

The glycopatterns of *Pseudomonas aeruginosa* as a potential biomarker for its carbapenem resistance

Jing Dang,¹ Jian Shu,¹ Ruiying Wang,² Hanjie Yu,¹ Zhuo Chen,¹ Wenbo Yan,¹ Bingxiang Zhao,¹ Li Ding,¹ Yuzi Wang,¹ Huizheng Hu,² Zheng Li¹

AUTHOR AFFILIATIONS See affiliation list on p. 15.

ABSTRACT Multidrug-resistant (MDR) *Pseudomonas aeruginosa* infections pose a significant challenge to effective treatment. Although carbapenems were once considered the primary therapeutic option for MDR *P. aeruginosa*, the clinical use of these antibiotics has become increasingly limited, and the exploration of alternative antimicrobial strategies remains necessary. Bacterial surface glycans are critical in response to antibiotics and represent an attractive therapeutic target. However, the understanding of the role of glycan structures in bacterial resistance is currently quite limited. In this study, we used lectin microarrays to analyze the differences in glycan alterations between 53 drug-sensitive *P. aeruginosa* (DSPA) strains and 57 carbapenem-resistant *P. aeruginosa* (CRPA) strains obtained from clinical isolates, with the goal of identifying important glycopatterns associated with carbapenem resistance. The results revealed significant differences in the expression levels of glycan structures [e.g., Fuca1–6GlcNAc, α -D-Man, and Fuca1–3(Gal β 1–4)GlcNAc] recognized by 20 lectins (e.g., LCA, PSA, and AAL) on the bacterial surface. Furthermore, we applied *K*-fold cross-validation to determine the optimal parameters and constructed the DSPA and CRPA models using gradient boosting decision tree (GBDT) algorithm. The GBDT model presented the best performance in an independent test cohort, demonstrating that the screened glycopatterns could serve as potential biomarkers for differentiating bacterial resistance. Our study is the first to apply lectin microarrays to monitor carbapenem resistance in clinical *P. aeruginosa* isolates, and we hope that the methodology used in this work will serve as an important tool for exploring the association of bacterial resistance with glycosylation mechanisms.

IMPORTANCE Bacterial surface glycans are an attractive therapeutic target in response to antibiotics; however, current knowledge of the corresponding mechanisms is rather limited. Antimicrobial susceptibility testing, genome sequencing, and MALDI-TOF MS, commonly used in recent years to analyze bacterial resistance, are unable to rapidly and efficiently establish associations between glycans and resistance. The discovery of new antimicrobial strategies still requires the introduction of promising analytical methods. In this study, we applied lectin microarray technology and a machine-learning model to screen for important glycan structures associated with carbapenem-resistant *P. aeruginosa*. This work highlights that specific glycopatterns can be important biomarkers associated with bacterial antibiotic resistance, which promises to provide a rapid entry point for exploring new resistance mechanisms in pathogens.

KEYWORDS carbapenem-resistant *P. aeruginosa* (CRPA), lectin microarrays, glycan structures, bacterial cell surface, machine learning

Multidrug-resistant (MDR) bacterial strains are becoming increasingly prevalent, particularly *Pseudomonas aeruginosa*, the most common pathogenic organism in human clinical isolates (1). *P. aeruginosa* is an adaptable opportunistic pathogen capable

Editor Monika Kumaraswamy, University of California, San Diego, La Jolla, California, USA

Address correspondence to Jian Shu, shujian@nwwu.edu.cn, Huizheng Hu, 1224328711@qq.com, or Zheng Li, zhengli@nwwu.edu.cn.

Jing Dang and Jian Shu contributed equally to this article. The order of names is decided by the first letter of family name.

The authors declare no conflict of interest.

See the funding table on p. 16.

