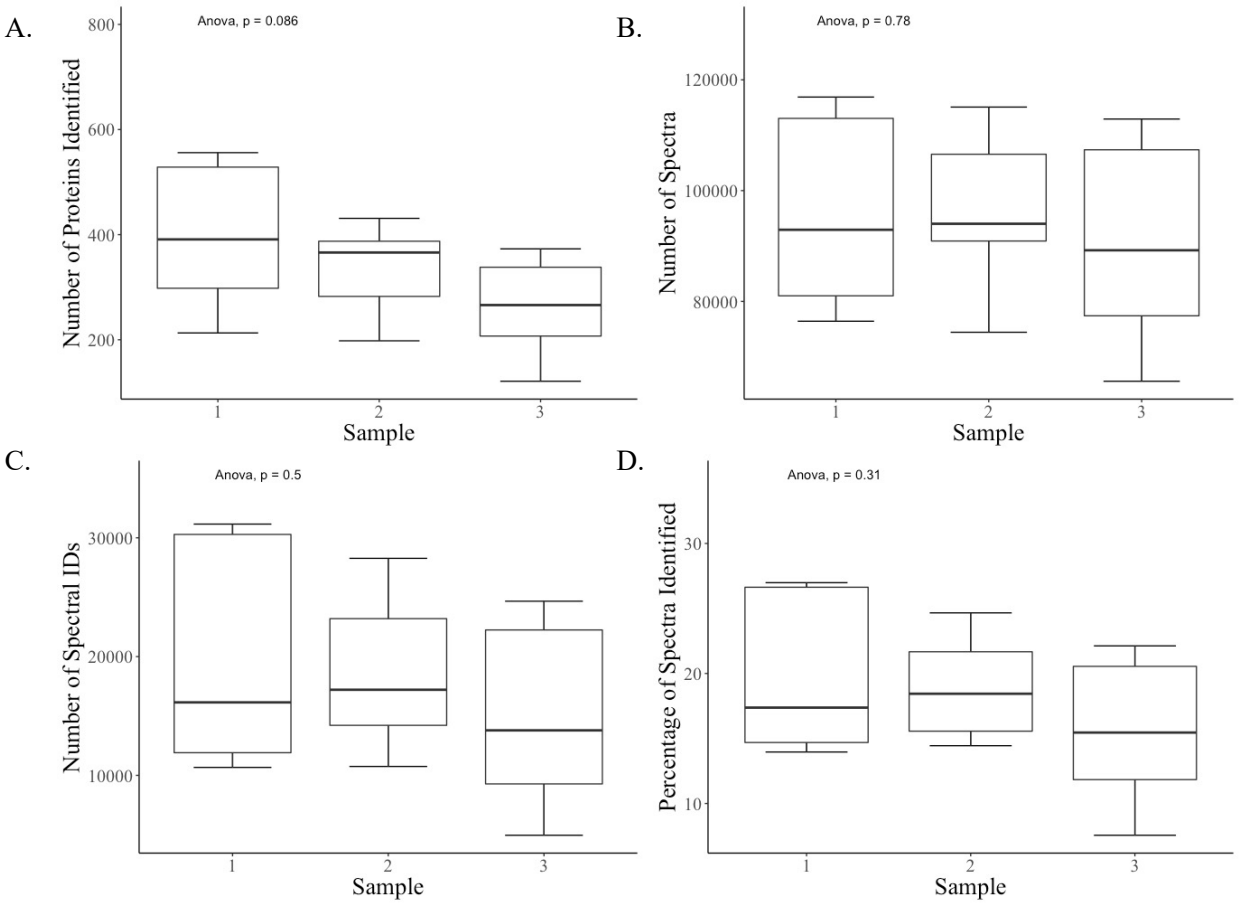


Figure 3.8 Boxplots of combined injection statistics by sample. Boxplots based on (A) number of identified proteins, (B) number of spectra, (C) number of spectral identifications, and (D) percentage of spectra identified for each sample.

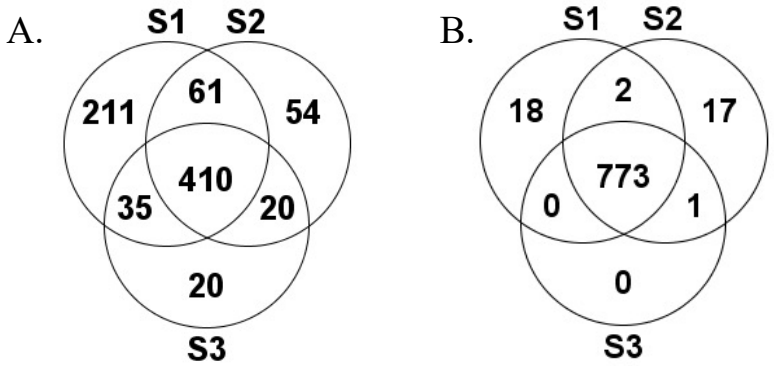


Figure 3.9 Number of proteins by sample. (A) Proteins found in each sample based on presence/absence and (B) proteins with a statistically significantly higher NSAF by ANOVA with the Benjamini-Hochberg multiple tests correction. S1 = Sample 1, S2 = Sample 2, and S3 = Sample 3.

Of the 38 proteins that vary significantly in NSAF across the samples (Appendix II **Table AII7.4**), 9 are only present in one sample (1.1% of 811), 8 are present in two out of three samples (0.98% of 811), and 21 are present in all three samples (2.6% of 811). No proteins are significantly higher in sample 3 than in either of the other two samples. When comparing the NSAF profiles for statistically significant proteins by sample (**Figure 3.10**), samples 1 and 3 are more similar to each other than either is to sample 2. The majority of the fractions cluster by sample, with sample 2 qEV as an outgroup of the sample 3 cluster and sample 1 fraction 5 clustering in the middle of sample 3.

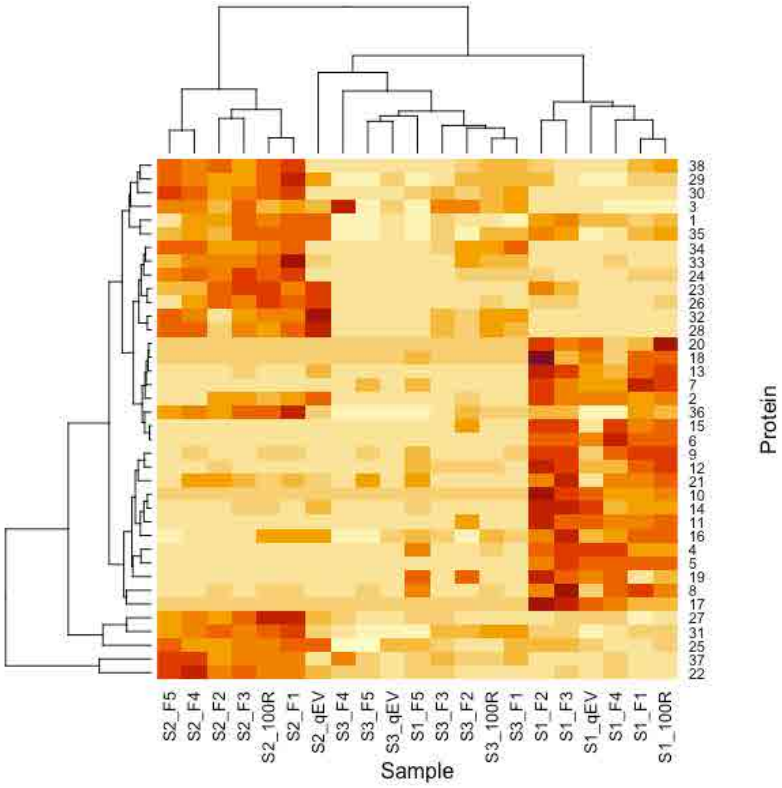


Figure 3.10 Heat map of proteins that vary across samples. The samples and statistically significant proteins were clustered based on NSAF, with the darker color indicating a higher protein abundance.

Sample 1 proteins that were high compared to the other samples are primarily located in the cell wall (61.1%), cytosol (83.3%) and membrane (77.8%) with very few secreted proteins (20.0%). Sample 2 also had proteins in the cell wall (47.1%), cytosol (41.2%), and membrane (70.6%) but included a higher proportion of secreted proteins (41.2%). Based on the cellular locations and functional categories associated with the proteins that differ across samples, it is likely that there was more cellular lysis

occurring in sample 1 than the other two samples. This lysis was not significant enough to be detected in the CFP QC western blots. Comparing the NSAF of three chaperonins, GroES, GroEL1, and GroEL2 (**Figure 3.11**) individual injections supports this hypothesis, with sample 1 have a higher proportion of these proteins. GroEL2 (**Figure 3.11 C**), in particular, is an indicator of cell lysis.

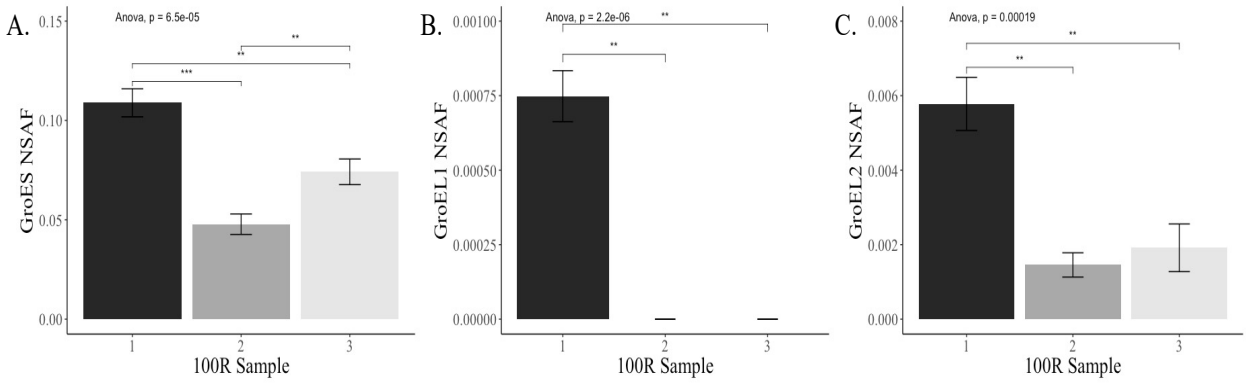


Figure 3.11 Chaperonin NSAF by sample. The mean NSAF for each 100R injection per sample is shown for (A) GroES, (B) GroEL1, and (C) GroEL2. Error bars represent standard deviation. Significance from Student's t-test is indicated as $p<0.01$** and $p<0.001$***.

Despite differential abundance in 38 proteins, the vast majority (95.3%) of identified proteins are present at similar levels across the three samples. The biological variation across the protein profiles of each sample are at an expected level and do not suggest that combining samples within fractions for analysis would be problematic. Instead, a relevant amount of biological variation is present which provides a test for the consistency of AF4 as a fractionation technique.

3.3.6 Proteomic Variation Across Fractions

In contrast to the comparison across samples, individual injection statistics evaluated by fraction contain a lot of statistically significant variation (Appendix II **Figure AII7.5**). While this variation decreases following MudPIT combination, there are still many significant differences (**Figure 3.12**). MudPIT reduces the complexity of the analysis, shifting the focus to each biological sample rather than technical injection replicates, which reduces the total number of samples to 3 per fraction. This reduction in sample size means the data cannot be evaluated using the same boxplot analyses and are presented as bar charts (**Figure 3.12**).

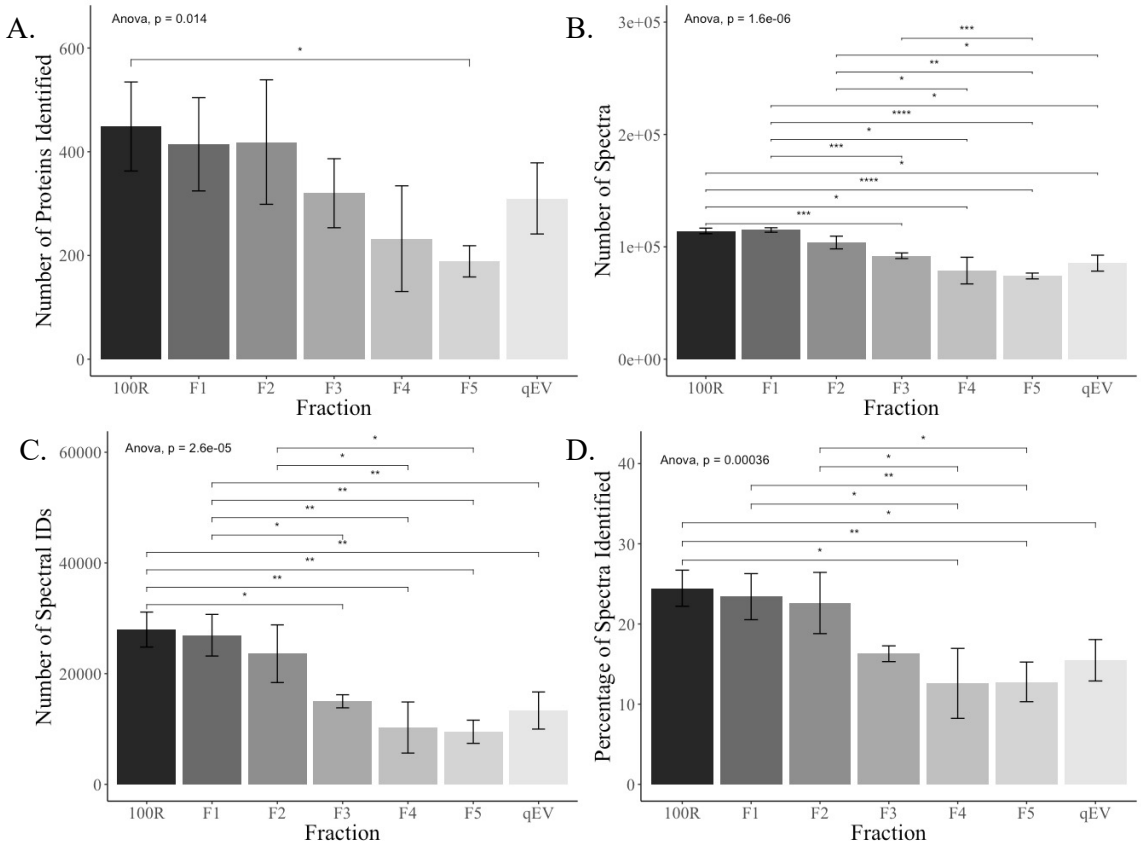


Figure 3.12 Bar charts of combined injection statistics by fraction. Bar charts are plotted based on (A) number of identified proteins, (B) number of spectra, (C) number of spectral identifications, and (D) percentage of spectra identified for each sample by fraction. Error bars represent the standard deviation. Significance from Student's t-test are indicated as $p<0.05$*, $p<0.01$**, $p<0.001$***, and $p<0.0001$****.

The number of proteins identified decreases as fraction number increases, which is expected since the soluble proteins elute first (**Figure 3.12**). The lack of significance between fractions for the number of proteins identified is likely due to the reduction in sample size (n = 3) and larger standard deviations as compared to the individual injection statistics (n = 9) (Appendix II **Figure AII7.5**). Additionally, the number of spectra, number of spectra identified, and percentage of spectra identified decreases (**Figure 3.12**). This could be due to an increase in lipids present in the sample because larger vesicles have more lipids and we see a higher particle to protein ratio in these fractions (**Table 3.1**). High lipid content may result in protein-lipid micelles, inhibiting proteolytic digestion. Although most lipids are non-polar and do not ionize well for LC-MS/MS, phosphatidylethanolamine (PE) is zwitterionic and has been shown to ionize well in positive ion mode [21]. PE is a major lipid in Mtb EVs, and an increase in this lipid may

cause ionization suppression of peptides [9]. Lipids can also interfere with A205 nm absorption (Appendix II **Table AII7.1**) which may be why the peptide digestion efficiencies were still similar (**Figure 3.7**) since samples loaded by protein had less lipid and samples loaded by vesicles had less protein. Overall, the combined LC-MS/MS statistics indicate compositional variation across the fractions.

ANOVA testing of protein NSAF by fraction demonstrates 139 proteins (17.1% of 811) with statistically significant variation (Appendix II **Table AII7.5**). Fifty-one of these proteins (36.7% of 139) are high in 100R and F1 but low in F2-F5 and qEV indicating that these proteins are not enriched in Mtb EVs. The second largest category are high in 100R, F1, and F2 but low in F3-F5 and qEV (20 proteins, 14.4% of 139). It is also likely that these proteins are not enriched in Mtb EVs. The cut off between F1 and F2 does not occur exactly where Mtb EVs begin to elute as more advanced optimization of the AF4 fractionation would be required to establish the exact time. This supports efficient separation of soluble proteins from MtbEVs by AF4, which can be visualized by heatmap in **Figure 3.13** where all 100R and F1 samples cluster together after the first major split from all the other fractions.

Fractions 2-5 and qEV do not cluster cleanly even when only considering the proteins that vary significantly in NSAF by fraction (**Figure 3.13**). Following the first split, only fraction 3 for all three samples cluster in the same node of the dendrogram (after 3 additional splits). While the impact of biological variation cannot be ruled out as a source of the inconsistent clustering, this data suggests that the global proteomic differences among fractions 2-5 and qEV are not very distinct by fraction. When 100R, F1, and qEV are removed from the ANOVA testing, only 15 proteins (2.3% of 641 proteins identified with 1.6% decoy FDR) are significantly different across fractions 2-5. A heatmap of all the proteins across the fractions (**Figure 3.14**) further supports that AF4 separates soluble proteins (100R and F1 cluster together), but the proteomic differences beyond those fractions do not result in distinct populations by fraction when considering the total protein.

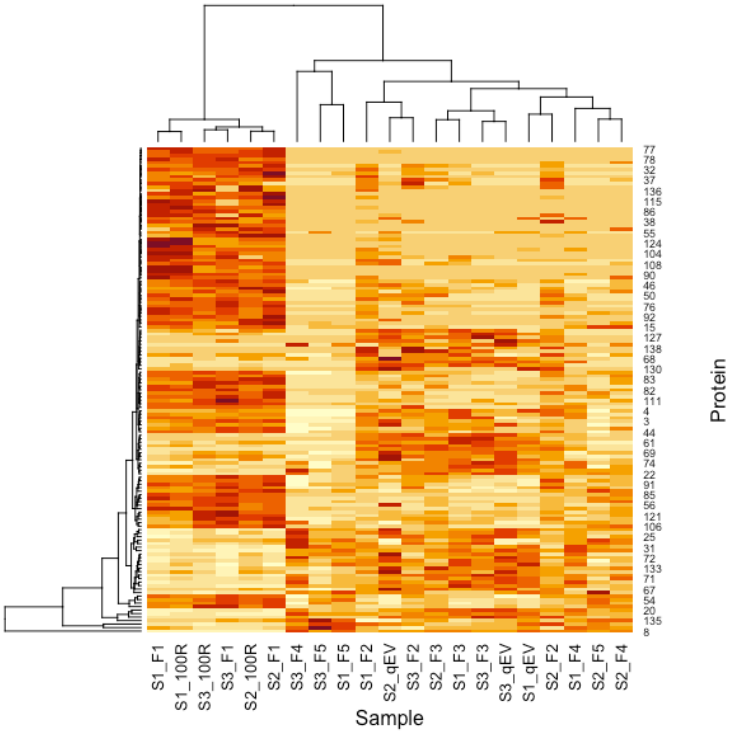


Figure 3.13 Heat map of protein subset that vary significantly across fractions. The samples and statistically significant proteins were clustered based on NSAF, with the darker color indicating a higher protein abundance.

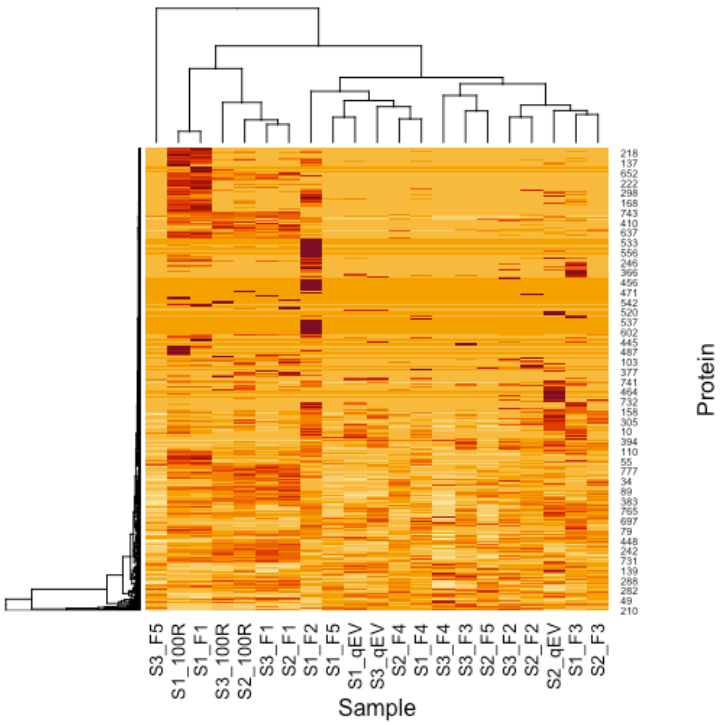


Figure 3.14 Heat map of all identified proteins. All identified Mtb proteins are clustered according to NSAF.

Although the total protein content across all the biological samples did not vary significantly by fraction, the silver stain and Western blot data (**Figure 3.4**, **Figure 3.5**, and **Figure 3.6**) still strongly indicate differences across the fractions. Lipoproteins, including LpqH and SodC, have been shown to be consistently enriched in Mtb EVs, and the Western blots suggest a higher concentration of those proteins in F2-F4. This led to the hypothesis that Mtb EVs separated by AF4 contain proteins that differ by physiochemical properties, such as post translational modifications (PTMs) like acylation and glycosylation. Many proteins secreted through the Sec pathway (and some through TAT) are known to be acylated and glycosylated, including LpqH and SodC [22]. Searching this proteomic dataset for proteins predicted to have a signal peptide using SignalP revealed 87 proteins with a predicted signal sequence (Appendix II **Table AII7.6**). Of those, 25 were identified as significantly different across the fractions (**Table 3.2**). When comparing all the fractions based on these 25 proteins (**Figure 3.15**), an important pattern emerges: F1 and F5 cluster apart from F2-F4. This pattern holds true when looking at all proteins identified with a lipobox motif just across AF4 fractions, even when considering proteins not significantly different across all the fractions (**Figure 3.16**). These data support the hypothesis that Mtb EVs have physiochemically distinct proteins based on their size.

Table 3.2 Significantly different proteins with signal peptide by SignalP. Proteins identified by SignalP that vary across all the fractions significantly are listed. LIPO = lipobox motif, TAT = twin-arginine motif, SP = signal peptide.

SignalP	Protein	Mycobrowser Function	Mycobrowser Functional Category	Molecular Weight	ANOVA Test (p-value): *(p < 0.00863)
100R low, F1 low, F2 high, F3 high, F4 high, F5 low, qEV high					
LIPO	LprF (Rv1368)	Probable conserved lipoprotein	Cell wall and cell processes	27 kDa	< 0.00010
LIPO	LprG (Rv1411c)	Conserved lipoprotein	Cell wall and cell processes	25 kDa	< 0.00010
LIPO	LpqH (Rv3763)	19 kDa lipoprotein antigen precursor	Cell wall and cell processes	15 kDa	< 0.00010
LIPO	Rv1006	Unknown	Conserved hypotheticals	61 kDa	0.0013
100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV high					
LIPO	LppK (Rv2116)	Probable conserved lipoprotein	Cell wall and cell processes	20 kDa	< 0.00010
LIPO	LppX (Rv2945c)	Possible conserved lipoprotein	Cell wall and cell processes	24 kDa	0.00046
LIPO	Rv0999	Unknown	Conserved hypotheticals	26 kDa	0.0013
100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV low					
LIPO	LpqG (Rv3623)	Probable conserved lipoprotein	Cell wall and cell processes	25 kDa	0.0003
100R high, F1 high, F2 high, F3 low, F4 low, F5 low, qEV low					
TAT	LprQ (Rv0483)	Probable conserved lipoprotein	Cell wall and cell processes	48 kDa	< 0.00010
LIPO	PstS3 (Rv0928)	Periplasmic phosphate-binding lipoprotein	Cell wall and cell processes	38 kDa	< 0.00010
LIPO	CaeA (Rv2224c)	Probable carbosylesterase	Cell wall and cell processes	56 kDa	< 0.00010
SP	FbpC (Rv0129c)	Secreted antigen 85-C, mycolytransferase	Lipid metabolism	37 kDa	0.0043
100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV high					
SP	MycP3 (Rv0291)	Probable membrane-anchored mycosin	Intermediary metabolism and respiration	46 kDa	0.00067
SP	Mpt53 (Rv2878c)	Soluble secreted antigen Mpt53 precursor	Cell wall and cell processes	18 kDa	0.0007
100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low					
TAT	Rv0315	Possible beta-1,3-glucanase precursor	Intermediary metabolism and respiration	32 kDa	0.0053
LIPO	DppA (Rv3666c)	Probable periplasmic dipeptide-binding lipoprotein, active transport (import)	Cell wall and cell processes	58 kDa	< 0.00010
LIPO	SubI (Rv2400c)	Probable sulfate-binding lipoprotein, active transport (import)	Cell wall and cell processes	37 kDa	0.0018
LIPO	Rv0526	Possible thioredoxin protein	Intermediary metabolism and respiration	23 kDa	0.0025
SP	Rv3668c	Possible protease/surface lipoprotein	Intermediary metabolism and respiration	23 kDa	< 0.00010
SP	Rv0398c	Possible secreted protein	Cell wall and cell processes	22 kDa	< 0.00010
SP	Mpt70 (Rv2875)	Major secreted immunogenic protein	Cell wall and cell processes	19 kDa	< 0.00010
SP	TB22 (Rv3036c)	Probable conserved secreted protein TB22.2	Cell wall and cell processes	24 kDa	0.0026
SP	Mpt64 (Rv1980)	Immunogenic protein Mpt64	Cell wall and cell processes	25 kDa	0.0026
SP	Mtc28 (Rv0040c)	Secreted proline rich protein	Cell wall and cell processes	32 kDa	0.0032
100R high, F1 low, F2 low, F3 low, F4 low, F5 low, qEV low					
SP	PepA (Rv0125)	Probable serine protease	Intermediary metabolism and respiration	35 kDa	0.0039

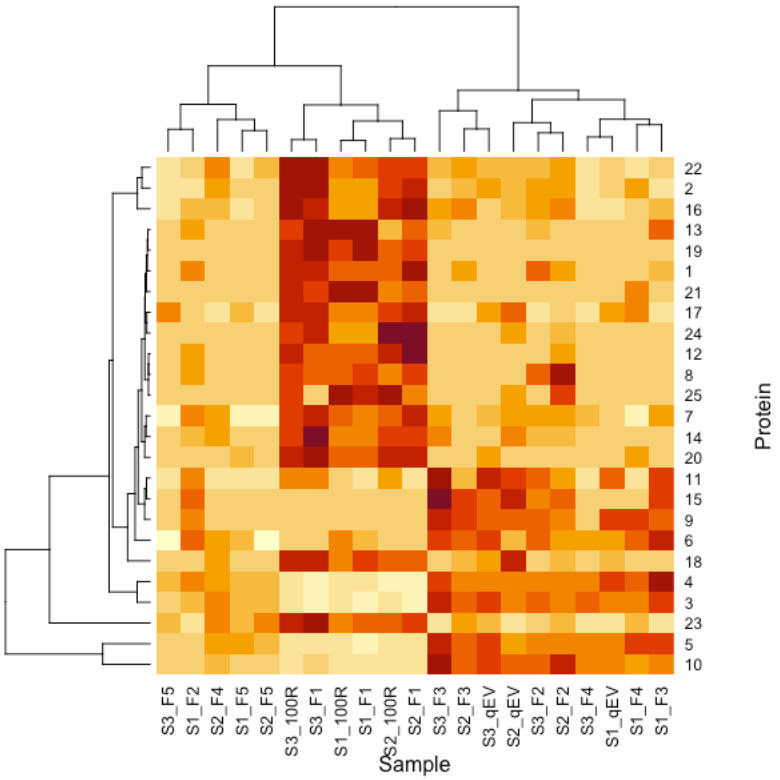


Figure 3.15 Heat map of SignalP proteins with significant difference across the fractions. Proteins known to contain a signal peptide that were identified as significantly different across individual fractions are clustered by NSAF.

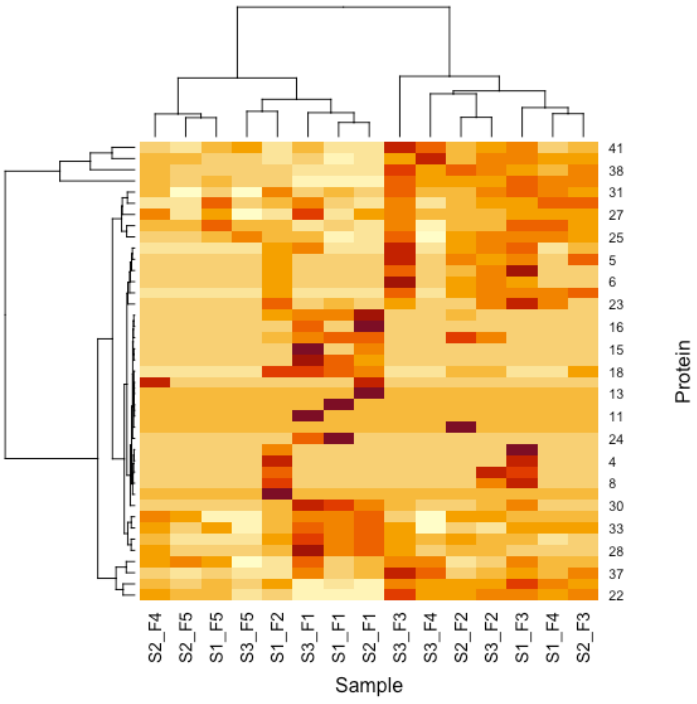


Figure 3.16 Heat map of proteins containing Lipobox motif by AF4 fraction. Proteins found in F1-F5 known to contain the lipobox motif are clustered based on NSAF.

