Journal of Proteomics 289 (2023) 105012

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Beyond the identifiable proteome: Delving into the proteomics of polymyxin-resistant and non-resistant *Acinetobacter baumannii* from Brazilian hospitals

Amanda Dal Lin [a,b,1], Juliana de S. da G. Fischer [a,1], Marlon D.M. Santos [a,c], Amanda Caroline Camillo-Andrade [a,c], Louise Ulrich Kurt [a], Tatiana A.C.B. Souza [a], Ana Beatriz Lyrio Lajas [a], Bernardina Rivera [c], Magdalena Portela [c], Rosario Duran [c], Marcelo Távora Mira [b], Marcelo Pillonetto [b,d,*], Paulo Costa Carvalho [a,**]

[a] *Laboratory for Structural and Computational Proteomics, Carlos Chagas Institute, Fiocruz – Paraná, Brazil*
[b] *Laboratório Experimental Multiuso, Pontifícia Universidade Católica do Paraná, Brazil*
[c] *Analytical Biochemistry and Proteomics Unit, Institut Pasteur de Montevideo/IIBCE, Montevideo, Uruguay*
[d] *Laboratório Central do Estado do Paraná, Brazil*

ARTICLE INFO

Keywords:
DiagnoMass
Polymyxin
Acinetobacter baumannii
Resistant bacteria
ARM

ABSTRACT

This work discloses a unique, comprehensive proteomic dataset of *Acinetobacter baumannii* strains, both resistant and non-resistant to polymyxin B, isolated in Brazil generated using Orbitrap Fusion Lumos. From nearly 4 million tandem mass spectra, the software DiagnoMass produced 240,685 quality-filtered mass spectral clusters, of which PatternLab for proteomics identified 44,553 peptides mapping to 3479 proteins. Crucially, DiagnoMass shortlisted 3550 and 1408 unique mass spectral clusters for the resistant and non-resistant strains, respectively, with only about a third with sequences (and PTMs) identified by PatternLab. Further open-search attempts via FragPipe yielded an additional ~20% identifications, suggesting the remaining unidentified spectra likely arise from complex combinations of post-translational modifications and amino-acid substitutions. This highlights the untapped potential of the dataset for future discoveries, particularly given the importance of PTMs, which remain elusive to nucleotide sequencing approaches but are crucial for understanding biological mechanisms. Our innovative approach extends beyond the identifications that are typically subjected to the bias of a search engine; we discern which spectral clusters are differential and subject them to increased scrutiny, akin to spectral library matching by comparing captured spectra to themselves. Our analysis reveals adaptations in the resistant strain, including enhanced detoxification, altered protein synthesis, and metabolic adjustments.

Significance: We present comprehensive proteomic profiles of non-resistant and resistant *Acinetobacter baumannii* from Brazilian Hospitals strains, and highlight the presence of discriminative and yet unidentified mass spectral clusters. Our work emphasizes the importance of exploring this overlooked data, as it could hold the key to understanding the complex dynamics of antibiotic resistance. This approach not only informs antimicrobial stewardship efforts but also paves the way for the development of innovative diagnostic tools. Thus, our findings have profound implications for the field, as far as methods for providing a new perspective on diagnosing antibiotic resistance as well as classifying proteomes in general.

1. Introduction

Acinetobacter baumannii is a gram-negative coccobacillary bacterium, increasingly recognized as a significant public health concern. *A. baumannii* is commonly found in hospital environments and can cause various infections, including hospital-acquired pneumonia (HAP),

* Correspondence to: Marcelo Pillonetto, Rua Imaculada Conceição 1155, Prado Velho, Curitiba, PR 80215-901, Brazil.
** Correspondence to: Paulo Costa Carvalho, R. Professor Algacyr Munhoz Mader 3775, Cidade Industrial de Curitiba, PR 81350-010, Brazil.
E-mail addresses: m.pilonetto@pucpr.br (M. Pillonetto), paulo@pcarvalho.com (P.C. Carvalho).
[1] Equal contributions

https://doi.org/10.1016/j.jprot.2023.105012
Received 19 July 2023; Received in revised form 21 September 2023; Accepted 21 September 2023
Available online 23 September 2023

ventilator-associated pneumonia (VAP), and bacteremia [1]. The emergence and spread of antibiotic resistance in *A. baumannii* are a complex phenomenon influenced by genetic and environmental factors. This microorganism is notorious for its ability to rapidly develop resistance to antibiotics, making it a challenging pathogen to treat [2].

Addressing bacterial resistance to antibiotics is crucial; therefore, continued research into new antibiotics, efficient vaccines, alternative treatments, faster and more sensitive diagnoses, and promoting the appropriate use of antibiotics (Antimicrobial Stewardship) to delay the development of resistance are essential [3]. Among the various types of antibiotics, Carbapenem was once considered the last therapeutic option for treating multidrug-resistant *A. baumannii* (MDR) infections. However, in recent years, isolates resistant to this class have emerged, leading to the clinical application of polymyxins. With the emergence of extensively drug-resistant and even pan-drug resistance, clinicians have even been forced to use polymyxin B and E (colistin), an antibiotic group that can do kidney and nerve damage, to combat these infections [4]. Brazilian researchers described the first outbreak of Carbapenem-Resistant *Acinetobacter baumannii* – CRAB, as early as 1999, which produced OXA-23 enzyme [5]. It's important to note that CRABs are always MDR isolates, having developed resistance to multiple antibiotic classes before becoming resistant to carbapenems. As an aggravator, the COVID-19 pandemic has led to an increase in infections by multidrug-resistant bacteria, further impacting the global scenario. Unfortunately, during the pandemic, the uncontrolled use of antimicrobials has accelerated the decrease of efficiency of these antibiotic classes as well [6].

In 2019, Murray's published data estimated deaths associated with bacterial antimicrobial resistance reaching 4.95 million, a figure seven times higher than the reports described by the Centers for Disease Control and Prevention in 2013 and O'Neill in 2016 [7,8]. Particularly alarming is the situation in the southern region of Brazil, where recent research revealed an extensive presence of drug-resistant *A. baumannii*; from 2017 to 2020, 659 isolates were analyzed [9]. Polymyxin B remained the only antimicrobial with in vitro efficacy. In another report, an outbreak of carbapenem-resistant strains in an intensive care unit was detailed [10], thus deepening our concerns. With escalating resistance, the necessity for stringent surveillance and conscientious antimicrobial stewardship becomes even more pressing.

Brazil, with its extensive population, diverse healthcare infrastructure, including the Unified Health System (SUS), and its network of Central Public Health Laboratories (LACENs), provides a unique setting for exploring the emergence and spread of antibiotic resistance in *A. baumannii*. This intricate assembly of public health services and diagnostic laboratories nationwide offers a comprehensive and nuanced perspective on the country's ongoing efforts to combat antibiotic resistance. With this as motivation, our data descriptor originates from the joint efforts of these systems and aims to contribute to a broader understanding of the molecular underpinnings of resistance and the drivers behind the emergence of multidrug-resistant strains within the country. Here, we provide and explore proteomic profiles of *A. baumannii* strains originating from the states of Paraná and Santa Catarina. It is important to note the escalating prevalence of polymyxin-resistant *A. baumannii* strains in Brazil, especially in this region. This trend urges a deeper understanding of these strains and underscores the need for novel therapeutic options. Such knowledge is crucial in devising innovative therapeutic strategies, surveillance control, and preserving the efficacy of our current antibiotic arsenal and thus should serve as an example for similar investigations worldwide [11].

