In silico secretome characterization of clinical *Mycobacterium abscessus* isolates provides insights into antigenic differences

Fernanda Cornejo-Granados[1], Thomas A. Kohl[2,3], Flor Vásquez Sotomayor[4], Sönke Andres[4], Rogelio Hernández-Pando[5], Juan Manuel Hurtado-Ramírez[1], Christian Utpatel[2,3], Stefan Niemann[2,3], Florian P. Maurer[4,6,# *] and Adrian Ochoa-Leyva[1,#,*]

[1] Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autonoma de México, Cuernavaca, Morelos, México.
[2] Molecular and Experimental Mycobacteriology, Research Center Borstel, Borstel, Germany
[3] German Center for Infection Research (DZIF), Partner site Hamburg-Lübeck-Borstel, Borstel, Germany
[4] National and WHO Supranational Reference Center for Mycobacteria, Research Center Borstel, Leibniz Lung Center, Borstel, Germany
[5] Experimental Pathology Section, National Institute of Medical Sciences and Nutrition Salvador Zubirán, Mexico City, Mexico.
[6] Institute of Medical Microbiology, Virology and Hospital Hygiene, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[#]equal contribution

Correspondance:
[*]co-corresponding authors:
Adrian Ochoa-Leyva: aochoa@ibt.unam.mx
Florian P. Maurer: fmaurer@fz-borstel.de

Keywords: bioinformatics, antigenicity, M. abscessus subspecies, in silico analysis, vaccinology.

Abstract

Mycobacterium abscessus (MAB) is a widely disseminated pathogenic non-tuberculous mycobacterium (NTM). Like with *M. tuberculosis* complex (MTBC), excreted / secreted (ES) proteins play an essential role for its virulence and survival inside the host. ES proteins contain highly immunogenic proteins, which are of interest for novel diagnostic assays and vaccines. Here, we used a robust bioinformatics pipeline to predict the secretome of the *M. abscessus* ATCC 19977 reference strain and fifteen clinical isolates belonging to all three MAB subspecies, *M. abscessus* subsp. *abscessus*, *M. abscessus* subsp. *bolletii*, and *M. abscessus* subsp. *massiliense*. We found that ~18% of the proteins encoded in the MAB genomes were predicted as secreted and that the three MAB subspecies shared > 85 % of the predicted secretomes. MAB isolates with a rough (R) colony morphotype showed larger predicted secretomes than isolates with a smooth (S) morphotype. Additionally, proteins exclusive to the secretomes of MAB R variants had higher antigenic densities than those exclusive to S variants, independently of the subspecies. For all investigated isolates, ES proteins had a significantly higher antigenic density than non-ES proteins. We identified 337 MAB ES proteins with homologues in previously investigated *M. tuberculosis* secretomes. Among these, 222 have previous experimental support of secretion, and some proteins showed homology with protein drug targets reported in the DrugBank database. The predicted MAB secretomes showed a higher abundance of proteins related to quorum-sensing and Mce domains as compared to MTBC indicating the importance of these pathways for MAB pathogenicity and virulence. Comparison of the predicted secretome of *M. abscessus* ATCC 19977 with the list of essential genes revealed that 99 secreted proteins corresponded to essential proteins required for *in vitro* growth. All predicted secretomes were deposited in the Secret-AAR web-server (http://microbiomics.ibt.unam.mx/tools/aar/index.php).

Introduction

Non-tuberculous mycobacteria (NTM) are widely disseminated, mostly saprophytic and partly opportunistic bacteria. The prevalence of NTM in clinical specimens has increased globally, and in some industrialized countries, infections caused by NTM are becoming more common than tuberculosis (TB). Infections caused by *M. abscessus* (MAB) are particularly challenging to manage due to the extensive innate resistance of MAB against a wide spectrum of clinically available antimicrobials (Nessar et al., 2012). MAB causes mostly pulmonary and occasionally extrapulmonary infections that can affect all organs in the human body (Lee et al., 2015). Current treatments for MAB induced pulmonary disease are long, associated with severe side effects and a cure rate below 50 % (Chen et al., 2019; Jarand et al., 2011; Sanguinetti et al., 2001). MAB is comprised of three subspecies, *M. abscessus* subsp. *abscessus, M. abscessus* subsp. *bolletii* and *M. abscessus* subsp. *massiliense*, hereafter referred to as $MAB_{A,}$ MAB_{B}, and MAB_{M}, respectively (Tortoli et al., 2016). MAB isolates can show smooth (S) and rough (R) colony morphotypes, a trait that relies on the presence (S) or absence (R) of surface- associated glycopeptidolipids (GPLs) and that correlates with the virulence of the strain (Abeles and Pride, 2014; Howard et al., 2006; Ripoll et al., 2007) (Gutiérrez et al., 2018)). Transitioning from high-GPL to low-GPL production is observed in sequential MAB isolates obtained from patients with chronic underlying pulmonary disease. In these patients, S-to-R conversion is thought to present a selective advantage as the aggregative properties of MAB R variants strongly affect intracellular survival. In addition, a propensity to grow as extracellular cords allows these low-GPL producing bacilli to escape innate immune defenses (Gutiérrez et al., 2018).

The complete set of proteins excreted / secreted (ES) by a bacterial cell is referred to as its secretome. The secretome is involved in critical biological processes such as cell adhesion, migration, cell-to-cell communication and signal transduction (Tjalsma et al., 2004) ES proteins are considered an important source of molecules for serological diagnosis. Also, secreted proteins can be highly antigenic due to their immediate availability to the host immune system and are thus of interest in vaccinology (Daugelat et al., 1992; Zheng et al., 2013). So far, there have been few efforts to experimentally determine the secretome of MAB, and in particular, the secretomes of clinical MAB isolates (Gupta et al., 2009; Laencina et al., 2018; Shin et al., 2010; Yadav and Gupta, 2012). Nowadays, sequencing and bioinformatics strategies can be explored for the systematized prediction of ES proteins from bacterial genomes (Cornejo-Granados et al., 2017; Gomez et al., 2015). Recently, a robust bioinformatics pipeline for predicting and analyzing the complete *in silico* secretome of two clinical *M. tuberculosis* (MTB) genomes was published in 2017 by our group showing higher overall agreement with an experimental secretome compiled from literature than two previously reported secretomes for *M. tuberculosis* H37Rv (Cornejo-Granados et al., 2017).

To gain further insights into MAB ES proteins and their association with virulence and pathogenicity we sequenced and assembled the genomes of fifteen clinical MAB isolates belonging to all three subspecies including S and R morphotypes. We then adapted the bioinformatics strategy previously established for MTB to predict and analyze the complete set of ES proteins encoded in these isolates and in the *M. abscessus* ATCC 19977 type strain, and compared it with our previous findings for MTB (Cornejo-Granados et al., 2017).

Materials and methods

Ethics statement

Ethical review and approval was not required for the study as all work was performed on bacterial isolates archived at the strain repository of the National Reference Center for Mycobacteria in Borstel, Germany, in accordance with local legislation and institutional requirements. In particular, no data allowing identification of the affected patients was shared or released and no human DNA was sequenced or analyzed.

Clinical isolates

We selected fifteen MAB clinical isolates comprising members of all MAB subspecies (MAB_A, n = 7; MAB_B, n = 4; MAB_M, n = 4) and both S (n = 8) and R (n = 6) morphotypes (not determined, n = 1). The strains were isolated from different biological sources representing both pulmonary colonization / infection (sputum, n = 10) and extrapulmonary samples (skin, n = 1; soft tissue, n = 1; lymph nodes, n = 2; blood, n = 1) (Table 1and S1). For routine diagnostic purposes, species identification was performed using GenoType NTM-DR line probe assays (HAIN Lifescience, Nehren, Germany) and sequencing of the 16S and *rpoB* genes as described previously (Adekambi et al., 2003).

Whole genome sequencing and genome assemblies

Genomic DNA (gDNA) of the 15 MAB clinical isolates was extracted from solid cultures using a Centrimonium bromide chloroform DNA extraction protocol as previously described (De Almeida et al., 2013). DNA libraries were constructed with the Nextera XT kit from Illumina and sequenced on the Illumina MiSeq benchtop platform with a v3 chemistry paired –end run and a read lenght of 2x300 bp. We processed the resulting reads with Trimmomatic (Bolger et al., 2014), clipping the Illumina adapter sequences and trimming the reads with a sliding window of 20 bp looking for quality >30 and discarding all reads shorter than 100 bp. Trimmed reads were used to construct *de novo* assemblies using SPADES (Nurk et al., 2013) with default parameters and the --careful option enabled. Then, each assembly was analyzed with RAST (Aziz et al., 2008) to obtain all the open reading frames (ORFs). Additionally, we predicted the ORFs from the deposited genome sequence of the *M. abscessus* ATCC 19977 type strain (GenBank CU458896.1) (Supplementary Table S1).