3.3.7 Mtb EV Enrichment by Size Exclusion Chromatography Versus AF4 Fractions

Size exclusion chromatography (SEC) using qEV is essentially the reverse of AF4: larger components elute first, and small soluble proteins elute later. The cut off for qEV fraction inclusion in this study was based on the presence of GroES, a marker of soluble protein contamination of Mtb EVs. It is almost certain that some Mtb EVs were excluded in the interest of purity. Because over 50% of the recovered particles are in F4 and F5 regardless of the sample (**Table 3.1**), the qEV preparations would be expected to cluster toward those fractions in **Figure 3.14**. This occurred for samples 1 and 3, but not sample 2 which is most similar to fraction 3 of samples 1 and 2. Contrastingly, when looking only at proteins that are significantly different by fraction, the qEV samples are most similar to fractions 2 and 3. Since the larger vesicles would elute first in SEC, it is possible that this procedure disrupted some of the larger EVs causing their associated components to elute later.

There are two proteins that vary in abundance in qEV compared to all the other fractions: Rv2416c (Eis), which is low in qEV, and Rv1001 (ArcA) which is high. Plots of the NSAF by sample can be seen in **Figure 3.17**. Because Rv2416c is fairly consistent from 100R through F3, and then a little bit higher in F4 and F5, it is likely that some part of the qEV process is impacting this protein. The enrichment of ArcA in qEV suggests that there is some bias in the enrichment technique, possibly due to disruption of larger vesicles.

Additional insight about how the qEV method may be missing and enriching certain proteins emerges when comparing the overall fractions as groups (100R and F1 low/no EVs, F2-F5 high EVs, qEV) (**Table 3.3**). There are quite a few proteins that are lower than expected in the qEV samples (18 with significant NSAF p-values, 12 non-significant included), with some having no spectral counts while the AF4 EV fractions have multiple spectral counts for those proteins in several fractions. Additionally, there are proteins present in qEV samples that are low or absent from AF4 F2-F5 (27 with significant NSAF p-values, 2 non-significant included). Some of these proteins overlap with proteins that vary by sample (**Table AII7.4**), and proteins that vary by fraction (**Table AII7.5**). Proteins that have higher or

lower spectral counts in F3-F5 compared to qEV are particularly surprising, as the resolution of AF4 provides confidence that those fractions have primarily Mtb EV associated proteins, so they are expected to be at similar levels in qEV preparations.

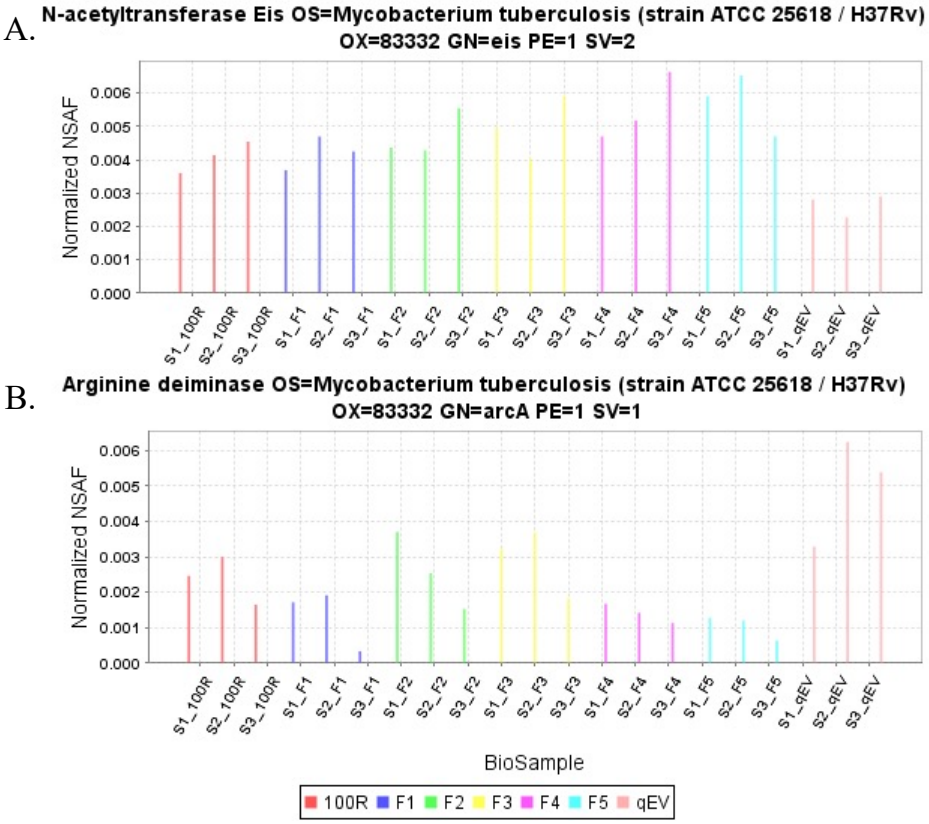


Figure 3.17 Proteins with significantly different abundance in qEV only. The NSAF for each fraction for proteins (A) Rv2416c (Eis) and (B) ArcA is plotted.

Table 3.3 Proteins lower and higher than expected in qEV based on AF4. Proteins listed include those with a significant p-value when NSAF for the groups 100R&F1, F2-F5, and qEV are compared. Several proteins are also included based on subjective comparison of the average total spectrum counts between qEV and Mtb EV containing AF4 fractions. Italicized proteins also appear in **Table AII7.4**, bolded proteins also appear in **Table AII7.5**.

Alternate ID	Molecular Weight	GO Cellular Component(s)	Mycobrowser Functional Category	ANOVA Test (p-value): *(p < 0.01178)	NSAF Quantitative Profile	Average Total Spectrum Count per Fraction (n = 3)						
						100R	F1	F2	F3	F4	F5	qEV
qEV lower than expected												
BfrA (Rv1876)	18 kDa	Cyt, PM, Sec	Intermediary metabolism and respiration	0.0043	100R&F1 low, F2-F5 high, qEV low	128.0	140.7	199.0	194.0	120.0	132.0	84.7
CanA/MtcA1 (Rv1284)	18 kDa	PM	Intermediary metabolism and respiration	0.0022	100R&F1 high, F2-F5 low, qEV low	27.7	30.7	15.0	4.3	8.0	4.3	0.0
Cfp32 (Rv0577)	27 kDa	Cyt, PM, Sec	Conserved hypotheticals	< 0.00010	100R&F1 high, F2-F5 low, qEV low	68.0	68.3	16.7	10.7	13.7	11.0	0.0
EchA17 (Rv3039c)	27 kDa	PM	Lipid metabolism	0.08	Not significant	14.0	14.0	13.3	11.7	5.7	7.7	0.0
EchA4 (Rv0673)	34 kDa	CW	Lipid metabolism	0.036	Not significant	51.3	62.3	47.7	32.7	14.7	14.3	0.00
Eis (Rv2416c)	44 kDa	Cyt, Sec	Virulence, detoxification, adaptation	< 0.00010	100R&F1 high, F2-F5 high, qEV low	133.7	129.7	134.3	88.3	64.7	67.7	43.3
Frr (Rv2882c)	21 kDa	Cyt, Sec	Information pathways	0.23	Not significant	22.7	25.0	6.0	6.0	12.0	12.0	0.0
GlcB (Rv1837)	80 kDa	CW, Cyt, PM, Sec	Intermediary metabolism and respiration	0.01	100R&F1 high, F2-F5 high, qEV low	485.7	518.7	299.0	149.7	173.3	193.7	111.0
GlgE (Rv1327c)	79 kDa	Cyt, PM	Intermediary metabolism and respiration	< 0.00010	100R&F1 low, F2-F5 high, qEV low	157.0	137.0	272.3	212.0	132.3	109.0	54.67
HsaD (Rv3569c)	32 kDa	CW, PM	Intermediary metabolism and respiration	0.0089	100R&F1 high, F2-F5 high, qEV low	110.7	122.3	116.7	55.3	31.7	31.3	12.7
HtdY (Rv3389c)	30 kDa	PM, Sec	Intermediary metabolism and respiration	0.024	Not significant	39.3	40.0	24.0	7.7	7.7	9.0	0.0
Icd2 (Rv0066c)	83 kDa	CW, Cyt, PM, Sec	Intermediary metabolism and respiration	0.14	Not significant	330.7	337.7	256.3	224.7	196.7	170.3	88.7
Icl (Rv0467)	47 kDa	Cyt, PM, Sec	Intermediary metabolism and respiration	0.071	Not significant	170.0	181.3	109.3	87.0	65.3	75.0	43.3
KshA (Rv3526)	44 kDa	Cyt	Intermediary metabolism and respiration	0.15	Not significant	33.0	37.0	24.0	25.0	10.7	11.0	0.0
LeuA (Rv3710)	70 kDa	PM, Sec	Intermediary metabolism and respiration	0.00032	100R&F1 high, F2-F5 high, qEV low	196.7	204.7	138.0	97.7	69.3	55.7	53.7
LpqB (Rv3244c)	62 kDa	CW, PM	Cell wall and cell processes	0.00036	100R&F1 high, F2-F5 low, qEV low	71.3	75.0	45.3	25.3	19.0	11.0	8.3
PepN (Rv2467)	94 kDa	Cyt, PM	Intermediary metabolism and respiration	0.00067	100R&F1 high, F2-F5 high, qEV low	189.3	190.0	109.7	55.7	61.0	60.7	37.0
Rv0248c	71 kDa	CW, PM	Intermediary metabolism and respiration	0.00094	100R&F1 high, F2-F5 low, qEV low	49.3	50.0	24.0	10.3	2.0	0.0	0.0
Rv0794c	52 kDa	PM, Sec	Intermediary metabolism and respiration	0.17	Not significant	53.0	54.7	63.0	28.0	14.0	8.7	4.33
Rv1906c	16 kDa	Sec	Conserved hypotheticals	0.057	Not significant	21.3	20.0	4.0	5.0	6.3	4.3	0.0

Alternate ID	Molecular Weight	GO Cellular Component(s)	Mycobrowser Functional Category	ANOVA Test (p-value): *(p < 0.01178)	NSAF Quantitative Profile	Average Total Spectrum Count per Fraction (n = 3)						
						100R	F1	F2	F3	F4	F5	qEV
Rv2251	50 kDa	PM, Sec	Intermediary metabolism and respiration	0.011	100R&F1 high, F2-F5 high, qEV low	75.3	81.3	58.0	42.7	29.3	18.7	16.3
Rv2971	30 kDa	Cyt, CW, PM	Intermediary metabolism and respiration	0.0014	100R&F1 high, F2-F5 low, qEV low	25.3	27.0	6.0	2.3	7.7	6.7	0.0
Rv3169	42 kDa	PM	Conserved hypotheticals	0.1	Not significant	188.3	197.7	163.7	120.0	78.0	97.3	63.3
Rv3368c	24 kDa	PM	Intermediary metabolism and respiration	0.002	100R&F1 high, F2-F5 low, qEV low	24.7	26.0	11.7	7.0	6.3	0.0	0.0
ScoA (2504c)	26 kDa	PM	Lipid metabolism	0.018	Not significant	30.3	32.0	27.0	15.7	11.7	10.0	1.7
ScoB (Rv2503c)	23 kDa	PM, Sec	Lipid metabolism	0.0083	100R&F1 high, F2-F5 high, qEV low	28.0	24.3	29.0	10.3	6.3	5.0	0.0
SirA (Rv2391)	62 kDa	CW, Cyt, PM, Sec	Intermediary metabolism and respiration	0.0018	100R&F1 high, F2-F5 high, qEV low	140.3	161.3	124.0	51.0	41.0	27.3	13.3
SodA/SodB (Rv3846)	23 kDa	Cyt, PM, Sec	Virulence, detoxification, adaptation	< 0.00010	100R&F1 high, F2-F5 high, qEV low	145.3	156.3	98.7	54.0	44.3	46.7	31.3
ThyX (Rv2754c)	28 kDa	PM	Intermediary metabolism and respiration	0.00013	100R&F1 high, F2-F5 low, qEV low	20.3	22.0	12.7	4.7	2.7	0.7	0.0
TreX/GlgX (Rv1564c)	81 kDa	Sec	Virulence, detoxification, adaptation	0.21	Not significant	44.3	47.7	25.3	22.0	17.3	16.7	1.0
qEV higher than expected												
AccA1 (Rv2501c)	71 kDa	Cyt, PM	Lipid metabolism	0.024	Not significant	16.0	5.7	15.7	12.7	3.0	0.0	24.3
ArcA (Rv1001)	43 kDa	Cyt, PM	Intermediary metabolism and respiration	0.001	100R&F1 low, F2-F5 low, qEV high	79.0	43.3	78.7	52.7	18.3	12.7	85.0
AroC (Rv2540c)	42 kDa	Cyt, Sec	Intermediary metabolism and respiration	0.024	Not significant	67.7	40.3	33.7	25.3	11.0	3.0	41.0
AtpA (Rv1308)	59 kDa	CW, PM	Intermediary metabolism and respiration	0.00078	100R&F1 low, F2-F5 low, qEV high	0.0	0.0	0.0	0.0	1.7	0.7	8.0
DapD (Rv1201c)	33 kDa	Cyt, PM	Intermediary metabolism and respiration	0.004	100R&F1 low, F2-F5 low, qEV high	47.3	28.0	58.0	56.0	26.7	14.7	58.3
EchA9 (Rv1071c)	36 kDa	Cyt, PM, Sec	Lipid metabolism	0.00039	100R&F1 high, F2-F5 low, qEV high	43.3	49.3	7.3	3.3	9.3	5.7	18.7
EsxB/CFP10 (Rv3874)	11 kDa	CW, PM, Sec	Cell wall and cell processes	0.0045	100R&F1 low, F2-F5 low, qEV high	0.0	0.0	0.0	1.0	0.0	0.0	7.7
FumC (Rv1098c)	50 kDa	Cyt, CW, PM, Sec	Intermediary metabolism and respiration	0.00019	100R&F1 low, F2-F5 low, qEV high	141.7	79.3	177.7	124.0	73.3	56.7	149.7
GatB (Rv3009c)	55 kDa	Cyt, PM	Information pathways	0.0053	100R&F1 low, F2-F5 low, qEV high	19.3	5.7	21.0	13.7	1.0	0.0	25.7
GlnA4 (Rv2860c)	50 kDa	PM, Sec	Intermediary metabolism and respiration	0.0024	100R&F1 low, F2-F5 low, qEV high	93.7	74.0	77.3	66.3	41.0	38.3	71.0
HspX (Rv2031c)	16 kDa	Cyt, CW, PM, Sec	Virulence, detoxification, adaptation	0.003	100R&F1 low, F2-F5 low, qEV high	46.7	34.0	23.0	18.3	17.0	9.7	35.3
IlvD (Rv0189c)	59 kDa	PM, Sec	Intermediary metabolism and respiration	0.011	100R&F1 low, F2-F5 low, qEV high	31.7	20.3	39.0	21.3	4.0	0.0	37.7

Alternate ID	Molecular Weight	GO Cellular Component(s)	Mycobrowser Functional Category	ANOVA Test (p-value): *(p < 0.01178)	NSAF Quantitative Profile	Average Total Spectrum Count per Fraction (n = 3)						
						100R	F1	F2	F3	F4	F5	qEV
LppL (Rv2138)	37 kDa	PM, Sec	Cell wall and cell processes	0.0031	100R&F1 high, F2-F5 low, qEV high	7.3	7.3	0.0	0.0	0.0	0.0	2.7
Mce2D (Rv0592)	55 kDa	CW	Virulence, detoxification, adaptation	0.00013	100R&F1 low, F2-F5 low, qEV high	0.0	0.0	0.0	0.0	0.0	0.0	1.3
MetC (Rv3340)	47 kDa	PM	Intermediary metabolism and respiration	0.0035	100R&F1 low, F2-F5 low, qEV high	69.7	52.7	78.0	52.0	33.3	22.0	57.0
Mpt53 (Rv2878c)	18 kDa	Cyt, Sec	Cell wall and cell processes	< 0.00010	100R&F1 high, F2-F5 low, qEV high	43.0	50.7	3.0	2.3	3.0	0.0	19.0
MycP3 (Rv0291)	46 kDa	CW, PM, Sec	Intermediary metabolism and respiration	< 0.00010	100R&F1 high, F2-F5 low, qEV high	25.3	27.7	1.7	0.0	2.7	3.0	8.7
PrcA (Rv2109c)	27 kDa	CW, PM	Intermediary metabolism and respiration	< 0.00010	100R&F1 low, F2-F5 high, qEV high	63.0	37.0	84.7	59.0	45.0	40.3	86.3
PstB2 (Rv0933)	30 kDa	PM	Cell wall and cell processes	0.0064	100R&F1 low, F2-F5 low, qEV high	0.0	0.0	0.0	0.0	0.0	0.0	4.3
Pta (Rv2174)	73 kDa	PM	Cell wall and cell processes	0.003	100R&F1 low, F2-F5 low, qEV high	19.0	8.0	37.7	29.0	7.0	3.3	41.3
PykA (Rv1617)	51 kDa	Cyt, PM	Intermediary metabolism and respiration	0.0026	100R&F1 low, F2-F5 low, qEV high	48.3	19.0	57.7	38.0	10.7	4.3	59.3
Rv0308	26 kDa	PM	Cell wall and cell processes	0.007	100R&F1 low, F2-F5 low, qEV high	0.0	0.0	0.0	12.0	4.0	2.0	14.7
Rv0787	34 kDa	Cyt, Sec	Conserved hypotheticals	0.0002	100R&F1 high, F2-F5 low, qEV high	5.3	6.3	0.0	0.0	1.0	0.0	2.3
Rv1265	25 kDa	Cyt, PM	Conserved hypotheticals	0.0027	100R&F1 low, F2-F5 low, qEV high	0.0	2.7	10.7	2.3	0.0	0.0	8.3
Rv1815	23 kDa	CW, Sec	Conserved hypotheticals	0.0082	100R&F1 high, F2-F5 low, qEV high	6.7	6.0	0.0	0.0	0.0	0.0	2.7
Rv3401	87 kDa	PM	Intermediary metabolism and respiration	0.0066	100R&F1 high, F2-F5 low, qEV high	104.7	79.0	79.7	45.7	13.0	8.7	66.3
SerC (Rv0884c)	40 kDa	PM, Sec	Intermediary metabolism and respiration	0.0013	100R&F1 high, F2-F5 low, qEV high	76.7	83.3	43.3	19.3	17.7	16.0	37.7
Tal (Rv1448c)	41 kDa	Cyt, PM, Sec	Intermediary metabolism and respiration	0.0049	100R&F1 high, F2-F5 low, qEV high	119.0	131.7	47.0	44.0	35.3	30.7	56.3
TrpC (Rv1611)	28 kDa	CW, PM	Intermediary metabolism and respiration	0.0034	100R&F1 high, F2-F5 low, qEV high	8.7	5.7	0.0	0.0	0.0	0.0	4.0

3.3.8 Conclusions

The data presented clearly support AF4 as an effective technique for Mtb EV enrichment. TEM demonstrates vesicle presence in fractions 2-5, which is supported by proteomic interrogation via Western blot for known Mtb EV markers (**Figure 3.5** and **Figure 3.6**) and by LC-MS/MS. These Mtb EV fractions have different size vesicles, both by relative sizing in NTA and by TEM visualization (**Figure 3.2** and **Figure 3.3**).

The proteomics confirm that the protein species present in Mtb EVs change in abundance across the AF4 fractions, 100R, and qEV samples. AF4 fractionation suggests that smaller Mtb EVs contain more acylated and glycosylated proteins, which aligns with the Western blot data. It is possible that other biomolecules like lipoglycans, lipids, and glycolipids vary either more drastically or in synergy with these proteomic changes. This is supported by the drastic drop off in LAM as fraction number increases (**Figure 3.5** and **Figure 3.6**). Although the total protein across AF4 fractions had only slight variation and did not cluster cleanly by fraction, fractionation of physiochemically distinct proteins could have biological significance. Many of the proteins known to have acylation and glycosylation identified in this data are involved in host-protein interactions, particularly regarding TLR-2. In combination with the higher concentration of LAM, it is possible that smaller Mtb EVs are more immunogenic than larger ones. This could be evaluated by exposing immune cells like macrophages to Mtb EVs from each fraction and evaluating their TLR-2 response.

AF4 fractionation of Mtb EVs also provides a new avenue for exploring potential biases in traditional separation techniques. The SEC method used in these experiments does influence at least some of the proteins expected to be in Mtb EVs based on comparison to AF4 fractionation. Additionally, AF4 provides a more highly resolved understanding of what proteins are and are not enriched in Mtb EVs. New proteins that may serve as marker of non-EV associated contaminants in a preparation are listed in **Table 3.4**. These proteins significantly were higher by NSAF for 100R and F1, but are very scarce in F2-F5 (less than 1 total spectrum count per fraction). Ideally, an antibody-based detection method would be

pursued. Alternatively, a more sensitive quantitation by multiple-reaction monitoring LC-MS/MS could be explored.

Table 3.4 Potential candidates for Mtb EV enrichment purity controls.

Protein	Molecular Weight	NSAF ANOVA Test (p-value): *(p < 0.00863)	Average Total Spectrum Count per Fraction (n = 3)						
			100R	F1	F2	F3	F4	F5	qEV
DhaA (Rv2579)	34 kDa	0.0005	23.67	26.33	0.00	0.00	0.00	0.00	0.00
Mtc28 (Rv0040c)	32 kDa	0.0032	23.67	21.33	0.67	0.00	0.00	0.00	1.67
Rv2575	31 kDa	< 0.00010	20.00	19.00	0.00	0.00	0.00	0.00	0.00
Rv1836c	70 kDa	< 0.00010	19.67	21.33	0.00	0.00	0.00	0.00	0.00
MmaA4 (Rv0642c)	35 kDa	< 0.00010	19.33	16.33	0.00	0.00	0.00	0.00	1.67
CmaA2 (Rv0503c)	35 kDa	0.0026	17.67	15.33	0.67	0.00	0.00	0.00	1.33
PunA (Rv3307)	28 kDa	< 0.00010	16.33	18.67	0.00	0.00	0.00	0.00	0.00
Rv0331	41 kDa	0.0068	15.67	16.00	0.00	0.00	0.00	0.00	0.00
Rv2765	26 kDa	0.0053	15.67	16.00	0.00	0.00	0.00	0.00	0.67
Hrp1 (Rv2626c)	16 kDa	0.00021	15.33	17.33	0.00	0.00	0.00	0.00	3.00
TrxB1 (Rv1471)	36 kDa	0.0014	12.33	11.33	0.00	0.00	0.00	0.00	0.00
Rv3668c	23 kDa	< 0.00010	11.33	14.33	0.00	0.00	0.00	0.00	0.00
Rv0546c	14 kDa	0.00086	10.67	12.67	0.00	0.00	0.00	0.00	0.00
Rv1700	23 kDa	< 0.00010	7.00	8.67	0.00	0.00	0.00	0.00	0.00
GpgP (Rv2419c)	24 kDa	0.0029	6.00	6.00	0.00	0.00	0.00	0.00	0.00
DeoC (Rv0468)	22 kDa	0.0083	5.67	6.00	0.00	0.00	0.00	0.00	0.00

3.4 References

[1] Phillips W, Willms E, Hill AF. Understanding extracellular vesicle and nanoparticle heterogeneity: Novel methods and considerations. PROTEOMICS 2021;21:2000118. https://doi.org/10.1002/PMIC.202000118.

[2] Brisson AR, Tan S, Linares R, Gounou C, Arraud N. Extracellular vesicles from activated platelets: a semiquantitative cryo-electron microscopy and immuno-gold labeling study. Platelets 2017;28:263–71. https://doi.org/10.1080/09537104.2016.1268255.

[3] Bordanaba-Florit G, Royo F, Kruglik SG, Falcón-Pérez JM. Using single-vesicle technologies to unravel the heterogeneity of extracellular vesicles. Nature Protocols 2021 16:7 2021;16:3163–85. https://doi.org/10.1038/s41596-021-00551-z.

[4] Zhang H, Freitas D, Kim HS, Fabijanic K, Li Z, Chen H, et al. Identification of distinct nanoparticles and subsets of extracellular vesicles by asymmetric flow field-flow fractionation. Nature Cell Biology 2018 20:3 2018;20:332–43. https://doi.org/10.1038/s41556-018-0040-4.

[5] Rath P, Huang C, Wang T, Wang T, Li H, Prados-Rosales R, et al. Genetic regulation of vesiculogenesis and immunomodulation in Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences of the United States of America 2013;110. https://doi.org/10.1073/pnas.1320118110.

[6] White DW, Elliott SR, Odean E, Bemis LT, Tischler AD. Mycobacterium tuberculosis Pst/SenX3-RegX3 regulates membrane vesicle production independently of ESX-5 activity. MBio 2018;9. https://doi.org/10.1128/MBIO.00778-18/SUPPL_FILE/MBO003183934ST2.DOCX.

[7] Gupta S, Palacios A, Khataokar A, Weinrick B, Lavín JL, Sampedro L, et al. Dynamin-like proteins are essential for vesicle biogenesis in Mycobacterium tuberculosis. BioRxiv 2020:2020.01.14.906362. https://doi.org/10.1101/2020.01.14.906362.

[8] Kikuchi Y, Obana N, Toyofuku M, Kodera N, Soma T, Ando T, et al. Diversity of physical properties of bacterial extracellular membrane vesicles revealed through atomic force microscopy phase imaging. Nanoscale 2020;12:7950–9. https://doi.org/10.1039/C9NR10850E.

[9] Prados-Rosales R, Baena A, Martinez LR, Luque-Garcia J, Kalscheuer R, Veeraraghavan U, et al. Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. Journal of Clinical Investigation 2011;121:1471–83. https://doi.org/10.1172/JCI44261.

[10] Liangsupree T, Multia E, Riekkola ML. Modern isolation and separation techniques for extracellular vesicles. Journal of Chromatography A 2021;1636:461773. https://doi.org/10.1016/J.CHROMA.2020.461773.

[11] Plavchak CL, Smith WC, Bria CRM, Williams SKR. New Advances and Applications in Field-Flow Fractionation. Https://DoiOrg/101146/Annurev-Anchem-091520-052742 2021;14:257–79. https://doi.org/10.1146/ANNUREV-ANCHEM-091520-052742.

[12] Quattrini F, Berrecoso G, Crecente-Campo J, Alonso MJ. Asymmetric flow field-flow fractionation as a multifunctional technique for the characterization of polymeric nanocarriers. Drug Delivery and Translational Research 2021;11:373–95. https://doi.org/10.1007/S13346-021-00918-5/FIGURES/6.

[13] Wallace E, Hendrickson D, Tolli N, Mehaffy C, Peña M, Nick JA, et al. Culturing Mycobacteria. Methods in Molecular Biology 2021;2314:1–58. https://doi.org/10.1007/978-1-0716-1460-0_1.

[14] Lucas M, Ryan JM, Watkins J, Early K, Kruh-Garcia NA, Mehaffy C, et al. Extraction and Separation of Mycobacterial Proteins. Methods in Molecular Biology 2021;2314:77–107. https://doi.org/10.1007/978-1-0716-1460-0_3.

[15] Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. Analytical Chemistry 2003;75:4646–58. https://doi.org/10.1021/AC0341261.

[16] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics 2000;25:25–9. https://doi.org/10.1038/75556.

[17] Zhang Y, Wen Z, Washburn MP, Florens L. Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. Analytical Chemistry 2010;82:2272–81. https://doi.org/10.1021/AC9023999.

[18] Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nature Biotechnology 2022 2022:1–3. https://doi.org/10.1038/s41587-021-01156-3.

[19] Bai K, Barnett G v., Kar SR, Das TK. Interference from Proteins and Surfactants on Particle Size Distributions Measured by Nanoparticle Tracking Analysis (NTA). Pharmaceutical Research 2017 34:4 2017;34:800–8. https://doi.org/10.1007/S11095-017-2109-3.

[20] Diaz G. Exosomes: A Potential Novel Source of Biomarkers for Tuberculosis. 2017. Doctoral dissertation, Colorado State University, Fort Collins, CO. Mountain Scholar.

[21] Sartain MJ, Dick DL, Rithner CD, Crick DC, Belisle JT. Lipidomic analyses of Mycobacterium tuberculosis based on accurate mass measurements and the novel "Mtb LipidDB." Journal of Lipid Research 2011;52:861–72. https://doi.org/10.1194/JLR.M010363/ATTACHMENT/CB902AAC-DCD1-4D05-AA38-FF6BCB370ABD/MMC1.ZIP.

[22] Mehaffy C, Belisle JT, Dobos KM. Mycobacteria and their sweet proteins: An overview of protein glycosylation and lipoglycosylation in M. tuberculosis. Tuberculosis 2019;115:1–13. https://doi.org/10.1016/J.TUBE.2019.01.001.

4.1 Defining Risk

If you query *Merriam-Webster* for the definition of the word "risk," you will be greeted with two entries, each containing multiple definitions, for a word that has been in use since the 1600s. Risk the transitive verb means "to expose to hazard or danger," and "to incur the risk or danger of." Risk the noun includes insurance and financial matters like "the chance that an investment (such as a stock or commodity) will lose value" as well as "possibility of loss or injury" and "someone or something that creates or suggests a hazard"[1]. What this word ultimately boils down to is the possibility of a consequence. Attributing a positive or negative value to the consequence depends entirely on the person perceiving the risk, as the specific consequence may benefit one while harming another. When considering risk in the context of biosafety and biosecurity, generally the consequence is negative (although the pathogen(s) or bad actor(s) may disagree).

In biosafety and biosecurity, risk (R) is viewed as a function of both the likelihood (L) of something happening and the consequences (C) of it occurring: R = f(L*C)[2]. This function for defining risk is often portrayed as a matrix (**Figure 4.1**). Evaluating the risk of something occurring is the basis for the primary tool in the biosafety and biosecurity professional's toolbox: risk assessment. People do risk assessments constantly without even thinking about it, from deciding what clothes to wear to watching a traffic light turn yellow and choosing to stop or proceed through an intersection. The challenges in biosafety and biosecurity are defining the risks in a quantifiable and reproducible manner, then determining how mitigation strategies may influence the risks, and ultimately establishing what the acceptable level of risk is. These tasks are not nearly as clear-cut as a calculation or matrix; biosafety and biosecurity risk management professionals operate in continuous shades of gray.