Proteomics is a powerful tool for understanding local antibiotic-resistant bacteria, providing a comprehensive view of their proteome [12]. This information can identify specific proteins involved in antibiotic resistance and understand cellular changes in response to antibiotic exposure. Proteomics can also serve as a diagnostic tool to detect resistance-related proteins, such as beta-lactamases and aminoglycoside-modifying enzymes [13]. However, the presence of these enzymes doesn't definitively indicate active resistance, which could be latent or suppressed under certain conditions [14]. Moreover, resistance could be conferred by post-translational modifications, which are undetectable by DNA sequencing approaches, and amino acid substitutions, which are frequently missed by proteomic search engines [15]. Current solutions, such as Bruker's MALDI Biotyper, exemplify the utility of mass spectrometry in bacterial diagnostics. However, these systems typically rely on single spectra, which limits their discriminative power for diagnosing resistant bacteria. In contrast, shotgun proteomics profiles provide a more expansive and detailed analysis of protein fragments, enhancing the discriminatory power when identifying and classifying bacterial strains. In this vein, the studies by Ping Wang et al. and Zhenbo Xu et al. provide comprehensive proteomic analyses of *Acinetobacter baumannii*, revealing significant differences between drug-resistant and drug-susceptible isolates, and identifying differentially expressed proteins that may influence drug resistance, particularly those located in the periplasmic or outer membrane of the cell [16,17].

This data descriptor presents proteomic profiles of polymyxin-resistant *A. baumannii* strains from Brazilian hospitals, highlighting regional variations. By cross-examining proteomic data across diverse strains and resistance profiles, our dataset uncovers unique evolutionary trends and adaptations under antibiotic pressure. These insights inform Antimicrobial Stewardship in Brazilian healthcare settings and can be instrumental in developing diagnostic tools and slowing the development of further resistance.

We utilized DiagnoMass [18] to investigate the proteomic data, grouping similar spectra into spectral clusters, transcending the limitations of traditional search engines that introduce bias towards identifiable portions of the provided sequence database. DiagnoMass interfaced with PatternLab for Proteomics V (PLV) for peptide spectrum matching [19] and FragPipe for open search [20], shortlisting spectral clusters unique to a biological condition. Notably, we aim to discover and highlight those unique mass spectral clusters that remain unidentified by both search engines, offering an unbiased perspective for comparing biological conditions and shedding light on potential biomarkers for understanding and combating drug resistance in *A. baumannii.* This methodology offers an unbiased perspective for comparing biological conditions and promises to shed light on potential biomarkers key in understanding and combating drug resistance in *A. baumannii*. As such, we argue that this is pivotal for surveillance purposes and can serve as a resource for the development and refinement of diagnostic tools. Such data can form the foundational blocks for integration with machine learning algorithms, predicting antibiotic resistance based on comprehensive proteomic profiles.

2. Methodology

2.1. Sample acquisition and preparation

We were granted permit A64AE41 from the Brazilian National System of Management of Genetic Heritage and Traditional Knowledge (SisGen), which authorized us to assess samples of *A. baumanni.* The bacterial strains (14 resistant and 14 non-resistant to polymyxin B) were obtained from the "*Laboratório Central do Estado do Paraná*". Stored at −80 °C in Brain Heart Infusion (BHI) medium, the strains were thawed and streaked onto MacConkey agar medium using a 10 μL bacteriological loop. Following incubation for 24 h at 36 °C, growth and isolate purity were assessed. Bacterial lysis was carried out following the SPEED protocol [21]. Subsequently, sample quantification was performed using the fluorometric Qubit assay, as per the manufacturer's instructions. One hundred micrograms of each sample were reduced with dithiothreitol (final concentration 10 mM) for 30 min at 60 °C, cooled to room temperature (20 °C) followed by alkylation with iodoacetamide (final concentration 30 mM) for 25 min. Finally, the samples were digested overnight with trypsin in a 1/50 (E/S) ratio at 37 °C. The reaction was interrupted with trifluoroacetic acid (TFA) 10% (final concentration of

1%), followed by centrifugation for 15 min at 18,000 x g and again quantified by the fluorometric Qubit assay. Ten micrograms of each sample were desalted with Stop and Go Extraction Tips (Stage Tips) as described by Rappislber and collaborators [22].

2.2. Mass spectrometry acquisition

The peptides were subjected to LC-MS/MS (Liquid chromatography with tandem mass spectrometry) analysis with an UltiMate 3000 (Thermo Fisher®)) ultra-high-performance liquid chromatography (UHPLC) system coupled with an Orbitrap Fusion™ Lumos™ mass spectrometer (Thermo, San José), as follows. The peptide mixtures were loaded into a column (75 mm i.d., 30 cm long) packed in-house with a 3 μm ReproSil-Pur C18-AQ resin (Dr. Maisch) with a flow of 250 nL/min and subsequently eluted with a flow of 250 nL/min from 5% to 40% ACN in 0.1% formic acid in a 140 min gradient [23]. The mass spectrometer was set in data-dependent acquisition (DDA) mode to automatically switch between full-scan (MS) and MS/MS (MS2) acquisition. Survey MS spectra (from *m/z* 300–1500) were acquired in the Orbitrap analyzer with a resolution of 120,000 at *m/z* 200. The most intense ions captured in 2 s cycle time were chosen, excluding unassigned ones that had a 1+ charge state. The ions were sequentially isolated and fragmented using higher-energy collisional dissociation (HCD) with normalized energy of 30. The fragment ions were analyzed with a resolution of 15,000 at 200 *m/z*. The general mass spectrometric conditions were spray voltage, 2.5 kV; no sheath and auxiliary gas flow; ion transfer tube temperature of 250 °C; predictive automatic gain control (AGC) enable, and S-lens RF level of 40% Mass spectrometer scan functions and nLC solvent gradients were regulated using the XCalibur 4.1 data system (Thermo, San José). Two technical replicates were acquired for each biological replicate.

2.3. Data analysis

2.3.1. Dataset quality control with rawvegetable 2.0

To ensure the quality control of our mass spectrometry data, especially between the technical replicates of the experiments, we utilized RawVegetable, a specialized software tool tailored for mass spectrometry data assessment [24]. This tool provided us with several key modules. The charge state chromatogram module, as well as the TopN density estimation module, enabled us to optimize chromatography. The TopN is particularly interesting, as it enables the control of how many MS/MS scans are being generated by cycle, which in turn helped us identify retention time intervals of under or over-sampling, facilitating necessary gradient adjustments. The chromatography reproducibility module allowed for pairwise comparisons between multiple experiments, ensuring the consistency and reliability of our data. And lastly, we were able to assess the quality of MS/MS spectra by looking at their Xrea scores [25] throughout the run and the precursor signal ratio distribution to analyze the fragmentation efficiency.

2.3.2. Protein and peptide identification with PatternLab for Proteomics V

We utilized PatternLab for Proteomics V (version 5.0.0.152) for peptide spectrum matching due to its proven reliability and accessibility [26]; the software is freely available at http://patternlabforproteomics.org. Data analysis followed the software's protocol [19], with the exception of the new search feature that enables multiplexed spectra identification through the Y.A.D.A. 3.0 deconvolution algorithm [26,27].