Received 23 May 2023

Accepted 8 September 2023

Published 20 October 2023

Copyright © 2023 Dang et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

of infecting various organs, including the respiratory, vascular system, urinary tract, and central nervous system, among others, causing significant morbidity and mortality (1), and is associated with prolonged hospital stays and increased health care costs (2). The adverse clinical outcomes associated with *P. aeruginosa* infection are largely due to its multiple drug resistance, which limits effective antibiotic treatment options and makes eradicating it increasingly difficult (2). Carbapenems used to be the most important therapeutic option against MDR *P. aeruginosa* (3). However, in recent years, carbapenem-resistant strains have become widespread and continue to expand in clinical practice (4, 5). The World Health Organization has listed carbapenem-resistant *P. aeruginosa* (CRPA) as one of the most serious pathogenic threats to humans, and addressing it has become an important issue in public health (6, 7). Further exploration of new antimicrobial strategies continues to be necessary.

Protein glycosylation is a ubiquitous post-translational modification that polymerizes or attaches carbohydrates to target macromolecules (proteins or lipids) and regulates many important biological processes. In bacteria, most glycosylation products are located on the cell surface, including oligosaccharides and polysaccharides such as lipopolysaccharides (LPS) and capsular polysaccharides (CPS), as well as glycoproteins such as pili, flagella, adhesin, self-transporter, and efflux pump (8). The glycosylation on the bacterial cell surface not only plays a crucial role in colonization (9–12), adhesion (13–18), symbiosis (10, 11), virulence factors (19), pathogenesis (20–22), cell-cell interactions, and immune evasion (10, 23–26) but also contributes directly to bacterial resistance (19, 27–33). Biofilm formation enables bacteria to adapt to the micro-environment under antibiotic selection pressure and to produce persister cells to enhance resistance. Biofilms are known to be complex entities consisting of bacterial cells aggregated in a secretory matrix containing polysaccharides, usually adhering to the surface of the organism (34, 35). Defective lipoprotein mannosylation alters the permeability of the cell envelope of *Mycobacteria abscessus*, making it more susceptible to macromolecular compounds acting on the peptidoglycan layer and to antibiotics such as penicillin G, ampicillin, and vancomycin (36). In *Campylobacter jejuni*, the CmeABC complex is a major multidrug efflux pump that is resistant to various antibiotics. When the *N*-glycosylation of CmeA alone was abolished, the function of the entire CmeABC efflux pump was impaired, resulting in a significant increase in sensitivity to azithromycin (37). These findings reflect the critical importance of bacterial surface glycosylation in response to antimicrobial agents, making it an attractive therapeutic target (8). However, compared to other biological processes, the direct impact of glycosylation on the mechanism of bacterial resistance is poorly understood.

Lectin is a non-immunogenic glycan-binding protein that recognizes glycan epitopes of free carbohydrates or glycoproteins with high specificity and is a primary means of glycosylation research. When adapted to microarrays, lectins can analyze a multiplicity of distinct binding glycan patterns and, thus, provide information on the carbohydrate composition of the sample (38, 39). Lectin microarrays have been developed and applied to analyze the glycosylation profiles of different bacterial species, allowing differentiation between given bacterial strains and monitoring changes in sugar signaling associated with environmental conditions (40). For instance, Hsu et al. used a microarray containing 21 lectins to observe significant differences in lectin-binding patterns and intensities between two closely related *Escherichia coli* strains, indicating the presence of different surface glycan structural libraries (41). The invasiveness of *E. coli* RS218 is known to be growth-dependent, with a general decrease in the positive signal intensity observed for 10 lectins when moving from exponential phase to stationary phase, suggesting a potential association between glycosylation and invasion (41). These results demonstrate that lectin microarray can differentiate between *E. coli* strains and monitor the dynamics of cell surface glycans. Furthermore, lectin microarrays have been used to differentiate between *Lactobacillus casei/paracasei* strains that could not be distinguished from each other by 16S ribosomal RNA gene sequencing. By using microarrays containing 44 lectins, half of the 16 strains were bound by only one or two lectins, while the