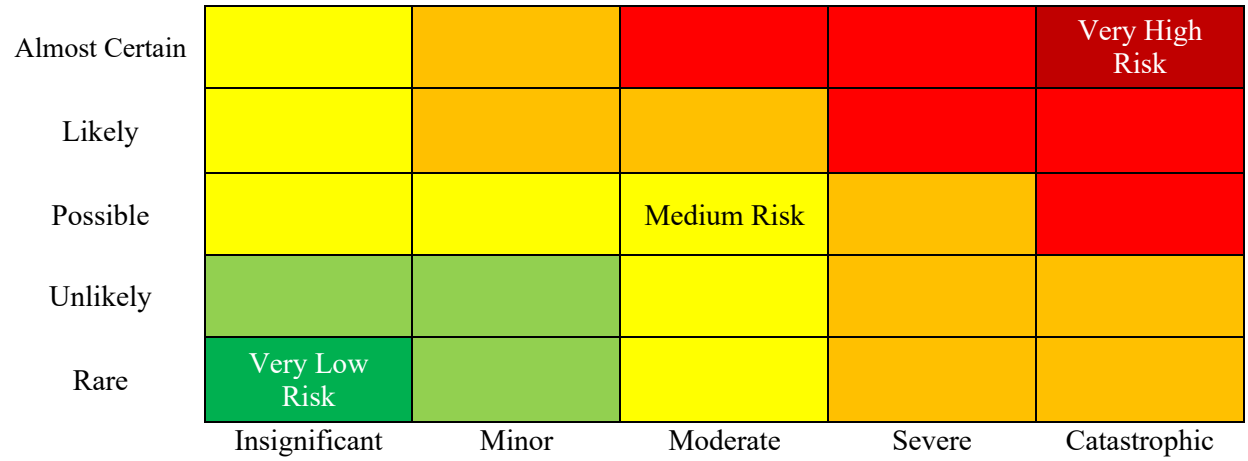


Figure 4.1 Risk matrix. The likelihood of a consequence increases going up while the severity of the consequence increases going right.

My graduate school education extended well beyond my laboratory experiments and the classroom and into the world of risk assessments. My advisors supported me in pursuing internships with the Institutional Biosafety Committee (IBC) and the Biosafety Office (BSO), experiences that have not only set me up to pursue a career in research safety and compliance, but also taught me the importance of blending the black and white to see and appreciate the shades of gray. The following chapter is a presentation of the skills and knowledge I gained beyond my laboratory studies, with a focus on biosafety and biosecurity risk assessment as the world entered into and continues to navigate the COVID-19 pandemic.

4.2 The NIH Guidelines and IBCs

I started working with the IBC just a few months before the most recent version of the "NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules (NIH Guidelines)" [3] was published in early 2019 (84 FR 17858). Although they are called guidelines, the name is misleading. Any institution or individual receiving support from the NIH must abide by them for all recombinant or synthetic nucleic acid (rsNA) research occurring at that institution, regardless of the individual project's funding source. Noncompliance at the institution, even on projects not funded by NIH, can result in censure and all NIH funding being revoked from an entity. These consequences not only delay or prevent important research from progressing, but also end careers and stain reputations

which is particularly harmful to mid and early career scientists. Fortunately, there have been very few instances where these types of consequences have been applied. This set of guidelines and how they were enacted is an interesting case study in risk assessment.

The first recombinant DNA (rDNA) experiment was published in 1972 when Jackson, Symons, and Berg reported their method for combining simian virus 40 (SV40) and lambda bacteriophage DNA into one plasmid with an *Escherichia coli* operon [4]. Shortly after this, construction of functional bacterial plasmids with antibiotic resistance markers joining DNA from two completely different organisms was described [5]. These plasmids could stably replicate in *E.coli* and allow the transcription of eukaryotic DNA in a bacterial cell [6]. Following concerns raised at the 1973 Gordon Research Conference on Nucleic Acids, several authors on the aforementioned publications served on the National Academy of Sciences Committee on Recombinant DNA Molecules, Assembly of Life Sciences which published a letter regarding "Potential Biohazards of Recombinant DNA Molecules" [7]. In it they state "new DNA elements introduced into *E. coli* might possibly become widely disseminated among human, bacterial, plant, or animal populations with unpredictable effects," while noting that "our concern is based on judgments of potential rather than demonstrated risk" [7]. At the time, scientists agreed to a self-imposed moratorium on rDNA research. Ultimately, their recommendation to create an NIH advisory committee for overseeing this type of work came to fruition following the second Asilomar conference in 1975. The NIH Recombinant DNA Advisory Committee (RAC) was formed and the NIH Guidelines published in the Federal Register in 1976 [8].

Although rDNA was gaining some attention in the press in the late 1970s, true public engagement in the discourse surrounding genetic engineering began surrounding the 1980 U.S. Supreme Court case *Diamond v. Chakrabarty*. In this case, the legality of patenting human-made microorganisms was established [9]. The microorganism in question was a *Pseudomonas* bacterium with two plasmids, causing it to be able to break down crude oil. This was anticipated to have value for oil spill treatment. Chakrabarty was not only attempting to patent the method for producing the bacteria, but also the bacteria

themselves. This sparked widespread public debate regarding the creation and control of new life forms through genetic engineering. The general secretaries from the National Council of Churches, Synagogue Council of America, and United States Catholic Conference wrote to President Carter requesting immediate examination of genetic engineering "for representatives of a broad spectrum of our society to consider these matters and advise the government on its necessary role" [10]. In response, the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavior Research published a report titled "Splicing Life" [10]. Ultimately the report indicated that some of the concerns regarding oversight needed to be addressed, but others, like the fear of remaking human beings, were exaggerated. Following this report, the NIH RAC became heavily focused on reviewing human gene therapy (HGT) related experiments and the HGT subcommittee was formed.

The NIH Guidelines and the RAC have greatly evolved since their creation, but some of the original frameworks and early changes remain in place. The responsibilities of the RAC have largely shifted to institutional IBCs and the NIH Office of Science Policy, while the recently established NExTRAC reviews matters involving synthetic biology. Consistent with what was discussed at Asilomar in 1975, certain types of experiments fall under higher levels of scrutiny: introducing antimicrobial resistance into a pathogenic bacterium has more levels of review than putting microbes with genes for fluorescing proteins into an animal model. Additionally, the local oversight of rsNA falls to individual IBCs, which must include at least two members not affiliated with the institution to represent the interests of the community [3]. Like the original scientific debates around rDNA, IBC meetings and records are open to the public. The administration and management of IBCs and their processes, however, are not dictated in the NIH Guidelines, allowing for significant flexibility across institutions. Because the research portfolios across institutions vary greatly, the composition of individual IBCs changes based on the required expertise. A 2019 benchmarking study of IBCs revealed that although the NIH Guidelines are specific to rsNA work, the majority of IBCs review research with other biohazards and many still require registration of rsNA work designated as exempt from the guidelines [11]. The fact that NIH does

not require review of all work with human pathogens remains baffling to me, as I can think of many experiments without rsNA in risk group 2 and 3 organisms that would have risks of higher likelihood and greater consequence than many experiments with rsNA regularly reviewed by our IBC. It is comforting to know that many institutions (including CSU) value biosafety and biosecurity such that they go above and beyond the minimum requirements by reviewing research that falls outside the NIH Guidelines.

When considering the evolution of the NIH Guidelines and their current implementation, I see several interesting themes around risk assessment emerge—particularly regarding uncertainty and collaboration. More often than not, at least one part of the risk equation (likelihood or consequence) is unknown. Originally there was much anxiety around the potential for common bacteria to become new vectors for cancer causing viruses or other harmful toxins and traits due to manmade changes in the ability to share and alter DNA [8]. In reality, exchanges of genetic information occur naturally between both lower and higher life forms and this was not a novel manmade phenomenon. The potential consequence of immediate and vast spread of harmful new viruses was extremely unlikely, and had a moratorium on rDNA research remained, many important scientific breakthroughs (like human insulin production by bacteria) would never have occurred. When performing a risk assessment where information is missing, there must be balance: fear and uncertainty cannot blanketly prevent investigation into the unknown.

The NIH Guidelines also highlight the collaborative nature of risk assessment. Many of the scientists involved in the original rDNA debates were molecular biologists. I suspect that had the group included more scientists trained specifically in microbiology, infectious disease, epidemiology, and public health, some of the worst-case scenario fears would have been easily diverted from the start. Today the IBC includes a variety of experts, with the ability to include *ad hoc* consultants based on the research being reviewed. Like science, risk assessment does not occur in a vacuum: obtaining perspectives from experts in the specific research field, other scientists, safety professionals, and the community encourages a well-rounded and thorough consideration of the risks. Although not all risk assessments can (or should)

be subject to numerous reviewers, the successful biosafety and biosecurity professional seeks input from appropriate stakeholders in a collaborative process in the interest of facilitating safe and secure research.

4.3 CSU IBC and the Start of COVID-19

The CSU IBC saw its first two applications for SARS-CoV-2 in late January 2020, before the virus was even officially named, months ahead of the world-wide COVID-19 shutdowns. These applications came from principal investigators who had years of experience working with a variety of pathogens including other risk group 3 coronaviruses and often worked on newly identified viruses. Extra care was taken to consider all aspects of these projects due to growing public concern, but ultimately the procedures in place for work under high containment were—and continue to be—sufficient for containment of this virus. Approval of those projects was relatively straightforward since the groups had extensive knowledge regarding respiratory pathogen containment.

As cases rose and shutdowns began, the IBC became inundated with requests to work with SARS-CoV-2. These reviews were not nearly as straightforward as the first few. Guidelines from organizations like Centers for Disease Control and Colorado Department of Public Health & Environment for handling things like diagnostic samples and working in laboratories with others were changing frequently. There were very little data regarding the mechanisms of transmission, susceptibility to disinfectants, or environmental stability for this virus. Determining the appropriate biosafety level and protocols for handling certain kinds of samples was a real-time rapid example of the iterative nature of risk assessment (**Figure 4.2**). Work involving intentional propagation of the virus was straightforward in terms of risk, but projects surrounding new detection methods and/or diagnostic samples were more difficult to assess. At the same time, severe shortages of personal protective equipment (PPE) limited standard risk mitigation approaches, requiring creative solutions to allow the research to continue [12].

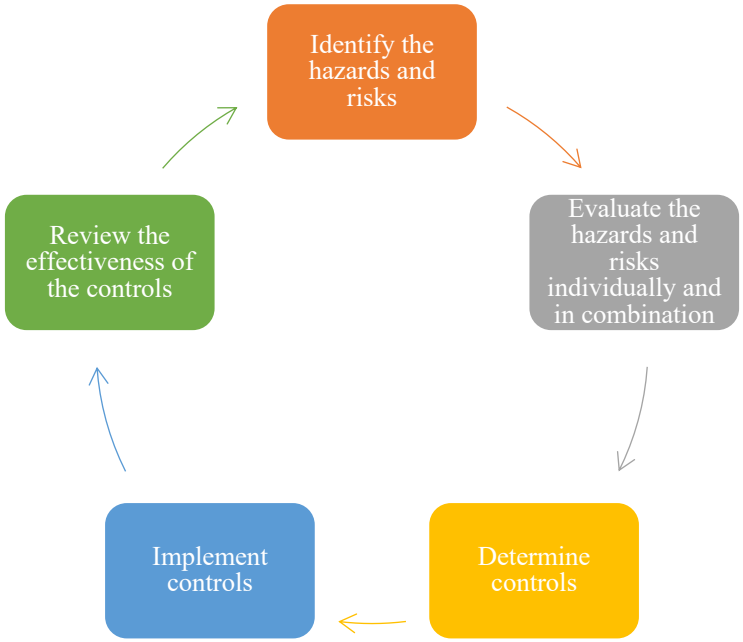


Figure 4.2 Risk Assessment Cycle. The steps of risk assessment are portrayed as a cycle, where the process is repeated throughout the life of a research project.

Because high-containment laboratories regularly deal with decontamination of items that require re-use (for example, lab equipment that needs to leave containment for service), many groups across the country pitched in to address the N95 respirator shortage by establishing methods for disinfection and re-use [13,14]. In addition to respirators, disposable gowns became extremely difficult to acquire and this PPE was essential for both healthcare professionals and researchers. After determining the efficacy of vaporized hydrogen peroxide decontamination and reuse in collaboration with another institution [15], members of the CSU research community established procedures for respirators and disposable gown decontamination and re-use. This effort, which was reviewed by the IBC, went beyond supporting PPE needs for CSU researchers and extended into the community through partnerships with healthcare facilities. Assessing risk for this type of large, coordinated effort was complicated, as potentially SARS-CoV-2 contaminated PPE is not a typical research or diagnostic specimen seen in CSU research labs since it is a type of medical waste. There were also many people involved in the effort, from the healthcare workers using the PPE, to the individuals transporting both used and decontaminated PPE, and the group performing the decontamination and re-packaging; everyone needed to be aware of the risks and have the

relevant training for their role. The leading investigators were diligent in communicating with CSU IBC, biosafety, occupational health, and the healthcare facilities to establish appropriate procedures based on the knowledge at the time. As new information arose and the actual rates of SARS-CoV-2 infection in the specific healthcare facilities were established, re-evaluation of the risks allowed for changes in the processes to improve efficiency and throughput.

Every person participating in the research or contributing to the risk assessment brings their own biases to the process. This often manifests as individual risk perception, a very subjective gray area influencing risk assessment. I recall a specific IBC review that highlighted the importance of considering individual risk perception. The project involved wastewater sampling for COVID-19 monitoring from places like residence halls across campus. One of the main concerns from the researchers was getting SARS-CoV-2 from the wastewater. However, the greater risk was exposure to one of the many other pathogenic organisms found in wastewater since by the time SARS-CoV-2 got to the collection point it was highly unlikely to be infectious, while many bacteria and some hardier viruses would still be viable. This was a great lesson in managing individual risk perception. CSU research safety professionals had to ensure the researchers that their concerns were taken seriously while providing education around the more probable risks associated with the work. The wastewater testing became another example of the CSU research community shifting priorities and forming new collaborations to meet public health needs for COVID-19 response.

4.4 Return to Campus Risk Assessments with CSU Biosafety

I transitioned from the IBC to the BSO in July 2020. At this point the campus was starting to return research operations to limited in-person capacity, beyond what was considered essential and allowed to continue during the shutdown (mostly SARS-CoV-2 work). This was a fascinating time to join the BSO. My experience with risk assessment was primarily considering BSL-3 containment as a higher risk environment than BSL-1 and BSL-2, but at this time it was almost the opposite. Yes, the organisms worked with at BSL-3 have greater consequences and in some cases higher likelihood of infecting a

person, but because of the engineering designs and the required PPE at the time it was actually less of a COVID-19 risk than working with other people at lower containment. It is kind of bizarre to be more concerned about someone catching SARS-CoV-2 (a risk group 3 organism) from working with other people on a risk group 1 or 2 organism in BSL-1 or 2 environments than those working directly with SARS-CoV-2 in a BSL-3. Conversations with lab groups around how to work safely in their environments became just as much about containing their organism as it was about keeping them from exposing each other to SARS-CoV-2.

Another risk assessment that reinforced the influence of risk perception was around returning researchers to work in a BSL-3 lab that primarily focused on arboviruses. In my experience as a tuberculosis researcher, BSL-3 always meant a respiratory protection requirement and it was my perception that aerosol risks were an inherent part of BSL-3 work. Because arboviruses are not spread via aerosol, I learned that using PPE such as an N95 is typically unnecessary. However, for research to resume in the space, personnel needed to be protected from COVID-19 exposure by one another. At this point, PPE like surgical masks was difficult to come by, and we did not want to be taking resources away from healthcare professionals if it was not necessary. Ultimately a process for personnel to wear cloth masks, changing in and out of them before leaving the space as to not take anything in the lab out on them, was developed. The research groups and department worked to sew hundreds of cloth facemasks for use in this space and many other areas as people began to return to campus at reduced capacity.

As people returned to campus, I got to see many risk assessments in action. Through the IBC internship I learned a lot about looking at a project and identifying and mitigating risks, but very infrequently did I see the locations of work and meet the people performing the research. The transition of labs back to in-person work came with many requests for biosafety to visit spaces and help the researchers develop new workflows for work in their spaces. The biosafety group also worked on non-laboratory related risk assessments. On-campus COVID testing was an important program to allow students and staff to return in-person. Biosafety worked with occupational health and the pandemic

response group to train student employees on appropriate PPE and disinfection protocols for COVID-19 testing sites. The risk perceptions of each student employee influenced their questions and comfort levels surrounding these protocols. When drafting standard operating procedures for their safety manual, their comments and concerns became sources of insight that I otherwise may not have considered including based on my personal risk perceptions.

4.5 Biosecurity and COVID-19

Biosecurity is a multidisciplinary concept that can have different meanings to different communities. In the laboratory context, it is the prevention of theft, loss, or misuse of biological material, technology, or research-related information [2]. Effective biosecurity programs consider physical security, information security, personnel suitability and training, dual-use research, responsible communication, and synergy with biosafety practices. Pinpointing the exact origin of SARS-CoV-2 became a hot topic in the media, which brought biosecurity and dual-use research back into the public spotlight. This stemmed from the theory that SARS-CoV was worked with in a laboratory where gain of function studies were being performed with US funding, and that an "enhanced" virus exited the lab and began circulating in the population [16]. To be clear, there is no evidence supporting this theory and the majority of scientist agree that this is a highly unlikely scenario. The news of a "lab leak" of an "enhanced" virus, however, spread rapidly through traditional and social media outlets, reigniting the gain of function and dual-use research debate.

The original gain of function debate surrounded research conducted in response to the H5N1 avian flu infection of humans in the early 2000s [17]. Although H5N1 is not transmissible between mammals, the high mortality rate in humans raised alarms over the potential for mutations causing changes in transmissibility. Two independent laboratories (Kawoka and Fouchier) conducted experiments using different methods to investigate this possibility in the interest of public health preparedness for potential flu pandemics, and both established ferret models for aerosol transmission. Upon attempts to publish their work, the journal editors requested a review of the security implications by the US National

Science Advisory Board for Biosecurity prior to public release [18]. Their primary concern was that bad actors would access the information and use it to generate influenza-based biological weapons. Like the early rDNA debates, an international group of influenza researchers declared a self-imposed 60-day pause on research regarding H5N1 transmissibility. More than 8 months later, both groups were able to proceed with publication of their articles, the Kawaoka lab in full and the Fouchier lab with some revisions [19,20]. This very public, very controversial, debate followed by several (unrelated) biosafety mishaps ultimately resulted in the United States Government Policy for Institutional Oversight of Life Sciences Dual Use Research of Concern (DURC) in 2012 focusing on funding agency review of ongoing research, which was updated to the current policy focused on institutional review of ongoing and future research in 2015 [21].

It is kind of ironic that SARS-CoV-2 brought gain of function back into the limelight because under the US DURC policy, SARS-CoV is not covered and therefore not automatically subject to dual-use review. This is one of the problems with list-based systems like the US DURC policy and the US Federal Select Agent Program (FSAP): experiments and agents with potential for a major impact on the world are completely excluded. At the same time, additional regulatory burden from programs like FSAP can hinder research progress over things that ultimately do not increase safety or security. Additionally, country specific legislation and regulation varies broadly with 66% of countries lacking adequate biosafety policies, 81% lacking adequate biosecurity policies, and only 1% having appropriate dual-use research oversight according to the 2019 Global Health Security Index [22]. I do not have the answer for how to ensure the world prioritizes biosafety and biosecurity. I do believe that instilling the fundamental concepts in the research community as part of the culture would go a long way to help prevent biosecurity failures. Making sure all scientists understand that they have no control over what another person does with their data, whether acquired by theft or through reading a publication, may make one reconsider submitting pre-print publication with the entire genome of a pathogen and how to synthesize it *de novo*.

4.6 Resilience in a Global Crisis

The world-wide response to the COVID-19 pandemic demonstrated how challenging swift and effective response to pathogens with pandemic potential truly is. This reinforces the importance of prioritizing biosafety and biosecurity in research to protect the researcher, the community, and the environment. At CSU, researchers and safety professionals worked together to ensure the highest level of safety possible while responding to critical scientific and societal needs. Research teams were forming, combining expertise in new ways to ensure the personnel were safe. For example, many groups with experience in *Mycobacterium tuberculosis*, a respiratory transmitted bacteria, teamed up with arbovirologists, non-respiratory transmitted virology experts. Lab groups collaborated with community members and businesses to provide things like hand sanitizer, which was difficult to come by. High-containment researchers and lab animal resources professionals teamed up to serve the healthcare community by decontaminating much needed PPE during the height of the shortages. The Veterinary Diagnostic Lab shifted to obtain Clinical Lab Improvement Amendments (CLIA) certification so they could become a human COVID-19 testing center for the state. Other researchers put their expertise toward saliva based screening to help answer the demand for COVID-19 testing at CSU. All of these things, and many more, were incredible feats; the IBC, BSO, and Occupational Health groups at CSU had a hand in many of them.

I spent a lot of time as the only graduate student in the room, both in my lab and in my internships. It highlighted the very unique experience that graduate school is. I have found that even people who have been through it before tend to forget certain things about what it is like or do not realize how it has changed. I am extremely grateful to have had the internship opportunities I have had, especially when everything shut down for COVID-19. I had something to do and I had a way to help, a nice distraction from the anxiety around trying to finish graduate school. Looking back on the amazing things that happened in response to the pandemic, it is important to appreciate just how vital it was that people stopped their work, and they stayed home. In some ways I think that was a more difficult job than

participating in response efforts. I believe that graduate students who went through COVID-19 learned a different kind of resilience, and I hope that ultimately the world becomes better because of it.

One of the most challenging parts of risk assessment for me lies in considering the uncertainty. As scientists we are trained to rely on data to make informed decisions, but in many cases there simply are no data. When considering controls to implement, one of the most important questions to ask is: does this actually reduce risk and make the research more safe or secure? Sometimes we simply do not know. In those situations, it is vital to listen to the stakeholders and revisit matters frequently as hopefully more information emerges. Combining my BSL-3 research experience with my biosafety and biosecurity education has helped me understand how to define risk while tolerating and even appreciating the gray. My appreciation for this was heightened because I entered biosafety and biosecurity during an extremely uncertain time, and that experience will stay with me forever.

4.7 References

[1] Risk Definition & Meaning - Merriam-Webster n.d. https://www.merriam-webster.com/dictionary/risk (accessed January 25, 2022).

[2] CDC, Nih. Biosafety in Microbiological and Biomedical Laboratories 6th Edition Centers for Disease Control and Prevention National Institutes of Health. 2020.

[3] National Institutes of Health, Department of Health and Human Services. NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules (NIH Guidelines). 2019.

[4] Jackson DA, Symonst RH, Berg - P, Lobban DP, Kaiser AD. Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of Escherichia coli. Proceedings of the National Academy of Sciences 1972;69:2904–9. https://doi.org/10.1073/PNAS.69.10.2904.

[5] Cohen SN, Chang ACY, Boyert HW, Hellingt RB. Construction of Biologically Functional Bacterial Plasmids In Vitro (R factor/restriction enzyme/transformation/endonuclease/antibiotic resistance) 1973;70:3240–4.

[6] Morrow JF, Cohen SN, Chang ACY, Boyer HW, Goodman HM, Helling RB. Replication and transcription of eukaryotic DNA in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 1974;71:1743–7. https://doi.org/10.1073/PNAS.71.5.1743.

[7] Berg P, Baltimore D, Boyer HW, Cohen SN, Davis RW, Hogness DS, et al. Potential biohazards of recombinant DNA molecules. Science 1974;185:303. https://doi.org/10.1126/SCIENCE.185.4148.303/ASSET/0E7197FA-6440-486B-80F7-BBBBC5568C32/ASSETS/SCIENCE.185.4148.303.FP.PNG.

[8] Wivel NA. Historical Perspectives Pertaining to the NIH Recombinant DNA Advisory Committee 2014. https://doi.org/10.1089/hum.2013.2524.

[9] United States Reports. Diamond v. Chakrabarty, 447 US 303 - Supreme Court 1980.

[10] Splicing Life The Social and Ethical Issues of Genetic Engineering with Human Beings. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research 1982.

[11] Johnson CM, Dobos KM. The Evolving Landscape of Institutional Biosafety Committees and Biosafety Programs: Results from a National Survey on Organizational Structure, Resources, and Practices. Applied Biosafety 2019;24:213–9. https://doi.org/10.1177/1535676019886175.

[12] Ranney ML, Griffeth V, Jha AK. Critical Supply Shortages — The Need for Ventilators and Personal Protective Equipment during the Covid-19 Pandemic. New England Journal of Medicine 2020;382:e41. https://doi.org/10.1056/NEJMP2006141/SUPPL_FILE/NEJMP2006141_DISCLOSURES.PDF.

[13] Schwartz A, Stiegel M, Greeson N, Vogel A, Thomann W, Brown M, et al. Decontamination and Reuse of N95 Respirators with Hydrogen Peroxide Vapor to Address Worldwide Personal Protective Equipment Shortages During the SARS-CoV-2 (COVID-19) Pandemic. Applied Biosafety 2020;25:67–70. https://doi.org/10.1177/1535676020919932/ASSET/IMAGES/LARGE/10.1177_1535676020919932-FIG1.JPEG.

[14] Smith JS, Hanseler H, Welle J, Rattray R, Campbell M, Brotherton T, et al. Effect of various decontamination procedures on disposable N95 mask integrity and SARS-CoV-2 infectivity. Journal of Clinical and Translational Science 2021;5:10–1. https://doi.org/10.1017/CTS.2020.494.

[15] Collins TA, Sparks AE, Walker MM, Kendall L v., Dobos KM, Bergdall VK, et al. Reuse of disposable isolation gowns in rodent facilities during a pandemic. Journal of the American Association for Laboratory Animal Science 2021;60:431–41. https://doi.org/10.30802/AALAS-JAALAS-20-000130.

[16] Zimmer C, Gorman J. Covid Lab-Leak Theory Renews "Gain-of-Function" Research Debate - The New York Times. 2021.

[17] Warmbrod KL, Montague MG, Gronvall GK. COVID-19 and the gain of function debates. EMBO Reports 2021;22:e53739. https://doi.org/10.15252/EMBR.202153739.

[18] National Academies Press. Perspective on Research with H5N1 Avian Influenza: Scientific Inquiry, Communication, Controversy: Summary of a Workshop. 2013.

[19] Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, et al. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature 2012;486:420–8. https://doi.org/10.1038/NATURE10831.

[20] Herfst S, Schrauwen EJA, Linster M, Chutinimitkul S, de Wit E, Munster VJ, et al. Airborne transmission of influenza A/H5N1 virus between ferrets. Science (New York, NY) 2012;336:1534–41. https://doi.org/10.1126/SCIENCE.1213362.

[21] United States Government Policy for Institutional Oversight of Life Sciences Dual Use Research of Concern. Federal Register; 2014.

[22] APP3 Statement on Biosecurity and Biosafety During the COVID-19 Pandemic – Global Health Security Agenda n.d. https://ghsagenda.org/2020/07/30/app3-statement-on-biosecurity-and-biosafety-during-the-covid-19-pandemic/ (accessed February 10, 2022).

Chapter 5 Conclusions and Future Directions

5.1 Concluding Remarks for Mtb EV Enrichment Explorations

Mycobacterium tuberculosis (Mtb) extracellular vesicles (EVs) are involved in the host-pathogen interaction and play a role in the bacterium's response to environmental stress [1]. Understanding more about Mtb EVs in the context of bacterial physiology and pathogenesis may contribute to the development of tuberculosis (TB) management and prevention strategies. These strategies are desperately needed as COVID-19 dampens progress on the control of an already devastating disease [2]. In this dissertation, we explored common Mtb EV enrichment techniques in order to clarify how various methods influence Mtb EV preparations. Asymmetric flow field-flow fractionation (AF4) was then utilized for high-resolution separation of Mtb EVs from soluble proteins. This work also defined proteomic differences and similarities in Mtb EVs of various sizes.

The size-exclusion chromatography (SEC) and ultracentrifugation (UC) based methods investigated in Chapter 2 all resulted in the enrichment of Mtb EVs with similar size and morphologies. This instills confidence that, despite sometimes conflicting results among publications, the material used in previous studies likely included Mtb EVs, even when visual confirmation by TEM or evidence of positive protein markers was not provided. From the simplest *in vitro* culture system with adequate nutrients, Mtb EVs are enriched most reproducibly by SEC. Commercially available qEV SEC using an automatic fraction collector provided Mtb EV preparations of moderate yield and high purity. This method is the most ideal for experiments where the goal is to understand the function of Mtb EVs themselves, separate from other secreted molecules.

Density gradient (DG) ultracentrifugation is the most commonly used Mtb EV enrichment technique in the literature to date. While ultracentrifugation remains the most utilized EV enrichment method in general (**Figure 5.1**), many other techniques are becoming increasingly popular, with SEC becoming more common than DG [3]. The reason for this shift can be partially attributed to the

documented drawbacks of UC-based methods, including low yield and reproducibility challenges [4,5]. Our results for Mtb EV enrichment aligned with the literature: DG preparations had the lowest particle yield for all biological replicates (**Figure 2.4**), and biological variation negatively impacted the reproducibility (**Figure 2.6**). While some biological variation is expected in any living system, our bacterial culture methods were kept as consistent as possible. Even so, DG-EV yield was so inconsistent that not all analyses were performed for all three biological replicates (**Figure 2.8** and **Figure AI.7**).