A. baumannii sequence (TG28175–5909 in total) were sourced from UniProt in March 2023, supplemented with 127 common mass spectrometry contaminants and searched using PLV's embedded Comet [28]. The search included semi-tryptic and fully-tryptic peptide candidates, allowing up to 2 missed cleavage sites. Variable modifications included oxidation of methionine and deamidation of asparagine and glutamine, while carbamidomethylation of cysteine was considered a fixed modification.

Validation of peptide-spectrum matches (PSMs) was conducted using the Search Engine Processor (SEPro) integrated into PLV [29]. This aimed to generate a list of identifications with <1% false discovery rate (FDR) at the protein level. Identifications were grouped by charge state (2+ and ≥ 3+) and tryptic status, forming four distinct subgroups. Each group was sorted based on a Bayesian discriminator generated from XCorr, DeltaCN, DeltaPPM, and Peaks Matched values. A cutoff score was set to accept a 2% FDR at the peptide level, based on the number of labeled decoys. This process was independently performed on each data subset, resulting in an FDR independent of charge state or tryptic status. Additional requirements included a minimum sequence length of six amino-acid residues and PSMs with <10 ppm from the global identification average. One-peptide identifications (proteins identified with only one mass spectrum) with an XCorr of <2 were discarded, leading to FDRs at the protein level of <1% for all search results.

2.3.3. Differential proteomics

We utilized the "Peptide-Centric Differential Proteomic Analyzer" module, integrated within the XIC Explorer from PLV, to identify differentially abundant proteins. This module initiates the process by normalizing peptide quantification values according to the respective Total XIC for each run. For our analysis, we only considered peptides that were found in at least eight biological replicates (more than half) of both condition and exhibited a minimum absolute 2.3-fold change ($Log_2(1.2)$). Subsequently, a single-sample *t*-test was applied to the fold changes of peptides within each protein. By treating each peptide as an independent observation, we minimized error propagation, thereby enhancing the reliability of our analysis. Proteins with <3 peptides were not considered. The method also incorporates the principle of maximum parsimony for protein inference. To control the false discovery rate at the protein level, we employed the Benjamini-Hochberg procedure with a *q*-value threshold of 0.05. This rigorous approach ensures the reliable identification of differentially abundant proteins.

For proteins uniquely identified to a single condition, we only considered those present in at least five biological replicates.

2.3.4. Leveraging DiagnoMass for exploratory and condition-specific spectral clustering

Shotgun proteomic data was processed using DiagnoMass, a tool for spectral clustering. Parameters were set to optimize results: *BinOffset* at 112.456, *BinSize* at 1.0005, and *BinMaxMZ* at 1612.2055 defined the binning strategy. A *SimilarityThreshold* of 0.75 ensured high similarity within clusters. Only tandem mass spectra derived from precursors with charge states from 2+ to 5+ and having 10 and more peaks above 500 *m/z* were considered. *PrecursorTolerance* was set at 0.002 *m/z* to accommodate minor variations in precursor m/z values. RetentionTimeTolerance was set at 10 min, accounting for variability in liquid chromatography separation times. These parameters culminated in spectral clusters, which were then annotated using PatternLab V. Only spectral clusters belonging to five or more biological replicates of a given biological condition and not identified in any other biological replicate from the other condition were shortlisted for further evaluation using open-search. This threshold (5 out of 14) is justifiable based on a chi-squared statistical test, resulting in a *p*-value of 0.0135.

2.3.5. Annotating discriminative and unidentified spectral clusters through open-search

Open-search is a versatile method that allows for the identification of unexpected post-translational modifications and sequence variants. For this purpose, we employed FragPipe, a comprehensive tool for large-scale proteomic data analysis [20]. We adhered to its default parameters for open-search, using the same FASTA sequence database as previously mentioned. This approach aimed to maximize the identification and annotation of unidentified spectral clusters, thereby enhancing our understanding of the proteomic landscape under investigation.

3. Results

3.1. Mass spectral quality control

Initial comparisons of the full chromatograms between the technical replicates, as exemplified by Fig. 1, show not only the great similarity between the runs, but also very well-defined peaks, which indicates good reproducibility between the samples and great separation of the molecules. A full quality control analysis done in RawVegetable [24] also presents charged chromatograms, which indicate when most of the species of each individual charge eluted. In these samples, the charged chromatograms showed an abundance of 2+ and 3+ species throughout the whole chromatography, as expected of linear peptides, which also indicates a successful digestion of the samples.

The TopN density estimation of this data presented a satisfactory spectra acquisition throughout the experiment, with an average of 30 to 35 MS/MS scans per MS spectrum for most of the chromatography run, which translates in a great amount of spectra generated for the peptide identification step.

In terms of the quality of the MS/MS spectra generated, the distribution of Xrea scores showed that most MS/MS spectra had a good overall quality and indicated that the higher quality of spectrum was generated by the first 70 min of the run. The distribution of precursor signal ratio in the MS/MS scans shows that, in all samples, most of the scans had <2.5% of precursor signal left, which indicates a satisfactory fragmentation of the peptides, thus possibly generating better signal for the fragments and a better spectrum for the identification step.

All of these analyses demonstrate that the samples had very efficient LC-MS runs and no major issues were detected, which makes the data ready for further steps in the study.

A full summary of the quality control analysis, complete with images for each pair of replicates can be found in the supplementary material.

3.2. Protein identification and differential abundance analysis using PatternLab for proteomics

Our analysis with PatternLab for Proteomics V yielded 928,831 identified tandem mass spectra, which corresponded to 44,553 peptides mapping to 3479 proteins across all analyses. The comprehensive list of proteins and peptides is available at 'All Proteins' and 'All Peptides' tabs of the Supplementary Spreadsheet, respectively. We identified 151 proteins that were exclusively identified in the resistant strains, considering only those found in five or more biological replicates specific to that condition (see 'Unique Resistant' tab). Conversely, a single protein (A0A0D5YDM4, Signal peptide) was exclusively identified in the non-resistant condition. We shortlisted 241 proteins identified in both conditions but exhibiting increased abundance in the resistant strains ('Up Resistant' tab), and 3 with increased abundance in the non-resistant strains ('Up Non-Resistant' tab).

3.3. Discriminant clusters via DiagnoMass

DiagnoMass was employed to cluster spectra based on precursor's charge state, *m/z*, and spectral angle. Not all spectra were assigned to clusters; for instance, only spectra with 10 or more *m/z* peaks above 500 were considered, and clusters had to contain at least two mass spectra. Our full dataset comprised 3,884,700 tandem mass spectra generated from both resistant and non-resistant strains. Of these, DiagnoMass considered 1,535,653 for generating the clusters, resulting in a total of 240,685 spectral clusters. Given that PatternLab identified a total of 44,553 peptides across all experiments, which could theoretically translate to around 89,106 spectral clusters (assuming that typically the same peptide is observed with 2 charge states), it's evident that the identified peptides explain only a little more than a third of the spectral clusters. This leaves a significant amount of information unexplored.

We propose that the most relevant information, not only in our work but also in other proteomic studies addressing conditions like cancer, might not be present in public databases. Alternatively, it might be tied up in post-translational modifications (PTMs) that are often missed by search engines due to their complexity or a combination of PTMs. These overlooked PTMs could play crucial roles in biological phenomena, such as enhancing bacterial resistance mechanisms as very well depicted in the work from Abouelhadid et al. [30].