rest were recognized by multiple lectins with different carbohydrate-binding specificity, indicating the diversity of glycan structures (42). The assay allows strains to be differentiated while providing information on the carbohydrates as determinants available for recognition on the bacterial surface. Additionally, lectin microarrays have been applied to analyze the glycosylation profiles of different bacteria and to monitor changes in glycoforms associated with changes in culture conditions. Two clinical isolates of *C. jejuni* were determined using microarrays printed with a set of 41 lectins. When cultured at 42°C, a significantly reduced binding signal for lectins specifically recognizing Gal, lactose[Gal β (1–4)Glc], or GlcNAc was observed in strain 81116. These glycans were present in the LPS-like structure of the strain, indicating reduced or altered expression of the LPS structure, whereas when cultured at 37°C, the cell surface structure was mainly recognized as an LPS-like molecule. In contrast, the hypervirulent strain 81–176 grown at 42°C showed only minor changes in the binding intensity of a fraction of the lectins, implying that the glycan structures that could be recognized were relatively constant; whereas at 37°C incubation, the surface CPS and lipooligosaccharides, which are known to play different roles in the adhesion and invasion of epithelial cells and evasion of the immune system, were better recognized by lectin microarray (43).

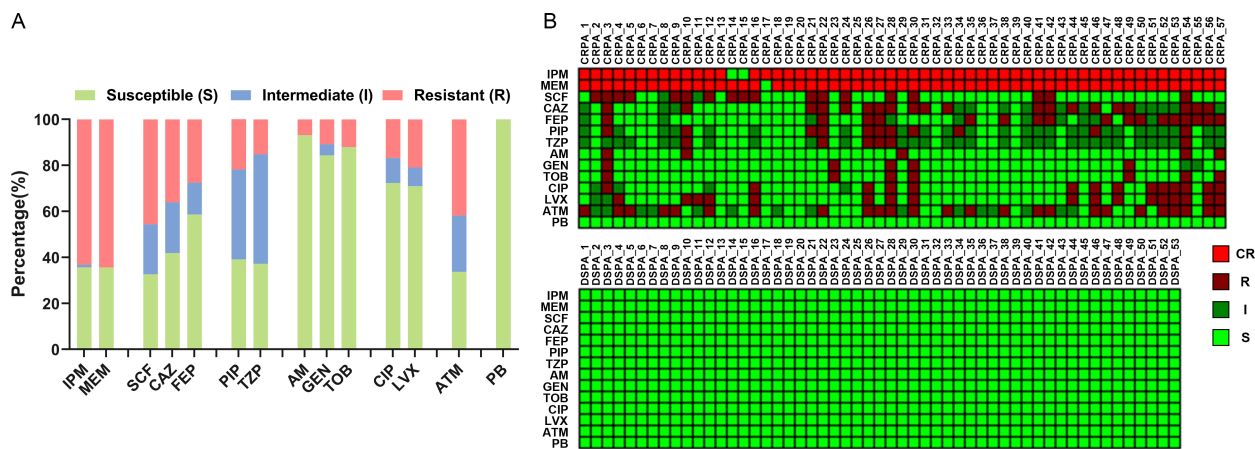
In this study, we used lectin microarrays to analyze the differences in the expression levels of glycan structures between 110 clinical isolates of DSPA (DSPA = 53 strains) and CRPA (CRPA = 57 strains). The results showed that the expression levels of multiple glycan structures on the bacterial cell surface recognized by lectins differed significantly between the two groups. Subsequently, we constructed a machine-learning model using a gradient boosting decision tree (GBDT) algorithm in an attempt to identify crucial glycopatterns that could serve as a potential biomarker for the differentiation of clinical DSPA and CRPA. We assessed the effectiveness of the classifier model based on a series of performance parameters, which also exhibited the feasibility of employing the high-throughput technology, lectin microarrays, to identify antibiotic resistance of clinical infectious pathogen. We hope that the methodology used in this work will serve as an important tool for establishing the association of bacterial resistance with glycosylation mechanisms.