Secretome prediction

The complete set of predicted ORFs was independently analyzed for each genome using the bioinformatics pipeline previously reported by Cornejo-Granados et al., 2017 and summarized in Supplementary Fig. S1. Briefly, we used six different feature-based tools (SignalP, SecretomeP, LipoP, TatP, TMHMM and Phobius) (Bendtsen et al., 2005a; 2005b; Petersen et al., 2011; Sonnhammer et al., 1998) (Juncker et al., 2003; Käll et al., 2007) to identify ES proteins by the different secretion pathways and to remove the ones that had

transmembrane domains (Supplementary Fig. S1). The proteins assigned as not-secreted (non-ES) were further classified into transmembrane proteins (TM) if they showed the presence of transmembrane domains with TMHMM 2.0 (Sonnhammer et al., 1998), and into intracellular proteins (incell) if they did not contain any transmembrane domains.

Annotation and comparative analysis of secreted proteins

To assign functional annotations to the proteins present in our genomes, we performed a BLASTP query of those proteins against the non-redundant (nr) complete database using Blast2GO (Conesa and Götz, 2008) with an E-value cut-off set at 1.0E-3. Furthermore, all proteins were associated with protein families through InterProScan (Zdobnov and Apweiler, 2001) and functionally mapped to Gene Ontology (GO) terms by setting the following parameters: E-value-hi-filter: 1.0E-3; Annotation cut-off: 55; GO weight: 5 and Hsp-Hit Coverage cut-off: 0. Blast2GO was then used to identify over- and under-represented GO and Enzyme Commission (EC) numbers in the ES proteins by setting the significance filter p-value to ≤ 0.05. Also, we used the KEGG Automatic Annotation Server (KAAS) database (Moriya et al., 2007) to assign the pathway annotation to the secreted proteins using the BBH (bidirectional best hit) method and the gene data set assigned to *Mycobacterium*.

To determine differences between the predicted secretomes in relation to MAB subspecies and morphotype, we established core secretomes by performing a bidirectional best-hit BLASTP search (E-value 1.0E−3) between the ES proteins of all genomes belonging to the respective subspecies and morphotypes. Then, we identified the shared and unique proteins for each comparison. Additionally, we determined the ES proteins shared between the MAB reference strain ATCC 19977 and *M. tuberculosis* H37Rv predicted and experimental secretomes (Cornejo-Granados et al., 2017). The resulting shared proteins were further investigated for sequence similarities against known drug targets available on the Drug Bank database (http://www.drugbank.ca/), setting the E-value to 1.0E-3 and all other options to default. In Supplementary Table S2, we show all proteins that have similarity with an approved drug target, as well as the drugs that can affect said target.

Additionally, we analyzed the presence of the core secretomes in twenty *M. abscessus* genomes per subspecies downloaded from NCBI (Supplementary Table S3). To this end, each downloaded genome was analyzed with RAST to obtain all the open reading frames (ORFs). Next, we performed a BLASTP search (E-value 1.0E−3) of each core secretome against each genome of the corresponding subspecies, and all hit proteins were considered homologs.

Calculation of the Abundance of Antigenic Regions

The Abundance of Antigenic Regions (AAR) value is used to estimate the antigenic density of a protein by calculating the number of antigenic regions and normalizing it to the sequence length (Gomez et al., 2015). Of note, proteins with higher antigenic densities have lower AAR values. For this study, we calculated the AAR value for each protein in each data set using the Secret-AAR web-server (http://microbiomics.ibt.unam.mx/tools/aar/index.php) and reported the average unless stated otherwise (Cornejo-Granados et al., 2018). Then, we used a Mann-Whitney

statistical test to establish any significant differences between the AAR values of the different protein data sets.

Data availability

The reference genomes analyzed for *M. abscessus* ATCC19977 and *M. tuberculosis* H37Rv were taken from NCBI, under GenBank IDs CU458896.1 and NC_000962.3, respectively. The Whole Genome Shotgun project has been deposited at NCBI, under BioProject PRJNA646278. All the predicted secretomes were deposited in the Secret-AAR web-server (http://microbiomics.ibt.unam.mx/tools/aar/index.php).

Results

Genome assembly, secretome prediction and annotation

We sequenced the genomes of fifteen pulmonary and extrapulmonary (skin, tissue, and lymph node, blood) MAB isolates obtained from patients in Germany comprising all three MAB subspecies (Table 1 and S1). For each genome, we obtained an average of 2,601,444 quality-filtered reads. After *de novo* assembly, we obtained from 38 to 78 contigs (mean = 58 contigs) with genome coverage of 217- to 368-fold (mean = 310-fold) and with an average of 5,082 total proteins per genome (Supplementary Table S4).

We used a bioinformatics pipeline previously reported by our group (Cornejo-Granados et al., 2017) to predict the full secretome of all MAB clinical isolates and the widely used reference strain *M. abscessus* ATCC 19977 (GenBank CU458896.1) (Supplementary Fig. S1). We obtained an average of 939 ES proteins per genome, representing ∼18% of the total proteome (Table 1). The predicted secretome for the MAB reference strain consisted of 886 proteins. From these, all the proteins showed a BLASTP hit against the NR database, and only 494 (55.8%) could be annotated with GO terms.

We analyzed the over-representation of GO terms in the secretome of *M. abscessus* ATCC 19977 as compared to the whole genome. The most significantly enriched GO-terms were: “lytic vacuole” (p = 9.37E-04) and “fungal-type vacuole” (p = 0.004) in Cellular Component (Fig. 1A), “serine-type carboxypeptidase” (p = 1.83E-04), and “serine-type D-Ala-D-Ala carboxypeptidase” (p = 1.83E-04) activities in Molecular Function (Fig. 1B) and “response to inorganic substance” (p = 5.68E-04) and “cellular response to oxygen radical” (p = 0.001) in the Biological Process category (Fig. 1C). The KEGG pathway mapping of the ES proteins showed that 214 proteins (24.2 %) could be assigned to 100 different KEGG pathways (Table 2), with the ABC transporter pathway being the most abundant (n= 13, 1.47 %). Additionally, serine-type D-Ala-D-Ala carboxypeptidases (p = 1.83E-04) and peptidases (p = 8.40E-04) were the most significantly abundant enzymes according to the Enzyme Commission (EC) Classes (Figure S2), while the Mce/MiaD and PknH-like extracellular domains were the most enriched protein domains (Table 3). Of note, comparably few sequences were assigned to the PE/PPE category (n = 3). Notably, after comparing the predicted secretome of *M. abscessus* ATCC 19977 with a list of essential genes published by (Laencina et al., 2018), we found that 99 (11.17 %) of the predicted ES proteins, corresponded to essential proteins required for *in vitro* growth.

Comparison of* M. abscessus *subspecies core secretomes

We analyzed the differences between the predicted secretomes of the three MAB subspecies. To this end, we defined the core secretome of each subspecies as the set of proteins shared between all secretomes of isolates belonging to MAB_A, MAB_B, and MAB_M, respectively. The resulting core secretomes contained 735 (MAB_A), 794 (MAB_B), and 813 (MAB_M) proteins (Fig 2A).

Given that our study considered a limited number of de novo assembled genomes, we additionally compared the predicted core secretomes to sixty additional MAB genomes available in NCBI (Supplementary Table S3). We found that an average of 99.78%, 99.12%, and 98.59% of our core secretomes was also present in the investigated additional MAB_A, MAB_B, and MAB_M genomes, respectively, further corroborating the validity of the predicted subspecies core secretomes for other MAB isolates.

We then determined the respective AAR values to estimate antigenic densities for the protein sets in each core secretome. The average AAR values from most to least antigenic was: 40.24 for MAB_A, 40.75 for MAB_B, and 41.38 for MAB_M with no statistically significant difference between these values.

Next, we identified the ES proteins shared between the MAB_A, MAB_B, and MAB_M core secretomes. We found that 704 proteins (86.5 %) were shared among MAB_A, MAB_B, and MAB_M with an AAR value of 41.17 (Fig. 2B). The AAR values for the protein sets exclusively found in the MAB_A, MAB_B, or MAB_M secretome were 33.58, 41.22, and 43.13, respectively, with the MAB_A dataset showing a significantly lower AAR value indicating higher antigenicity than the others ($p < 0.1$; Fig. 2B).