Our dataset supports this hypothesis. For instance, the PCAs generated by PatternLab and DiagnoMass show a clear difference (Fig. 2). While PatternLab only considers information from identified peptides,

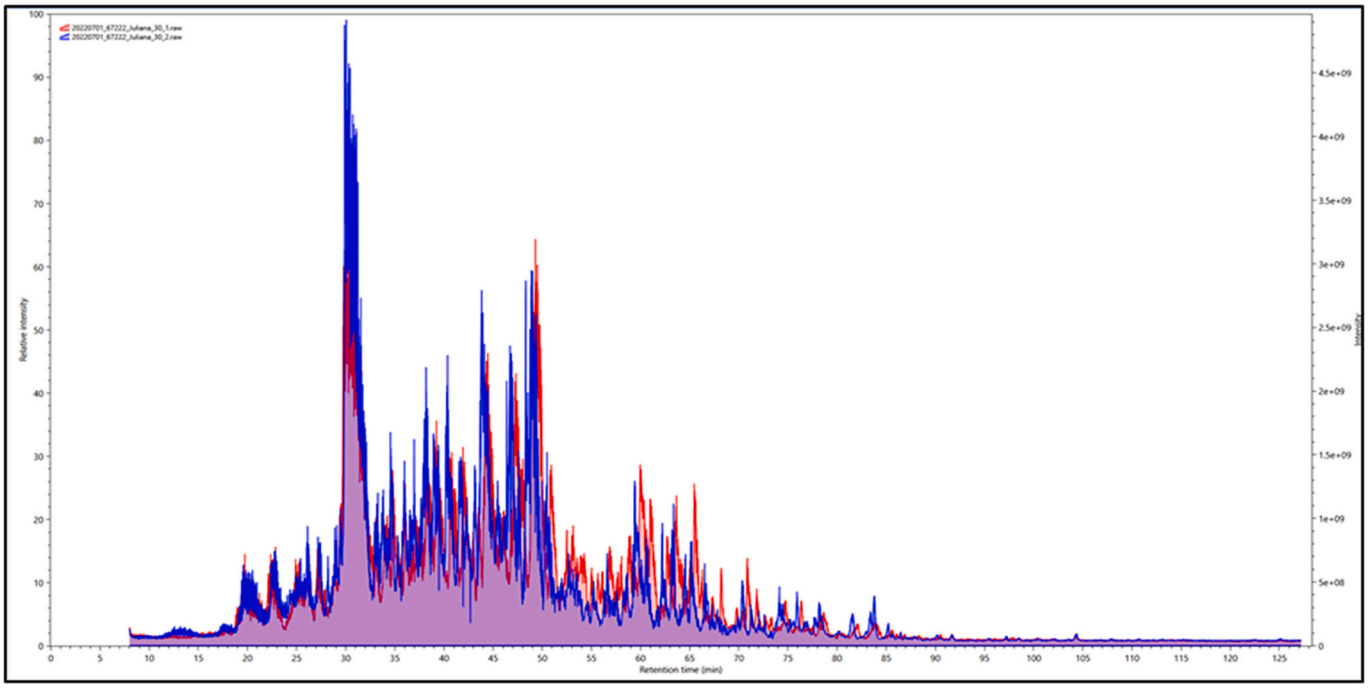

Fig. 1. Ion chromatogram comparison of technical replicates of Acinetocaber baumannii resistant to polymyxin B.

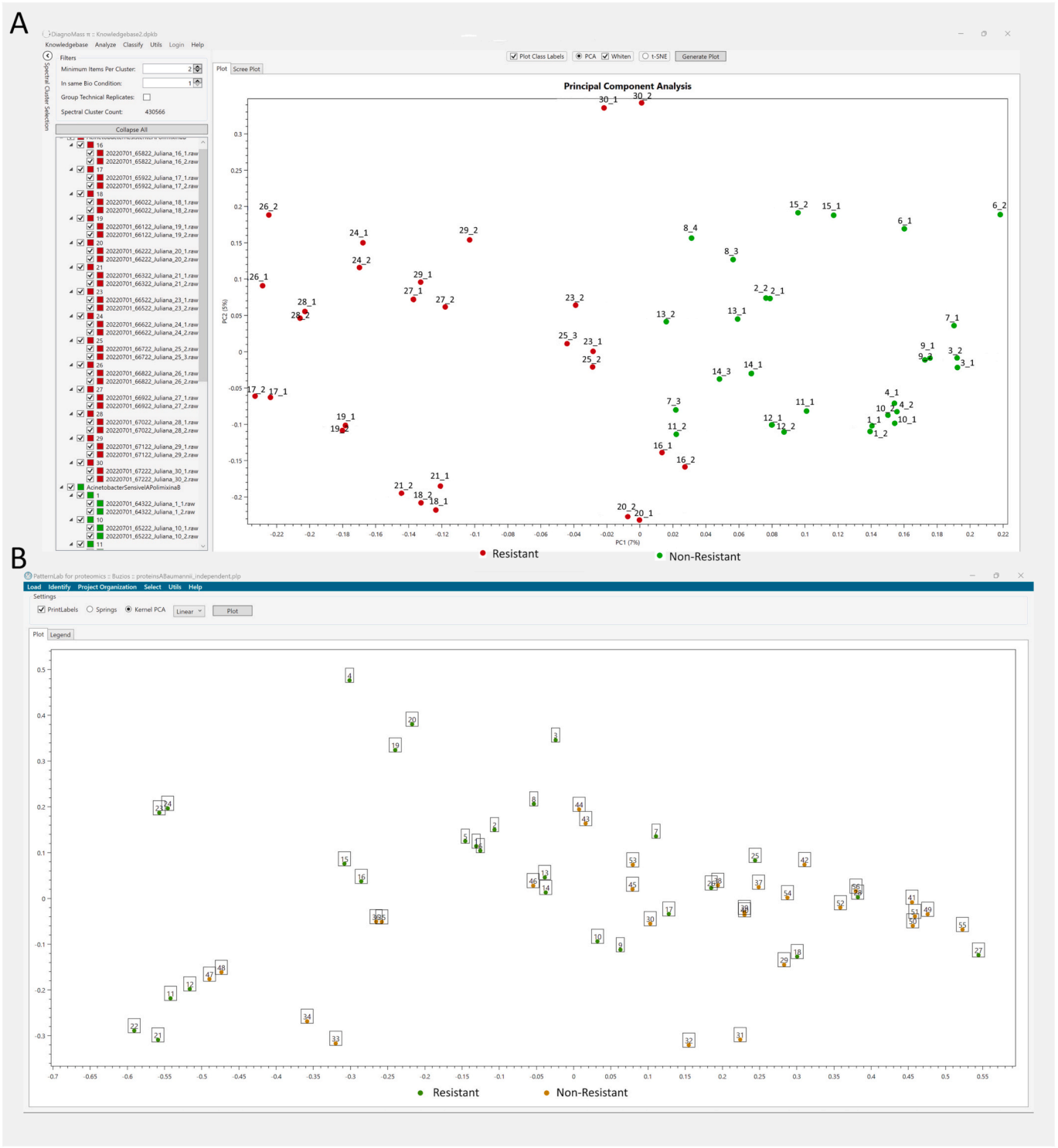

Fig. 2. DiagnoMass vs PatternLab PCA plots of Polymyxin B Resistant *A. baumannii.* - This figure presents a side-by-side comparison of *A. baumannii* strains that are resistant and non-resistant to polymyxin B, as visualized through screenshots of DiagnoMass (top) and PatternLab (bottom). The upper panel displays a DiagnoMass-generated PCA scatterplot, generated from spectral clusters from peptides (identified or not) depicted as red (resistant) and green (non-resistant) dots, indicating a clear proteomic distinction between the strains. Conversely, the lower panel shows a PatternLab-generated plot, based solely on identified peptides, where the resistant and non-resistant strains are represented by green and light-brown dots, respectively, with no evident separation. The numbers in the upper panel are associated with file names. The numbers associated with the dots in the lower panel are consecutive numbers, serving as a visual aid to indicate the grouping of technical replicates. Specifically, pairs of sequential numbers are expected to group closely together, representing technical replicates of the same sample. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