RESULTS

Differential expression of glycan structures on the cell surface

Imipenem (IPM) and meropenem (MEM) are members of carbapenems and are routinely used to treat *P. aeruginosa* infections (44). In the present study, *P. aeruginosa* tested in the clinical laboratory for resistance to IPM and/or MEM was defined as CRPA. A total of 140 isolates from the clinic were analyzed for antimicrobial susceptibility, of which 53 isolates (37.86%) were susceptible to antimicrobial agents and 87 isolates (62.14%) were resistant strains. A large proportion of these resistant isolates showed high-level resistance to carbapenems (Fig. 1A), with 63.22% resistance to IPM and 64.37% resistance to MEM, and a total of 57 strains were CRPA, almost all of which were resistant to both IPM and MEM (Fig. 1B).

To determine whether and which association exists between carbapenem resistance and glycan structures, we first examined differences in the expression levels of glycan structures on the bacterial cell surface between DSPA and CRPA strains using lectin microarrays. The abundance of carbohydrates and glycoconjugates distributed on the bacterial cell surface allowed the whole bacteria fluorescently labeled with Cy3 dyes to bind to the microarrays through the interaction of cell surface glycan with lectin on a glass slide (28, 40) (Fig. 2A). One hundred and ten samples of bacterial strains (DSPA = 53, CRPA = 57) were examined independently to obtain the corresponding microarray fluorescence signal (Fig. 2B). After normalized fluorescence intensities (NFIs), a heat map was created using hierarchical clustering to visualize the lectin signal intensity distribution (Fig. 2C), confirming significant differences in bacterial surface glycan expression between DSPA and CRPA. Further between-group comparisons of each of the 37 lectins were performed. The relationship between DSPA and CRPA was visualized by principal



component analysis (PCA) (Fig. 2D), which was performed on the data generated collectively for the NFIs of the 20 altered lectins in the 110 samples. The PCA results displayed a tendency for the subjects assigned to the scatterplot to cluster separately to form DSPA and CRPA pools. The presence of the partial non-overlapping region between the two pools suggests that it is possible to differentiate between DSPA and CRPA based on specific variations in bacterial glycopatterns. Compared to the DSPA group, 13 lectins of GSL-II, MAL-II, PHA-E, PTL-I, SJA, WGA, LEL, BS-I, DSA, PHA-E+L, STL, MAL-I, and UEA-I showed increased changes of NFIs in CRPA group (Fig. 3A), and among these lectins,