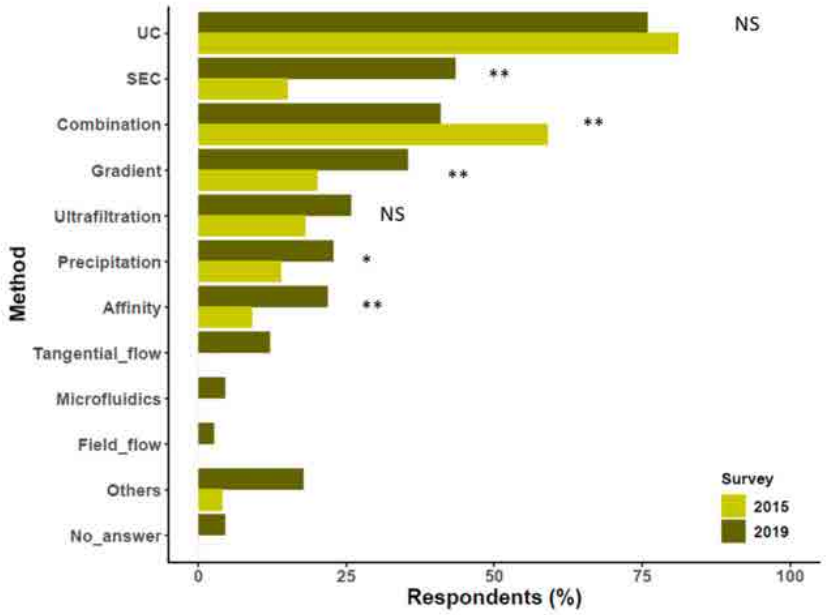


Figure 5.1 EV enrichment methods survey comparison. Results from two different surveys of scientists in the EV field conducted in 2015 [6] and 2019 [3]. Method options included ultracentrifugation (UC), size-exclusion chromatography (SEC), combination of methods (Combination), density gradient (Gradient), ultrafiltration (Ultrafiltration), precipitation-based methods (Precipitation), affinity capture (Affinity), tangential flow filtration (Tangential_flow), microfluidics (Microfluidics), field flow fractionation (Field_flow), other not listed (Others), and no answer (No_answer). Z- and chi-square tests were used to evaluate proportion changes between the surveys (NS = not significant, $p<0.05$*, $p<0.01$**, missing bars indicate option not included in 2015 survey). This figure is adapted from [3]; the final version is free according to the Creative Commons CC-BY license.

In an early study of DG-enriched Mtb EVs as potential vaccines, the major pitfall was lack of reproducibility [7]. Two-thirds of the trials performed showed that Mtb EV vaccination produced a protective immune response similar to standard BCG vaccination. Unfortunately, the third trial did not have the same results. The authors attribute this variation to slight changes in culture conditions, even though they used the same methods throughout. There was no report of quality control parameters for

evaluating their preparations. While biological variation may have been the cause of this inconsistency, our results suggest that an SEC-based enrichment method would provide more consistent Mtb EV preparations and may improve the overall reproducibility. Our study reinforces the importance of investigating the impact of a separation technique on the resulting EV preparation. Defining markers of EV purity, such as the presence of GroES in concentrated culture filtrate but not SEC or DG-enriched EVs by Western blot (**Figure AI.2**), provides vital quality control metrics for Mtb EV preparations. Technical and biological variation influence all Mtb EV enrichment methods, but the magnitude of this impact should be considered and empirically evaluated during experimental design and optimization.

In Chapter 3, the first reported application of AF4 for Mtb EV separation resulted in successful fractionation of Mtb EVs based on size (**Figure 3.2** and **Table 3.1**). Although we did not find clearly distinct proteomic signatures between smaller and larger Mtb EVs when considering the total protein content, we did uncover a more specific type of variation: post translational modification. Proteins known to be acylated and glycosylated cluster based on AF4 fractionation (**Figure 3.15**, **Figure 3.16**). We also saw differences in some of these proteins by Western blot, and we discovered a similar pattern with lipoarabinomannan (LAM) (**Figure 3.5** and **Figure 3.6**). This suggests that Athman et al. (2015) primarily separated smaller Mtb EVs in their density gradient of mixed origin EVs [8]. They found strong TLR-2 response to that population, and we see many known TLR-2 agonists, including LAM, more abundantly in the F2-F3 AF4 fractions. The fact that their LAM-rich EVs were more dense than the host EVs remains perplexing; however, we did see a lower particle to protein ratio in F2 and F3 (**Table 3.1**) and a higher number of proteins identified in those fractions (**Table AII7.3**), suggesting that these small EVs are more protein packed, increasing their density. In addition to proteomic interrogation of the Athman et al. (2015) density gradient fractions, repeating TEM with additional monoclonal antibodies specific to proteins more concentrated in smaller and larger Mtb EVs (rather than a LAM/LM biased polyclonal) would help clarify if there is as complete of a separation as the authors suggest. It would also

be beneficial to repeat that work with an upward displacement density gradient to achieve a more confident separation based on density.

Although we cannot definitively determine if non-reproducible and/or conflicting results in the literature can be attributed specifically to enrichment method, our experiments suggest that there are biologically relevant impacts due to the technique. We demonstrated that even the best technique evaluated in Chapter 2 (qEV) was not completely representative of the Mtb EVs resolved by AF4 (**Figure 3.17** and **Table 3.4**). Additionally, the impact of biological variation on DG-enriched Mtb EVs combined with the proteomic differences across Mtb EVs of different sizes suggests that the variation in the vaccine study was due to inconsistency in the density gradient preparations [7]. It is unclear what metrics were used to evaluate their Mtb EV preparations, and there is no mention of evaluating the presence or absence of certain biomolecules. Perhaps in their third experiment the Mtb EVs collected were towards the larger end of the spectrum, containing less LAM, LpqH, and SodC. In our DG method development (**Figure AI.4**) we demonstrate that markers for Mtb EVs vary slightly by density, with the highest and lowest density fractions containing LpqH but not LAM. This further supports reproducibility challenges as a cause for the inconsistent protection in the vaccination study. Even if they performed the density gradient in exactly the same manner, our study shows that DG-EVs are difficult to obtain in a highly reproducible manner.

Finally, evaluating our AF4 fractions proteomic data for proteins previously reported to be mycobacterial EV associated provides both agreement and conflict with the literature (**Table 5.1**) [1]. Of the 89 proteins listed, 32 were not detected in this study and 39 were detected but did not vary significantly across the fractions. Ten of the remaining 18 proteins agree with the literature and appear to be enriched in Mtb EVs. Two of the proteins (LprQ and PstS3) are high in F2 but drop off (particularly LprQ) and may not be truly enriched in Mtb EVs. Six proteins are not enriched in Mtb EV AF4 fractions. SubI and DppA drop off quite drastically. DppA is only present in S1 and S2 F2 but strong in the 100R and F1 for all samples, suggesting it might be a negative purity marker for Mtb EVs. Although Mpt53 and

Mpt64 conflict in terms of Mtb EV enrichment, our findings align with their detection in human serum EVs that are more likely to be loaded with non-MtbEV associated proteins [9,10]. Although the proteins presented in the table are related to all mycobacterial EVs and not Mtb EV specific, this comparison highlights the value of AF4 for establishing Mtb EV associated proteins.

Table 5.1 Comparing proteins between AF4 and Mtb EV literature. Proteins reported in a recent Mtb EV review [1] are compared to AF4 fraction NSAF ANOVA values. NS = not significant, NTD = not detected. Green indicates agreement with the literature, yellow indicates slight conflict, and orange indicates conflict.

Protein	P-value	NSAF Quantitative Profile
Virulence associated proteins		
ClpB	NS	
Cfp29	< 0.00010	100R low, F1 low, F2 high, F3 low, F4 high, F5 high, qEV high
Eis	0.0033	100R low, F1 high, F2 high, F3 high, F4 high, F5 high, qEV low
EphG	NS	
HspX	NS	
HtpG	NS	
GroEL1	NS	
GroES	NS	
GrpE	NS	
GlgX	NS	
KatG	NS	
MymT	NTD	
Tpx	NS	
TreS	NS	
SodB	0.00067	100R high, F1 high, F2 high, F3 low, F4 low, F5 high, qEV low
SodC	NS	
VapC10	NS	
VapC11	0.00016	100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV high
Membrane remodeling proteins		
Cut2	NTD	
CrgA	NTD	
CFP21	NTD	
CwsA	NS	
DacB2	NS	
EmbC	NTD	
FtsE	NS	
FtsH	NTD	
FtsK	NTD	
FtsY	NS	
PonA1	NS	
PonA2	NS	
ParA-related	NTD	
ParB-like	NTD	
SppA	NTD	
Tig	NS	
Wag31	NS	
Transport Associated proteins		
CtpC	NTD	
DppA	< 0.00010	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
DppD	NTD	
MMPL3	NS	
Mkl	NS	
MscL	NTD	
Rv1410c	NTD	
Rv1747	NTD	
SecA	NTD	
SecD	NTD	
TatA	< 0.00010	100R low, F1 low, F2 high, F3 high, F4 high, F5 high, qEV high
TrkB	NTD	

Protein	P-value	NSAF Quantitative Profile
Lipoproteins		
FecB	0.00098	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
LpqD	NTD	
LpqE	NS	
LpqH	< 0.00010	100R low, F1 low, F2 high, F3 high, F4 high, F5 low, qEV high
LpqI	NS	
LpqJ	NTD	
LpqL	NS	
LpqN	NS	
LpqT	NS	
LprA	NS	
LprC	NTD	
LprF	< 0.00010	100R low, F1 low, F2 high, F3 high, F4 high, F5 low, qEV high
LprG	< 0.00010	100R low, F1 low, F2 high, F3 high, F4 high, F5 low, qEV high
LprQ	< 0.00010	100R high, F1 high, F2 high, F3 low, F4 low, F5 low, qEV low
LppI	NTD	
LppO	NS	
LppX	0.00046	100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV high
LppZ	NS	
SubI	0.0018	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
ModA	NTD	
PstS1	NS	
PstS2	NS	
PstS3	< 0.00010	100R high, F1 high, F2 high, F3 low, F4 low, F5 low, qEV low
Adehsins		
Apa	NS	
DnaK	0.003	100R low, F1 low, F2 low, F3 low, F4 high, F5 high, qEV low
GroEL2	NS	
HBHA	NS	
Secretion associated proteins		
EsxA	NTD	
EsxB	NS	
EsxM	NTD	
EsxN	NTD	
EsxO	NTD	
EspA	NTD	
EspC	NTD	
EspK	NTD	
Immunogenic proteins		
Atg85A	NS	
Atg85B	NS	
Mpt53	0.0007	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV high
Mpt63	NTD	
Mpt64	0.0026	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
Mtc28	0.0032	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low

5.2 Mtb EV Enrichment Future Directions

There are many ways this work can be expanded to provide additional insight on Mtb EV composition and enrichment. Proteomic comparisions of the samples generated in Chapter 2 would solidify our understanding of the reproducibility of Mtb EV enrichment techniques. We hope to pursue this again, as the first attempt endured significant sample loss during peptide clean-up prior to MS. Fortunately, there is enough material to repeat LC-MS/MS of those samples in the future. Data generated from those samples could also be compared to the AF4 proteomics. Although it would not be a direct comparision because the starting 100R is different, we could still get a general picture of where Mtb EVs purified by CC, DG, and UC fall in comparision to the qEV and AF4 fractions. In particular, evaluating whether proteins that fall in AF4 F1 are in high or low abundance based on enrichment technique provides a better understanding of sample purity.

AF4 is a valuable tool for continuing to investigate Mtb EV composition. Based on the changes in LAM across the AF4 fractions (**Figure 3.5** and **Figure 3.6**), exploring lipoglycan, lipid, and glycolipid variation across the fractions may reveal additional heterogeneity in Mtb EVs based on their size. Mtb lipoglycans and lipids previously identified in Mtb EVs like LAM, phosphatidylinnositol (PI), and phosphatidylinnositol mannosidases (PIMs) are known to influence the host immune response and have even been used in vaccine adjuvants [11,12]. Further exploring the physiochemical differences found in the proteins based on AF4 fraction would also help solidify the potential implications of Mtb EV size. As previously mentioned, AF4 also highlights proteins that are not Mtb EV associated. Development of antibody-based detection methods or even a targeted mass spectrometry assay for purity evaluation of Mtb EV preparations would help ensure that any future attempts to evaluate Mtb EV functions are not skewed by co-purifying non-EV assocaited proteins.

Finally, the impact of environmental variation on Mtb EVs can also be evaluated using AF4. Since studies have shown that Mtb EV composition changes in response to the environment, it is important to recognize the limitations of just looking at the simplest system of Mtb EV culture in these

studies. Using AF4 to evaluate Mtb EVs from various conditions would allow a more thorough look at how the proteins change with confidence that the observed variation is Mtb EV associated and not soluble proteins. Also, AF4 is gentle compared to the other Mtb EV enrichment techniques in this study, as it does not involve high forces or movement through a resin. This increases the confidence that all Mtb EVs are being captured and evaluated. Although producing large quantities of Mtb EVs with this technology would be difficult at this time, AF4 is a promising technology for a thougough and high confidence evaluation of the true composition of Mtb EVs regardless of the culture conditions.

5.3 Risk Assessment and Biosafety and Biosecurity Closing Remarks

COVID-19 has shaken the world, and uniquely impacted the major components of my graduate education: tuberculosis research, biosafety, and biosecurity. My hope is that researchers who may have shifted priorities from TB to COVID-19 due to the urgent need for scientists experienced with respiratory pathogens find their way back to studying TB. The incredible speed at which SARS-CoV-2 vaccinations and treatments were developed shows just how quickly scientists can make a positive impact on public health crises. It would be wonderful if that kind of support can be given to fighting the TB crisis. In contrast to TB, the COVID-19 pandemic has brought a lot of positive opportunities to the biosafety and biosecurity profession. When I explained what I wanted to do before the pandemic began, only people involved in infectious disease research really understood what I was talking about. Now, PPE is a household term and I find that many more people are aware of the importance of biosafety and biosecurity, even if they don't know those specific words. I hope that this increased awareness of biosafety and biosecurity will spur conversations about how we approach biosafety and biosecurity from a global perspective. Additionally, I hope that researchers of all levels reflect on what they can learn from the pandemic, including the rampant misinformation and the revival of the dual-use research debate, and apply it to communication of their own work for years to come.

5.4 References

[1] Palacios A, Gupta S, Rodriguez GM, Prados-Rosales R. Extracellular vesicles in the context of Mycobacterium tuberculosis infection. Molecular Immunology 2021;133:175–81. https://doi.org/10.1016/J.MOLIMM.2021.02.010.

[2] World Health Organization. Global Tuberculosis Report 2021.

[3] Royo F, Théry C, Falcón-Pérez JM, Nieuwland R, Witwer KW. Methods for Separation and Characterization of Extracellular Vesicles: Results of a Worldwide Survey Performed by the ISEV Rigor and Standardization Subcommittee. Cells 2020, Vol 9, Page 1955 2020;9:1955. https://doi.org/10.3390/CELLS9091955.

[4] Torres Crigna A, Fricke F, Nitschke K, Worst T, Erb U, Karremann M, et al. Inter-Laboratory Comparison of Extracellular Vesicle Isolation Based on Ultracentrifugation. Transfusion Medicine and Hemotherapy : Offizielles Organ Der Deutschen Gesellschaft Fur Transfusionsmedizin Und Immunhamatologie 2021;48:48–59. https://doi.org/10.1159/000508712.

[5] Dauros Singorenko P, Chang V, Whitcombe A, Simonov D, Hong J, Phillips A, et al. Isolation of membrane vesicles from prokaryotes: a technical and biological comparison reveals heterogeneity. Journal of Extracellular Vesicles 2017;6. https://doi.org/10.1080/20013078.2017.1324731.

[6] Gardiner C, Vizio D di, Sahoo S, Théry C, Witwer KW, Wauben M, et al. Techniques used for the isolation and characterization of extracellular vesicles: Results of a worldwide survey. Journal of Extracellular Vesicles 2016;5. https://doi.org/10.3402/JEV.V5.32945/SUPPL_FILE/ZJEV_A_11821259_SM0001.PDF.

[7] Prados-Rosales R, Carreño LJ, Batista-Gonzalez A, Baena A, Venkataswamy MM, Xu J, et al. Mycobacterial membrane vesicles administered systemically in mice induce a protective immune response to surface compartments of mycobacterium tuberculosis. MBio 2014;5. https://doi.org/10.1128/mBio.01921-14.

[8] Athman JJ, Wang Y, McDonald DJ, Boom WH, Harding C v., Wearsch PA. Bacterial Membrane Vesicles Mediate the Release of Mycobacterium tuberculosis Lipoglycans and Lipoproteins from Infected Macrophages . The Journal of Immunology 2015;195:1044–53. https://doi.org/10.4049/jimmunol.1402894.

[9] Mehaffy C, Dobos KM, Nahid P, Kruh-Garcia NA. Second generation multiple reaction monitoring assays for enhanced detection of ultra-low abundance Mycobacterium tuberculosis peptides in human serum. Clinical Proteomics 2017;14. https://doi.org/10.1186/s12014-017-9156-y.

[10] Mehaffy C, Kruh-Garcia NA, Graham B, Jarlsberg LG, Willyerd CE, Borisov A, et al. Identification of Mycobacterium tuberculosis Peptides in Serum Extracellular Vesicles from Persons with Latent Tuberculosis. Journal of Clinical Microbiology 2020;58. https://doi.org/10.1128/JCM.00393-20.

[11] Prados-Rosales R, Baena A, Martinez LR, Luque-Garcia J, Kalscheuer R, Veeraraghavan U, et al. Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. Journal of Clinical Investigation 2011;121:1471–83. https://doi.org/10.1172/JCI44261.

[12] Ishikawa E, Mori D, Yamasaki S. Recognition of Mycobacterial Lipids by Immune Receptors. Trends in Immunology 2017;38:66–76. https://doi.org/10.1016/J.IT.2016.10.009.

Appendix I: Chapter 2 Supplementary Information

6.1 Selection of Mtb EV Enrichment Workflows

Primary research articles and methods publications involving mycobacterial EV enrichment vary significantly. First, the overall processes for each experiment were compared by general enrichment methods which is presented in **Table AI.1**. Only studies where EVs were enriched from strictly bacterial culture were used, as this is the simplest source of mycobacterial EVs. Experiments involving mycobacterial infection of eukaryotic tissue cultures or animals add a layer of complexity in determining eukaryotic vs prokaryotic EV contributions. Overall, five techniques including ultrafiltration (UF), differential centrifugation (DC), ultracentrifugation (UC), density gradient (DG), and size exclusion chromatography (SEC), were used alone or in tandem. The purpose of each method was challenged prior to inclusion/exclusion from the final EV enrichment workflow. Ultrafiltration (MWCO 100 kDa) was performed in the majority of the procedures (68.4%, n=19) for the removal of small soluble biomolecules and volume reduction; therefore, UF was selected for inclusion as the first step of EV enrichment on concentrated CFP.

Table AI.1 Method summary for mycobacterial EV publications. The over-arching techniques used for mycobacterial EV separation in each publication are indicated including the species and strains used if the information was available in the publication. UF: ultrafiltration, DC: differential centrifugation, UC: ultracentrifugation, DG: density gradient UC, SEC: size-exclusion chromatography

Year	Species/Strains	UF	DC	UC	DG	SEC	Reference
2007	*M. ulcerans*	N	Y	Y	N	N	[1]
2011	*M. bovis* BCG, Mtb H37Rv, Mtb H37Ra, *M. kansasii, M. avium, M. smegmatis, M. phlei*	Y	Y	Y	Y	N	[2]
2013	Mtb H37Rv, mutants	Y	Y	Y	Y	N	[3]
2013	*M. bovis* BCG, Mtb H37Rv	Y	N	Y	N	N	[4]
2014	Mtb mutants	Y	Y	Y	Y	N	[5]
2014	Gram positive and mycobacteria	Y	Y	Y	Y	N	[6]
2014	*M. bovis* BCG, Mtb H37Rv	Y	Y	Y	Y	N	[7]
2015	Mtb H37Rv	Y	Y	Y	Y	N	[8]
2015	Mtb H37Ra	N	Y	Y	Y	N	[9]

Year	Species/Strains	UF	DC	UC	DG	SEC	Reference
2015	*M. smegmatis*	N	N	Y	N	N	[10]
2016	Mtb	N	Y	Y	N	Y	[11]
2017	Mtb H37Rv, Mtb H37Ra	N	Y	Y	N	Y	[12]
2017	*M. smegmatis*	Y	Y	Y	Y	Y	[13]
2018	Mtb Erdman, mutants	Y	N	Y	N	N	[14]
2019	*M. bovis* NEIKER1403, *M. avium subsp. paratuberculosis* K10, Mtb H37Rv	Y	Y	Y	Y	N	[15]
2019	*M. avium subsp. hominissuis* 104	Y	Y	Y	Y	N	[16]
2019	Iron-limited mycobacteria	Y	Y	Y	Y	N	[17]
2020	Mtb H37Rv, mutants	N	Y	Y	Y	N	[18]
2021	*Mycobacterium spp.*	Y	N	N	N	Y	[19]

6.1.1 Exclusion of Differential Centrifugation

Differential centrifugation was also widely used (73.7%, n=19); the most common differential centrifugation technique after 100 kDa ultrafiltration used 4,000 x g then 15,000 x g on the supernatant for 15 min at 4 °C [2–4,6,7,15]. To evaluate the necessity of those steps, 1 mL of approximately 1.5 mg/mL of 100R, in triplicate, was centrifuged for those intervals, with a portion of the material retained for BCA after each step. Importantly, no pellet was visible in any of the tubes after either centrifugation round, and a small amount of supernatant was left in each tube to prevent disrupting any non-visible pelleted material.

Because CFP harvest is achieved by filtration, we hypothesized that several steps included in the published DC process are unnecessary, as whole cells and large cell debris are eliminated by the 0.2 micron filter. Based on the BCA results presented in **Table AI.2**, there is only a minor difference in protein content following the differential centrifugation steps; thus, this process was not included in the final workflow. Since all methods involve the exact same pre-processing steps for this comparison, the impact of not including differential centrifugation, if any, is present in the starting material for each separation technique.

Table AI.2 Impact of differential centrifugation on 100R protein concentration. Three replicates of 100R were subjected to differential centrifugation. The protein concentration of the supernatant was evaluated and compared to 100R that was not centrifuged.

Sample ID	Concentration after 4,000 x g then 15,000 x g centrifugation (mg/mL)	Concentration after sitting at 4°C for the centrifugation time periods	Difference (mg/mL)
Replicate 1	1.353	1.351	0.002
Replicate 2	1.397	1.340	0.057
Replicate 3	1.452	1.316	0.136
Average	1.401	1.337	0.065
Standard Deviation	0.050	0.018	0.068

6.1.2 Selection of Ultracentrifugation Parameters

After elimination of differential centrifugation, the methods involving just ultracentrifugation were compared (21.1%, n = 19) [1,4,10,14]. The force and time of centrifugation as reported in each publication included four variations: (1) 40,000 x g for 3 h [1], (2) 60,000 rpm for 1 h at 4 °C [4], (3) 150, 000 x g for 2 h at 4 °C [10], and (4) 100,000 x g for 75 min at 4 °C [14]. Only one article specified the type of rotor used and none included important parameters such as the rotor k-factor. Because ultracentrifugation at 100,000 x g for 1 h at 4 °C was commonly used prior to density gradient ultracentrifugation, those parameters were chosen so that the material resulting from the ultracentrifugation alone could be more directly compared to the density gradient enrichment technique. Additionally, the advantages of a swinging bucket vs fixed angle rotor for EV preparation have been previously described and where the rotor is mentioned for density gradient from these publications a swinging bucket was used; thus, all ultracentrifugation steps use the same swinging bucket rotor, tubes, and ultracentrifuge.

6.1.3 Selection of Density Gradient Ultracentrifugation Parameters

Defining the density gradient ultracentrifugation parameters was also challenging due to slight variations in procedures and lack of specific reporting. Of the twelve publications using density gradient ultracentrifugation (63.2%, n=19), only one uses sucrose as the gradient material [9]. The remaining

publications used iodixanol which is the ingredient used in OptiPrep™. Because the gradient step ranges varied among protocols, the highest and lowest iodixanol percentages across all the publications were used, with 40% to 5% overlayed on the UC-EV pellet suspended in 45% iodixanol. The most common centrifugal force of 100,000 x g and time of 16 h was selected. All methods that reported a temperature for this step used 4 °C and the majority took 1 mL fractions from the top of the gradient. While ultrafiltration for buffer exchange was not the most commonly used technique, it was successfully demonstrated for mycobacterial EV recovery [13] and prevented an additional ultracentrifugation step for concentration after traditional dialysis.

6.1.4 Selection of Size-Exclusion Chromatography Parameters

Selecting the size exclusion chromatography methods was more straightforward. The majority of publications using SEC employed a commercially available ready to use column from Izon Science, the qEV [11,12,20]. The 35 nm resin and original column size was selected based on the manufacturer recommendations. Based on previous work from the Dobos and Kruh-Garcia Laboratories—including the development of Mtb EVs as a product for the BEI repository—the second SEC protocol, using Capto™Core 700 slurry, was adapted from [19,21,22]. This slurry is slightly different than traditional SEC, as it is a multimodal chromatography resin. The beads are cross-linked agarose, 90 μm in diameter, with holes containing a molecular weight cut-off of 700 kDa. The inner core of the beads contains octylamine which captures soluble proteins that can enter the core (<700 kDa), so that larger Mtb EVs elute and small contaminating proteins remain in the column. The protein loading capacity per the manufacturer is 14 mg/mL, well above the 5 mg/mL upper limit used in our laboratory for Mtb EV enrichment. A visual summary of the methods for comparison is presented in **Figure AI.1**.

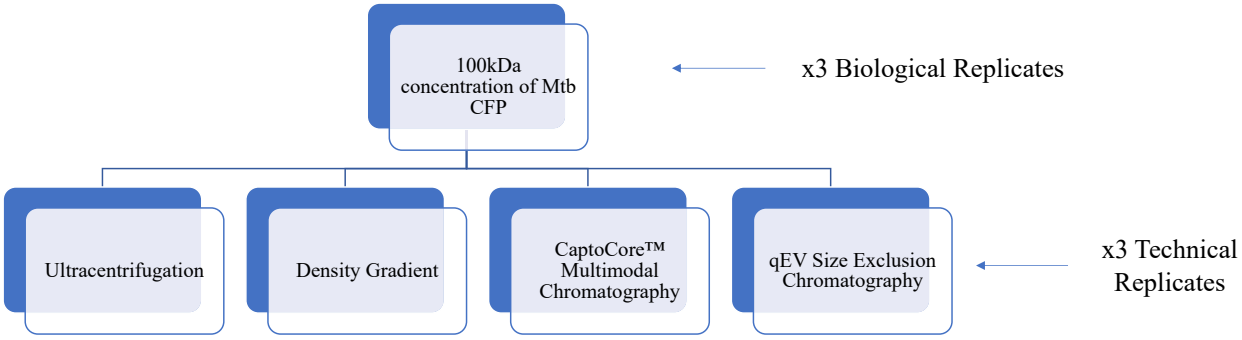


Figure AI.1 Mtb EV enrichment methods. 100 kDa concentrated CFP from three separate batches of CFP were put through each enrichment method workflow three times for a total of 12 Mtb EV preparations from each CFP batch.

6.2 Optimization of Selected Mtb EV Enrichment and Analysis Methods

Sixteen liters of unconcentrated Mtb H37Rv CFP were generously provided by Anne Simpson (Dobos laboratory) for optimization of these techniques. Bacterial cultivation was performed as described in section **2.2.2.1** Bacterial culture with double the quantity (2 Fernbach cultures seeding 40 roller bottles). CFP concentration and qualification procedures were performed as described in section **2.2.2.3** CFP ultrafiltration, with 16 L of CFP added to the pressure vessel instead of 8 L.

In order to directly compare the Mtb EV yields across enrichment methods, the same amount of 100R was needed for each technical replicate. Because variation in total EV output was expected [20], each method was evaluated to determine the minimum amount of 100R required to obtain enough EV material for analysis. The analyses performed were selected to address specific aspects of technical reproducibility. First, the vesicle yield, size, and phenotype should be similar from the same amount of 100R which was evaluated by NTA and TEM. Protein yield should also be similar from the same amount of 100R, which was evaluated by microBCA. Finally, the protein content should be similar from the same amount of 100R, which was evaluated by silver stained protein gel and Western blot.

6.2.1 Establishing SEC Methods and EV Analysis Process

Based on our previous work with CC-EVs (data not shown), we estimated at least 65 µg of EV material was required for each technical replicate. Previous preparations of CC-EVs also suggested that 5-

10% of 100R protein was vesicle associated (data not shown). With this information, 1.5 mg of 100R was used for CC-EV generation as a starting point. **Table AI.3** describes the protein and particle yield by BCA and NTA respectively for three replicates of CC-EVs. The first attempt at NTA for this material involved running 5 µg/mL on the instrument, which was well beyond the limit of detection. Instead, diluting (1:100) 10 µL CC-EV in 990 µL of 1X PBS was optimal for measurement. For these three replicates, 10 µL is approximately 1.22 µg of CC-EV. All subsequent NTA analysis presented was analyzed based on a volumetric dilution, not protein concentration.

Table AI.3 Evaluation of 100R loading quantity for EV-CC. Protein and vesicle recovery for three technical replicates of CC-EV preparation averaged to over 65 µg of protein from 1.5 mg of 100R.

Replicate	Total protein (µg)	Total particles	Particles per µg protein
1	69.329	3.16E+10	4.56E+8
2	67.610	4.77E+10	7.06E+8
3	77.365	6.10E+10	7.88E+8
Average	71.435	4.68E+10	6.50E+8

Normalizing the amount of material loaded on an SDS-PAGE gel for silver stains and Western blots allows direct comparison between samples. In EV biology, the values used to achieve normalization can be based on a variety of measurements including protein concentration, lipid concentration, vesicle counts, and various ratios. Our previous work has demonstrated that normalization by protein or vesicle count allows for comparisons among samples (data not shown) with human biofluid-derived EVs. Based on that work, a comparison of Mtb EVs normalized by vesicle count versus protein was performed. The CC-EV replicates were used to determine the best strategy to visualize the results (**Figure AI.2**). Western blot of the CC-EVs demonstrated that GlcB and GroES are absent, while LpqH and LAM are present with consistent intensity across replicates when normalized by either strategy. The silver stain shows distinguishable banding when CC-EVs are loaded based on particle count rather than protein. Most likely, the smearing that occurs when 5 µg is loaded is due to overloading of vesicles. The BCA assay reagents likely do not efficiently penetrate the membrane of EVs thus there is more protein present than measurable by BCA. Additionally, these replicates have an average of 6.50E+8 particles per µg protein,

meaning approximately three times as many vesicles are present in lanes 8 – 10 compared to lanes 5-7; the additional lipids may prevent efficient SDS-PAGE resolution.