DiagnoMass bases its analysis on spectral clusters. This difference in approach allows DiagnoMass to reveal information that is typically missed by conventional proteomics workflows, thereby providing a more comprehensive understanding of the biology at hand. Importantly, both PatternLab and DiagnoMass were developed by our group, utilizing the same underlying libraries. This ensures consistency in the analytical approach and gives us confidence in the reliability and validity of the DiagnoMass approach.

Utilizing DiagnoMass, we generated a heatmap and histogram, focusing on clusters derived from spectra found in at least five instances of the same biological condition (Fig. 3). This analysis suggests a correlation within the proteomic profiles of the resistant and non-resistant strains. As indicated in Fig. 3, DiagnoMass shortlisted 3550 and 1408 clusters exclusive to the resistant and non-resistant strains, respectively. Out of these, PatternLab could confidently identify spectra from only 1281 and 193 clusters, respectively. A detailed manual review indicates the presence of numerous high-quality, yet unidentified, mass spectra. For instance, we present an example of a consensus mass spectrum from cluster 12,192, originating from 15 tandem mass spectra distributed among 10 biological replicates of the Resistant form and found in no non-resistant form. This consensus spectrum originates from parent ions with a 3+ charge and an m/z of 478.608; all of them remained unidentified (Fig. 4). Motivated by these findings, we exported all unidentified spectra from the unique clusters for further analysis using FragPipe, an open-search tool designed to search for variants not found in our sequence database. The exported mass spectra are provided in Supplementary2-UnidentifiedResistant.mgf and Supplementary3-UnidentifiedNonResistant.mgf.

3.4. *Unveiling hidden Proteomic profiles: the power of open-search*

Employing FragPipe's open-search capabilities, we identified 1004 peptides originating from the 3550 clusters exclusive to the resistant strains. These results underscore the value of integrating complementary search strategies alongside the prevalent peptide spectrum matching approach in proteomic identification. However, they also highlight the potential for further refinement in current proteomic identification tools. We speculate that the unidentified mass spectra may predominantly stem from peptides bearing a combination of post-translational modifications (PTMs) and amino acid substitutions. The peptides identified by FragPipe are made available in the 'Open-Search Resistant' tab of the Supplementary Spreadsheet.

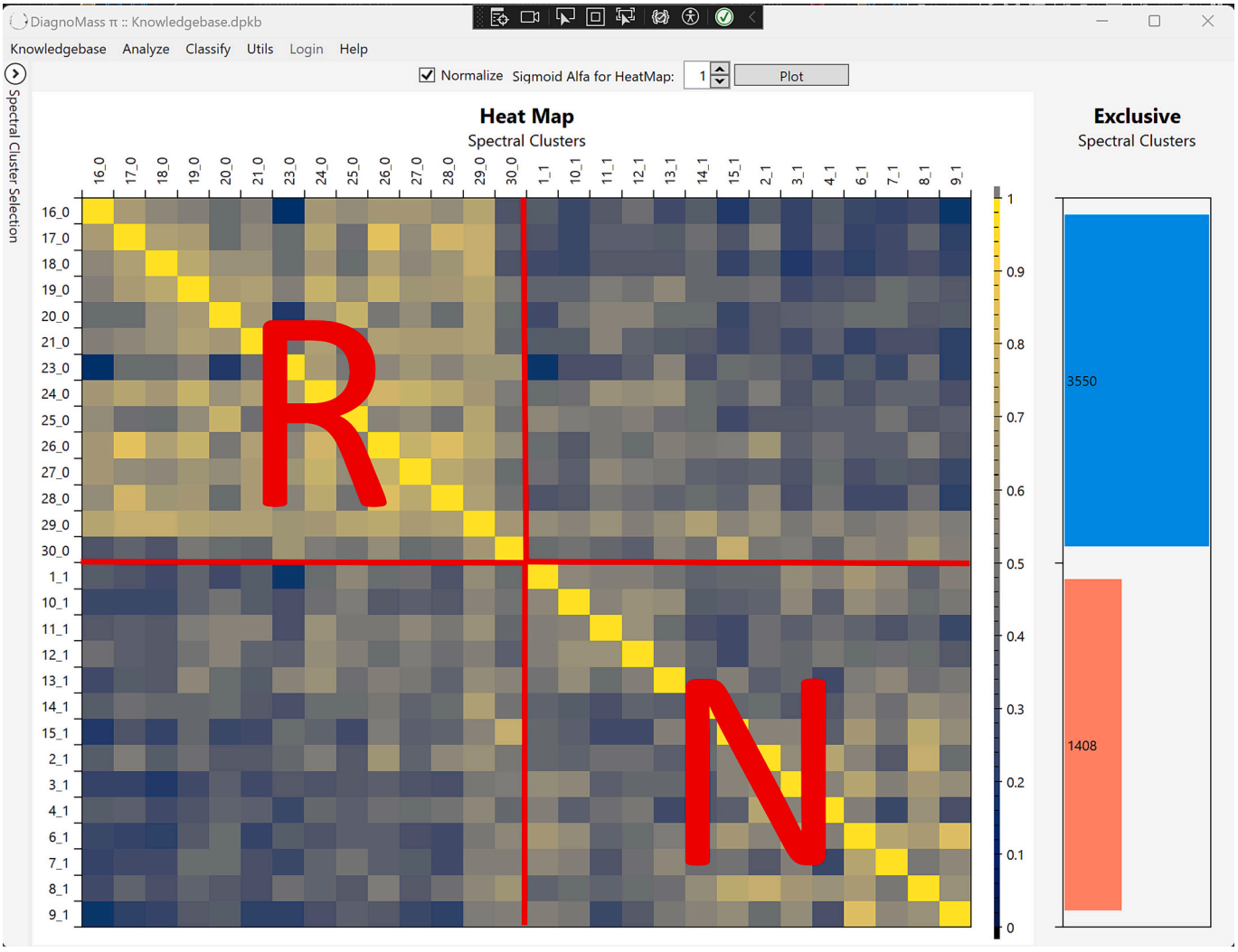

Fig. 3. Heatmap and Histogram of Spectral Clusters - This figure presents a heatmap of all bacterial samples analyzed in the study, generated using DiagnoMass. The heatmap visualizes the similarity in proteomic profiles between each pair of samples, with colour intensity indicating the degree of similarity. On the right, two bars represent the number of spectral clusters unique to each condition, i.e., the resistant (blue) and non-resistant (red) bacteria. The figure illustrates the significant differences in the proteomic profiles, especially the greater abundancy in unique spectral clusters in the resistant condition, highlighting the presence of numerous discriminative and yet unidentified mass spectral clusters. The results underscore the potential of this overlooked data in understanding antibiotic resistance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