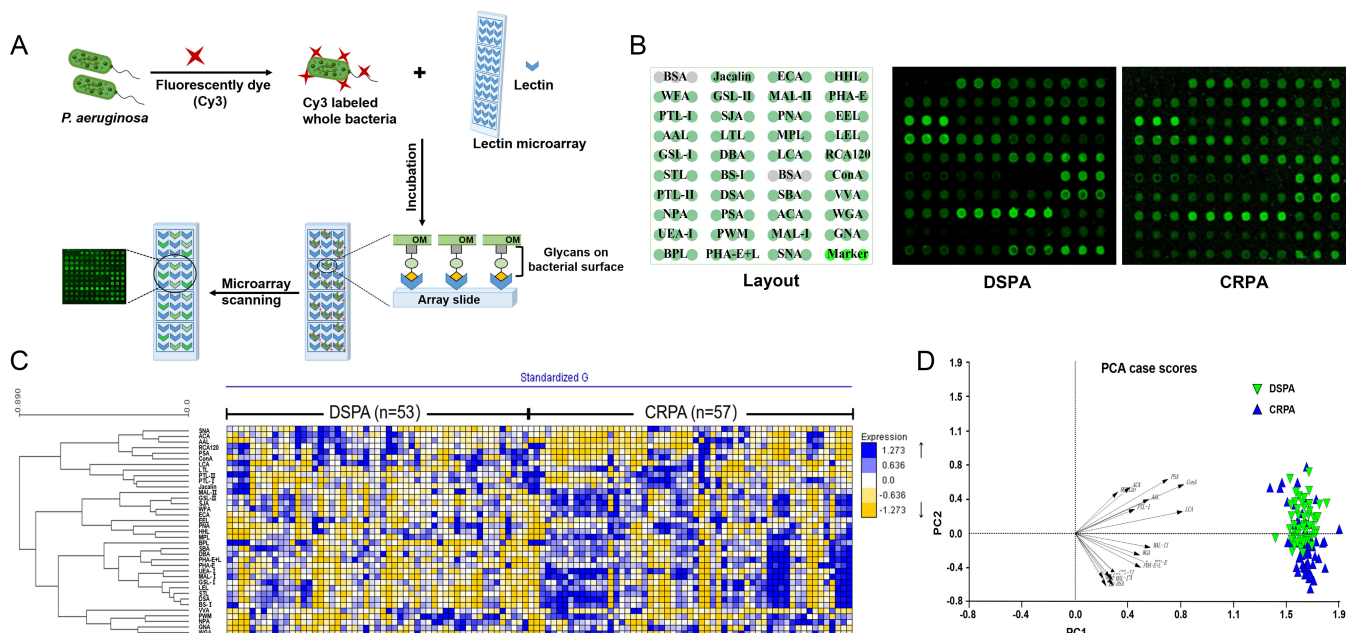


FIG 2 Lectin microarray analysis of whole bacteria. (A) The process diagram of Cy3-labeled whole bacteria binding to lectin microarray through the interaction between bacterial surface glycans and lectins. (B) (Left) Lectin microarray format. Glycan-binding specificities of the lectins are listed in Table 1. (Right) Glycopatterns of bacterial surface of DSPA and CRPA according to lectin microarray. A representative microarray scanning image was selected from each group. (C) Heat map of comparison of lectin microarray data for 53 DSPA and 57 CRPA isolates. The levels of lectin binding signals were indicated by the color change from yellow (low binding level) to blue (high binding level). (D) Principal component analysis of 20 lectins microarray data with relative changes in DSPA and CRPA groups.