Differences in core secretomes between R and S morphotypes

As MAB isolates with R and S morphotypes show differences in virulence and pathogenicity, we compared the predicted core secretomes of R and S isolates (Fig. 3). We observed that the core secretomes of R variants were larger (840, 924 and 845 proteins for MAB_A, MAB_M, and MAB_B) than those of the investigated S variants (764, 872 and 833 proteins, respectively) with no significant differences in antigenic densities as per mean AAR values (Fig. 3). Intra-subspecies comparison of S and R secretomes revealed that 96.4 %, 90.7% and 95% of the identified ES proteins were found in both R and S morphotypes for MAB_A, MAB_M and MAB_B respectively. The number of unique proteins was larger in the core secretome of the R morphotypes (n = 93, 109, and 48 for MAB_A, MAB_M, and MAB_B) as compared to the S morphotypes (n = 9, 76, and 35, respectively; Fig.3).
Interestingly, antigenic densities for the unique ES proteins of the R morphotypes were higher (AAR = 40.84, 36.71, and 35.59 for MAB_A, MAB_M, and MAB_B) than for the proteins exclusive to the S morphotypes irrespective of the subspecies (AAR = 45.43, 37.72, and 42.14; Fig. 3). To assess if the AAR values of these specific protein sets were different from same-sized protein sets randomly chosen from the respective core secretomes, we created 1000 random sets of 109, 93, 76, 48, 35 and 9 proteins and calculated the AAR value for each set. Then, we determined an empirical p-value based on the number of random protein sets that equaled or exceeded the AAR value for each protein dataset as was previously suggested by Cornejo-Granados et al., 2017. We found that the ES proteins exclusive to the R morphotypes of MAB_M and MAB_B had significantly ($p < 0.05$) higher antigenic densities than randomly constructed protein sets (Supplementary Table S5).

Finally, we determined the MAB core secretomes by sample origin (pulmonary, extrapulmonary, blood). This resulted in 706 ES proteins shared among the ten pulmonary isolates, 758 proteins shared among the four extrapulmonary isolates, and 885 proteins for the single isolate grown from a blood sample. However, as per the GO, KEGG, and antigenicity analyses, we did not find any distinct characteristics specific to either sample source and, hence, type of infection.

Antigenicity of ES and non-ES proteins

It has previously been reported for different microorganisms including MTB that ES proteins tend to be more antigenic than non-ES proteins (Cornejo-Granados et al., 2017; Gomez et al., 2015; Wang et al., 2015). We thus tested if this was also true for the investigated MAB isolates. First, we found that the antigenic densities as indicated by mean AAR values were very similar among all isolates irrespective of subspecies or morphotype within the same cell compartment, i.e. for ES, non-ES, intracellular (incell) and transmembrane (TM) proteins (Fig. 4). Second, we found that antigenic densities were significantly higher in ES proteins as compared to non-ES proteins in all isolates (AAR = 40.57 and 43.60, respectively; p-value < 0.0001) (Fig. 4). However, within the non-ES category, incell proteins showed even higher antigenic densities (AAR = 39.04) than the predicted ES proteins ($p < 0.0001$) while the lowest overall antigenic densities were observed for the TM category (AAR = 59.23; $p < 0.0001$).

Comparison of M. abscessus *and* M. tuberculosis *secretomes*

Lastly, we compared the predicted secretome of *M. abscessus* ATCC 19977 against the previously reported secretome of *M. tuberculosis* H37Rv (Cornejo-Granados et al., 2017). We observed that the *M. abscessus* secretome was predicted to be almost equally antigenic (AAR=40.78) than the *M. tuberculosis* secretome (AAR=40.63) (Fig. 5). We found 337 MAB ES proteins (38.04%) with homology to proteins in the predicted MTB secretome (Fig. 5). In addition, 222 of these proteins had sequence homology with proteins experimentally reported as secreted in MTB (comparable experimental secretome data for MAB was not available to us) (Cornejo-Granados et al., 2017) (Supplementary Table S6). Furthermore, we determined the average AAR value of the 680 ES proteins shared among the fifteen MAB isolates (AAR = 41.53). This value means that antigenic density was lower than for the predicted secretome of *M. tuberculosis* H37Rv (AAR = 40.63) and two clinical *M. tuberculosis* isolates belonging to the Beijing lineage (isolate C3 AAR = 37.51 and isolate C4 AAR = 37.54,) (Table 4) (Cornejo-Granados et al., 2017). Finally, we identified 17 ES proteins with homologues in both MAB and *M. tuberculosis*, which are listed as targets for various FDA approved drugs (Supplementary Table S2).

Discussion

This is the first study that proposes a method for prediction of MAB secretomes based on fifteen clinical MAB isolates and the *M. abscessus* ATCC 19977 reference strain. Our results show that an average of 18% (939 proteins) of the total proteins encoded in the MAB core genome carry sequence patterns indicative of secretion. Notably, this percentage

is 6% larger than the proportion previously reported for several MTB isolates (~12%) (Cornejo-Granados et al., 2017). Nearly 200 species of mycobacteria have been identified with diverse lifestyles and a high degree of morphological, biochemical, and physiological diversity and a comparative genome analysis suggests that only a relatively small number of genes (1080) are shared between several *Mycobacterium* species (Malhotra et al., 2017; Tortoli et al., 2017). Moreover, loss of ancestral genes is a well described phenomenon in slowly growing mycobacteria such as MTB and, in particular, *M. leprae* (Bachmann et al., 2019). In contrast, rapidly growing NTM such as MAB are considered to represent a more ancient evolutionary state, with larger genomes than those of MTB (Bachmann et al., 2019; Malhotra et al., 2017). Thus, it is not surprising that we found a larger number of ES proteins in MAB than MTB. Furthermore, the increased abundance of ES proteins in MAB as compared to MTB could be related to the ability of MAB to cause a different spectrum of disease and to adapt to different environmental settings requiring frequent interaction with a wide variety of host cells and organisms competing for the same ecological niche, likely involving cross species exchange of genetic information, for example by plasmid transfer (Ripoll et al., 2009; Ryan and Byrd, 2018; Waman et al., 2019). A similar hypothesis has been suggested for fungal secretomes (O'Toole et al., 2013).

The GO and KEGG pathway annotations of the secretomes of *M. abscessus* ATCC 19977 and the MAB clinical isolates showed enrichment consistent with the characterization of previously reported mycobacterial secretomes (Cornejo-Granados et al., 2017; Gomez et al., 2015). Interestingly and in line with the increased secretome size as compared to MTB, the KEGG pathway analysis showed a high abundance of the Quorum sensing pathway for the predicted MAB secretomes, which was not present in our previous MTB secretome pathway analysis (Cornejo-Granados et al., 2017). The presence of a Quorum sensing pathway would be another similarity shared between MAB and non-mycobacterial pathogens commonly affecting patients with chronic lung disease such as *Pseudomonas aeruginosa* (Mukherjee and Bassler, 2019). In addition, it could be related to the ability of MAB to form biofilms (Clary et al., 2018; Orme and Ordway, 2014), further contributing to the capacity of MAB to tolerate antibiotics and to persist over long periods in the environment (Faria et al., 2015; Hunt-Serracin et al., 2019; Kulka et al., 2012; Maurer et al., 2014a).

The InterPro annotation showed that Mce domains were the most abundant (2.14 %) domains in the MAB reference secretome, while PPE and PE-PGRS domains only corresponded to 0.3 % of the ES protein sequences. This tendency is contrary to our observations for MTB (Cornejo-Granados et al., 2017), where the PPE and PE-PGRS domains accounted for ~12% of the secreted proteins and the Mce domains for only 0.5%. The lower quantity of predicted PE/PPE proteins in MAB was somewhat expected. *M. tuberculosis* has five ESX secretion systems, four of which encode PE/PPE proteins, while MAB has only two (ESX-3 and ESX-4) of which only the ESX-3 operon includes PE/PPE genes (Dumas et al., 2016). In contrast, Mce domains are known for participating in host cell entry by mycobacteria (Kumar et al., 2005). Thus, their higher abundance in MAB as compared to MTB highlights the importance of this pathway for MAB survival within the host. It needs to be mentioned though that Kumar et al., 2005, also suggested that in low virulence bacteria, transport activities could be the primary function of Mce operons (Kumar et al., 2005).

To compare the predicted secretomes according to colony morphotype, we first established the core secretome for the R and S variants per subspecies, thus eliminating individualities among the different isolates (Fig. 3). The high overall agreement between the core secretomes for both morphotypes of approximately 90 % was expected, considering the fact that R variants can arise from the S morphotypes during persistent infection by loss of surface-exposed GPLs caused by mutations in the GPL synthesis pathway (Bernut et al., 2014; Catherinot et al., 2007; Roux et al., 2016; Ryan and Byrd, 2018). However, both the higher number and the higher antigenic densities (lower AAR values) of the ES proteins exclusively found in R variants indicate that additional genetic changes may evolve during S-to-R conversion. Moreover, this observation raises the question whether some strains with additional genetic traits associated with virulence are able to undergo S-to-R conversion and cause disease due to R variants more easily than others. Genomic studies involving sequentially isolated S and R variants of the same strain obtained from individual patients over time will be required to better characterize the microevolution of MAB strains within the chronically infected host.