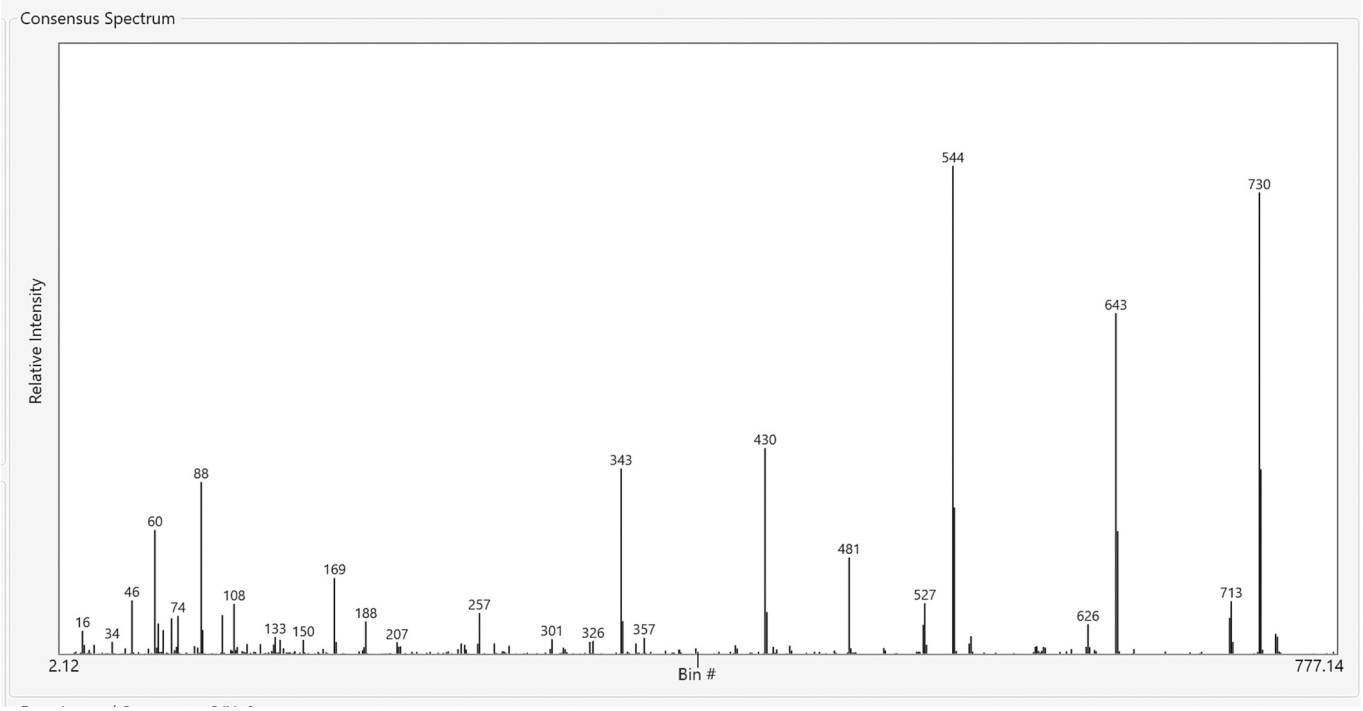

Fig. 4. Consensus Spectrum Illustration - The figure above presents a consensus binned spectrum, as provided by DiagnoMass, generated through the averaging of intensity values for all mass spectral peaks within a specific spectral cluster. The *y*-axis denotes relative intensity, while the *x*-axis represents the corresponding mass spectral bin number. This consensus spectrum provides a comprehensive representation of the spectral cluster's characteristic features.

4. Discussion

4.1. Importance of polymyxin-resistance in CRAB

Carbapenem drugs have been broadly used to treat serious gram-negative infections in the 1990's and 2000's. Unfortunately, carbapenem-resistant strains have been increasingly isolated in Brazil and worldwide [31,32]. One of the last-resort treatments for CRABs is polymyxin and cefiderocol. However, the latter is not available in Brazil and many other low- and middle-income countries. So, polymyxin is of prominent importance to treat this multi-drug resistant organism. The empirical use and abuse of polymyxin during COVID-19 to treat CRABS, precluded outbreaks of polymyxin-resistant *Acinetobacter baumannii*. Since the laboratory test to evaluate polymyxin susceptibility needs a broth microdilution method, not easily available at microbiology labs in LMIC, alterantive methods are welcome.

4.2. Proteins exclusively identified or more abundant in the resistant strains

The intricate network of proteins in polymyxin B-resistant *A. baumannii* presents a multifaceted portrait of the mechanisms underlying antibiotic resistance. Each protein, with its unique identifier, contributes to a narrative of adaptation and resilience in the face of antibiotic stress.

Among the proteins exclusively identified in the resistant strain we spotlight the Aldehyde dehydrogenase family protein (A0A429MLH1). By enhancing the bacterium's detoxification capabilities, it serves as a protective shield against the stress induced by polymyxin B. This shield is further fortified by the Glutamate-cysteine ligase (A0A3R9TSX1), which synthesizes glutathione, a molecule crucial for maintaining cellular redox balance and detoxifying reactive oxygen species. Together, these proteins form the bacterium's first line of defense against the antibiotic.

Simultaneously, the bacterium is also adjusting its protein synthesis, as suggested by the presence of the 16S rRNA (Cytosine(1402)-N (4))-methyltransferase (A0A3R9RYT7). This alteration in protein synthesis could be a response to the changes in the bacterium's metabolic and transport processes, indicated by the presence of the D-lactate dehydrogenase (A0A429MKP4) and the ATP-binding cassette domain-containing protein (A0A3R9TP97). By adjusting these processes, the bacterium is likely optimizing its resources to withstand the effects of polymyxin B.

Further supporting this optimization are the Bifunctional riboflavin kinase/FMN adenylyltransferase (A0A429MRU0) and the Malate dehydrogenase (quinone) (A0A429MNJ3). These enzymes could be indicative of metabolic adjustments in the resistant strain, fine-tuning its energy production to cope with the antibiotic stress.

These mechanisms, notwithstanding, is further supported by several other proteins in our results that were shortlisted as having increased statistical abundance in the resistance strain and are also henceforth discussed. The ATP synthase gamma chain (V5VHZ7) and ATP synthase subunit b (V5VHG5) are integral components of the cell's energy production machinery. As part of the ATP synthase complex, these molecular motors drive the synthesis of ATP, the primary energy currency of the cell. Their increased abundance in the resistant strain suggests an increased demand for energy, likely to fuel various resistance mechanisms. This is not an isolated phenomenon, but rather a component of a broader response involving numerous other proteins.

For instance, the long-chain fatty acid transporter (A0A3R9RY45) and the RND transporter (A0A3R9S1X2) exhibit increased abundance in the resistant strain. These proteins facilitate the transport of molecules across the cell membrane. Their increased abundance could be associated with the efflux of polymyxin B from the bacterial cell, suggesting a concerted effort by the bacterium to physically expel the antibiotic from its interior. This strategy likely requires a significant amount of energy, hence the increased abundance of ATP synthase proteins.