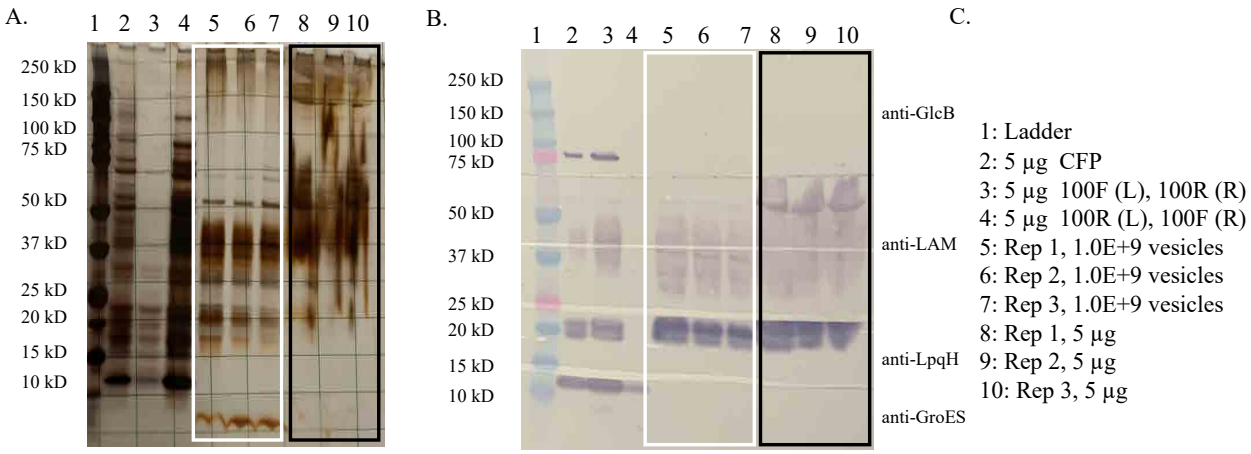


Figure AI.2 Evaluation of gel loading normalization strategies. CC-EV loaded in 1.0E+9 vesicles or 5 µg quantities as compared to 5 µg of the original CFP, 100F, and 100R for (A) silver stain and (B) Western blot evaluation. The sample layout key is provided (C), and CC-EV samples loaded by vesicle count is outlined in white (lanes 5-7) while those loaded by protein are outlined in black (lanes 8-10). Note: lanes 3 and 4 are 100F then 100R on the silver stain but loaded opposite (100R then 100F) on the Western blots.

The confirmation of positive and negative markers for Mtb EVs by Western blot is also demonstrated in **Figure AI.2 B**. Mtb EVs are enriched in lipoproteins, and LpqH has consistently been reported as a dominant component [2,7,9,14,19]. **Figure AI.2 B** shows that LpqH is present in CFP, enriched in the 100R, and absent in the 100F. LpqH intensity is consistent with the enrichment of EVs from the 100R, making it an ideal positive control. GroES is present in CFP, slightly enriched in the 100R, and appears in the 100F. In contrast to LpqH, GroES is absent after the enrichment of EVs despite the increased intensity in the 100R fraction and is an ideal negative control.

To optimize qEV SEC, the fractions containing Mtb EVs needed to be identified. The initial loading quantity for qEV was 1.5 mg because this resulted in sufficient CC-EV for analysis. The qEV fractions were analyzed by Western blot (**Figure AI.3**) following microBCA and NTA. Ultimately fractions 1-3 were pooled based on the strong presence of LpqH and LAM and absence of GroES to obtain the purest EV population possible.

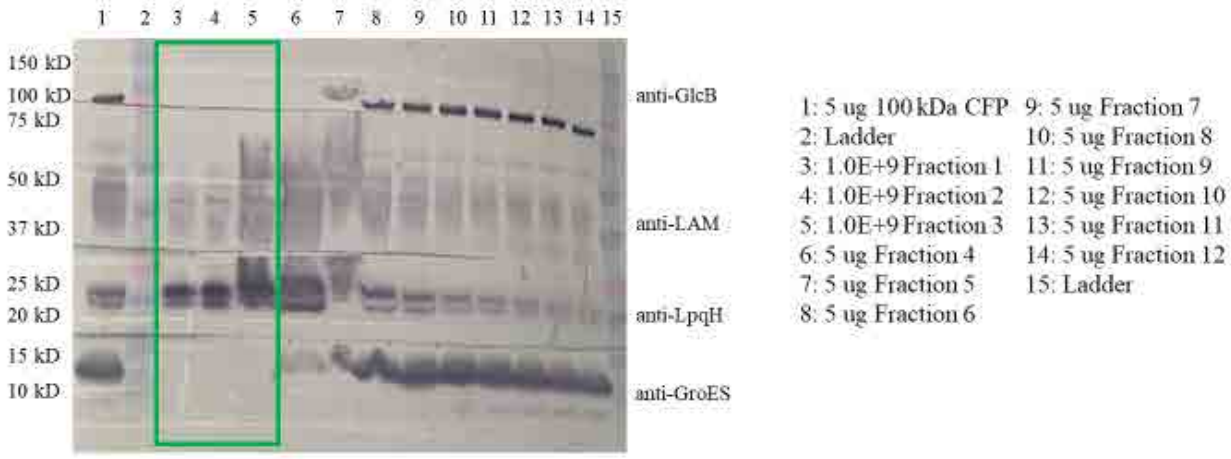


Figure AI.3 Determining qEV fraction pooling. qEV fractions loaded by protein concentration or vesicle count compared to the 100R starting material. The sample layout key is provided. Fractions 1-3 (lanes 3-5) were ultimately chosen for pooling.

The protein yield per fraction from qEV was much lower than Capto™Core. Pooling the first three fractions resulted in approximately 25 µg of total protein (n = 3); however, the particle counts by NTA were high, with the total averaging 7.3E+10 (n = 3), a twice as much as CC-EV (**Table AI.3**). The previous goal of generating 65 µg of EVs for analysis was therefore updated based on qEV-EV yield and comparison of normalization strategies in **Figure AI.2**. In order to perform all the desired analyses, each replicate needed to have at least 3E+10 particles.

6.2.2 Establishing Ultracentrifugation Methods

Optimizing ultracentrifugation as a method required confirmation that MtbEVs were recovered and that the 100R input was sufficient to yield enough UC-EVs for analysis. Staring with 1.5 mg of 100R yielded an average of 15 µg of protein and 7.7E+10 particles by NTA (n = 3). Confirmation of Mtb EV presence is shown in **Figure AI.4**.

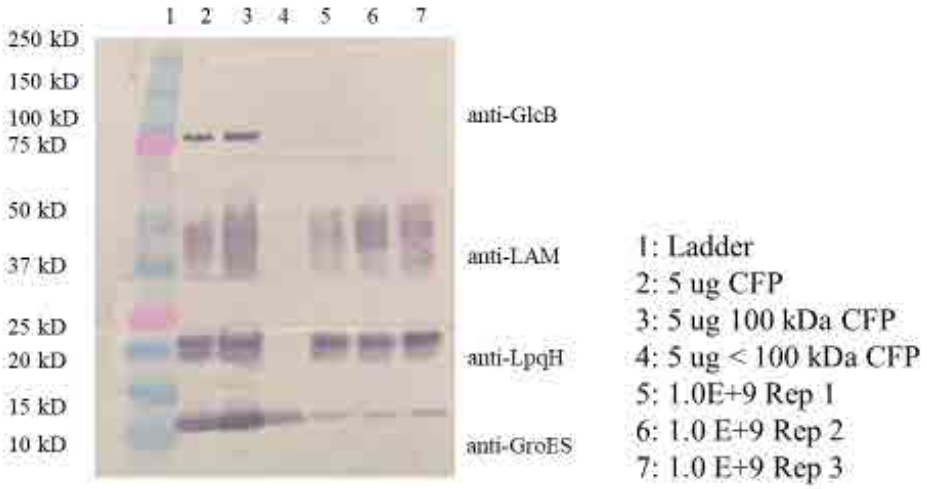


Figure AI.4 Mtb EV markers found after ultracentrifugation. UC-EVs were compared to CFP, 100R, and 100F for the presence of positive markers of Mtb EVs (LpqH, LAM) and contaminant markers (GroES).

Density gradient UC optimization involved determining the EV positive fractions and confirming the starting quantity. Because the UC-EV just matched the recovery for CC-EV and qEV, the starting 100R was increased due to anticipation of a decreased yield compared to UC-EV due to additional manipulation. After analyzing all 17 fractions by microBCA and NTA (data not shown), we expected Mtb EVs to be present somewhere in fractions 6-12. Based on Western blot analysis (**Figure AI.5**), fractions 8-12 were identified as Mtb EV positive fractions. As a pool, fractions 8-12 yield an average of 18 µg of protein and 3.6E+10 particles by NTA (n = 3).

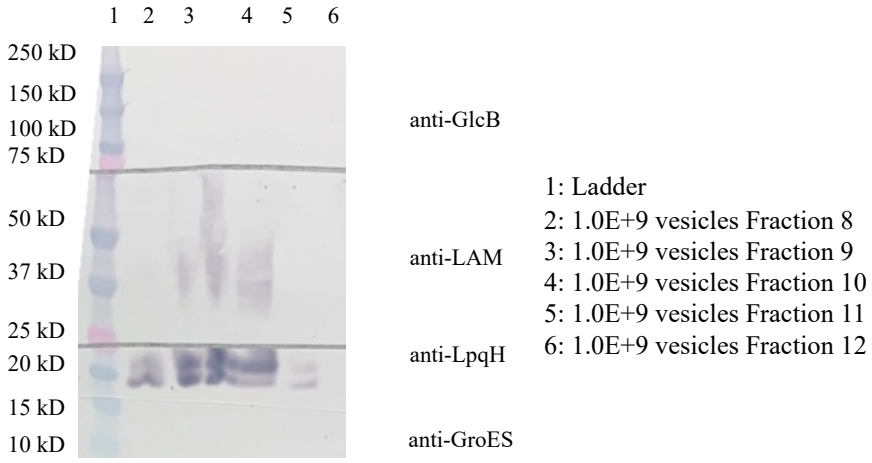
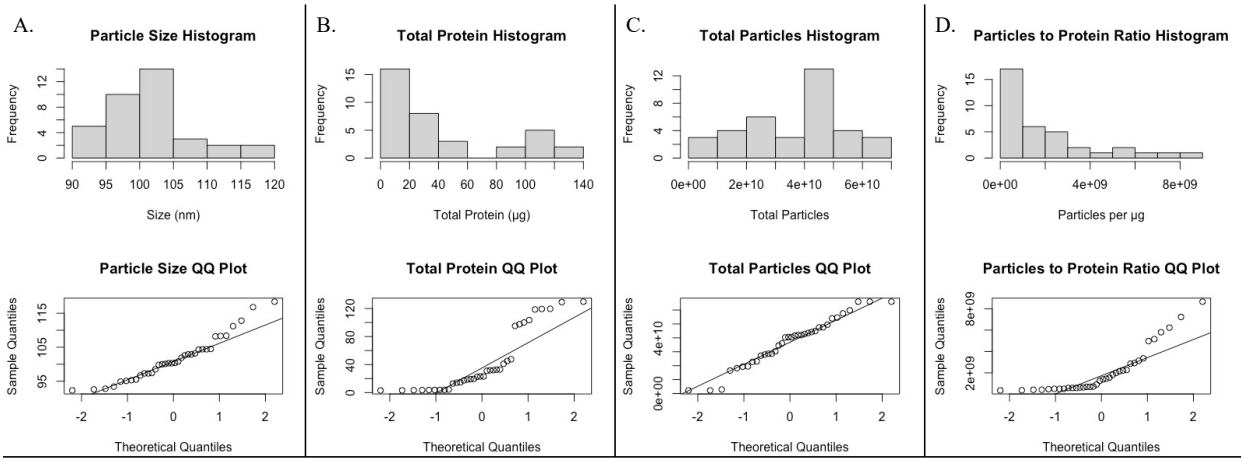


Figure AI.5 Mtb EV markers in density gradient fractions. Density gradient fractions were evaluated for the presence of positive markers of Mtb EVs (LpqH, LAM), and contaminant markers (GroES).

6.3 Chapter 2 Supporting Tables and Figures



E.

Shapiro-wilks test	Particle Size	Total Protein	Total Particles	Particles to Protein Ratio
Statistic	0.94947	0.79031	0.9644	0.779
P-value	0.05253	1.051E-5	0.2926	6.4E-6

Figure AI.6 Normality Testing for all Mtb EV enrichment methods. The histograms and QQ plots are displayed for (A) particle size, (B) protein recovery, (C) particle recovery, and (D) particle to protein ratio for all technical and biological replicates. (E) The statistic and p-value for Shapiro-Wilks test of these data is also included. Significant p-values (<0.05) are in red font.

Table AI.4 Protein recovery by method.

Method	Biological Replicate	Technical Replicate Average Total Protein (µg)	Technical Replicate Standard Deviation	Technical Replicate CoV (%)	Average Technical Replicate CoV (%)	Method Average Total Protein (µg)	Method Standard Deviation	Method CoV (%)
CC	1	113.01	11.20	9.91	6.40	112.54	13.54	12.03
CC	2	98.77	4.31	4.36				
CC	3	125.85	6.22	4.94				
qEV	1	20.01	2.86	14.31	10.22	18.51	3.95	21.33
qEV	2	13.95	0.91	6.49				
qEV	3	21.59	2.13	9.86				
UC	1	41.87	7.78	18.59	20.52	34.80	8.69	24.97
UC	2	27.52	7.67	27.87				
UC	3	34.99	5.29	15.12				
DG	1	3.97	0.42	10.46	8.76	3.68	0.39	10.52
DG	2	3.41	0.41	12.07				
DG	3	3.66	0.14	3.77				

Table AI.5 Particle recovery by method.

Method	Biological Replicate	Technical Replicate Average Total Particles	Technical Replicate Standard Deviation	Technical Replicate CoV (%)	Average Technical Replicate CoV (%)	Method Average Total Particles	Method Standard Deviation	Method CoV (%)
CC	1	4.51E+10	2.41E+09	5.34	6.23	5.48E+10	8.50E+09	15.51
CC	2	6.32E+10	4.85E+09	7.67				
CC	3	5.61E+10	3.19E+09	5.68				
qEV	1	3.95E+10	4.33E+09	10.96	18.37	4.38E+10	9.36E+09	21.36
qEV	2	4.35E+10	5.20E+09	11.95				
qEV	3	4.85E+10	1.56E+10	32.19				
UC	1	2.34E+10	4.29E+09	18.29	21.49	3.35E+10	1.01E+10	30.10
UC	2	3.92E+10	9.67E+09	24.65				
UC	3	3.77E+10	8.12E+09	21.53				
DG	1	2.45E+09	3.58E+08	14.63	11.97	1.58E+10	1.09E+10	68.99
DG	2	1.80E+10	1.41E+09	7.84				
DG	3	2.69E+10	3.61E+09	13.44				

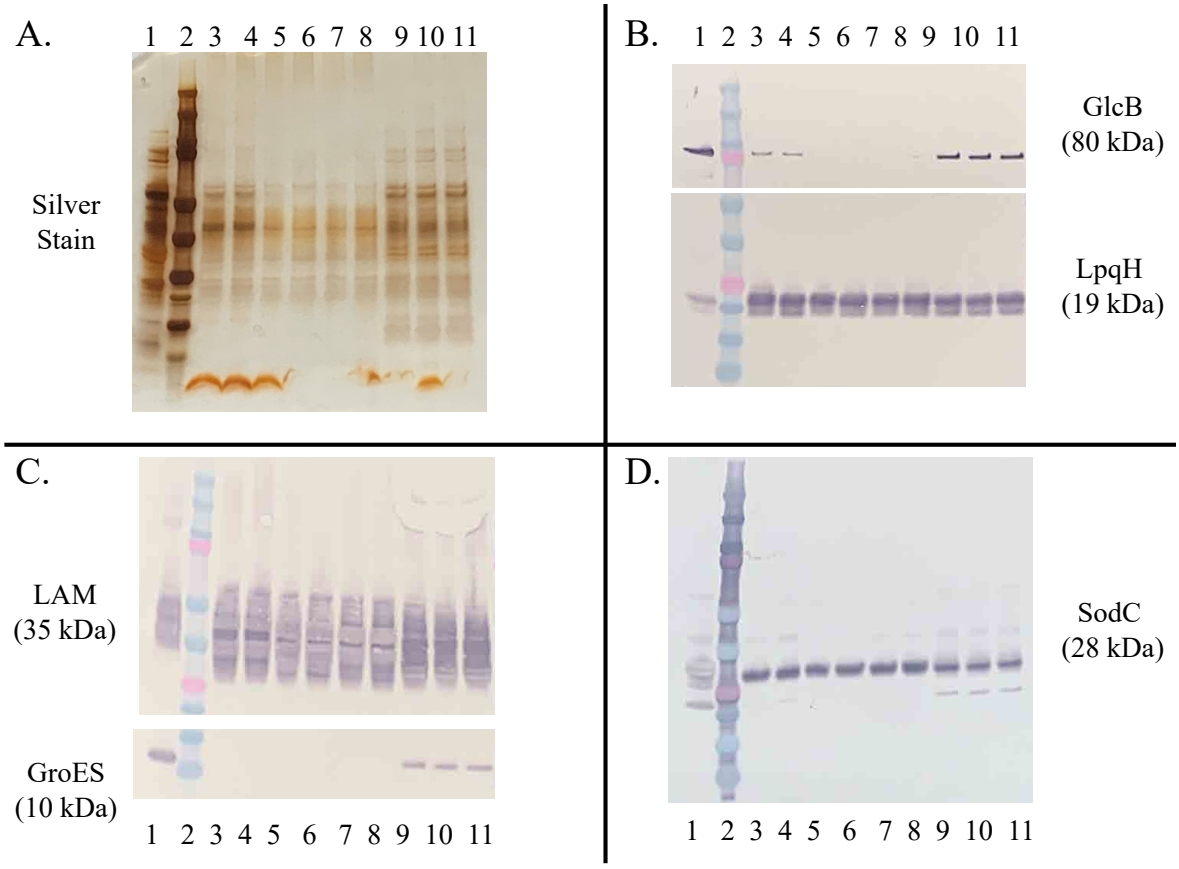


Figure AI.7 Silver stain and Western blots of biological replicate 1. (A) Silver stain and (B-D) Western blots for each technical replicate of biological replicate 1. All are loaded in the following format: 1 = 5 µg 100R, 2 = ladder, 3-5 = 1E9 CC-EV 1.1-1.3, 6-8 = 1E9 qEV-EV 1.1-1.3, and 9-11 = 1E9 UC 1.1-1.3. Note: due to recovery biological replicate 1 DG-EVs were not included for silver stain and Western blots.

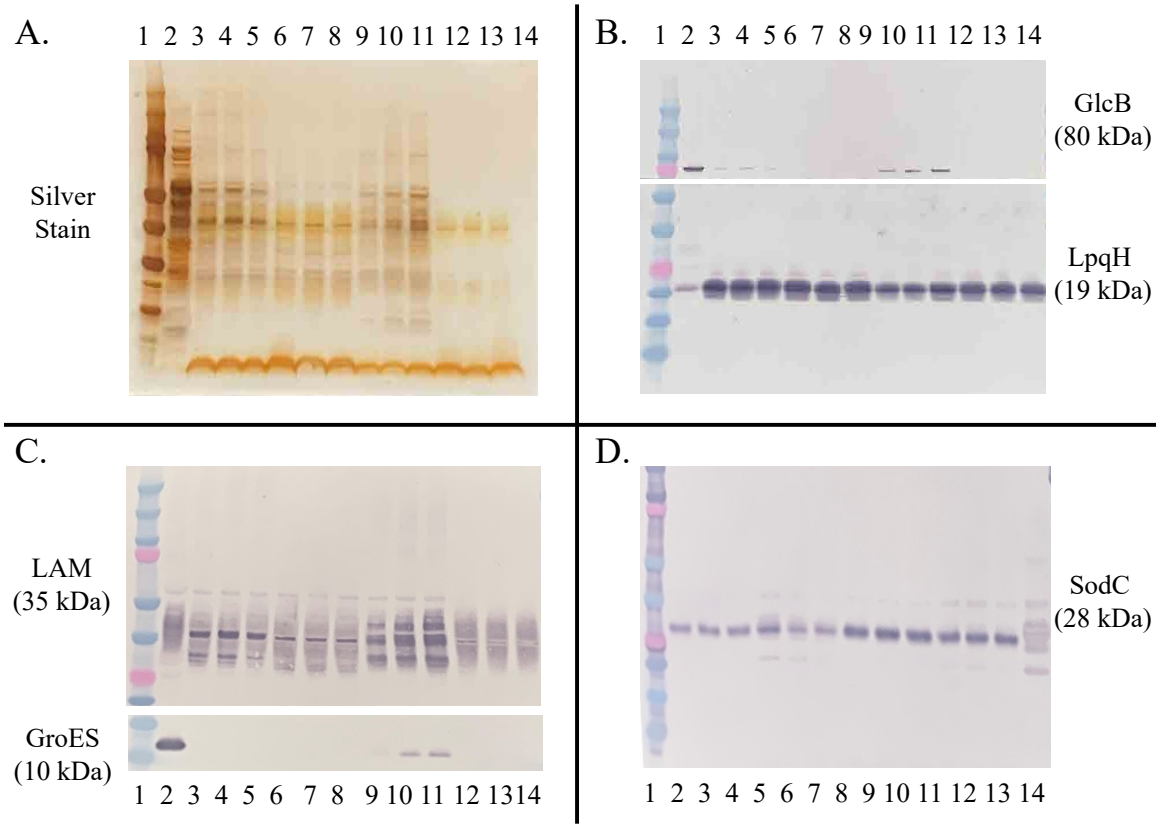


Figure AI.8 Silver stain and Western blots of biological replicate 3. (A) Silver stain and (B-D) Western blots for each technical replicate of biological replicate 3. All are loaded in the following format: 1 = ladder, 2 = 5 µg 100R, 3-5 = 1E9 CC-EV3.1-3.3, 6-8 = 1E9 qEV-EV 3.1-3.3, 9-11 = 1E9 UC 3.1-3.3, and 12-14 = 1E9 DG 3.1-3.3 except for D where 2-4 = 1E9 CC-EV 3.1-3.3, 5-7 = 1E9 qEV-EV 3.1-3.3, 8-10 = 1E9 UC 3.1-3.3, 11-13 = DG 3.1-3.3 and 14 = 5 µg 100R.

6.4 R Code

R script is available at: https://gist.github.com/jmryan12/511c325342858769233e1f2abf3e33b2

6.4.1 Set up libraries

```
#Load libraries
library(datasets)
library(ggplot2)

## Warning in register(): Can't find generic `scale_type` in package ggplot2 to
## register S3 method.

library(multcompView)
library(ggsignif)
library(dplyr)

##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

library(ggpubr)
library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✓ tibble  3.1.6     ✓ purrr   0.3.4
## ✓ tidyr   1.1.4     ✓ stringr 1.4.0
## ✓ readr   2.1.1     ✓ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(rmarkdown)
library(knitr)
knitr::opts_chunk$set(
    echo = TRUE,
    message = TRUE,
    warning = TRUE
)
```

6.4.2 Load and format data

```
Ch2MethodComp <- read.csv("~/Documents/R/Ch2MethodComp.csv")
Ch2MethodComp$Biological.Replicate <- as.factor(Ch2MethodComp$Biological.Replicate)
print(Ch2MethodComp)

##    Method Biological.Replicate Technical.Replicate Total.Protein
## 1      CC                    1                   1        119.60
## 2      CC                    1                   2        119.35
## 3      CC                    1                   3        100.07
## 4      CC                    2                   1         97.55
## 5      CC                    2                   2        103.55
## 6      CC                    2                   3         95.20
## 7      CC                    3                   1        129.76
## 8      CC                    3                   2        129.12
## 9      CC                    3                   3        118.68
## 10    qEV                    1                   1         17.38
## 11    qEV                    1                   2         19.58
## 12    qEV                    1                   3         23.06
## 13    qEV                    2                   1         12.96
## 14    qEV                    2                   2         14.14
## 15    qEV                    2                   3         14.74
## 16    qEV                    3                   1         19.14
## 17    qEV                    3                   2         23.01
## 18    qEV                    3                   3         22.61
## 19     UC                    1                   1         33.00
## 20     UC                    1                   2         47.54
## 21     UC                    1                   3         45.08
## 22     UC                    2                   1         18.71
## 23     UC                    2                   2         32.70
## 24     UC                    2                   3         31.16
## 25     UC                    3                   1         32.03
## 26     UC                    3                   2         31.85
## 27     UC                    3                   3         41.10
## 28     DG                    1                   1          3.56
## 29     DG                    1                   2          4.39
## 30     DG                    1                   3          3.96
## 31     DG                    2                   1          3.14
```

```
## 32     DG                    2                   2          3.88
## 33     DG                    2                   3          3.20
## 34     DG                    3                   1          3.50
## 35     DG                    3                   2          3.71
## 36     DG                    3                   3          3.76
##    Total.Particles Particles.to.Protein  Size
## 1         4.49e+10             3.75e+08  96.7
## 2         4.28e+10             3.59e+08  98.5
## 3         4.76e+10             4.76e+08  97.3
## 4         5.76e+10             5.90e+08 100.4
## 5         6.60e+10             6.37e+08  92.3
## 6         6.60e+10             6.93e+08  92.6
## 7         5.40e+10             4.16e+08 101.9
## 8         5.46e+10             4.23e+08 102.9
## 9         5.98e+10             5.04e+08 100.3
## 10        3.45e+10             1.99e+09  95.3
## 11        4.20e+10             2.15e+09 100.0
## 12        4.20e+10             1.82e+09 102.6
## 13        4.05e+10             3.13e+09 100.8
## 14        4.05e+10             2.86e+09 100.3
## 15        4.95e+10             3.36e+09 100.1
## 16        4.35e+10             2.27e+09 112.8
## 17        3.60e+10             1.56e+09 116.8
## 18        6.60e+10             2.92e+09 104.4
## 19        1.95e+10             5.91e+08 108.4
## 20        2.80e+10             5.89e+08 104.5
## 21        2.28e+10             5.06e+08 104.4
## 22        4.16e+10             2.22e+09  92.8
## 23        4.75e+10             1.45e+09  94.9
## 24        2.86e+10             9.18e+08 102.9
## 25        4.03e+10             1.26e+09 108.2
## 26        4.42e+10             1.39e+09 104.3
## 27        2.86e+10             6.96e+08 118.4
## 28        2.24e+09             6.29e+08  97.4
## 29        2.24e+09             5.10e+08  93.4
## 30        2.86e+09             7.22e+08  99.8
## 31        1.82e+10             5.80e+09 103.2
## 32        1.93e+10             4.97e+09 111.2
## 33        1.65e+10             5.16e+09 108.3
## 34        3.03e+10             8.66e+09  95.5
## 35        2.31e+10             6.23e+09  97.3
## 36        2.72e+10             7.23e+09  95.1
```

6.4.3 Assess univariate normality (Figure 6.6)

```
#assess normality for particle size from NTA
attach(Ch2MethodComp)
hist(Size, main = "Particle Size Histogram", xlab = "Size (nm)")

qqnorm(Size, main = "Particle Size QQ Plot")
qqline(Size)

shapiro.test(Ch2MethodComp$Size)

##
##  Shapiro-Wilk normality test
##
## data:  Ch2MethodComp$Size
## W = 0.94047, p-value = 0.05253

#assess normality for total protein
hist(Total.Protein, main = "Total Protein Histogram", xlab = "Total Protein (µg)")

qqnorm(Total.Protein, main = "Total Protein QQ Plot")
qqline(Total.Protein)

shapiro.test(Ch2MethodComp$Total.Protein)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Ch2MethodComp$Total.Protein
## W = 0.79031, p-value = 1.051e-05

#assess normality for total particles
hist(Total.Particles, main = "Total Particles Histogram", xlab = "Total Particles")

qqnorm(Total.Particles, main = "Total Particles QQ Plot")
qqline(Total.Particles)

shapiro.test(Ch2MethodComp$Total.Particles)

##
##  Shapiro-Wilk normality test
##
## data:  Ch2MethodComp$Total.Particles
## W = 0.9644, p-value = 0.2926

#assess normality for particle to protein ratio
hist(Particles.to.Protein, main = "Particles to Protein Ratio Histogram", xlab = "Particles per µg" )

qqnorm(Particles.to.Protein, main = "Particles to Protein Ratio QQ Plot")
qqline(Particles.to.Protein)

shapiro.test(Ch2MethodComp$Particles.to.Protein)

##
##  Shapiro-Wilk normality test
##
## data:  Ch2MethodComp$Particles.to.Protein
## W = 0.779, p-value = 6.4e-06
```

Particle size and total particles are normally distributed. Total protein and particle to protein ratio are not normally distributed and need to be compared with nonparametric Kruskal-walis and Wilcoxon signed-Rank tests.

6.4.4 Assess equal variance for normally distributed data

```
bartlett.test(Size ~ Method, data = Ch2MethodComp)

##
##  Bartlett test of homogeneity of variances
##
## data:  Size by Method
## Bartlett's K-squared = 3.5548, df = 3, p-value = 0.3137

bartlett.test(Total.Particles ~ Method, data = Ch2MethodComp)

##
##  Bartlett test of homogeneity of variances
##
## data:  Total.Particles by Method
## Bartlett's K-squared = 0.49903, df = 3, p-value = 0.9191
```

Particle size and total particles have equal variance and can be compared using ANOVA and Student's t-test or TukeyHSD.