Similarly, the fact that MAB causes both chronic pulmonary disease (with R variants sometimes increasing over time) and extrapulmonary manifestations (mostly caused by S variants) led us to investigate whether differences exist in the predicted secretomes of isolates related to these clinical presentations. The absence of major differences in the GO, KEGG, and antigenicity analyses suggest that secretome variations do not influence MAB tissue tropism. Consequently, host characteristics such as severe immunosuppression may be the main driver for invasive MAB infections. Likewise, in the case of tissue infections, which often occur following surgical interventions, insufficient hygiene procedures and sterilization protocols for surgical equipment appear to be more relevant than pathobiological traits such as the secretome intrinsic to the causative MAB isolate (Maurer et al., 2014b)

Lastly, we observed that the predicted secretomes of all investigated clinical MAB isolates were less antigenic than the secretomes of *M. tuberculosis* H37Rv and two clinical *M. tuberculosis* isolates. Additionally, although there was no statistical difference, the isolates with a rough phenotype tended to be more antigenic that the isolates with smooth phenotype. Previous evidence with *M. tuberculosis* (Cornejo-Granados et al., 2017) showed that clinical isolates from the Beijing phenotype showed increased virulence and less antigenic secretomes than the reference strain H37Rv. Thus, the diminished antigenicity of MAB could be viewed as a virulence trait in itself as it would support colonization of the host for extended time periods without immediate progression into clinical disease. However, further experimental tests on antigenicity are needed to demonstrate this observation.

This study represents the first systematic prediction and *in silico* characterization of the MAB secretome. We acknowledge that an important constraint of this study is the limited total number of genomes analyzed per subspecies and biological source. Thus, care must be taken to not overinterpret the findings related to sample subcategories such as subspecies and morphotypes. Also, published experimental data on MAB secretomes are very limited and no systematic validation of the *in silico* findings reported herein could be performed against such datasets. Although, more research will be needed to determine experimental secretomes in NTM, our study demonstrates that using bioinformatics strategies can help to broadly explore mycobacterial secretomes including those of clinical isolates and to tailor subsequent, complex and time-consuming experimental approaches

accordingly. This approach can support a systematic investigation of mycobacterial secretomes exploring candidate proteins suitable for developing new vaccines and diagnostic markers to distinguish between colonization and infection.

Tables

Table 1| Clinical isolates metadata and number of ES proteins.

Strain	Genome ID	Origin	Phenotype	Total predicted proteins	ES proteins	% ES proteins
M. abscessus subsp. abscessus	4549-15	sputum	rough	5,105	929	18
	11351-15	sputum	rough	5,138	966	19
	8844-15	skin	smooth	4,854	956	20
	3563-15	sputum	smooth	5,239	968	18
	12389-15	sputum	smooth	5,276	990	19
	2677-16	sputum	smooth	4,900	919	19
	2572-17	tissue (breast implant)	NA	4,847	874	18
M. abscessus subsp. massiliense	14479-15	sputum	rough	5,120	962	19
	10896-16	sputum	rough	5,109	950	19
	10003-15	sputum	smooth	4,835	891	18
	16155-15	sputum	smooth	4,884	898	18
M. abscessus subsp. bolletii	11702-16	sputum	rough	5,079	931	18
	713-16	lymph node	rough	5,456	1,037	19
	7742-15	blood culture	smooth	4,913	885	18
	13116-16	lymph node	smooth	5,305	990	19
M. abscessus subsp. abscessus	reference strain ATCC19977 (GenBank CU458896.1)	-	-	4,942	886	18
M. tuberculosis H37Rv	reference strain (GenBank AL123456.3)	-	-	4,337	548	13

Table 2. | Top 10 KEGG pathways assigned for *M. abscessus* ATCC19977 ES proteins.

Ranking	Pathway name	Number of represented ES proteins (%)
1	ABC transporters	13 (1.47)
2	Two-component system	9 (1.02)
3	Quorum sensing	6 (0.68)
4	Oxidative phosphorylation	4 (0.45)
5	Sulfur metabolism	4 (0.45)
6	Glycerolipid metabolism	4 (0.45)
7	Peptidoglycan biosynthesis	4 (0.45)
8	Protein export	4 (0.45)
9	Starch and sucrose metabolism	3 (0.34)
10	Glyoxylate and dicarboxylate metabolism	3 (0.34)

Table 3| Top 10 most represented protein domains in *M. abscessus* ATCC19977 secretome.

InterProcode	InterPro description	Number of ES proteins (%)
IPR003399	Mce/MlaD	19 (2.14)
IPR026954	PknH-like extracellular domain	15 (1.69)
IPR032407	Haemophore, haem-binding	10 (1.13)
IPR020846	Major facilitator superfamily domain	7 (0.79)
IPR013766	Thioredoxin domain	6 (0.68)
IPR000064	Endopeptidase, NLPC/P60 domain	6 (0.68)
IPR001638	Solute-binding protein family 3/N-terminal domain of MltF	6 (0.68)
IPR000675	Cutinase/acetylxylan esterase	6 (0.68)
IPR005490	L,D-transpeptidase catalytic domain	5 (0.56)
IPR000073	Alpha/beta hydrolase fold-1	5 (0.56)

Table 4 | Abundance of Antigenic Regions (AAR) for *M. abscessus* and *M. tuberculosis* strains.

Strain	Number of proteins in the dataset	Average AAR value	
M. tuberculosis Beijing isolate C3*	553	37.52	
M. tuberculosis Beijing isolate C4*	519	37.55	
M. bovis BCG Pasteur	564	38.99	
***M. abscessus* ATCC 19977**	**886**	**40.78**	
M. tuberculosis H37Rv	548	40.63	
M. abscessus clinical isolates	680	41.54	

* both Beijing isolates were previously reported in Cornejo-Granados et al., 2017.

Figure Legends

Figure 1. GO enrichment analysis for the *M. abscessus* ATCC 19977 reference strain. Top 10 most enriched GO terms for the *M. abscessus* ATCC 19977 secretome (blue) and complete genome (red) in three categories: A) Cellular Component, B) Molecular Function and C) Biological Process.

Figure 2. Venn diagram between the core secretomes of the three *M. abscessus* *subspecies*. A) Number of total proteins contained in the core secretome of each subspecies. B) Shared and unique proteins between the three subspecies as per BLASTP (E-value 1.0E-3).

Figure 3. Venn diagram between the core secretomes of the three *M. abscessus* subspecies by colony morphotype. We used BLASTP (E-value 1.0E-3) to assess the core secretomes for isolates with rough and smooth colony morphotypes A) *M. abscessus* subsp. *abscessus*, B) *M. abscessus* subsp. *massiliense* and C) *M. abscessus* subsp. *bolletii*.

Figure 4. Comparison between AAR values for Excreted/Secreted (ES), non Excreted/Secreted (non-ES), intracellular (incell) and transmembrane (TM) proteins. AAR values were calculated for each of the 15 genomes sequenced. The X-axis shows the cellular compartment and the Y-axis shows AAR values for the genomes of each subspecies: *M. abscessus* subsp. *abscessus* (green), *M. abscessus* subsp. *bolletii* (blue), *M. abscessus* subsp. *massiliense* (purple), *M. abscessus* ATCC19977 (red) and *M. tuberculosis* H37Rv (orange). Mann-Whitney tests were performed to compare the AAR of each group with a confidence level of 99% (***, $p < 0.001$).

Figure 5. Venn diagram between the predicted secretomes of *M. tuberculosis* H37Rv and *M. abscessus* ATCC 19977. We used BLASTP (E-value 1.0E-3) to compare the complete secretomes of both species.

Acknowledgements

We would like to thank Julia Zallet and Vanessa Mohr for excellent technical assistance and Henrik Nielsen for his assistance with the SecrtomeP software. Parts of this work have been supported by the German Center for Infection Research and Grants by Joachim Herz Foundation, Hamburg, and Mukoviszidose Institut gGmbH, Bonn, the research and development arm of the German Cystic Fibrosis Association Mukoviszidose e.V. to F.P.M. We acknowledge the support provided by CONACyT grant CB- 2013-223279 and SALUD-2014-C01-234188 to A.O.L. This research also received support by the DGAPA PAPIIT UNAM (IN215520) to A.O.L. F.C.G. acknowledges the support of CONACyT as a Postgraduate fellow.