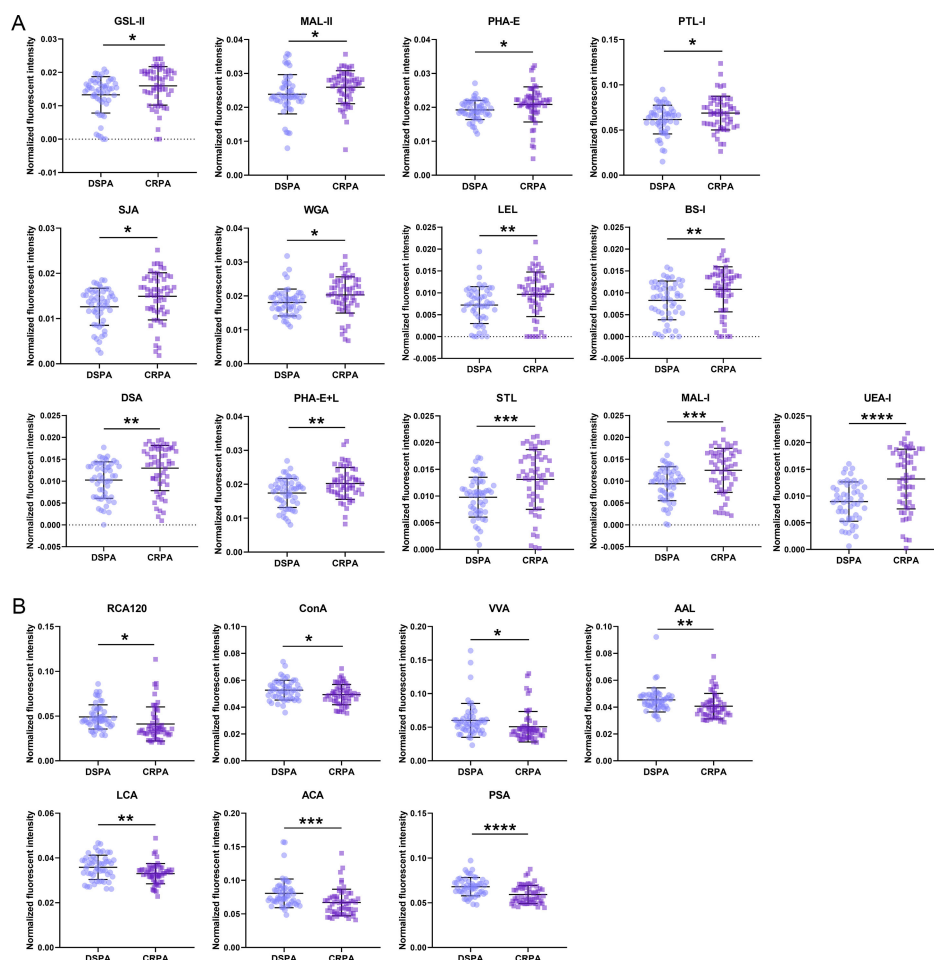


FIG 3 Comparison of glycan expression on the bacterial surface between DSPA and CRPA. (A and B) Glycan expression is up-regulated (A) and down-regulated (B) in CRPA, respectively. Statistical significance was calculated using an unpaired two-tailed *t*-test; **P* < 0.05; ***P* < 0.01; ****P* < 0.001; *****P* < 0.0001. Error bars represent standard deviation.

high mannose-type *N*-glycans recognized by LEL; Galactose recognized by BS-I; GlcNAc recognized by DSA, PHA-E+L, and STL; Gal β 1-4GlcNAc and Gal β 1-3GlcNAc recognized by MAL-I; and Fuca1-2Gal β 1-4Glc(Nac) recognized by UEA-I increased significantly (*P* value < 0.01); the NFIs of 7 lectins, RCA120, ConA, VVA, AAL, LCA, ACA, and PSA, were decreased in CRPA group (Fig. 3B), where the Fuca1-6GlcNAc (core fucose) and Fuca1-3(Gal β 1-4)GlcNAc recognized by AAL, α -D-Man and Fuca1-6GlcNAc recognized by LCA and PSA, and Gal β 1-3GalNAc α -Ser/Thr(Tn) recognized by ACA were significantly reduced (*P* value < 0.01) (Table 1). These results show that the expression levels of multiple glycan structures on the bacterial surface were altered between the DSPA and CRPA groups.

Differential expression of glycan structures in the intracellular space

In bacteria, glycoproteins are located in diverse cellular compartments, ranging from the cell wall and cytoplasmic membrane proteins to structural components of ribosomes (27). Whether the *en bloc* or a sequential addition glycosylation pathway, the abundant presence of glycoconjugates on the bacterial cell surface starts in the cytoplasm (nucleotide sugars), subsequently undergoes synthesis (or even extension) and is eventually transferred to the periplasm or outer membrane (10, 13). Whole-bacteria analysis relies on the direct and specific recognition and binding of glycan structures on the bacterial cell surface to lectins on microarrays to capture bacteria onto glass slides, and the intensity of the collected fluorescent signals reflects the expression level of the