6.4.5 Run tests for comparing means

```
#NTA size
sizeanova <- aov(Size ~ Method, data = Ch2MethodComp)
summary(sizeanova)

##             Df Sum Sq Mean Sq F value Pr(>F)
## Method       3  234.5   78.18   2.006  0.133
## Residuals   32 1247.3   38.98

#Total Protein
kruskal.test(Total.Protein ~ Method, data = Ch2MethodComp)

##
##  Kruskal-Wallis rank sum test
##
## data:  Total.Protein by Method
## Kruskal-Wallis chi-squared = 32.077, df = 3, p-value = 5.041e-07

pairwise.wilcox.test(Ch2MethodComp$Total.Protein, Ch2MethodComp$Method,
                     p.adjust.method = "BH")

##
##  Pairwise comparisons using Wilcoxon rank sum exact test
##
## data:  Ch2MethodComp$Total.Protein and Ch2MethodComp$Method
##
##     CC      DG      qEV
## DG  4.9e-05 -       -
## qEV 4.9e-05 4.9e-05 -
## UC  4.9e-05 4.9e-05 0.00078
##
## P value adjustment method: BH

#Total Particles
particleanova <- aov(Total.Particles ~ Method, data = Ch2MethodComp)
summary(particleanova)

##             Df    Sum Sq   Mean Sq F value   Pr(>F)
## Method       3 7.444e+21 2.481e+21   26.14 9.74e-09 ***
## Residuals   32 3.038e+21 9.492e+19
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

tukey <- TukeyHSD(particleanova)
print(tukey)

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Total.Particles ~ Method, data = Ch2MethodComp)
##
## $Method
##                diff          lwr          upr     p adj
## DG-CC  -39040000000 -51483567520 -26596432480 0.0000000
## qEV-CC -10977777778 -23421345297   1465789742 0.0994181
## UC-CC  -21355555556 -33799123075  -8911988036 0.0003066
## qEV-DG  28062222222  15618654703  40505789742 0.0000046
## UC-DG   17684444444   5240876925  30128011964 0.0028487
## UC-qEV -10377777778 -22821345297   2065789742 0.1291716

#Particles to protein ratio
kruskal.test(Particles.to.Protein ~ Method, data = Ch2MethodComp)

##
##  Kruskal-Wallis rank sum test
##
```

```
## data:  Particles.to.Protein by Method
## Kruskal-Wallis chi-squared = 21.446, df = 3, p-value = 8.503e-05

pairwise.wilcox.test(Ch2MethodComp$Particles.to.Protein, Ch2MethodComp$Method,
                     p.adjust.method = "BH")

##
##  Pairwise comparisons using Wilcoxon rank sum exact test
##
## data:  Ch2MethodComp$Particles.to.Protein and Ch2MethodComp$Method
##
##     CC      DG      qEV
## DG  0.00156 -       -
## qEV 0.00025 0.25808 -
## UC  0.00413 0.07503 0.00148
##
## P value adjustment method: BH
```

6.4.6 Set up data summary function for mean comparisons on graphs

```
data_summary <- function(data, varname, groupnames){
  require(plyr)
  summary_func <- function(x, col){
    c(mean = mean(x[[col]], na.rm = TRUE),
      sd = sd(x[[col]], na.rm = TRUE))
  }
  data_sum <- ddply(data, groupnames, .fun = summary_func,
                    varname)
    data_sum <- rename(data_sum, c("mean" = varname))
  return(data_sum)}
```

6.4.7 Plot NTA size (Figure 2.2)

```
ggplot(Ch2MethodComp, aes(Method, Size))+
  stat_boxplot(geom = "errorbar", linetype = 1, width = 0.5) +
  geom_boxplot(outlier.shape = 1) +
  theme_classic() +
  labs( x="Method", y="Particle Size (nm)") +
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.title = element_text(size = 15))+
  theme(legend.text = element_text(size = 15))
```

6.4.8 Plot protein recovery (Figure 2.3)

```
df3 <- data_summary(Ch2MethodComp, varname = 'Total.Protein',
                    groupnames = c("Method", "Biological.Replicate"))

## Loading required package: plyr

## ------------------------------------------------------------------------------

## You have loaded plyr after dplyr - this is likely to cause problems.
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:
## library(plyr); library(dplyr)

## ------------------------------------------------------------------------------

##
## Attaching package: 'plyr'

## The following object is masked from 'package:purrr':
##
##     compact
```

```
## The following object is masked from 'package:ggpubr':
##
##     mutate

## The following objects are masked from 'package:dplyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize

protein_comparisions = list(c("CC", "DG"), c("CC", "qEV"), c("CC", "UC"),
                            c("DG", "qEV"), c("qEV", "UC"), c("DG", "UC"))
ggplot(df3, aes(x = Method, y = Total.Protein, fill = Biological.Replicate))+
  geom_bar(stat = "identity", color = "black", position = position_dodge())+
  geom_errorbar(aes(ymin = Total.Protein - sd, ymax = Total.Protein + sd),
                width = .2, position = position_dodge(.9))+
  stat_compare_means(data = Ch2MethodComp, comparisons = protein_comparisions,
                     label = "p.signif", method = "wilcox.test")+
  theme_classic()+
  labs (x = "Method", y = "Total Protein (µg)", fill = "Biological
Replicate")+
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.title = element_text(size = 15))+
  theme(legend.text = element_text(size = 15))
```

6.4.9 Plot particle recovery (Figure 2.4)

```
df4 <- data_summary(Ch2MethodComp, varname = 'Total.Particles',
                    groupnames = c("Method", "Biological.Replicate"))
particle_comparisions = list(c("CC", "DG"), c("CC", "UC"),
                             c("DG", "qEV"), c("DG", "UC"))
ggplot(df4, aes(x = Method, y = Total.Particles, fill = Biological.Replicate))+
  geom_bar(stat = "identity", color = "black", position = position_dodge())+
  geom_errorbar(aes(ymin = Total.Particles - sd, ymax = Total.Particles + sd),
                width = .2, position = position_dodge(.9))+
  stat_compare_means(data = Ch2MethodComp, comparisons = particle_comparisions,
                     label = "p.signif", method = "t.test")+
  theme_classic()+
  labs (x = "Method", y = "Total Particles", fill = "Biological
Replicate")+
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.title = element_text(size = 15))+
  theme(legend.text = element_text(size = 15))
```

6.4.10 Scatter plot particles by protein for each method (Figure 2.5)

```
ggplot(Ch2MethodComp, aes(Total.Protein, Total.Particles))+
  geom_point(size=3, aes(color=Method))+
  facet_wrap(~Biological.Replicate, ncol=3)+
  labs(x="Total Protein (µg)",
       y="Total Particles")+
  theme(panel.background = element_rect(fill = "NA"),
        axis.line=element_line(color="black"))+
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.title = element_text(size = 15))+
  theme(legend.text = element_text(size = 15))+
  theme(strip.text = element_text(size = 15))
```

6.4.11 Plot particle to protein ratio (Figure 2.6)

```
df2 <- data_summary(Ch2MethodComp, varname = 'Particles.to.Protein',
                    groupnames = c("Method", "Biological.Replicate"))
ggplot(df2, aes(x = Biological.Replicate, y = Particles.to.Protein, fill = Method))+
  geom_bar(stat = "identity",color = "black", position = position_dodge())+
```

```
  geom_errorbar(aes(ymin = Particles.to.Protein - sd, ymax = Particles.to.Protein + sd),
                width = .2, position = position_dodge(.9))+
  theme_classic()+
  labs (x = "Biological Replicate", y = "Particles per µg Protein", fill = "Method")+
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.title = element_text(size = 15))+
  theme(legend.text = element_text(size = 15))
```

6.4.12 Western blots load data

```
Ch2WB <- read.csv("~/Documents/R/Ch2WBIntensities.csv")
print(Ch2WB)
```

```
##    Method Biological.Replicate Technical.Replicate   LpqH  GroES   GlcB    LAM
## 1    100R                    1                   1 11.493 16.185 31.000 13.383
## 2    100R                    1                   1 11.493 16.185 31.000 13.383
## 3    100R                    1                   1 11.439 16.185 31.000 13.383
## 4      CC                    1                   1 45.752  0.000 11.913 29.309
## 5      CC                    1                   2 43.796  0.761 12.574 32.700
## 6      CC                    1                   3 39.264  1.944 10.636 25.544
## 7     qEV                    1                   1 41.616  3.033 11.182 29.749
## 8     qEV                    1                   2 38.338  2.714 11.368 29.219
## 9     qEV                    1                   3 42.348  3.290 10.912 32.377
## 10     UC                    1                   1 38.862 11.078 24.041 38.400
## 11     UC                    1                   2 34.365 10.393 22.838 31.771
## 12     UC                    1                   3 36.143 10.508 23.089 38.626
## 13   100R                    2                   1  8.361 41.209 37.948 26.313
## 14   100R                    2                   1  8.361 41.209 37.948 26.313
## 15   100R                    2                   1  8.361 41.209 37.948 26.313
## 16     CC                    2                   1 59.133  0.000  4.686 26.381
## 17     CC                    2                   2 59.985  0.000  4.686 26.173
## 18     CC                    2                   3 61.752  1.601  2.116 20.919
## 19    qEV                    2                   1 66.434  2.936  0.533 26.483
## 20    qEV                    2                   2 61.233  3.711  1.097 26.125
## 21    qEV                    2                   3 64.426  3.753  1.710 25.461
## 22     UC                    2                   1 43.757  4.825  9.576 26.579
## 23     UC                    2                   2 41.705  8.261 12.012 36.122
## 24     UC                    2                   3 57.112  9.969 19.301 44.393
## 25     DG                    2                   1 56.623  2.857  1.835 24.911
## 26     DG                    2                   2 51.956  2.002  1.299 20.165
## 27     DG                    2                   3 48.364  0.000  1.199 18.556
## 28   100R                    3                   1 13.043 53.154 31.471 23.382
## 29   100R                    3                   1 13.043 53.154 31.471 23.382
## 30   100R                    3                   1 13.043 53.154 31.471 23.382
## 31     CC                    3                   1 59.299  1.154 10.580 57.006
## 32     CC                    3                   2 62.983  0.000  9.316 38.567
## 33     CC                    3                   3 60.838  0.000  6.105 47.975
## 34    qEV                    3                   1 71.378  0.000  1.073 41.624
## 35    qEV                    3                   2 74.482  0.000  0.453 55.508
## 36    qEV                    3                   3 63.628  0.347  0.873 37.653
## 37     UC                    3                   1 57.178 12.365 12.899 41.101
## 38     UC                    3                   2 54.583 12.730 11.353 43.269
## 39     UC                    3                   3 66.068 24.573 18.701 55.155
## 40     DG                    3                   1 52.661  0.179  0.627 29.113
## 41     DG                    3                   2 50.540  0.000  0.000 33.641
## 42     DG                    3                   3 45.885  1.018  0.303 29.084
##      SodC
## 1  27.566
## 2  27.566
## 3  27.566
## 4  57.890
## 5  54.377
## 6  50.244
## 7  59.333
## 8  59.203
## 9  60.382
## 10 49.877
```

```
## 11 42.605
## 12 44.446
## 13 27.156
## 14 27.156
## 15 27.156
## 16 41.436
## 17 40.967
## 18 38.353
## 19 46.884
## 20 47.511
## 21 43.199
## 22 24.191
## 23 21.327
## 24 34.105
## 25 24.449
## 26 21.825
## 27 17.935
## 28 36.161
## 29 36.161
## 30 36.161
## 31 56.320
## 32 52.725
## 33 48.334
## 34 53.566
## 35 60.226
## 36 43.654
## 37 37.117
## 38 35.905
## 39 45.275
## 40 30.969
## 41 35.335
## 42 34.524

Ch2WB$Biological.Replicate <- as.factor(Ch2WB$Biological.Replicate)
```

6.4.13 Plot Western blots (Figure 2.8)

```
#Lpqh
plotlpqh <- data_summary(Ch2WB, varname = 'LpqH',
                         groupnames = c("Method", "Biological.Replicate"))

ggplot(plotlpqh, aes(x = Method, y = LpqH))+
  geom_bar(aes(fill = Method), stat = "identity", color = "black", position = position_dodge())+
  facet_wrap(~Biological.Replicate, ncol=3)+
  geom_errorbar(aes(ymin = LpqH - sd, ymax = LpqH + sd),
                width = .2, position = position_dodge(.9))+
  labs(x="Method",
       y="LpqH Pixel Intensity")+
  scale_fill_grey(start = 0, end = 0.9)+
  scale_y_continuous(limits = c(0, 80))+
  theme(strip.text.x = element_text(size = 15))+
  theme(panel.background = element_rect(fill = "NA"),
        axis.line=element_line(color="black"))+
  theme(axis.text.x = element_text(angle = 90))+
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.position = "none")

#GroES
groes <- data_summary(Ch2WB, varname = 'GroES',
                      groupnames = c("Method", "Biological.Replicate"))

ggplot(groes, aes(x = Method, y = GroES))+
  geom_bar(aes(fill = Method),stat = "identity", color = "black", position = position_dodge())+
  facet_wrap(~Biological.Replicate, ncol=3)+
  geom_errorbar(aes(ymin = GroES - sd, ymax = GroES + sd),
                width = .2, position = position_dodge(.9))+
  labs(x="Method",
```

```
       y="GroES Pixel Intensity", fill = "Biological
Replicate")+
  scale_fill_grey(start = 0, end = 0.9)+
  theme(strip.text.x = element_text(size = 15))+
  scale_y_continuous(limits = c(0, 80))+
  theme(panel.background = element_rect(fill = "NA"),
        axis.line=element_line(color="black"))+
  theme(axis.text.x = element_text(angle = 90))+
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.position = "none")

#glcb
glcb <- data_summary(Ch2WB, varname = 'GlcB',
                     groupnames = c("Method", "Biological.Replicate"))

ggplot(glcb, aes(x = Method, y = GlcB))+
  geom_bar(aes(fill = Method), stat = "identity", color = "black", position = position_dodge())+
  facet_wrap(~Biological.Replicate, ncol=3)+
  geom_errorbar(aes(ymin = GlcB - sd, ymax = GlcB + sd),
                width = .2, position = position_dodge(.9))+
  labs(x="Method",
       y="GlcB Pixel Intensity", fill = "Biological
Replicate")+
  scale_fill_grey(start = 0, end = 0.9)+
  theme(strip.text.x = element_text(size = 15))+
  scale_y_continuous(limits = c(0, 80))+
  theme(panel.background = element_rect(fill = "NA"),
        axis.line=element_line(color="black"))+
  theme(axis.text.x = element_text(angle = 90))+
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.position = "none")

#LAM
lam <- data_summary(Ch2WB, varname = 'LAM',
                    groupnames = c("Method", "Biological.Replicate"))

ggplot(lam, aes(x = Method, y = LAM,))+
  geom_bar(aes(fill = Method), stat = "identity", color = "black", position = position_dodge())+
  facet_wrap(~Biological.Replicate, ncol=3)+
  geom_errorbar(aes(ymin = LAM - sd, ymax = LAM + sd),
                width = .2, position = position_dodge(.9))+
  labs(x="Method",
       y="LAM Pixel Intensity", fill = "Biological
Replicate")+
  scale_fill_grey(start = 0, end = 0.9)+
  theme(strip.text.x = element_text(size = 15))+
  scale_y_continuous(limits = c(0, 80))+
  theme(panel.background = element_rect(fill = "NA"),
        axis.line=element_line(color="black"))+
  theme(axis.text.x = element_text(angle = 90))+
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.position = "none")

#sodc
sodc <- data_summary(Ch2WB, varname = 'SodC',
                     groupnames = c("Method", "Biological.Replicate"))

ggplot(sodc, aes(x = Method, y = SodC))+
  geom_bar(aes(fill = Method), stat = "identity", color = "black", position = position_dodge())+
  facet_wrap(~Biological.Replicate, ncol=3)+
  geom_errorbar(aes(ymin = SodC - sd, ymax = SodC + sd),
                width = .2, position = position_dodge(.9))+
  labs(x="Method",
       y="SodC Pixel Intensity", fill = "Biological
Replicate")+
```

```
  scale_fill_grey(start = 0, end = 0.9)+
  theme(strip.text.x = element_text(size = 15))+
  scale_y_continuous(limits = c(0, 80))+
  theme(panel.background = element_rect(fill = "NA"),
        axis.line=element_line(color="black"))+
  theme(axis.text.x = element_text(angle = 90))+
  theme(axis.text = element_text(size = 15))+
  theme(axis.title = element_text(size = 20))+
  theme(legend.position = "none")
```

6.5 References

[1] Marsollier L, Brodin P, Jackson M, Korduláková J, Tafelmeyer P, Carbonnelle E, et al. Impact of Mycobacterium ulcerans biofilm on transmissibility to ecological niches and Buruli ulcer pathogenesis. PLoS Pathogens 2007;3:0582–94. https://doi.org/10.1371/journal.ppat.0030062.

[2] Prados-Rosales R, Baena A, Martinez LR, Luque-Garcia J, Kalscheuer R, Veeraraghavan U, et al. Mycobacteria release active membrane vesicles that modulate immune responses in a TLR2-dependent manner in mice. Journal of Clinical Investigation 2011;121:1471–83. https://doi.org/10.1172/JCI44261.

[3] Rath P, Huang C, Wang T, Wang T, Li H, Prados-Rosales R, et al. Genetic regulation of vesiculogenesis and immunomodulation in Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences of the United States of America 2013;110. https://doi.org/10.1073/pnas.1320118110.

[4] Ziegenbalg A, Prados-Rosales R, Jenny-Avital ER, Kim RS, Casadevall A, Achkar JM. Immunogenicity of mycobacterial vesicles in humans: Identification of a new tuberculosis antibody biomarker. Tuberculosis 2013;93:448–55. https://doi.org/10.1016/j.tube.2013.03.001.

[5] Prados-Rosales R, Weinrick BC, Piqué DG, Jacobs WR, Casadevall A, Rodriguez GM. Role for mycobacterium tuberculosis membrane vesicles in iron acquisition. Journal of Bacteriology 2014;196:1250–6. https://doi.org/10.1128/JB.01090-13.

[6] Prados-Rosales R, Brown L, Casadevall A, Montalvo-Quirós S, Luque-Garcia JL. Isolation and identification of membrane vesicle-associated proteins in Gram-positive bacteria and mycobacteria. MethodsX 2014;1:e124–9. https://doi.org/10.1016/j.mex.2014.08.001.

[7] Prados-Rosales R, Carreño LJ, Batista-Gonzalez A, Baena A, Venkataswamy MM, Xu J, et al. Mycobacterial membrane vesicles administered systemically in mice induce a protective immune response to surface compartments of mycobacterium tuberculosis. MBio 2014;5. https://doi.org/10.1128/mBio.01921-14.

[8] Lee J, Kim SH, Choi DS, Lee JS, Kim DK, Go G, et al. Proteomic analysis of extracellular vesicles derived from Mycobacterium tuberculosis. Proteomics 2015;15:3331–7. https://doi.org/10.1002/pmic.201500037.

[9] Athman JJ, Wang Y, McDonald DJ, Boom WH, Harding C v., Wearsch PA. Bacterial Membrane Vesicles Mediate the Release of Mycobacterium tuberculosis Lipoglycans and Lipoproteins from Infected Macrophages . The Journal of Immunology 2015;195:1044–53. https://doi.org/10.4049/jimmunol.1402894.

[10] Kumar S, Mittal E, Deore S, Kumar A, Rahman A, Krishnasastry M v. Mycobacterial tlyA gene product is localized to the cell-wall without signal sequence. Frontiers in Cellular and Infection Microbiology 2015;5:60. https://doi.org/10.3389/FCIMB.2015.00060.

[11] Jurkoshek KS, Wang Y, Athman JJ, Barton MR, Wearsch PA. Interspecies Communication between Pathogens and Immune Cells via Bacterial Membrane Vesicles. Frontiers in Cell and Developmental Biology 2016;4:125. https://doi.org/10.3389/fcell.2016.00125.

[12] Athman JJ, Sande OJ, Groft SG, Reba SM, Nagy N, Wearsch PA, et al. Mycobacterium tuberculosis Membrane Vesicles Inhibit T Cell Activation . The Journal of Immunology 2017;198:2028–37. https://doi.org/10.4049/jimmunol.1601199.

[13] Dauros Singorenko P, Chang V, Whitcombe A, Simonov D, Hong J, Phillips A, et al. Isolation of membrane vesicles from prokaryotes: a technical and biological comparison reveals heterogeneity. Journal of Extracellular Vesicles 2017;6. https://doi.org/10.1080/20013078.2017.1324731.

[14] White DW, Elliott SR, Odean E, Bemis LT, Tischler AD. Mycobacterium tuberculosis Pst/SenX3-RegX3 regulates membrane vesicle production independently of ESX-5 activity. MBio 2018;9. https://doi.org/10.1128/MBIO.00778-18/SUPPL_FILE/MBO003183934ST2.DOCX.

[15] Palacios A, Sampedro L, Sevilla IA, Molina E, Gil D, Azkargorta M, et al. Mycobacterium tuberculosis extracellular vesicle-associated lipoprotein LpqH as a potential biomarker to distinguish paratuberculosis infection or vaccination from tuberculosis infection. BMC Veterinary Research 2019;15:188. https://doi.org/10.1186/s12917-019-1941-6.

[16] Chiplunkar SS, Silva CA, Bermudez LE, Danelishvili L. Characterization of membrane vesicles released by Mycobacterium avium in response to environment mimicking the macrophage phagosome. Future Microbiology 2019;14:293–313. https://doi.org/10.2217/FMB-2018-0249/ASSET/IMAGES/LARGE/FIGURE6.JPEG.

[17] Gupta S, Rodriguez GM. Isolation and Characterization of Extracellular Vesicles Produced by Iron-limited Mycobacteria. JoVE (Journal of Visualized Experiments) 2019;2019:e60359. https://doi.org/10.3791/60359.

[18] Gupta S, Palacios A, Khataokar A, Weinrick B, Lavín JL, Sampedro L, et al. Dynamin-like proteins are essential for vesicle biogenesis in Mycobacterium tuberculosis. BioRxiv 2020:2020.01.14.906362. https://doi.org/10.1101/2020.01.14.906362.

[19] Lucas M, Ryan JM, Watkins J, Early K, Kruh-Garcia NA, Mehaffy C, et al. Extraction and Separation of Mycobacterial Proteins. Methods in Molecular Biology 2021;2314:77–107. https://doi.org/10.1007/978-1-0716-1460-0_3.

[20] Dauros Singorenko P, Chang V, Whitcombe A, Simonov D, Hong J, Phillips A, et al. Isolation of membrane vesicles from prokaryotes: a technical and biological comparison reveals heterogeneity. Journal of Extracellular Vesicles 2017;6:1324731. https://doi.org/10.1080/20013078.2017.1324731.

[21] Diaz G. Exosomes: A Potential Novel Source of Biomarkers for Tuberculosis. 2017. Doctoral dissertation, Colorado State University, Fort Collins. Mountain Scholar.

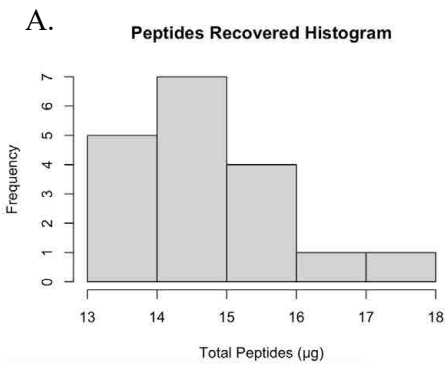
[22] Diaz G, Wolfe LM, Kruh-Garcia NA, Dobos KM. Changes in the Membrane-Associated Proteins of Exosomes Released from Human Macrophages after Mycobacterium tuberculosis Infection. Scientific Reports 2016;6. https://doi.org/10.1038/SREP37975.

7.1 Peptide Digestion

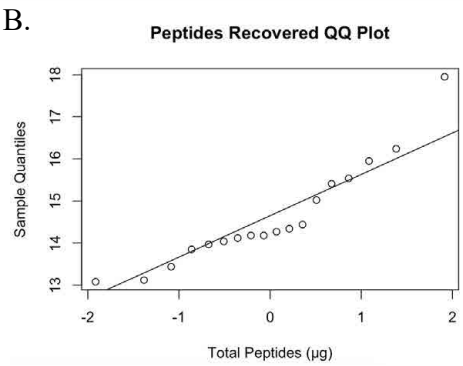
Table AII.1 Verification of normalization strategy based on peptide recovery. The total peptide mass recovered following digestion of the indicated amount of sample is reported. The mass of the peptides recovered is extrapolated from the average concentration by A205 of three NanoDrop measurements.

Sample	Fraction	Quantity Digested	Peptides Recovered (µg)
1	100R	15 µg	14.34
1	1	15 µg	14.44
1	2	3.0E+09	15.02
1	3	3.0E+09	14.81
1	4	3.0E+09	13.12
1	5	3.0E+09	13.97
1	qEV	3.0E+09	14.04
2	100R	15 µg	15.54
2	1	15 µg	14.27
2	2	1.5E+09	13.85
2	3	3.0E+09	14.12
2	4	3.0E+09	14.18
2	5	3.0E+09	13.44
2	qEV	3.0E+09	13.08
3	100R	15 µg	17.95
3	1	15 µg	15.41
3	2	3.0E+09	16.24
3	3	3.0E+09	15.95
3	4	3.0E+09	14.91
3	5	3.0E+09	13.05
3	qEV	3.0E+09	14.16

A.



B.



C.

Shapiro-wilks test	Total Peptides
Statistic	0.85397
P-value	0.0445

Figure AII.1 Peptide digest recovery normality assessment. (A) Histogram, (B) QQ plot, and (C) Shapiro-wilks testing for peptide recovery. Significant p-value (<0.05) is indicated in red font.

7.2 Assessment of Individual LC-MS/MS Injections

Because each sample was injected in triplicate, the individual injections were evaluated for quality and consistency based on the data presented in **Table AII7.2**. There were 779 proteins identified including 14 decoys for a 1.8% decoy FDR from 395,494 spectra with a 0.1% FDR. **Figure AII7.2** shows that the injection order did not cause any trends in the number of proteins identified, the number of spectra generated, the number of spectral identifications or the percent of spectra identified. This confirms that the LC-MS/MS instrument performance was consistent over time.

Table AII.2 LC-MS/MS statistics for individual injections. Highlighted cells indicate outliers or CV over 10% requiring manual TIC validation.

Sample	Fraction	Injection	Injection Order	# Proteins	Mean	StdDev	CoV (%)	# Spectral IDs	Mean	StdDev	CoV (%)	# Spectra	Mean	StdDev	CoV (%)	% Spectra ID'd	Mean	StdDev	CoV (%)
1	100R	1	4	477				10706				38248				27.99			
1	100R	2	27	464	463.67	11.03	2.38%	10453	10438.67	224.36	2.15%	38983	38777.00	377.10	0.97%	26.81	26.93	0.83	3.07%
1	100R	3	55	450				10157				39100				25.98			
1	F1	1	15	456				10568				38704				27.30			
1	F1	2	28	444	447.67	5.91	1.32%	10629	10511.00	126.22	1.20%	39489	38964.67	370.76	0.95%	26.92	26.98	0.25	0.92%
1	F1	3	48	443				10336				38701				26.71			
1	F2	1	12	480				10214				37052				27.57			
1	F2	2	29	458	464.33	11.15	2.40%	9850	9956.33	183.12	1.84%	35760	36583.00	583.84	1.60%	27.54	27.22	0.48	1.75%
1	F2	3	44	455				9805				36937				26.55			
1	F3	1	9	331				5813				30894				18.82			
1	F3	2	23	304	314.67	11.73	3.73%	5152	5412.33	287.52	5.31%	30304	30976.33	585.47	1.89%	17.00	17.48	0.96	5.49%
1	F3	3	58	309				5272				31731				16.61			
1	F4	1	5	271				4618				28873				15.99			
1	F4	2	25	199	220.33	35.98	16.33%	4111	4130.00	390.92	9.47%	26218	27500.00	1085.81	3.95%	15.68	15.01	1.18	7.84%
1	F4	3	47	191				3661				27409				13.36			
1	F5	1	20	143				2639				23769				11.10			
1	F5	2	30	171	155.33	11.67	7.51%	4215	3559.67	670.20	18.83%	27309	25475.00	1448.03	5.68%	15.43	13.88	1.97	14.17%
1	F5	3	62	152				3825				25347				15.09			
1	qEV	1	17	234				4146				27641				15.00			
1	qEV	2	33	194	206.33	19.60	9.50%	4129	3812.67	459.44	12.05%	26474	26506.00	913.94	3.45%	15.60	14.35	1.36	9.51%
1	qEV	3	59	191				3163				25403				12.45			
2	100R	1	13	373				9734				38532				25.26			
2	100R	2	37	365	365.67	5.73	1.57%	9588	9548.33	170.12	1.78%	38326	38221.67	305.04	0.80%	25.02	24.98	0.25	0.99%
2	100R	3	61	359				9323				37807				24.66			
2	F1	1	2	340				9048				37926				23.86			
2	F1	2	24	323	329.67	7.41	2.25%	8715	8731.00	252.55	2.89%	38468	38361.33	320.89	0.84%	22.66	22.77	0.85	3.73%
2	F1	3	54	326				8430				38690				21.79			
2	F2	1	14	322				7236				33576				21.55			
2	F2	2	42	289	301.33	14.70	4.88%	6849	6918.67	235.86	3.41%	32939	32821.33	669.41	2.04%	20.79	21.07	0.34	1.61%
2	F2	3	45	293				6671				31949				20.88			
2	F3	1	8	279				5158				30521				16.90			
2	F3	2	40	241	254.00	17.68	6.96%	5050	5075.33	59.90	1.18%	31529	31338.33	604.73	1.93%	16.02	16.21	0.51	3.14%
2	F3	3	52	242				5018				31965				15.70			
2	F4	1	18	224				4810				30113				15.97			
2	F4	2	34	206	211.33	8.99	4.26%	4412	4491.67	234.27	5.22%	29205	29491.67	439.79	1.49%	15.11	15.22	0.57	3.76%
2	F4	3	57	204				4253				29157				14.59			
2	F5	1	19	171				3577				24484				14.61			
2	F5	2	26	172	173.00	2.16	1.25%	3822	3607.33	164.30	4.55%	25349	24808.33	384.84	1.55%	15.08	14.54	0.48	3.27%
2	F5	3	43	176				3423				24592				13.92			