References

Abeles, S. R., and Pride, D. T. (2014). Molecular Bases and Role of Viruses in the Human Microbiome. *Journal of Molecular Biology* 426, 3892–3906. doi:10.1016/j.jmb.2014.07.002.

Adekambi, T., Colson, P., and Drancourt, M. (2003). rpoB-based identification of nonpigmented and late-pigmenting rapidly growing mycobacteria. *Journal of Clinical Microbiology* 41, 5699–5708. doi:10.1128/jcm.41.12.5699-5708.2003.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9, 75. doi:10.1186/1471-2164-9-75.

Bachmann, N. L., Salamzade, R., Manson, A. L., Whittington, R., Sintchenko, V., Earl, A. M., et al. (2019). Key Transitions in the Evolution of Rapid and Slow Growing Mycobacteria Identified by Comparative Genomics. *Front. Microbiol.* 10, 3019. doi:10.3389/fmicb.2019.03019.

Bendtsen, J. D., Kiemer, L., Fausbøll, A., and Brunak, S. (2005a). Non-classical protein secretion in bacteria. *BMC Microbiol* 5, 58. doi:10.1186/1471-2180-5-58.

Bendtsen, J. D., Nielsen, H., Widdick, D., Palmer, T., and Brunak, S. (2005b). Prediction of twin-arginine signal peptides. *BMC Bioinformatics* 6, 167. doi:10.1186/1471-2105-6-167.

Bernut, A., Herrmann, J.-L., Kissa, K., Dubremetz, J.-F., Gaillard, J.-L., Lutfalla, G., et al. (2014). Mycobacterium abscessus cording prevents phagocytosis and promotes abscess formation. *Proc. Natl. Acad. Sci. U.S.A.* 111, E943–52. doi:10.1073/pnas.1321390111.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.

607 doi:10.1093/bioinformatics/btu170.

608 Catherinot, E., Clarissou, J., Etienne, G., Ripoll, F., Emile, J. F., Daffé, M., et al. (2007).
609 Hypervirulence of a rough variant of the Mycobacterium abscessus type strain.
610 *Infection and Immunity* 75, 1055–1058. doi:10.1128/IAI.00835-06.

611 Chen, J., Zhao, L., Mao, Y., Ye, M., Guo, Q., Zhang, Y., et al. (2019). Clinical Efficacy
612 and Adverse Effects of Antibiotics Used to Treat Mycobacterium abscessus Pulmonary
613 Disease. *Front. Microbiol.* 10, 1977. doi:10.3389/fmicb.2019.01977.

614 Clary, G., Sasindran, S. J., Nesbitt, N., Mason, L., Cole, S., Azad, A., et al. (2018).
615 Mycobacterium abscessus Smooth and Rough Morphotypes Form Antimicrobial-
616 Tolerant Biofilm Phenotypes but Are Killed by Acetic Acid. *Antimicrobial Agents and*
617 *Chemotherapy* 62, 117. doi:10.1128/AAC.01782-17.

618 Conesa, A., and Götz, S. (2008). Blast2GO: A comprehensive suite for functional analysis
619 in plant genomics. *Int J Plant Genomics* 2008, 619832. doi:10.1155/2008/619832.

620 Cornejo-Granados, F., Hurtado-Ramírez, J. M., Hernandez-Pando, R., and Ochoa-Leyva, A.
621 (2018). Secret-AAR: a web server to assess the antigenic density of proteins and
622 homology search against bacterial and parasite secretome proteins. *Genomics*.
623 doi:10.1016/j.ygeno.2018.10.007.

624 Cornejo-Granados, F., Zatarain-Barrón, Z. L., Cantu-Robles, V. A., Mendoza-Vargas, A.,
625 Molina-Romero, C., Sánchez, F., et al. (2017). Secretome Prediction of Two M.
626 tuberculosis Clinical Isolates Reveals Their High Antigenic Density and Potential Drug
627 Targets. *Front. Microbiol.* 8, 128. doi:10.3389/fmicb.2017.00128.

628 Daugelat, S., Guile, H., Schoel, B., and Kaufmann, S. H. E. (1992). Secreted Antigens of
629 Mycobacterium tuberculosis: Characterization with T Lymphocytes from Patients and
630 Contacts after Two-Dimensional Separation. *Journal of Infectious Diseases* 166, 186–
631 190. doi:10.1093/infdis/166.1.186.

632 De Almeida, I. N., Da Silva Carvalho, W., Rossetti, M. L., Costa, E. R. D., and De Miranda,
633 S. S. (2013). Evaluation of six different DNA extraction methods for detection of
634 Mycobacterium tuberculosis by means of PCR-IS6110: preliminary study. *BMC*
635 *Research Notes* 6, 561–6. doi:10.1186/1756-0500-6-561.

636 Dumas, E., Christina Boritsch, E., Vandenbogaert, M., Rodríguez de la Vega, R. C.,
637 Thiberge, J.-M., Caro, V., et al. (2016). Mycobacterial Pan-Genome Analysis Suggests
638 Important Role of Plasmids in the Radiation of Type VII Secretion Systems. *Genome*
639 *Biol Evol* 8, 387–402. doi:10.1093/gbe/evw001.

640 Faria, S., Joao, I., and Jordao, L. (2015). General Overview on Nontuberculous
641 Mycobacteria, Biofilms, and Human Infection. *J Pathog* 2015, 809014–10.
642 doi:10.1155/2015/809014.

643 Gomez, S., Adalid-Peralta, L., Palafox-Fonseca, H., Cantu-Robles, V. A., Soberón, X.,

644 Sciutto, E., et al. (2015). Genome analysis of Excretory/Secretory proteins in Taenia
645 solium reveals their Abundance of Antigenic Regions (AAR). *Sci. Rep.* 5, 9683.
646 doi:10.1038/srep09683.

647 Gupta, M. K., Subramanian, V., and Yadav, J. S. (2009). Immunoproteomic identification
648 of secretory and subcellular protein antigens and functional evaluation of the secretome
649 fraction of Mycobacterium immunogenum, a newly recognized species of the
650 Mycobacterium chelonae-Mycobacterium abscessus group. *J. Proteome Res.* 8, 2319–
651 2330. doi:10.1021/pr8009462.

652 Gutiérrez, A. V., Viljoen, A., Ghigo, E., Herrmann, J.-L., and Kremer, L. (2018).
653 Glycopeptidolipids, a Double-Edged Sword of the Mycobacterium abscessus Complex.
654 *Front. Microbiol.* 9, 1145. doi:10.3389/fmicb.2018.01145.

655 Howard, S. T., Rhoades, E., Recht, J., Pang, X., Alsup, A., Kolter, R., et al. (2006).
656 Spontaneous reversion of Mycobacterium abscessus from a smooth to a rough
657 morphotype is associated with reduced expression of glycopeptidolipid and
658 reacquisition of an invasive phenotype. *Microbiology* 152, 1581–1590.
659 doi:10.1099/mic.0.28625-0.

660 Hunt-Serracin, A. C., Parks, B. J., Boll, J., and Boutte, C. C. (2019). Mycobacterium
661 abscessus Cells Have Altered Antibiotic Tolerance and Surface Glycolipids in
662 Artificial Cystic Fibrosis Sputum Medium. *Antimicrobial Agents and Chemotherapy* 63,
663 1370. doi:10.1128/AAC.02488-18.

664 Jarand, J., Levin, A., Zhang, L., Huitt, G., Mitchell, J. D., and Daley, C. L. (2011). Clinical
665 and microbiologic outcomes in patients receiving treatment for Mycobacterium
666 abscessus pulmonary disease. *Clin. Infect. Dis.* 52, 565–571. doi:10.1093/cid/ciq237.

667 Juncker, A. S., Willenbrock, H., Heijne, von, G., Brunak, S., Nielsen, H., and Krogh, A.
668 (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci.*
669 12, 1652–1662. doi:10.1110/ps.0303703.

670 Käll, L., Krogh, A., and Sonnhammer, E. L. L. (2007). Advantages of combined
671 transmembrane topology and signal peptide prediction--the Phobius web server.
672 *Nucleic Acids Research* 35, W429–32. doi:10.1093/nar/gkm256.

673 Kulka, K., Hatfull, G., and Ojha, A. K. (2012). Growth of Mycobacterium tuberculosis
674 biofilms. *JoVE*. doi:10.3791/3820.

675 Kumar, A., Chandolia, A., Chaudhry, U., Brahmachari, V., and Bose, M. (2005).
676 Comparison of mammalian cell entry operons of mycobacteria: in silico analysis and
677 expression profiling. *FEMS Immunol. Med. Microbiol.* 43, 185–195.
678 doi:10.1016/j.femsim.2004.08.013.