In addition to these transporters, several proteins involved in protein synthesis and folding also exhibit increased abundance in the resistant strain. These include various ribosomal proteins such as 50S ribosomal

protein L5 (A0A3R9U554), 50S ribosomal protein L4 (V5V9M5), and 30S ribosomal protein S1 (V5VDL1). The ribosome, the cellular machinery responsible for protein synthesis, is likely operating at an increased rate in the resistant strain. This could be a response to the need for more resistance-related proteins, or a general increase in protein synthesis as part of the stress response.

Molecular chaperones like DnaK (A0A3R9S7E3, A0A3R9S0W4) and chaperonin GroEL (V5VAH2) also exhibit increased abundance in the resistant strain. These proteins assist in the folding of newly synthesized proteins and help maintain protein homeostasis under stress conditions. Their up-regulation suggests an increased need for protein quality control mechanisms in the resistant strain, possibly due to the exacerbated rate of protein synthesis.

The cell wall and membrane of the bacterium also play a pivotal role in resistance. Several proteins involved in cell wall/membrane biogenesis exhibit increased abundance in the resistant strain, including the cell division protein FtsZ (A0A429MC99) and the lytic transglycosylase (A0A3R9S3W6). Changes in the cell wall/membrane could affect the binding of polymyxin B, thus contributing to resistance.

In addition to these, proteins involved in amino acid synthesis and metabolism, such as tryptophan-tRNA ligase (A0A3R9S100), cysteine synthase (A0A3R9TNT7), andthreonine synthase (A0A429MJ25), also exhibit increased abundance in the resistant strain. This could be due to an increased demand for amino acids for protein synthesis, or it could be part of a broader metabolic adjustment in response to antibiotic stress.

The resistant strain also shows an increased abundance of proteins with yet uncharacterized functions, such as the DUF3108 domain-containing protein (A0A3R9UE05, A0A3R9TM67) and the DUF4198 domain-containing protein (A0A0D5YHU6). The roles of these proteins in resistance are not yet clear, but their increased abundance suggests that they might be involved in some way. Further studies are needed to elucidate their functions.

In conclusion, the increased abundance of these proteins in polymyxin B-resistant *A. baumannii* is a testament to the organism's resilience and adaptability. It underscores the intricate interplay of various cellular processes - from energy production and protein synthesis to transport mechanisms, stress response, cell wall/membrane biogenesis, and amino acid metabolism. Each protein, acting not in isolation but in concert with others, contributes to a finely tuned symphony of survival mechanisms. This complex and coordinated response highlights the bacterium's strategic adaptation to antibiotic stress, reminding us of the multifaceted nature of antibiotic resistance. It underscores the need for a holistic approach in the study of antibiotic resistance, one that appreciates the interconnectedness of various proteins and cellular processes. As we continue to unravel this complex web of interactions, we move closer to understanding the full breadth of strategies employed by bacteria in their fight for survival.

4.3. Proteins more abundant or exclusively identified in the non-resistant strains

In the non-resistant strain of *A. baumannii*, several proteins were exclusively identified, suggesting a different set of mechanisms at play compared to the resistant strain. The ExeM/NucH family extracellular endonuclease (A0A3R9S704) might be involved in DNA degradation or repair processes, potentially contributing to the bacterium's ability to maintain genomic integrity under non-stressful conditions.

The Lipoprotein-34 (NlpB) (A0A3R9SE74) and two proteins identified as signal peptides (A0A098SJD0 and A0A0D5YDM4) could be involved in protein localization and transport, ensuring that proteins are correctly delivered to their functional locations within the cell.

Interestingly, several proteins in the non-resistant strain are yet to be characterized (A0A0D5YLE7, A0A429MLR8, and V5VCQ2). Their roles in the non-resistant strain are not yet clear, but their exclusive presence suggests that they might be involved in some way. Further studies are needed to elucidate their functions.

4.4. Final remarks

In conclusion, while Peptide Spectrum Matching (PSM) continues to be the most widely adopted method for proteomic data analysis, providing a robust foundation for the identification of proteins, it's important to recognize the value of complementary approaches such as open-search methods and other search engine unbiased approaches for analyzing data such as DiagnoMass. The latter can help uncover additional identifications that might otherwise be overlooked, albeit can be more error-prone. Despite these advancements, a significant portion of mass spectra remains unidentified. These elusive spectra, potentially carrying a combination of post-translational modifications and amino acid substitutions, may hold the key to understanding specific biological states, including mechanisms of resistance. As we continue to refine and develop our analytical tools, our ability to decode these spectra will undoubtedly enhance our understanding of complex biological systems and their responses to environmental challenges.

Declaration of Competing Interest

The authors declare to have no conflict of interests.

Data availability

Data available in Pride. Project accession: PXD043538. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [33] partner repository with the dataset identifier PXD043538.

Acknowledgements

The authors acknowledge Fiocruz ideias inovadoras (VPPIS-004-FIO-22-2-58), Fundação Araucária (421691/2022-0), CNPq, CAPES, Fondo Sectorial de Salud ANII FSS_X_2022_1_173332.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jprot.2023.105012.

References

[1] I. Kyriakidis, E. Vasileiou, Z.D. Pana, A. Tragiannidis, Acinetobacter baumannii antibiotic resistance mechanisms, Pathog. Basel Switz. 10 (3) (2021 Mar 19) 373.
[2] M. Petrillo, A. Angers-Loustau, J. Kreysa, Possible genetic events producing colistin resistance gene mcr-1, Lancet Infect. Dis. 16 (3) (2016 Mar) 280.
[3] D.O. Guimarães, L.S. da Momesso, M.T. Pupo, Antibióticos: importância terapêutica e perspectivas para a descoberta e desenvolvimento de novos agentes, Quím Nova. 33 (3) (2010) 667–679.
[4] G. Da Silva, S. Domingues, Interplay between Colistin resistance, virulence and fitness in Acinetobacter baumannii, Antibiotics. 6 (4) (2017 Nov 21) 28.
[5] L.M. Dalla-Costa, J.M. Coelho, H.A.P.H.M. Souza, M.E.S. Castro, C.J.N. Stier, K. L. Bragagnolo, et al., Outbreak of carbapenem-resistant Acinetobacter baumannii producing the OXA-23 enzyme in Curitiba, Brazil, J. Clin. Microbiol. 41 (7) (2003 Jul) 3403–3406.
[6] B.J. Langford, M. So, M. Simeonova, V. Leung, J. Lo, T. Kan, et al., Antimicrobial resistance in patients with COVID-19: a systematic review and meta-analysis, Lancet Microbe, Lancet Microbe 4 (2023) e179–e191.
[7] C.J. Murray, K.S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, et al., Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis, Lancet 399 (10325) (2022 Feb) 629–655.
[8] N.A. Rizk, N. Zahreddine, N. Haddad, R. Ahmadieh, A. Hannun, S. Bou Harb, et al., The impact of antimicrobial stewardship and infection control interventions on Acinetobacter baumannii resistance rates in the ICU of a tertiary Care Center in Lebanon, Antibiotics. 11 (7) (2022 Jul 7) 911.
[9] L.R.R. Perez, E. Carniel, G. Dalpiaz, M. Vetter, G.A. Narvaez, C.G. Dias, A four-year follow-up survey of antimicrobial resistance among Acinetobacter baumannii complex from inpatients in southern Brazil, Am. J. Infect. Control 49 (12) (2021 Dec) 1503–1505.
[10] K.E. da Silva, W.G. Maciel, J. Croda, R. Cayô, A.C. Ramos, R.O. de Sales, et al., A high mortality rate associated with multidrug-resistant Acinetobacter baumannii ST79 and ST25 carrying OXA-23 in a Brazilian intensive care unit, PLoS One 13 (12) (2018), e0209367.