Sample	Fraction	Injection	Injection Order	# Proteins	Mean	StdDev	CoV (%)	# Spectral IDs	Mean	StdDev	CoV (%)	# Spectra	Mean	StdDev	CoV (%)	% Spectra ID'd	Mean	StdDev	CoV (%)
2	qEV	1	3	348				6391				31909				20.03			
2	qEV	2	41	289	313.00	25.31	8.09%	5296	5802.67	450.75	7.77%	30122	31107.00	740.93	2.38%	17.58	18.63	1.03	5.52%
2	qEV	3	51	302				5721				31290				18.28			
3	100R	1	7	328				8403				37442				22.44			
3	100R	2	35	317	319.33	6.34	1.99%	8513	8320.33	199.80	2.40%	37738	37156.33	625.09	1.68%	22.56	22.39	0.16	0.73%
3	100R	3	63	313				8045				36289				22.17			
3	F1	1	16	305				8294				37695				22.00			
3	F1	2	38	287	296.67	7.41	2.50%	8282	8144.67	202.76	2.49%	37803	37639.67	160.78	0.43%	21.91	21.64	0.45	2.09%
3	F1	3	50	298				7858				37421				21.00			
3	F2	1	1	289				6681				33969				19.67			
3	F2	2	32	277	280.00	6.48	2.31%	7260	6941.00	240.03	3.46%	34733	34420.33	326.95	0.95%	20.90	20.16	0.53	2.65%
3	F2	3	53	274				6882				34559				19.91			
3	F3	1	21	227				4514				29332				15.39			
3	F3	2	36	222	222.00	4.08	1.84%	4531	4640.33	166.79	3.59%	29282	29743.00	616.93	2.07%	15.47	15.60	0.24	1.51%
3	F3	3	46	217				4876				30615				15.93			
3	F4	1	11	99				1649				21216				7.77			
3	F4	2	60	84	93.67	6.85	7.31%	1354	1639.00	228.73	13.96%	21201	21859.00	919.97	4.21%	6.39	7.47	0.80	10.64%
3	F4	3	22	98				1914				23160				8.26			
3	F5	1	6	98				1799				21045				8.55			
3	F5	2	31	82	100.67	16.44	16.33%	2031	2313.00	570.77	24.68%	23249	23754.33	2444.72	10.29%	8.74	9.60	1.36	14.19%
3	F5	3	49	122				3109				26969				11.53			
3	qEV	1	10	215				3282				25879				12.68			
3	qEV	2	39	200	206.00	6.48	3.15%	4089	3816.00	377.63	9.90%	28764	27846.00	1391.91	5.00%	14.22	13.67	0.70	5.12%
3	qEV	3	56	203				4077				28895				14.11			

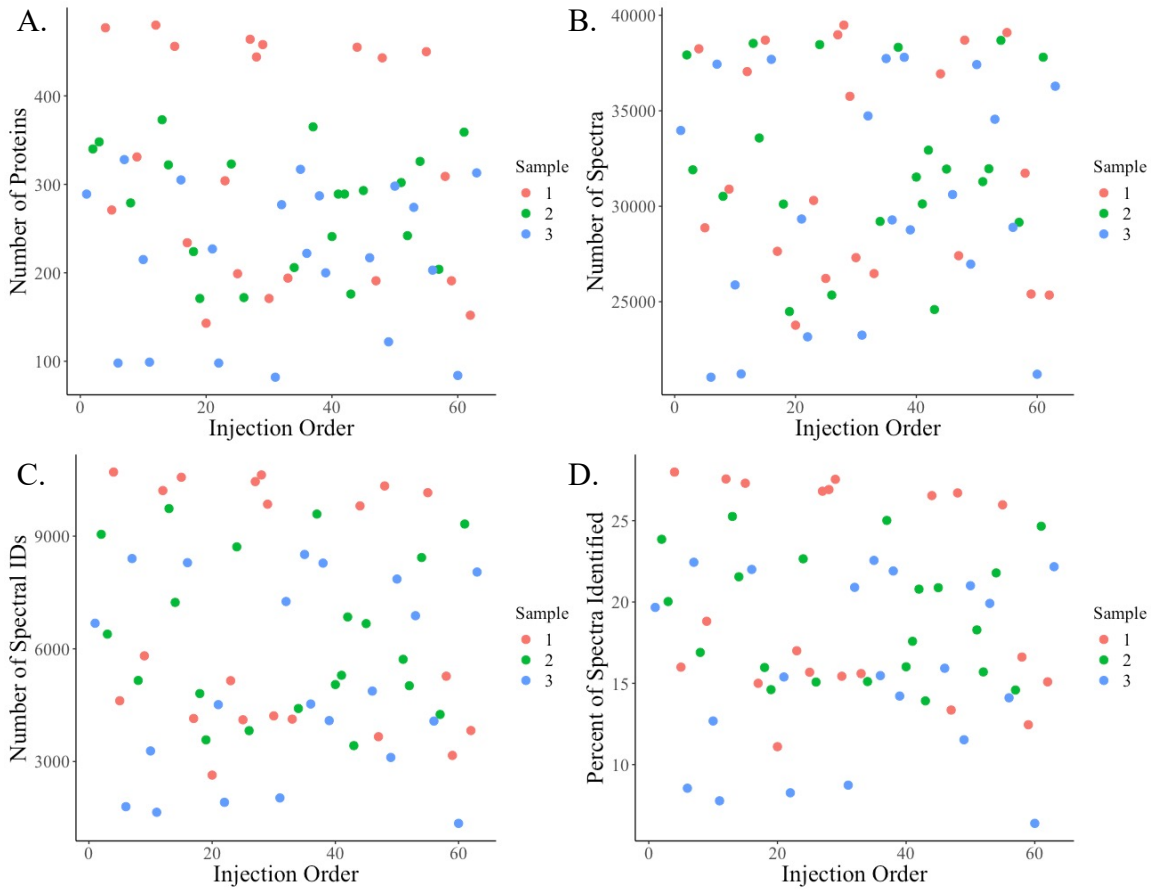
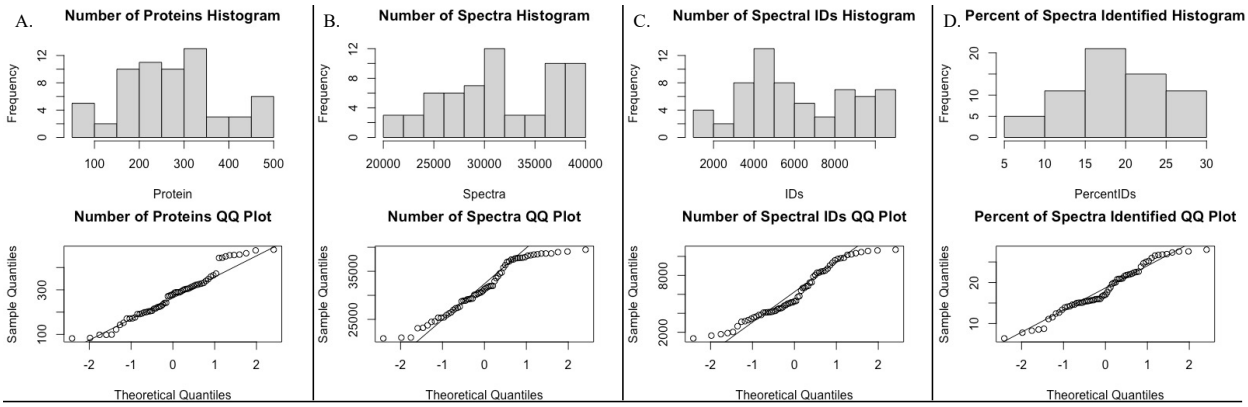


Figure AII.2 LC-MS/MS statistics by injection order. Individual injections were plotted by injection order on the x-axis and (A) number of identified proteins, (B) number of spectra, (C) number of spectral identifications, and (D) percentage of spectra identified on the y-axes.

The individual injection data was assessed for normality by histogram and QQ plot visualization followed by the Shapiro-Wilks test (**Figure AII7.3**). The number of proteins and the percentage of spectra identified follows a normal distribution. This data was then assessed for equal variance by Bartlett's test by sample and fraction. Number of proteins does not have equal variance by sample ($p = 0.013$). The number of proteins by sample and the percentage of spectra identified by sample and fraction have equal variance. These data also have independence, thus the percentage of spectra identified meets the assumptions for ANOVA with the post-hoc Tukey Test or Student's T-test. The other data (including number of proteins by fraction) must be evaluated by the non-parametric Kruskal-Wallis test and post-hoc pairwise Wilcoxon signed-rank test.



E.

Shapiro-wilks test	Number of Proteins	Number of Spectra	Number of Spectral IDs	% Spectra Identified
Statistic	0.9631344	0.9418334	0.935744	0.9634227
P-value	0.05623973	0.005042804	0.002650975	0.05818734

Figure AII.3 Normality testing for individual injections. The histogram and QQ plots are displayed for (A) number of proteins identified, (B) number of spectra, (C) number of spectral IDs, and (D) percent of spectra identified. (E) The statistic and p-value for Shapiro-Wilks test of these data is also included. Significant p-values (<0.05) are in red font.

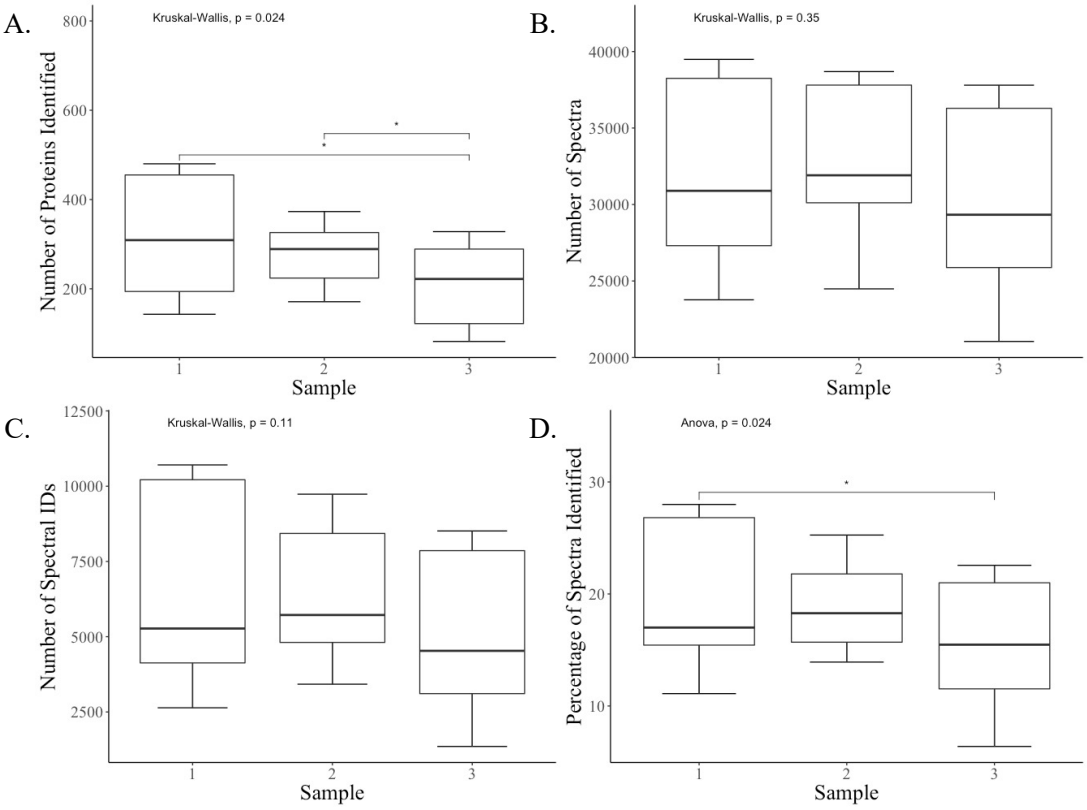


Figure AII.4 Individual injection statistics boxplots by sample. Boxplots of the (A) number of identified proteins, (B) number of spectra, (C) number of spectral identifications, and (D) percentage of spectra identified for each sample (n = 21). Significance from pairwise Wilcoxon signed-rank test (A-C) or Student's t-test (D) are indicated as p<0.05*.

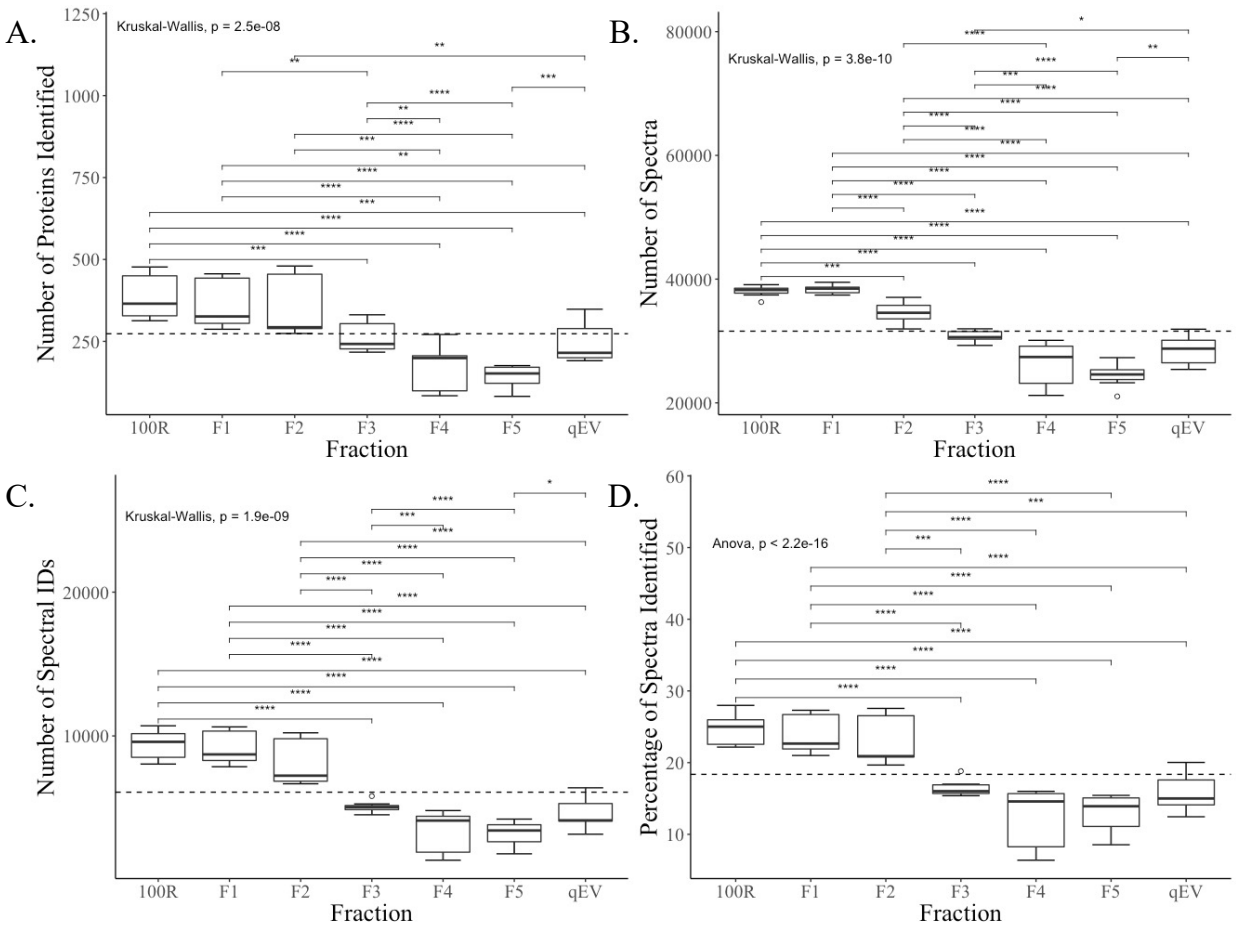


Figure AII.5 Boxplots of individual injection statistics by fraction. Boxplots of (A) number of identified proteins, (B) number of spectra, (C) number of spectral identifications, and (D) percentage of spectra identified for each fraction (n = 9). The dashed horizonal line represents the overall mean. Significance from pairwise Wilcoxon signed-rank test (A-C) or Student's t-test (D) is indicated as $p<0.05$*, $p<0.01$**, $p<0.001$***, and $p<0.0001$****.

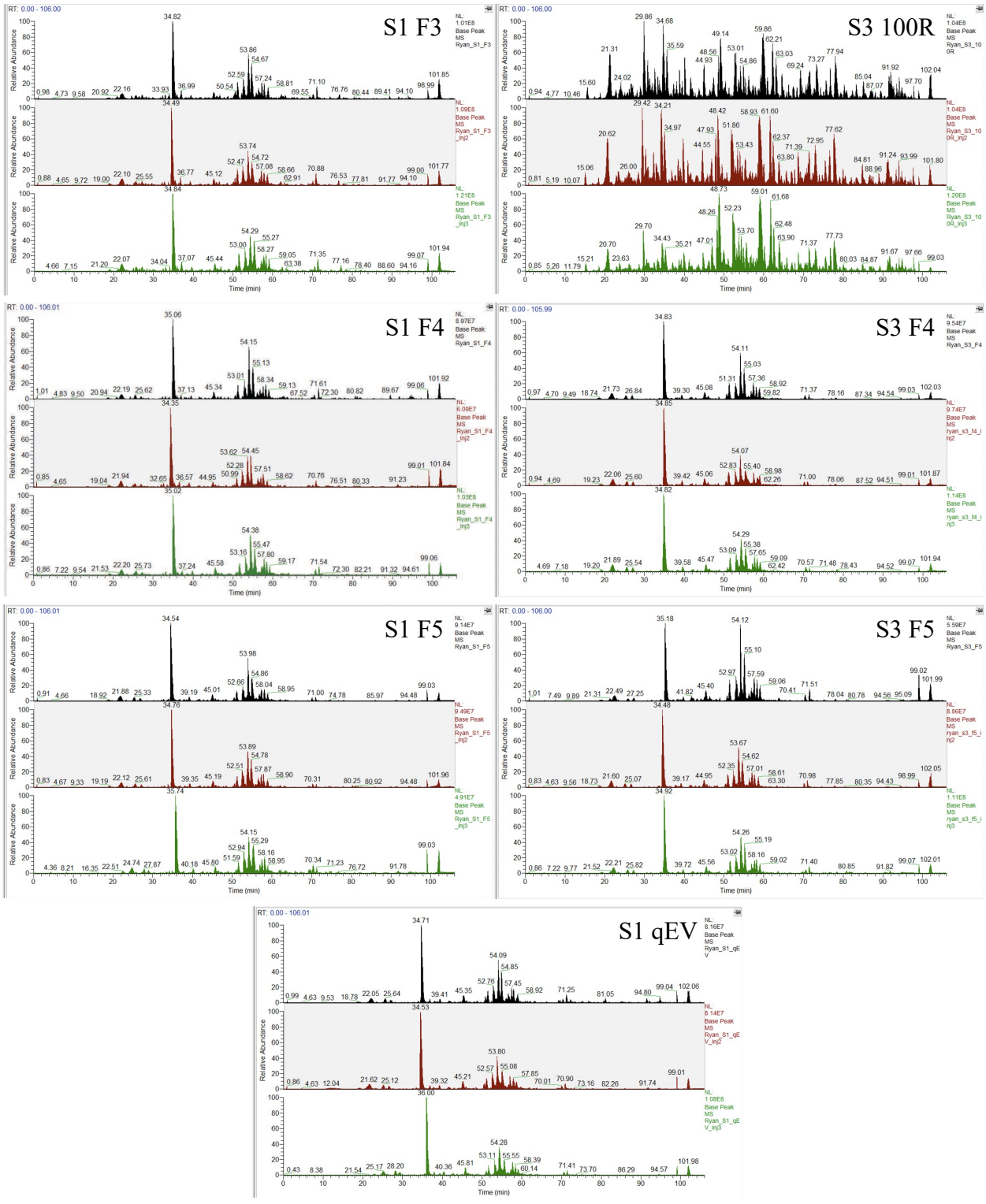


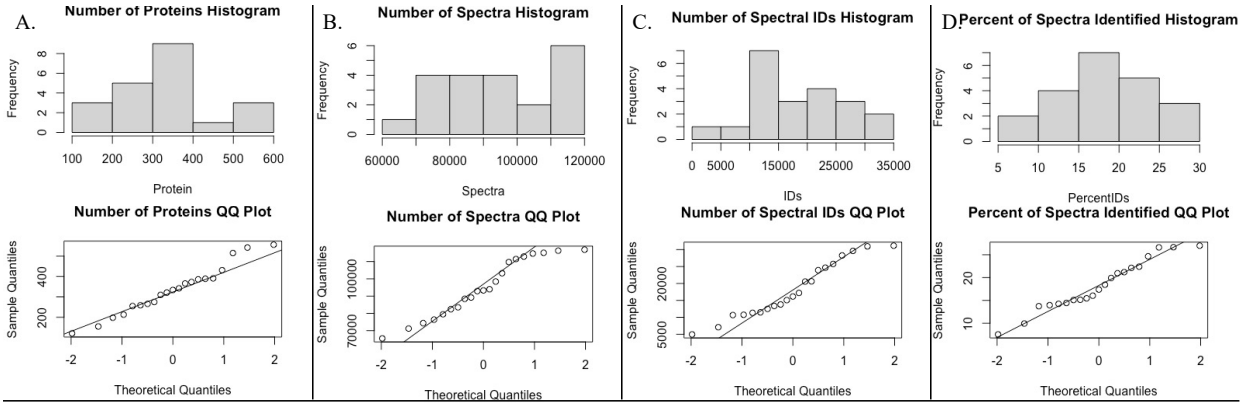
Figure AII.6 Manual validation of base-peak total ion chromatograms (TICs). The base-peak TIC is displayed for all three injections of samples with either an outlier when comparing LC-MS/MS statistics or a CV greater than 10% (highlighted in **Table AII7.2**)

7.3 Assessment of Samples After MudPIT Combination of Individual Injections

Following the MudPIT procedure, the combined injection data (**Table AII7.3**) was assessed for normality by histogram, QQ plot, and the Shapiro-Wilk's test (**Figure AII7.7**). All the data are normally distributed. This data was then assessed for equal variance by sample and fraction using Bartlett's test. All of the data have equal variance by sample and by fraction. This data meets the assumptions for ANOVA with the post-hoc Tukey test or Student's t-test.

Table AII.3 LC-MS/MS statistics for combined injections.

Sample	Fraction	# Proteins	# IDs	# Spectra	% IDs
1	100R	542	30959	116331	26.21
1	F1	515	31156	116894	26.65
1	F2	556	29615	109749	26.98
1	F3	391	16146	92929	17.37
1	F4	321	12489	82500	15.14
1	F5	213	10668	76425	13.96
1	qEV	275	11326	79518	14.24
2	100R	431	28275	114665	24.66
2	F1	386	25765	115084	22.39
2	F2	366	20626	98464	20.95
2	F3	310	15077	94015	16.04
2	F4	255	13349	88475	15.09
2	F5	198	10748	74425	14.44
2	qEV	389	17209	93321	18.44
3	100R	373	24662	111469	22.12
3	F1	342	23922	112919	21.19
3	F2	334	20561	103261	19.91
3	F3	259	13790	89229	15.45
3	F4	121	4955	65577	7.56
3	F5	155	7081	71263	9.94
3	qEV	266	11476	83538	13.74



E.

Shapiro-wilks test	Number of Proteins	Number of Spectra	Number of Spectral IDs	% Spectra Identified
Statistic	0.9718712	0.9324914	0.9393227	0.9559462
P-value	0.7740438	0.1544192	0.211282	0.4385181

Figure AII.7 Normality testing for combined injections. The histogram and QQ plots are displayed for (A) number of proteins identified, (B) number of spectra, (C) number of spectral IDs, and (D) percent of spectra identified. (E) The statistic and p-value for Shapiro-Wilks test of these data is also included.

Table AII.4 Proteins that vary in abundance by sample. The Go Cellular Components codes are the following: Cyt = cytosol, PM = plasma membrane, Sec = secreted/extracellular, CS = cell surface, and CW = cell wall

Protein	ANOVA Test (p-value): *(p < 0.00329)	GO Cellular Component(s)	Mycobrowser Function(s)	Mycobrowser Functional Category	Samples Detected In
High in S1 & S2, Low in S3					
Alr (Rv3423c)	0.00057	Cyt	Alanine racemase	Cell wall biosynthesis	S1, S2, S3
GuaB1 (Rv1843c)	0.0013	Cyt, PM	GMP biosynthesis	Intermediary metabolism and respiration	S1, S2, S3
High in S2 & S3, Low in S1					
GltB (Rv3859c)	0.0012	Cyt, PM	Probable ferredoxin-dependent glutamate synthase	Intermediary metabolism and respiration	S1, S2, S3
High in S1, Low in S2 & S3					
AccA3 (Rv3285)	0.00056	CW, PM	Long-chain fatty acid synthesis	Lipid metabolism	S1, S2, S3
ClpB (Rv0384c)	< 0.00010	CW, PM, Cyt	Probable endopeptidase ATP binding, heat shock protein	Virulence, detoxification, adaptation	S1, S2, S3
EttA (Rv2477c)	0.002	CW, Cyt, PM	Probable macrolide-transport ATP-binding protein	Cell wall and cell processes	S1
FabG4 (Rv0242c)	< 0.00010	CW, Cyt, PM, Sec	Fatty acid biosynthesis	Lipid metabolism	S1, S2, S3
GrpE (Rv0351)	< 0.00010	CW, Cyt	Heat shock protein 70 cofactor	Virulence, detoxification, adaptation	S1, S2, S3
HtpG (Rv2299c)	< 0.00010	CW, Cyt, PM, Sec	Probable chaperone, heat shock protein	Virulence, detoxification, adaptation	S1, S3
MenB (Rv0548c)	0.0022	PM	Naphthoate synthase, menaquinone biosynthesis	Intermediary metabolism and respiration	S1, S2, S3
MraZ (Rv2166c)	0.0016	Cyt, PM	Conserved protein	Conserved hypotheticals	S1, S3
Pca (Rv2967c)	< 0.00010	CW, Cyt, PM	Probable pyruvate carboxylase in gluconeogenesis and lipogenesis	Intermediary metabolism and respiration	S1
Rne (Rv2444c)	0.0015	CW, Cyt, PM	Possible ribonuclease	Information pathways	S1
RplO (Rv0723)	< 0.00010	Cyt, PM	50S ribosomal protein	Information pathways	S1, S3
RpoB (Rv0667)	0.00013	CW, Cyt, PM	DNA-directed RNA polymerase	Information pathways	S1, S2
RpoC (Rv0668)	0.00017	CW, Cyt, PM	DNA-directed RNA polymerase	Information pathways	S1, S2
Rv1738	< 0.00010	Cyt, PM	Conserved protein	Conserved hypotheticals	S1
Rv2226	< 0.00010	Cyt, PM	Conserved protein	Conserved hypotheticals	S1, S3
Rv3075c	< 0.00010	PM	Conserved protein	Conserved hypotheticals	S1
Tuf (Rv0685)	0.00048	Cyt, PM	Probable iron-regulated elongation factor	Information pathways	S1, S3
Wag31 (Rv2145c)	0.00056	CW, Cyt, PM	Unknown	Cell wall and cell processes	S1
High in S2, Low in S1 & S3					
AccA2 (Rv0973c)	< 0.00010	PM	Long-chain fatty acid synthesis	Lipid metabolism	S2
AccD2 (Rv0974c)	0.00039	Cyt	Fatty acid metabolism	Lipid metabolism	S2
Cbs (Rv1077)	< 0.00010	CW, Cyt, PM, Sec	Probable cystathionine beta-synthase	Intermediary metabolism and respiration	S1, S2, S3
DapB (Rv2773c)	< 0.00010	CW, Cyt, PM	Dihydrodipicolinate reductase	Intermediary metabolism and respiration	S1, S2
DesA1 (Rv0824c)	< 0.00010	CW, Cyt, PM, CS	Probable acyl-[acyl-carrier protein] desaturase	Lipid metabolism	S1, S2, S3
EchA21 (Rv3374)	< 0.00010	Not Reported	Probable enoyl-CoA hydratase (fragment)	Lipid metabolism	S1, S2, S3
EchA8 (Rv1070c)	< 0.00010	CW, PM	Probable enoyl-CoA hydratase	Lipid metabolism	S1, S2, S3
FadE22 (Rv3061c)	0.00094	CW, PM, Sec	Probable acyl-CoA dehydrogenase	Lipid metabolism	S1, S2, S3
FadE8 (Rv0672)	0.00039	CW, Sec	Probable acyl-CoA dehydrogenase	Lipid metabolism	S1, S2, S3
GltS/GltX (Rv2992c)	0.00093	Cyt, PM	Glutamyl-tRNA synthetase	Information pathways	S1, S2, S3
GlyA1 (Rv1093)	< 0.00010	CW, Cyt, PM	Serine hydroxymethyltransferase 1	Intermediary metabolism and respiration	S1, S2, S3
HtdY (Rv3389c)	0.002	PM, Sec	Fatty acid synthesis type II	Intermediary metabolism and respiration	S1, S2, S3
Icd2 (Rv0066c)	0.00097	CW, Cyt, PM, Sec	Probable isocitrate dehydrogenase	Intermediary metabolism and respiration	S1, S2, S3
Mtr (Rv3744)	0.00065	Not Reported	Metal sensor transcriptional regulator	Regulatory proteins	S2

Protein	ANOVA Test (p-value): *(p < 0.00329)	GO Cellular Component(s)	Mycobrowser Function(s)	Mycobrowser Functional Category	Samples Detected In
Rv0458	< 0.00010	PM, Sec	Probable aldehyde dehydrogenase	Intermediary metabolism and respiration	S1, S2, S3
Rv3161c	0.00032	PM	Possible dioxygenase	Intermediary metabolism and respiration	S1, S2, S3
TreX/GlgX (Rv1564c)	0.00017	Sec	Probable maltooligosyltrehalose synthase	Virulence, detoxification, adaptation	S1, S2, S3

Table AII.5 Proteins that vary in abundance across fractions. The Go Cellular Components codes are the following: Cyt = cytosol, PM = plasma membrane, Sec = secreted/extracellular, CS = cell surface, and CW = cell wall