679 Laencina, L., Dubois, V., Le Moigne, V., Viljoen, A., Majlessi, L., Pritchard, J., et al.
680 (2018). Identification of genes required for Mycobacterium abscessus growth in vivo
681 with a prominent role of the ESX-4 locus. *Proc. Natl. Acad. Sci. U.S.A.* 115, E1002–

682 E1011. doi:10.1073/pnas.1713195115.

683 Lee, M.-R., Sheng, W.-H., Hung, C.-C., Yu, C.-J., Lee, L.-N., and Hsueh, P.-R. (2015).
684 Mycobacterium abscessus Complex Infections in Humans. *Emerg. Infect. Dis.* 21,
685 1638–1646. doi:10.3201/2109.141634.

686 Malhotra, S., Vedithi, S. C., and Blundell, T. L. (2017). Decoding the similarities and
687 differences among mycobacterial species. *PLoS Negl Trop Dis* 11, e0005883.
688 doi:10.1371/journal.pntd.0005883.

689 Maurer, F. P., Bruderer, V. L., Ritter, C., Castelberg, C., Bloemberg, G. V., and Böttger, E.
690 C. (2014a). Lack of antimicrobial bactericidal activity in Mycobacterium abscessus.
691 *Antimicrobial Agents and Chemotherapy* 58, 3828–3836. doi:10.1128/AAC.02448-14.

692 Maurer, F., Castelberg, C., Braun, von, A., Wolfensberger, A., Bloemberg, G., Bottger, E.,
693 et al. (2014b). Postsurgical wound infections due to rapidly growing mycobacteria in
694 Swiss medical tourists following cosmetic surgery in Latin America between 2012 and
695 2014. *Euro Surveill.* 19, 20905. doi:10.2807/1560-7917.es2014.19.37.20905.

696 Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., and Kanehisa, M. (2007). KAAS: an
697 automatic genome annotation and pathway reconstruction server. *Nucleic Acids*
698 *Research* 35, W182–5. doi:10.1093/nar/gkm321.

699 Mukherjee, S., and Bassler, B. L. (2019). Bacterial quorum sensing in complex and
700 dynamically changing environments. *Nature Publishing Group* 17, 371–382.
701 doi:10.1038/s41579-019-0186-5.

702 Nessar, R., Cambau, E., Reyrat, J.-M., Murray, A., and Gicquel, B. (2012). Mycobacterium
703 abscessus: a new antibiotic nightmare. *J. Antimicrob. Chemother.* 67, 810–818.
704 doi:10.1093/jac/dkr578.

705 Nurk, S., Bankevich, A., Antipov, D., Gurevich, A. A., Korobeynikov, A., Lapidus, A., et
706 al. (2013). Assembling single-cell genomes and mini-metagenomes from chimeric
707 MDA products. *J. Comput. Biol.* 20, 714–737. doi:10.1089/cmb.2013.0084.

708 O'Toole, N., Min, X. J., Butler, G., Storms, R., and Tsang, A. (2013). *Sequence-Based*
709 *Analysis of Fungal Secretomes*. Elsevier B.V doi:10.1016/S1874-5334(06)80015-8.

710 Orme, I. M., and Ordway, D. J. (2014). Host response to nontuberculous mycobacterial
711 infections of current clinical importance. *Infection and Immunity* 82, 3516–3522.
712 doi:10.1128/IAI.01606-13.

713 Petersen, T. N., Brunak, S., Heijne, von, G., and Nielsen, H. (2011). SignalP 4.0:
714 discriminating signal peptides from transmembrane regions. *Nature Publishing Group*
715 8, 785–786. doi:10.1038/nmeth.1701.

716 Ripoll, F., Deshayes, C., Pasek, S., Laval, F., Beretti, J.-L., Biet, F., et al. (2007). Genomics
717 of glycopeptidolipid biosynthesis in Mycobacterium abscessus and M. chelonae. *BMC*

718 *Genomics* 8, 114–9. doi:10.1186/1471-2164-8-114.

719 Ripoll, F., Pasek, S., Schenowitz, C., Dossat, C., Barbe, V., Rottman, M., et al. (2009). Non
720 mycobacterial virulence genes in the genome of the emerging pathogen Mycobacterium
721 abscessus. *PLoS ONE* 4, e5660. doi:10.1371/journal.pone.0005660.

722 Roux, A.-L., Viljoen, A., Bah, A., Simeone, R., Bernut, A., Laencina, L., et al. (2016). The
723 distinct fate of smooth and rough Mycobacterium abscessus variants inside
724 macrophages. *Open Biol* 6, 160185. doi:10.1098/rsob.160185.

725 Ryan, K., and Byrd, T. F. (2018). Mycobacterium abscessus: Shapeshifter of the
726 Mycobacterial World. *Front. Microbiol.* 9, 2642. doi:10.3389/fmicb.2018.02642.

727 Sanguinetti, M., Ardito, F., Fiscarelli, E., La Sorda, M., D'Argenio, P., Ricciotti, G., et al.
728 (2001). Fatal pulmonary infection due to multidrug-resistant Mycobacterium abscessus
729 in a patient with cystic fibrosis. *Journal of Clinical Microbiology* 39, 816–819.
730 doi:10.1128/JCM.39.2.816-819.2001.

731 Shin, A.-R., Sohn, H., Won, C. J., Lee, B., Kim, W. S., Kang, H. B., et al. (2010).
732 Characterization and identification of distinct Mycobacterium massiliense extracellular
733 proteins from those of Mycobacterium abscessus. *J. Microbiol.* 48, 502–511.
734 doi:10.1007/s12275-010-0038-5.

735 Sonnhammer, E. L., Heijne, von, G., and Krogh, A. (1998). A hidden Markov model for
736 predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol*
737 *Biol* 6, 175–182.

738 Tjalsma, H., Antelmann, H., Jongbloed, J. D. H., Braun, P. G., Darmon, E., Dorenbos, R.,
739 et al. (2004). Proteomics of protein secretion by Bacillus subtilis: separating the
740 "secrets" of the secretome. *Microbiol. Mol. Biol. Rev.* 68, 207–233.
741 doi:10.1128/MMBR.68.2.207-233.2004.

742 Tortoli, E., Fedrizzi, T., Meehan, C. J., Trovato, A., Grottola, A., Giacobazzi, E., et al.
743 (2017). The new phylogeny of the genus Mycobacterium: The old and the news. *Infect.*
744 *Genet. Evol.* 56, 19–25. doi:10.1016/j.meegid.2017.10.013.

745 Tortoli, E., Kohl, T. A., Brown-Elliott, B. A., Trovato, A., Leão, S. C., Garcia, M. J., et al.
746 (2016). Emended description of Mycobacterium abscessus, Mycobacterium abscessus
747 subsp. abscessus and Mycobacteriumabscessus subsp. bolletii and designation of
748 Mycobacteriumabscessus subsp. massiliense comb. nov. *Int. J. Syst. Evol. Microbiol.*
749 66, 4471–4479. doi:10.1099/ijsem.0.001376.

750 Waman, V. P., Vedithi, S. C., Thomas, S. E., Bannerman, B. P., Munir, A., Skwark, M. J.,
751 et al. (2019). Mycobacterial genomics and structural bioinformatics: opportunities and
752 challenges in drug discovery. *Emerg Microbes Infect* 8, 109–118.
753 doi:10.1080/22221751.2018.1561158.