TABLE 1 Lectin information contained on the microarray

Abbreviation	Lectin	Specificity	Containing monosaccharide ^a	Two-tailed <i>t</i> test (<i>P</i> value) DSPA vs CRPA
Jacalin	<i>Artocarpus integrifolia</i> lectin	Galβ1–3GalNAcα-Ser/Thr(T), GalNAcα-Ser/Thr(Tn)	Galactose	0.5021, ns
ECA	<i>Erythrina cristagalli</i> lectin	Galβ1–4GlcNAc (type II), Galβ1–3GlcNAc (type I)	Galactose	0.1146, ns
HHL	<i>Hippeastrum</i> hybrid lectin	α1–3/1–6Man, Gala1–3Galβ1–4GlcNAc	Mannose	0.6192, ns
WFA	<i>Wisteria floribunda</i> lectin	GalNAcα/β1–3/6Gal	GalNAc	0.6036, ns
GSL-II	<i>Griffonia simplicifolia</i> lectin II	GlcNAc and agalactosylated tri/tetra antennary glycans	GlcNAc	0.0139, *
MAL-II	<i>Maackia amurensis</i> lectin II	Siaa2–3Galβ1–3GalNAc, Siaa2–3Galβ1–4Glc – (NAc)/Glc, Siaa2–3Gal, Siaa2–3, Siaa2–3GalNAc		0.0042, *
PHA-E	<i>Phaseolus vulgaris</i> erythroagglutinin	Galβ1–4GlcNAc, biantennary <i>N</i> -glycan	GlcNAc	0.0042, *
PTL-I	<i>Psophocarpus tetragonolobus</i> lectin I	GalNAc, GalNAcα1–3Gal, GalNAcα1–3Galβ1–3/1–4Glc	GalNAc	0.0037, *
SJA	<i>Sophora japonica</i> agglutinin	Terminal GalNAc/Gal	GalNAc	0.0115, *
PNA	Peanut agglutinin	Galβ1–3GalNAcα-Ser/Thr(T)	Galactose	0.9500, ns
EEL	<i>Euonymus europaeus</i> lectin	Gala1–3(Fuca1–2)Gal	Galactose	0.5757, ns
AAL	<i>Aleuria aurantia</i> lectin	Fuca1–6GlcNAc (core fucose), Fuca1–3(Galβ1–4)GlcNAc	Fucose	0.0088, **
LTL	<i>Lotus tetragonolobus</i> lectin	Fuca1–2Galβ1–4GlcNAc, Fuca1–3(Galβ1–4)GlcNAc	Fucose	0.2665, ns
MPL	<i>Maclura pomifera</i> lectin	Galβ1–3GalNAc, GalNAc	GalNAc	0.0641, ns
LEL	<i>Lycopersicon esculentum</i> lectin	(GlcNAc) _n , high mannose-type <i>N</i> -glycans	LacNAc	0.0072, **
GSL-I	<i>Griffonia simplicifolia</i> lectin I	αGal, αGalNAc	GalNAc	0.1731, ns
DBA	<i>Dolichos biflorus</i> agglutinin	GalNAc, GalNAcα1–3(Fuca1–2)Gal	GalNAc	0.5546, ns
LCA	<i>Lens culinaris</i> agglutinin	α-D-Man, Fuca1–6GlcNAc, α-D-Glc	Mannose	0.0036, **
RCA120	<i>Ricinus communis</i> agglutinin I	Gal, Galβ1–4GlcNAc, Galβ1–3GlcNAc	Galactose	0.0148, *
STL	<i>Solanum tuberosum</i> lectin	Core GlcNAc, oligosaccharide containing GlcNAc	GlcNAc	0.0004, ***
BS-I	<i>Bandeiraea simplicifolia</i> lectin I	α-Gal, α-GalNAc, Gala-1,3Gal, Gala-1,6Glc	Galactose	0.0072, **
ConA	<i>Concanavalin A</i>	High-Mannose, Mana1–6(Mana1–3)Man, terminal GlcNAc	Mannose	0.0206, *
PTL-II	<i>Psophocarpus tetragonolobus</i> lectin II	Gal	Galactose	0.9891, ns
DSA	<i>Datura stramonium</i> lectin	β-GlcNAc, Galβ1–4GlcNAc	GlcNAc	0.0029, **
SBA	Soybean agglutinin	α/β-terminal GalNAc, (GalNAc) _n	GalNAc	0.2134, ns
VVA	<i>Vicia villosa</i> lectin	Terminal GalNAc, GalNAcα-Ser/Thr(Tn), GalNAcα1–3Gal	GalNAc	0.0391, *
NPA	<i>Narcissus pseudonarcissus</i> lectin	High mannose	Mannose	0.9613, ns
PSA	<i>Pisum sativum</i> agglutinin	α-D-Man, Fuca1–6GlcNAc, α-D-Glc	Fucose	<0.0001, ****
ACA	<i>Amaranthus caudatus</i> lectin	Galβ1–3GalNAc	Galactose	0.0008, ***
WGA	Wheat germ agglutinin	Multivalent Sia and (GlcNAc) _n	GlcNAc	0.0150, *
UEA-I	<i>Ulex europaeus</i> agglutinin I	Fuca1–2Galβ1–4Glc(NAc)	Fucose	<0.0001, ****
PWM	<i>Phytolacca americana</i> lectin	GlcNAc, Galβ1–4GlcNAc	GlcNAc	0.6568, ns
MAL-I	<i>Maackia amurensis</i> lectin I	Galβ1–4GlcNAc, Galβ1–3GlcNAc	Galactose	0.0005, ***
GNA	<i>Galanthus nivalis</i> lectin	High-Mannose, Mana1–3Man	Mannose	0.6600, ns
BPL	<i>Bauhinia purpurea</i> lectin	Galβ1–3GalNAc, Terminal GalNAc	Galactose	0.1182, ns
PHA-E+L	<i>Phaseolus vulgaris</i> agglutinin	Bisecting GlcNAc, bi-antennary <i>N</i> -glycans, tri- and tetra-antennary complex-type <i>N</i> -glycan	GlcNAc	0.0013, **
SNA	<i>Sambucus nigra</i> lectin	Siaa2–6Gal/GalNAc	GlcNAc	0.7945, ns