Protein	ANOVA Test (p-value): *(p < 0.01067)	Molecular Weight	Go Cellular Component(s)	Mycobrowser Function(s)	Mycobrowser Functional Category
100R high, F1 high, F2 high, F3 high, F4 low, F5 low, qEV high					
AccD1 (Rv2502c)	0.0018	57 kDa	Cyt, PM	Probable acetyl-/propionyl-CoA carboxylase	Lipid metabolism
Dcd (Rv0321)	0.0029	21 kDa	PM, Sec	Probable deoxycytidine triphosphate deaminase	Intermediary metabolism and respiration
MetB (Rv1079)	0.00053	41 kDa	Cyt, PM	Cystathionine gamma-synthase, methionine biosynthesis	Intermediary metabolism and respiration
PfkA (Rv3010c)	0.0077	37 kDa	Cyt, CW, PM	Probable 6-phosphofructonkinase involved in glycolysis	Intermediary metabolism and respiration
Rv3401	0.00082	87 kDa	PM	Conserved protein	Intermediary metabolism and respiration
100R low, F1 high, F2 high, F3 high, F4 high, F5 high, qEV low					
Eis (Rv2416c)	0.0033	44 kDa	Cyt, Sec	Acetylation, substrate unknown	Virulence, detoxification, adaptation
100R low, F1 low, F2 high, F3 high, F4 high, F5 high, qEV high					
Gap (Rv1436)	0.0077	36 kDa	Cyt, CW, PM, Sec	Probable glyceraldehyde 3-phosphate dehydrogenase	Intermediary metabolism and respiration
Rv2744c	0.00044	29 kDa	Cyt, CW, PM	Conserved 35 kDa alanine rich protein	Conserved hypotheticals
Ssb (Rv0054)	< 0.00010	17 kDa	Cyt, PM, Sec	Single-strand binding protein	Information pathways
TatA (Rv2094c)	< 0.00010	9 kDa	PM	Sec-independent protein translocase	Cell wall and cell processes
100R high, F1 high, F2 high, F3 high, F4 low, F5 low, qEV low					
Pnp (Rv0535)	0.0049	80 kDa	Cyt	Probable 5'-methylthioadenosine phosphorylase	Intermediary metabolism and respiration
100R high, F1 high, F2 high, F3 low, F4 low, F5 high, qEV low					
SodA/SodB (Rv3846)	0.00067	23 kDa	Cyt, PM, Sec	Superoxide dismutase	Virulence, detoxification, adaptation
100R high, F1 high, F2 low, F3 high, F4 high, F5 low, qEV low					
LeuA (Rv3710)	0.0053	70 kDa	PM, Sec	2-isopropylmalate synthase, leucine biosynthesis	Intermediary metabolism and respiration
100R high, F1 high, F2 low, F3 low, F4 high, F5 high, qEV low					
PepN (Rv2467)	0.0056	94 kDa	Cyt, PM	Probable aminopeptidase N	Intermediary metabolism and respiration
Rv2971	0.0085	30 kDa	Cyt, CW, PM	Probable oxidoreductase	Intermediary metabolism and respiration
100R high, F1 low, F2 high, F3 high, F4 low, F5 low, qEV high					
PurA (Rv0357c)	0.0028	47 kDa	Cyt, PM, Sec	Probable adenylosuccinate synthetase	Intermediary metabolism and respiration
100R low, F1 low, F2 high, F3 high, F4 high, F5 high, qEV low					
GlgE (Rv1327c)	0.0037	79 kDa	Cyt, PM	M1P-dependent maltosyltransferase	Intermediary metabolism and respiration
100R low, F1 low, F2 high, F3 high, F4 high, F5 low, qEV high					
FumC (Rv1098c)	0.0019	50 kDa	Cyt, CW, PM, Sec	Probable fumarase, TCA cycle	Intermediary metabolism and respiration

Protein	ANOVA Test (p-value): *(p < 0.01067)	Molecular Weight	Go Cellular Component(s)	Mycobrowser Function(s)	Mycobrowser Functional Category
LpqH (Rv3763)	< 0.00010	15 kDa	CW, PM, Sec, CS	19 kDa lipoprotein antigen precursor	Cell wall and cell processes
LprF (Rv1368)	< 0.00010	27 kDa	CW, PM, Sec	Probable conserved lipoprotein	Cell wall and cell processes
LprG (Rv1411c)	< 0.00010	25 kDa	Cyt, CW, PM, Sec	Conserved lipoprotein	Cell wall and cell processes
MetC (Rv3340)	0.0077	47 kDa	PM	Probable O-acetylhomoserine sulfhydrylase	Intermediary metabolism and respiration
ProC (Rv0500)	0.0024	30 kDa	Cyt, PM	Probable pyrroline-5-carboxylate reductase, proline biosynthesis	Intermediary metabolism and respiration
RplB (Rv0704)	0.00099	31 kDa	Cyt, PM	50S ribosomal protein	Information pathways
RplV (Rv0706)	0.00099	20 kDa	Cyt, CW	50S ribosomal protein	Information pathways
Rv1006	0.0013	61 kDa	CW, PM	Unknown	Conserved hypotheticals
Rv2969c	0.0032	27 kDa	CW, PM, Sec	Possible conserved membrane or secreted protein	Cell wall and cell processes
100R low, F1 low, F2 high, F3 low, F4 high, F5 high, qEV high					
Cfp29 (Rv0798c)	< 0.00010	29 kDa	CW, PM, Sec	29 kDa antigen	Virulence, detoxification, adaptation
100R low, F1 low, F2 low, F3 high, F4 high, F5 high, qEV high					
GlnA4 (Rv2860c)	0.0048	50 kDa	PM, Sec	Probable glutamine synthetase	Intermediary metabolism and respiration
PrcA (Rv2109c)	< 0.00010	27 kDa	CW, PM	Proteasome alpha subunit	Intermediary metabolism and respiration
RplT (Rv1643)	0.0067	15 kDa	Cyt, PM	50S ribosomal protein	Information pathways
100R high, F1 high, F2 high, F3 low, F4 low, F5 low, qEV low					
CaeA (Rv2224c)	< 0.00010	56 kDa	Cyt, PM, Sec	Probable carbosylesterase	Cell wall and cell processes
DapA (Rv2753c)	0.0029	31 kDa	Cyt, PM	Probable dihydrodipicolinate synthase	Intermediary metabolism and respiration
FbpC (Rv0129c)	0.0043	37 kDa	CW, Sec	Secreted antigen 85-C, mycolytransferase	Lipid metabolism
FolB (Rv3607c)	0.0061	15 kDa	Cyt	Probable dihydroneopterin aldolase	Intermediary metabolism and respiration
GlpX (Rv1099c)	0.0065	38 kDa	Cyt, PM	Fructose 1,6-bisphosphatase	Intermediary metabolism and respiration
HisB (Rv1601)	0.00022	23 kDa	Cyt	Probable imidazole glycerol-phosphate dehydratase, histidine biosynthesis	Intermediary metabolism and respiration
LprQ (Rv0483)	< 0.00010	48 kDa	Sec	Probable conserved lipoprotein	Cell wall and cell processes
MoaE2 (Rv0866)	< 0.00010	15 kDa	Cyt, PM, Sec	Probable molybdenum cofactor biosynthesis protein	Intermediary metabolism and respiration
Nat (Rv3566c)	< 0.00010	31 kDa	PM	Arylamine N-acetyltransferase	Intermediary metabolism and respiration
NrdF2 (Rv3048c)	0.0018	37 kDa	Cyt	Ribonucleoside-diphosphate reductase	Information pathways
Ppa (Rv3628)	0.00047	18 kDa	Cyt, PM, Sec	Inorganic pyrophosphatase	Intermediary metabolism and respiration
PstS3 (Rv0928)	< 0.00010	38 kDa	CW, PM, Sec	Periplasmic phosphate-binding lipoprotein	Cell wall and cell processes
Rv0296c	0.00055	52 kDa	CW	Probable sulfatase	Intermediary metabolism and respiration
Rv2141c	< 0.00010	48 kDa	PM	Conserved protein	Intermediary metabolism and respiration
Rv2857c	< 0.00010	27 kDa	PM	Probable short-chain type dehydrogenase/reductase	Intermediary metabolism and respiration
Rv3210c	0.00025	25 kDa	Cyt, PM, Sec	Probable ATP-dependent DNA helicase	Information pathways
Rv3510c	< 0.00010	31 kDa	Cyt	Conserved protein	Conserved hypotheticals
Rv3722c	0.00011	47 kDa	PM, Sec	Conserved protein	Conserved hypotheticals
ThyX (Rv2754c)	0.00066	28 kDa	PM	Probable tymidylate synthase	Intermediary metabolism and respiration
ViuB (Rv2895c)	0.0021	31 kDa	PM	Possible mycobactin utilization protein	Intermediary metabolism and respiration
100R high, F1 high, F2 low, F3 high, F4 low, F5 low, qEV low					
AldC (Rv2858c)	0.0044	48 kDa	PM, Sec	Probable aldehyde dehydrogenase	Intermediary metabolism and respiration
100R high, F1 high, F2 low, F3 low, F4 low, F5 high, qEV low					
FadA3 (Rv1074c)	0.0011	43 kDa	CW, PM, Sec	Probable beta-ketoacyl CoA thiolase, lipid degeneration	Lipid metabolism
Pgi (Rv0946c)	0.00094	60 kDa	Cyt, PM	Probable glucose-6-phosphate isomerase, glycolysis and gluconeogenesis	Intermediary metabolism and respiration
100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV high					
Mpt53 (Rv2878c)	0.0007	18 kDa	Cyt, Sec	Soluble secreted antigen Mpt53 precursor	Cell wall and cell processes

Protein	ANOVA Test (p-value): *(p < 0.01067)	Molecular Weight	Go Cellular Component(s)	Mycobrowser Function(s)	Mycobrowser Functional Category
MycP3 (Rv0291)	0.00067	46 kDa	CW, PM, Sec	Probable membrane-anchored mycosin	Intermediary metabolism and respiration
Rv0787	0.0081	34 kDa	Cyt, Sec	Unknown	Conserved hypotheticals
100R low, F1 high, F2 high, F3 high, F4 low, F5 low, qEV low					
TyrS (Rv1689)	0.0038	46 kDa	Cyt, CW, PM	Probable tyrosyl-tRNA synthase	Information pathways
100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV high					
ArgC (Rv1652)	< 0.00010	36 kDa	Cyt, PM, Sec	Probable N-acetyl-gamma-glutamyl-phosphate reductase, arginine biosynthesis	Intermediary metabolism and respiration
LppK (Rv2116)	< 0.00010	20 kDa	PM	Probable conserved lipoprotein	Cell wall and cell processes
LppX (Rv2945c)	0.00046	24 kDa	CW, PM, Sec, CS	Possible conserved lipoprotein	Cell wall and cell processes
PrsA (Rv1017c)	0.002	35 kDa	Cyt, CW, PM	Probable ribose-phosphate pyrophosphokinase	Intermediary metabolism and respiration
Pta (Rv2174)	0.0035	73 kDa	PM	Alpha(1->6)mannosyltransferase	Cell wall and cell processes
PyrB (Rv1380)	0.0013	34 kDa	Cyt, CW, PM	Probable aspartate carbamoyltransferase	Intermediary metabolism and respiration
RplD (Rv0702)	< 0.00010	24 kDa	Cyt, PM	Probable ribose-phosphate pyrophosphokinase	Intermediary metabolism and respiration
Rv0968	< 0.00010	10 kDa	CW, PM	Conserved protein	Conserved hypotheticals
Rv0999	0.0013	26 kDa	PM, CS, Sec	Unknown	Conserved hypotheticals
VapC11 (Rv1561)	0.00016	15 kDa	PM	Possible toxin	Virulence, detoxification, adaptation
100R low, F1 low, F2 high, F3 low, F4 low, F5 high, qEV high					
Rv0799c	0.00062	36 kDa	Cyt, PM, Sec	Conserved protein	Conserved hypotheticals
100R low, F1 low, F2 low, F3 high, F4 high, F5 low, qEV high					
CsoR (Rv0967)	< 0.00010	13 kDa	Cyt, PM	Copper-sensitive operon repressor	Regulatory proteins
GlnA2 (Rv2222c)	0.00029	50 kDa	Cyt, PM	Probable glutamine synthetase, glutamine biosynthesis	Intermediary metabolism and respiration
PpiB (Rv2582)	0.0067	32 kDa	PM	Probable peptidyl-prolyl cis-trans isomerase	Information pathways
Rv0227c	0.004	46 kDa	CW, PM	Probable conserved membrane protein	Cell wall and cell processes
100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low					
Adk (Rv0733)	< 0.00010	20 kDa	Cyt, CW, PM	Adenylate kinase, nucleotide metabolism	Intermediary metabolism and respiration
Cfp32 (Rv0577)	0.0024	27 kDa	Cyt, PM, Sec	Conserved protein TB27.3	Conserved hypotheticals
CmaA2 (Rv0503c)	0.0026	35 kDa	Cyt, PM	Cyclopropane-fatty-acyl-phospholipid synthase 2	Lipid metabolism
Cyp142 (Rv3518c)	0.001	44 kDa	CW, Sec	Probable cytochrome P450 monooxygenase	Intermediary metabolism and respiration
DeoC (Rv0468)	0.0083	22 kDa	Cyt	Probable deoxyribose-phosphate aldolase	Intermediary metabolism and respiration
DhaA (Rv2579)	0.0005	34 kDa	CW, PM	Possible haloalkane dehalogenase	Intermediary metabolism and respiration
DppA (Rv3666c)	< 0.00010	58 kDa	PM, CS	Probable periplasmic dipeptide-binding lipoprotein, active transport (import)	Cell wall and cell processes
EccB5 (Rv1782)	< 0.00010	54 kDa	CW, PM, Sec	ESX conserved component ESX-5 type VII secretion system	Cell wall and cell processes
EphB (Rv1938)	< 0.00010	39 kDa	Sec	Probable epoxide hydrolase	Virulence, detoxification, adaptation
EspB (Rv3881c)	0.004	48 kDa	Sec	Secreted ESX-1 substrate protein B	Cell wall and cell processes
FecB (Rv3044)	0.00098	37 kDa	PM, Sec, CS	Probable FEIII-dicitrate-binding periplasmic lipoprotein	Cell wall and cell processes
GarA (Rv1827)	< 0.00010	17 kDa	CW, PM, Sec	Conserved protein with FHA domain	Conserved hypotheticals
GcvP/GcvB (Rv1832)	0.0051	100 kDa	Cyt, CW, PM	Probable glycine dehydrogenase	Intermediary metabolism and respiration
GlgB (Rv1326c)	0.0011	82 kDa	Cyt, PM	1,4-alpha-glucan branching enzyme, glycogen biosynthesis	Intermediary metabolism and respiration
GpgP (Rv2419c)	0.0029	24 kDa	Cyt	Glucosyl-3-phosphoglycerate phosphatase, MGLP biosynthesis	Intermediary metabolism and respiration
Gpm2 (Rv3214)	< 0.00010	22 kDa	PM, Sec	Possible phosphoglycerate mutase, glycolysis	Intermediary metabolism and respiration
Hrp1 (Rv2626c)	0.00021	16 kDa	Cyt, PM, Sec	Hypoxic response protein	Conserved hypotheticals
Icd1 (Rv3339c)	< 0.00010	46 kDa	PM, Sec	Probable isocitrate dehydrogenase	Intermediary metabolism and respiration
Mak (Rv0127)	0.0016	50 kDa	Cyt, PM	Maltokinase	Virulence, detoxification, adaptation
MapA (Rv0734)	0.0032	31 kDa	PM	Methionine aminopeptidase	Intermediary metabolism and respiration

Protein	ANOVA Test (p-value): *(p < 0.01067)	Molecular Weight	Go Cellular Component(s)	Mycobrowser Function(s)	Mycobrowser Functional Category
MmaA4 (Rv0642c)	< 0.00010	35 kDa	CW, PM	Methoxy mycolic acid synthase 4, mycolic acids modification	Lipid metabolism
MmpR5 (Rv0678)	< 0.00010	18 kDa	PM	Conserved protein	Conserved hypotheticals
Mpt64 (Rv1980)	0.0026	25 kDa	CW, Sec	Immunogenic protein Mpt64	Cell wall and cell processes
Mpt70 (Rv2875)	< 0.00010	19 kDa	Sec	Major secreted immunogenic protein	Cell wall and cell processes
Mtc28 (Rv0040c)	0.0032	32 kDa	Cyt, Sec	Secreted proline rich protein	Cell wall and cell processes
Pat (Rv3772)	< 0.00010	38 kDa	PM, Sec	Probable histidinol-phosphate aminotransferase, histidine biosynthesis	Intermediary metabolism and respiration
PckG (Rv0211)	0.0014	67 kDa	Cyt, CW, PM, Sec	Probable iron-regulated phosphoenolpyruvate carboxykinase	Intermediary metabolism and respiration
PunA (Rv3307)	< 0.00010	28 kDa	Cyt, PM	Probable purine nucleoside phosphorylase	Intermediary metabolism and respiration
RpiB (Rv2465c)	< 0.00010	17 kDa	PM, Sec	Ribose-5-phosphate isomerase	Intermediary metabolism and respiration
Rv0148	0.0067	30 kDa	CW, PM, Sec	Probable short-chain type dehydrogenase/reductase	Intermediary metabolism and respiration
Rv0161	< 0.00010	47 kDa	PM	Possible oxidoreductase	Intermediary metabolism and respiration
Rv0315	0.0053	32 kDa	PM, Sec	Possible beta-1,3-glucanase precursor	Intermediary metabolism and respiration
Rv0331	0.0068	41 kDa	Cyt, PM, Sec	Possible dehydrogenase/reductase	Intermediary metabolism and respiration
Rv0398c	< 0.00010	22 kDa	Sec	Possible secreted protein	Cell wall and cell processes
Rv0457c	0.00032	74 kDa	PM, Sec	Probable peptidase	Intermediary metabolism and respiration
Rv0526	0.0025	23 kDa	PM, Sec	Possible thioredoxin protein	Intermediary metabolism and respiration
Rv0546c	0.00086	14 kDa	PM, Sec	Conserved protein	Conserved hypotheticals
Rv1531	0.00069	21 kDa	Cyt, PM, Sec	Conserved protein	Conserved hypotheticals
Rv1700	< 0.00010	23 kDa	Cyt, Sec	NUDIX hydrolase	Information pathways
Rv1836c	< 0.00010	70 kDa	CW, PM, Sec	Conserved protein	Conserved hypotheticals
Rv2140c	< 0.00010	19 kDa	PM, Sec	Conserved protein TB18.6	Conserved hypotheticals
Rv2250A	< 0.00010	15 kDa	Not Reported	Possible flavoprotein, electron acceptor	Intermediary metabolism and respiration
Rv2575	< 0.00010	31 kDa	PM, Sec	Possible conserved membrane glycine rich protein	Cell wall and cell processes
Rv2765	0.0053	26 kDa	PM	Probable alanine rich hydrolase	Intermediary metabolism and respiration
Rv2799	< 0.00010	23 kDa	PM	Probable membrane protein	Cell wall and cell processes
Rv3668c	< 0.00010	23 kDa	CS, Sec	Possible protease/surface lipoprotein	Intermediary metabolism and respiration
SubI (Rv2400c)	0.0018	37 kDa	PM, Sec	Probable sulfate-binding lipoprotein, active transport (import)	Cell wall and cell processes
TB22 (Rv3036c)	0.0026	24 kDa	Sec	Probable conserved secreted protein TB22.2	Cell wall and cell processes
TpiA (Rv1438)	< 0.00010	27 kDa	Cyt, PM, Sec	Probable triosephosphate isomerase	Intermediary metabolism and respiration
TrxA (Rv1470)	0.004	13 kDa	Cyt, PM	Probable thioredoxin	Intermediary metabolism and respiration
TrxB1 (Rv1471)	0.0014	36 kDa	Cyt	Probable thioredoxin	Intermediary metabolism and respiration
100R high, F1 low, F2 high, F3 low, F4 low, F5 low, qEV low					
PurU (Rv2964)	< 0.00010	34 kDa	Cyt	Probable formyltetrahydrofolate deformylase, purine biosynthesis	Intermediary metabolism and respiration
100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV low					
LpqG (Rv3623)	0.0003	25 kDa	PM	Probable conserved lipoprotein	Cell wall and cell processes
VapC4 (Rv0595c)	0.0046	14 kDa	Cyt	Possible toxin	Virulence, detoxification, adaptation
100R low, F1 low, F2 high, F3 low, F4 low, F5 low, qEV high					
Rv1265	0.00052	25 kDa	Cyt, PM	Unknown protein	Conserved hypotheticals
Rv1762c	0.0017	29 kDa	Cyt, PM	Unknown protein	Conserved hypotheticals
100R low, F1 low, F2 low, F3 high, F4 low, F5 low, qEV high					
DapD (Rv1201c)	0.0017	33 kDa	Cyt, PM	Tetrahydrodipicolinate N-succinyltransferase, lysine biosynthesis	Intermediary metabolism and respiration
PykA (Rv1617)	0.0013	51 kDa	Cyt, PM	Probable pyruvate kinase	Intermediary metabolism and respiration
Rv1287	< 0.00010	17 kDa	Cyt	Conserved hypothetical protein	Conserved hypotheticals

Protein	ANOVA Test (p-value): *(p < 0.01067)	Molecular Weight	Go Cellular Component(s)	Mycobrowser Function(s)	Mycobrowser Functional Category
100R low, F1 low, F2 low, F3 low, F4 high, F5 high, qEV low					
DnaK (Rv0350)	0.003	67 kDa	Cyt, PM, CS, Sec	Probable chaperone protein, heat-shock response	Virulence, detoxification, adaptation
LpdC (Rv0462)	0.0073	49 kDa	Cyt, PM, Sec	Dihydrolipoamide dehydrogenase	Intermediary metabolism and respiration
100R high, F1 low, F2 low, F3 low, F4 low, F5 low, qEV low					
PepA (Rv0125)	0.0039	35 kDa	CW, Sec	Probable serine protease	Intermediary metabolism and respiration
SseC1 (Rv3118)	< 0.00010	10 kDa	Not Reported	Conserved hypothetical protein	Intermediary metabolism and respiration
100R low, F1 low, F2 high, F3 low, F4 low, F5 low, qEV low					
Rv0048c	< 0.00010	31 kDa	CW, PM	Possible membrane protein	Cell wall and cell processes
100R low, F1 low, F2 low, F3 low, F4 low, F5 low, qEV high					
ArcA (Rv1001)	0.0018	43 kDa	Cyt, PM	Probable arginine deiminase	Intermediary metabolism and respiration

Table AII.6 Proteins identified with predicted signal peptide. All proteins detected in this data set with a predicted signal sequence by signal P. TAT = twin-arginine motif, LIPO = lipobox motif, and SP = signal peptide

SignalP	Protein	ANOVA Test (p-value): *(p < 0.00863)	Quantitative Profile
TAT	lprQ	< 0.00010	100R high, F1 high, F2 high, F3 low, F4 low, F5 low, qEV low
TAT	Rv0315	0.0053	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
TAT	Rv3796	0.0089	NS
TAT	blaC	0.078	NS
TAT	Rv0063	0.093	NS
TAT	Rv0265c	0.11	NS
TAT	fbpB	0.11	NS
TAT	mmcO	0.28	NS
TAT	fbpA	0.32	NS
LIPO	lprF	< 0.00010	100R low, F1 low, F2 high, F3 high, F4 high, F5 low, qEV high
LIPO	lprG	< 0.00010	100R low, F1 low, F2 high, F3 high, F4 high, F5 low, qEV high
LIPO	lpqH	< 0.00010	100R low, F1 low, F2 high, F3 high, F4 high, F5 low, qEV high
LIPO	Rv1006	0.0013	100R low, F1 low, F2 high, F3 high, F4 high, F5 low, qEV high
LIPO	pstS3	< 0.00010	100R high, F1 high, F2 high, F3 low, F4 low, F5 low, qEV low
LIPO	caeA	< 0.00010	100R high, F1 high, F2 high, F3 low, F4 low, F5 low, qEV low
LIPO	lppK	< 0.00010	100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV high
LIPO	lppX	0.00046	100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV high
LIPO	Rv0999	0.0013	100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV high
LIPO	dppA	< 0.00010	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
LIPO	subI	0.0018	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
LIPO	Rv0526	0.0025	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
LIPO	lpqG	0.0003	100R low, F1 low, F2 high, F3 high, F4 low, F5 low, qEV low
LIPO	lpqB	0.011	NS
LIPO	lprA	0.011	NS
LIPO	lpqW	0.011	NS
LIPO	lppY	0.021	NS
LIPO	lpqN	0.036	NS
LIPO	lppM	0.037	NS
LIPO	pstS2	0.046	NS
LIPO	ggtB	0.056	NS
LIPO	sodC	0.074	NS
LIPO	Rv3033	0.11	NS
LIPO	lppL	0.12	NS
LIPO	Rv1922	0.17	NS
LIPO	proX	0.19	NS
LIPO	pstS1	0.23	NS
LIPO	lpqT	0.24	NS
LIPO	lpqK	0.24	NS
LIPO	lpqI	0.31	NS
LIPO	lppO	0.34	NS
LIPO	Rv2672	0.38	NS
LIPO	lppZ	0.38	NS
LIPO	lpqE	0.46	NS
LIPO	Rv0309	0.46	NS
LIPO	Rv0679c	0.46	NS
LIPO	lpqL	0.46	NS
LIPO	lpqO	0.46	NS
LIPO	Rv2585c	0.52	NS
LIPO	lpqM	0.56	NS
LIPO	mpt83	0.56	NS
LIPO	lppJ	0.56	NS
LIPO	lpqZ	0.57	NS
LIPO	msl3	1	NS
SP	fbpC	0.0043	100R high, F1 high, F2 high, F3 low, F4 low, F5 low, qEV low
SP	mycP3	0.00067	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV high
SP	mpt53	0.0007	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV high
SP	Rv3668c	< 0.00010	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
SP	Rv0398c	< 0.00010	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
SP	mpt70	< 0.00010	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
SP	TB22	0.0026	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low

SignalP	Protein	ANOVA Test (p-value): *(p < 0.00863)	Quantitative Profile
SP	mpt64	0.0026	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
SP	mtc28	0.0032	100R high, F1 high, F2 low, F3 low, F4 low, F5 low, qEV low
SP	pepA	0.0039	100R high, F1 low, F2 low, F3 low, F4 low, F5 low, qEV low
SP	dacB2	0.011	NS
SP	mycP5	0.012	NS
SP	mpt51	0.012	NS
SP	Rv1280c	0.02	NS
SP	mtb12	0.038	NS
SP	cut7	0.038	NS
SP	mpt63	0.062	NS
SP	Rv2721c	0.064	NS
SP	Rv3572	0.086	NS
SP	Rv2345	0.094	NS
SP	Rv1885c	0.1	NS
SP	Rv3732	0.2	NS
SP	Rv1815	0.21	NS
SP	sapM	0.24	NS
SP	mtp	0.27	NS
SP	mku	0.45	NS
SP	Rv1906c	0.46	NS
SP	Rv1987	0.46	NS
SP	Rv2576c	0.46	NS
SP	Rv1096	0.46	NS
SP	Rv0518	0.54	NS
SP	mycP2	0.56	NS
SP	Rv1566c	0.56	NS
SP	apa	0.81	NS

7.5 R Code Access

R script is available at:

https://gist.github.com/jmryan12/75e22caba63e2962d43a128300e91a91

R Markdown with data and script output will be available as a supplementary file.

LIST OF ABBREVIATIONS

ACN – acetonitrile

AFC – automatic fraction collector

AF4 – asymmetric flow field-flow fractionation

ANOVA – analysis of variance

BCA – bicinchoninic acid assay

BCG – Bacille Calmette-Guérin

BMM – bone marrow-derived macrophages

BSC – biosafety cabinet

BSL-# – biosafety level #

BSO – Biosafety Office

CC-EV – extracellular vesicles enriched with CaptoCore multimodal size exclusion chromatography

CFP – cellular filtrate protein

COVID-19 – coronavirus disease 2019

CS – cell surface

CSU – Colorado State University

CV – coefficient of variation

CW – cell wall

Cyt - cytosol

DNA – deoxyribonucleic acid

DC – dendritic cell

DLPs – dynamin-like proteins

DNA – deoxyribonucleic acid

DTT – dithiothreitol

DURC – dual use research of concern

ESCRT – endosomal sorting complex required for transport

EV – extracellular vesicle

F# - fraction #

FDR – false discovery rate

FSAP – federal select agent program

GAS – glycerol alanine salts medium

GO – gene ontology

H37Rv – the most studied strain of *Mycobacterium tuberculosis* isolated in 1905

HGT – human gene therapy

HIV – human immunodeficiency virus

IBC – institutional biosafety committee

IFN – interferon

IL-# – interleukin

INH – isoniazid

LAM – lipoarabinomannan

LC-MS/MS – liquid chromatography tandem mass spectrometry

LTBI – latent tuberculosis infection

MAH – *Mycobacterium avium* subspecies *hominissuis*

MALS – multi-angle light scattering

MAP – *Mycobacterium avium* subspecies *paratuberculosis*

MHC – major histocompatibility complex

MISEV – minimal information for studies of extracellular vesicles

Mtb – *Mycobacterium tuberculosis*

MWCO – molecular weight cut off

MS – mass spectrometry

MudPIT – multidimensional protein identification technology

NCBI – National Center for Biotechnology Information

NIH Guidelines - NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules

NSAF – normalized spectral abundance factor

NTA – nanoparticle tracking analysis

OADC – oleic albumin dextrose catalase

OMV – outer membrane vesicle

1XPBS – phosphate buffered saline

PE – phosphatidylethanolamine

PE/PPE- proteins with Pro-Glu (PE) or Pro-Pro-Glu (PPE) N-terminal motif

PES – polyethersulfone

PIM – phosphatidylinositol mannoside

PM – plasma membrane

PPD – purified protein derivative

PPE – personal protective equipment

PTM – post translational modification

qEV-EV – extracellular vesicles enriched with qEV size exclusion chromatography

QQ plot – quantile-quantile plot

RAC – NIH Recombinant DNA Advisory Committee

rDNA – recombinant deoxyribonucleic acid

RNA – ribonucleic acid

rsNA – recombinant or synthetic nucleic acid

S# - sample #

SARS-CoV-2 – severe acute respiratory syndrome coronavirus 2

SEC – size exclusion chromatography

SEM – scanning electron microscopy

SDS-PAGE – sodium dodecyl sulfate polyacrylamide gel electrophoresis

TAT – twin arginine translocation

TB – Tuberculosis disease

TBS – tris buffered saline

TBST – tris buffered saline with Tween 80

TEM – transmission electron microscopy

TGS – triacylglycerol synthase

Th1 – T helper type 1, lineage of CD4+ effector T cells

TIC – total ion chromatogram

TLR – toll-like receptor

TNF – tumor necrosis factor

T7S – type VII secretion

UC-EV – extracellular vesicles enriched by ultracentrifugation

WHO – World Health Organization

WT – wild type

100R – CFP concentrated with a 100 kDa ultrafiltration unit

100FT – flow through CFP from a 100 kDa ultrafiltration unit

UNITS

°C – degrees Celsius

rpm – rotations per minute

x g – relative centrifugal force

kDa – kilodalton

min – minute

h – hour

g – gram

mg – milligram

µg – microgram

L – liter

mL – milliliter

µL – microliter

cm – centimeter

mm – millimeter

µm – micrometer

nm – nanometer

M – molar

mM – millimolar

n = # – sample size

pH – power of hydrogen

$ USD – United States dollar

V – volt