754 Wang, S., Wei, W., and Cai, X. (2015). Genome-wide analysis of excretory/secretory

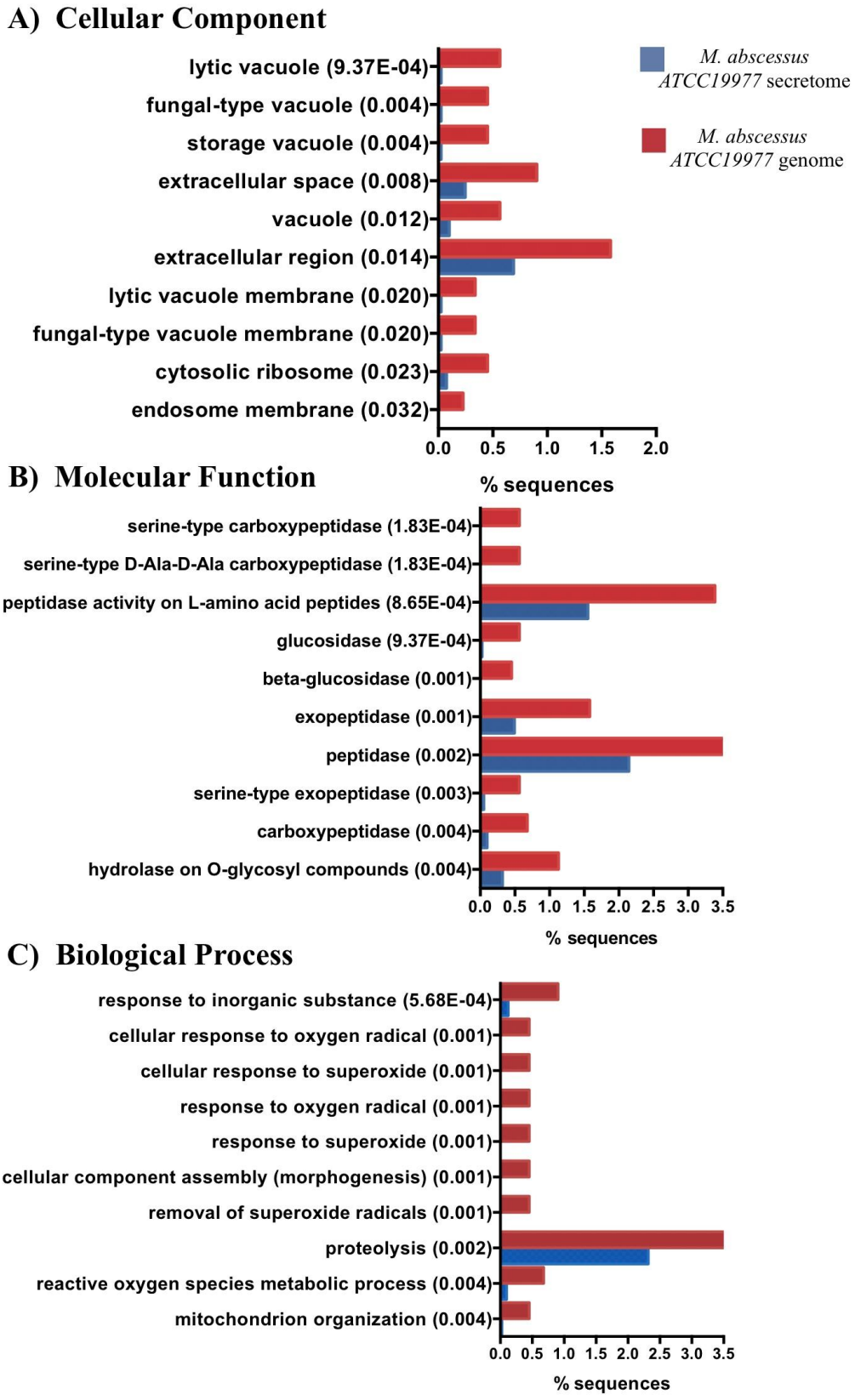
proteins in Echinococcus multilocularis: insights into functional characteristics of the tapeworm secretome. *Parasites & Vectors* 8, 666. doi:10.1186/s13071-015-1282-7.

Yadav, J. S., and Gupta, M. (2012). Secretome differences between the taxonomically related but clinically differing mycobacterial species Mycobacterium abscessus and M. chelonae. *JIOMICS* 2, 1–16. doi:10.5584/jiomics.v2i2.98.

Zdobnov, E. M., and Apweiler, R. (2001). InterProScan - an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17, 847–848. doi:10.1093/bioinformatics/17.9.847.

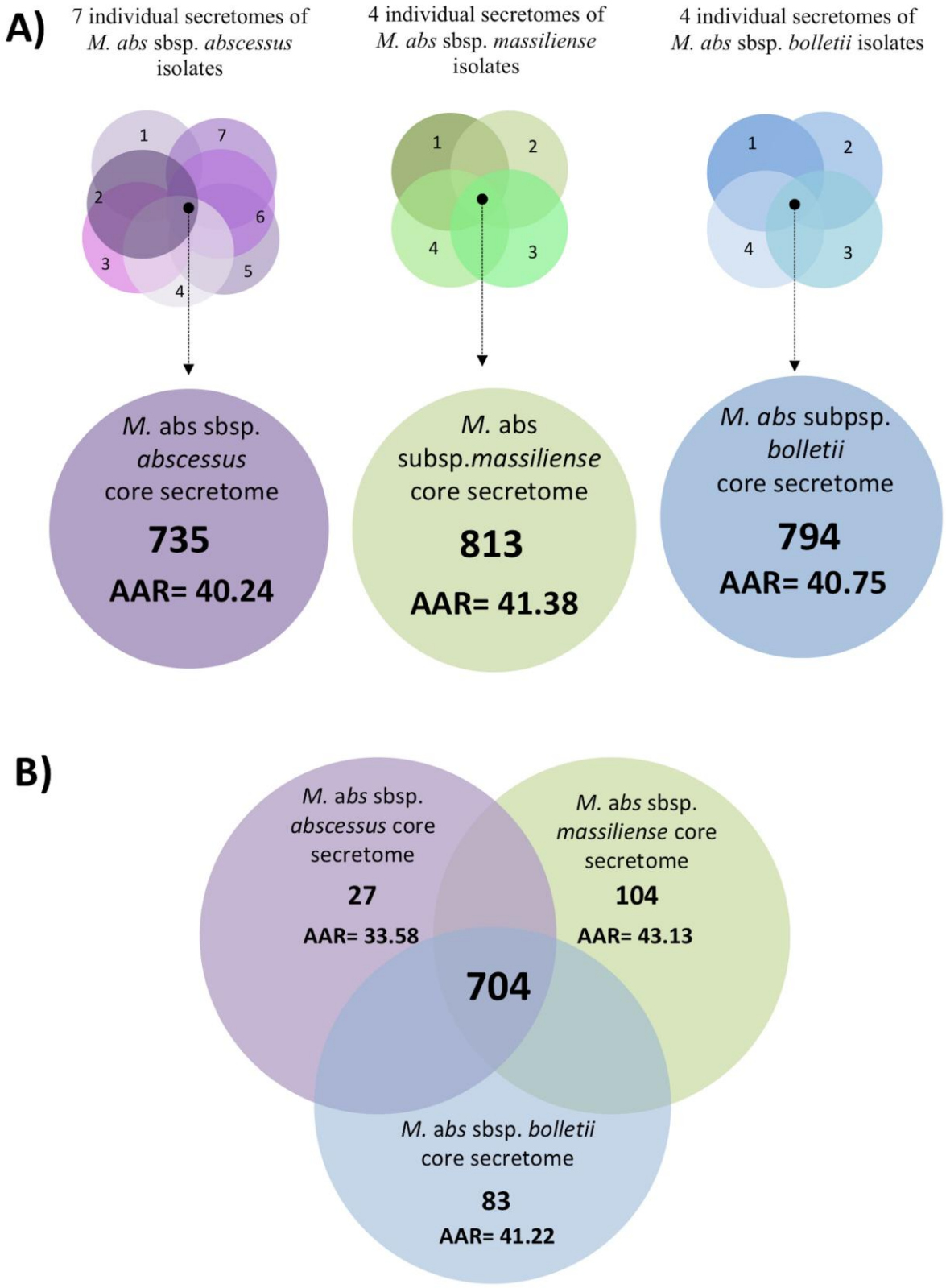
Zheng, J., Ren, X., Wei, C., Yang, J., Hu, Y., Liu, L., et al. (2013). Analysis of the secretome and identification of novel constituents from culture filtrate of bacillus Calmette-Guerin using high-resolution mass spectrometry. *Mol. Cell Proteomics* 12, 2081–2095. doi:10.1074/mcp.M113.027318.