[11] F. Tsakou, R. Jersie-Christensen, H. Jenssen, B. Mojsoska, The role of proteomics in bacterial response to antibiotics, Pharmaceuticals. 13 (9) (2020 Aug 27) 214.
[12] T.B. Lima, M.F.S. Pinto, S.M. Ribeiro, L.A. Lima, J.C. Viana, N.G. Júnior, et al., Bacterial resistance mechanism: what proteomics can elucidate, FASEB J. 27 (4) (2013 Apr) 1291–1303.
[13] D.E. Foudraine, N. Strepis, C. Stingl, M.T. ten Kate, A. Verbon, C.H.W. Klaassen, et al., Exploring antimicrobial resistance to beta-lactams, aminoglycosides and fluoroquinolones in *E. coli* and *K. pneumoniae* using proteogenomics, Sci. Rep. 11 (1) (2021 Jun 14) 12472.
[14] A.C. Palmer, R. Kishony, Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance, Nat. Rev. Genet. 14 (4) (2013 Apr) 243–248.
[15] S. Abouelhadid, J. Raynes, T. Bui, J. Cuccui, B.W. Wren, Characterization of Posttranslationally Modified Multidrug Efflux Pumps Reveals an Unexpected Link between Glycosylation and Antimicrobial Resistance. Parkhill J, editor, mBio 11 (6) (2020 Dec 22) e02604–20.
[16] P. Wang, R.Q. Li, L. Wang, W.T. Yang, Q.H. Zou, D. Xiao, Proteomic analyses of Acinetobacter baumannii clinical isolates to identify drug resistant mechanism, Front. Cell. Infect. Microbiol. 11 (2021 Feb 24) 625430.
[17] S.E. Cheah, M.D. Johnson, Y. Zhu, B.T. Tsuji, A. Forrest, J.B. Bulitta, et al., Polymyxin resistance in Acinetobacter baumannii: genetic mutations and transcriptomic changes in response to clinically relevant dosage regimens, Sci. Rep. 6 (1) (2016 May 19) 26233.
[18] M.D.M. Santos, A.C. Camillo-Andrade, D.B. Lima, T.A.C.B. Souza, J.S. de Fischer, G.R.H. da Valente, et al., DiagnoMass: A proteomics hub for pinpointing discriminative spectral clusters, J. Proteome 277 (2023 Apr) 104853.
[19] M.D.M. Santos, D.B. Lima, J.S.G. Fischer, M.A. Clasen, L.U. Kurt, A.C. Camillo-Andrade, et al., Simple, efficient and thorough shotgun proteomic analysis with PatternLab V, Nat. Protoc. 17 (7) (2022 Jul) 1553–1578.
[20] A.T. Kong, F.V. Leprevost, D.M. Avtonomov, D. Mellacheruvu, A.I. Nesvizhskii, MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics, Nat. Methods 14 (5) (2017 May) 513–520.
[21] J. Doellinger, A. Schneider, M. Hoeller, P. Lasch, Sample preparation by easy extraction and digestion (SPEED) - a universal, rapid, and detergent-free protocol for proteomics based on acid extraction, Mol. Cell Proteom. MCP. 19 (1) (2020 Jan) 209–222.
[22] J. Rappsilber, Y. Ishihama, M. Mann, Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics, Anal. Chem. 75 (3) (2003 Feb 1) 663–670.
[23] H. Hahne, F. Pachl, B. Ruprecht, S.K. Maier, S. Klaeger, D. Helm, et al., DMSO enhances electrospray response, boosting sensitivity of proteomic experiments, Nat. Methods 10 (10) (2013 Oct) 989–991.
[24] L.U. Kurt, M.A. Clasen, M.D.M. Santos, T.A.C.B. Souza, E.C. Andreassa, E.B. Lyra, et al., RawVegetable – a data assessment tool for proteomics and cross-linking mass spectrometry experiments, J. Proteome 225 (2020 Aug) 103864.
[25] S. Na, E. Paek, Quality assessment of tandem mass spectra based on cumulative intensity normalization, J. Proteome Res. 5 (12) (2006 Dec 1) 3241–3248.
[26] M.A. Clasen, L.U. Kurt, M.D.M. Santos, D.B. Lima, F. Liu, F.C. Gozzo, et al., Increasing confidence in proteomic spectral deconvolution through mass defect. Vitek O, editor, Bioinformatics. 38 (22) (2022 Nov 15) 5119–5120.
[27] M.A. Clasen, M.D.M. Santos, L.U. Kurt, J. Fischer, A.C. Camillo-Andrade, L.A. Sales, et al., PatternLab V handles multiplex spectra in shotgun proteomic searches and increases identification, J. Am. Soc. Mass Spectrom. 34 (4) (2023 Apr 5) 794–796.
[28] J.K. Eng, M.R. Hoopmann, T.A. Jahan, J.D. Egertson, W.S. Noble, M.J. MacCoss, A deeper look into comet–implementation and features, J. Am. Soc. Mass Spectrom. 26 (11) (2015 Nov) 1865–1874.
[29] P.C. Carvalho, J.S.G. Fischer, T. Xu, D. Cociorva, T.S. Balbuena, R.H. Valente, et al., Search engine processor: filtering and organizing peptide spectrum matches, Proteomics. 12 (7) (2012 Apr) 944–949.
[30] S. Abouelhadid, J. Raynes, T. Bui, J. Cuccui, B.W. Wren, Characterization of Posttranslationally Modified Multidrug Efflux Pumps Reveals an Unexpected Link between Glycosylation and Antimicrobial Resistance. Parkhill J, editor, mBio 11 (6) (2020 Dec 22) e02604–20.
[31] J.M. Cieslinski, L. Arend, F.F. Tuon, E.P. Silva, R.G.S. Ekermann, L.M. Dalla-Costa, et al., Molecular epidemiology characterization of OXA-23 carbapenemase-producing Acinetobacter baumannii isolated from 8 Brazilian hospitals using repetitive sequence–based PCR, Diagn. Microbiol. Infect. Dis. 77 (4) (2013 Dec) 337–340.
[32] C.R.V. Kiffer, T.F.T. Rezende, D.T. Costa-Nobre, A.S.S. Marinonio, L.H. Shiguenaga, D.N.O. Kulek, et al., A 7-Year Brazilian National Perspective on Plasmid-Mediated Carbapenem Resistance in Enterobacterales, *Pseudomonas aeruginosa* , and *Acinetobacter baumannii* Complex and the Impact of the Coronavirus Disease 2019 Pandemic on Their Occurrence, Clin. Infect. Dis. 77 (Supplement_1) (2023 Jul 5) S29–S37.
[33] Y. Perez-Riverol, J. Bai, C. Bandla, D. García-Seisdedos, S. Hewapathirana, S. Kamatchinathan, et al., The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences, Nucleic Acids Res. 50 (D1) (2022 Jan 7) D543–D552.