784 **Figure 1.**



785

786 **Figure 2.**



787

788 **Figure 3.**

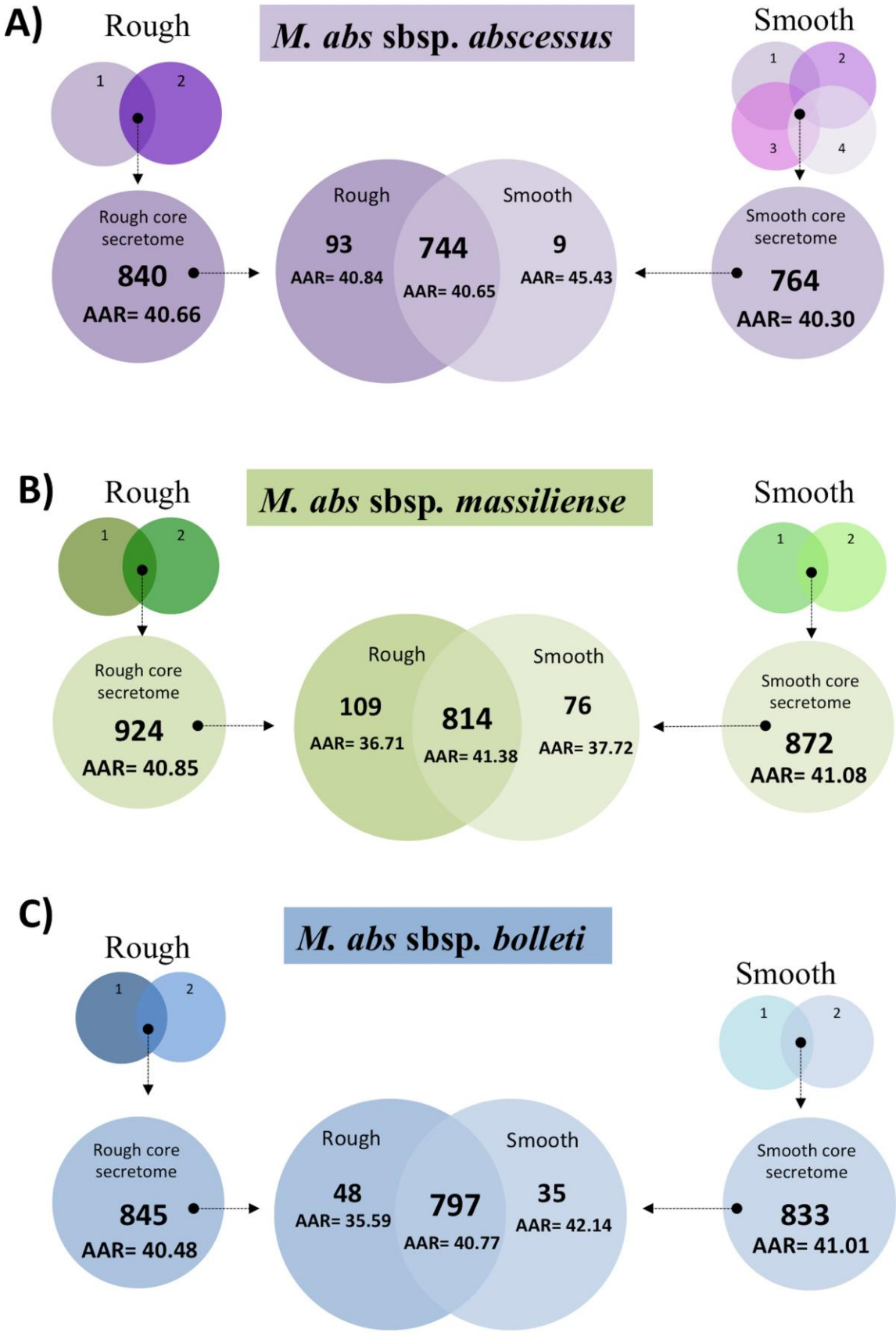
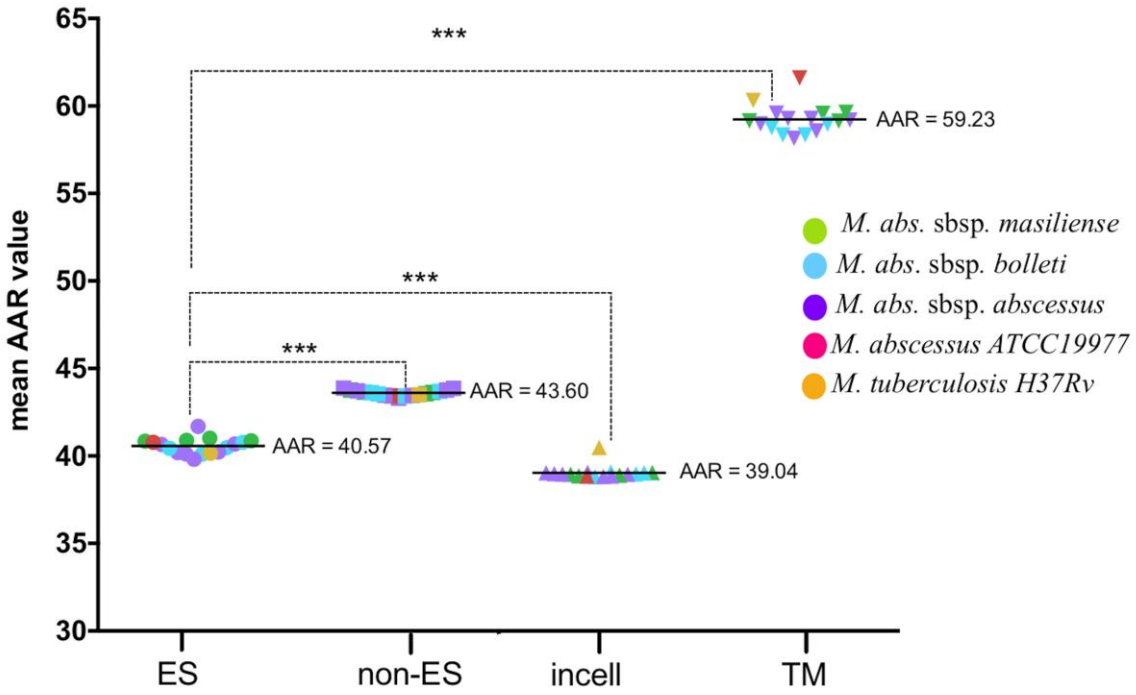
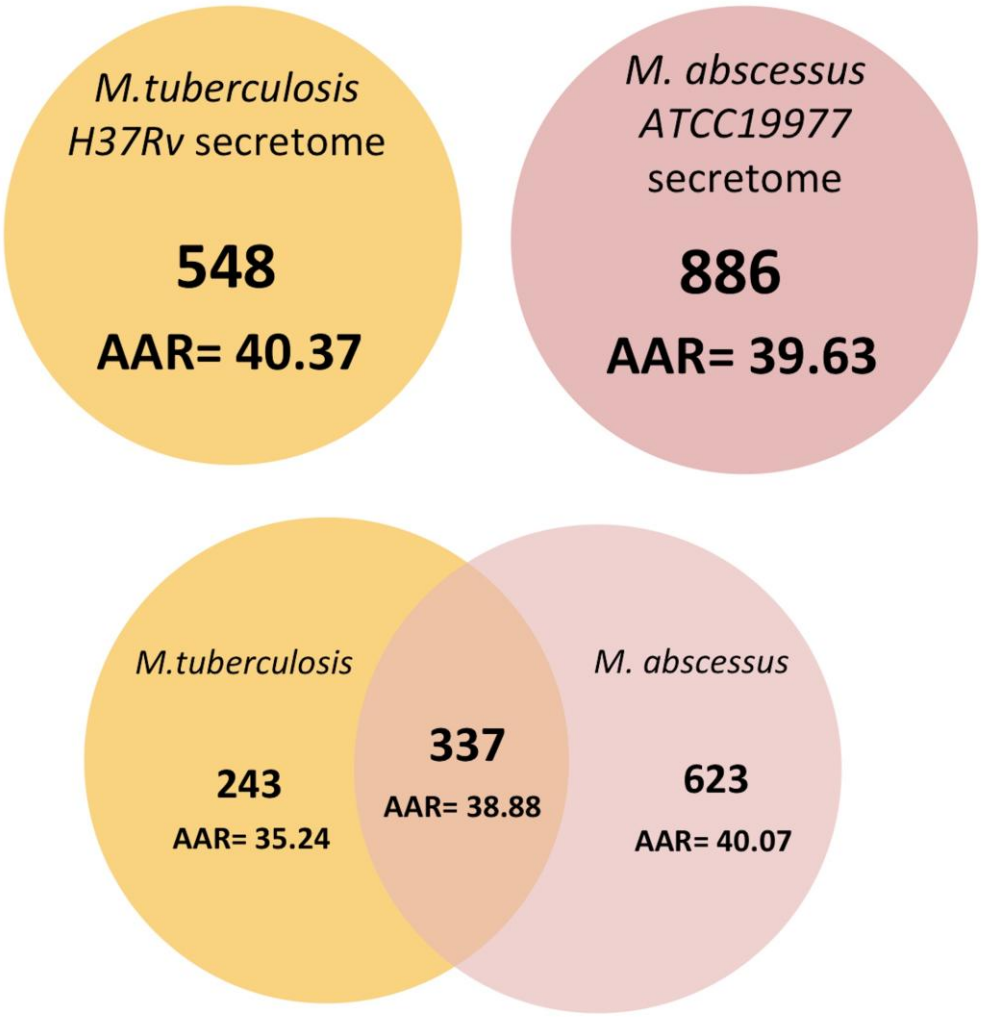


Figure 4.



805 **Figure 5.**



806