^aThe monosaccharide that should be added to the lectin solution before printing. Detailed information can also be found in the information provided by the manufacturer of the lectin. Fuc, fucose; Man, mannose; Gal, galactose; Glc, glucose; Sia, sialic acid; GlcNAc, N-acetyl-glucosamine; GalNAc, N-acetyl-galactosamine; LacNAc, N-acetyl-lactosamine; Ser, serine; Thr, threonine; Tn, Tn antigen. *P* value: *, *P* < 0.05; **, *P* < 0.01; ***, *P* < 0.001; ****, *P* < 0.0001; ns, not significant.

glycan structures on the bacterial surface. Changes in intracellular glycan structures are not available without a separate detection. The abundant glycan structures loaded on the bacterial surface originate from various intracellular glycosylation processes and their involved elements. To systematically detect changes in the glycan structures of entire bacteria, as well as to highlight the abundance of changes on the bacterial surface, it is also necessary to clarify whether the expression level of intracellular glycan structures has been changed. Therefore, we disrupted the bacterial cell wall by ultrasound to obtain intracellular soluble components (proteins, nucleic acids, etc.), and the determination of intracellular glycan changes proceeded as shown in Fig. 4A. Before fluorescent labeling, we pooled supernatants from each of the 53 DSPA strains and 57 CRPA strains in equal amounts to a final volume of 2 mL to perform microarray analysis. The layout of the lectin recognition pattern of Cy3-labeled pooled DSPA sample and pooled CRPA sample is shown in Fig. 4B. Statistical analysis of the NFIs for each lectin showed that in CRPA, the NFIs for three lectins, STL, WGA, and PWM, which recognize GlcNAc were significantly increased (Fig. 4C); the NFIs for two lectins, Galactose and GalNAc recognized by RCA120, GalNAc and GalNAcα-Ser/Thr(Tn) recognized by VVA, were significantly reduced (Fig. 4D). The other 32 of the 37 lectins showed no signal differences between DSPA and CRPA, indicating that alterations in intracellular glycan structures are relatively constant. Based on the above comparisons, we will follow by focusing on the changes in glycan structures on the bacterial surface.

Validation of the differential expression of glycan structures

To validate the differential expression of glycan structures on the bacterial surface, we performed lectin-based immunofluorescence analysis with LCA and STL between 3 DSPA and 3 CRPA (samples selected randomly from each group). Cy5-labeled LCA and STL were incubated with whole bacteria, respectively, and the expression levels of specific glycan structures on the cell surface were characterized by their interactions with the lectins. Immunofluorescence signals were visualized using confocal microscopy (Fig. 5A and C) and quantified in selected areas by the mean gray value (equal to the ratio of integrated density to the area) (Fig. 5B and D). The results showed that the expression level of glycan structures recognized by LCA was down-regulated in the CRPA group compared to the DSPA group (Fig. 5A and B), which was consistent with the results of the lectin microarrays; there was no significant difference in the expression of glycan

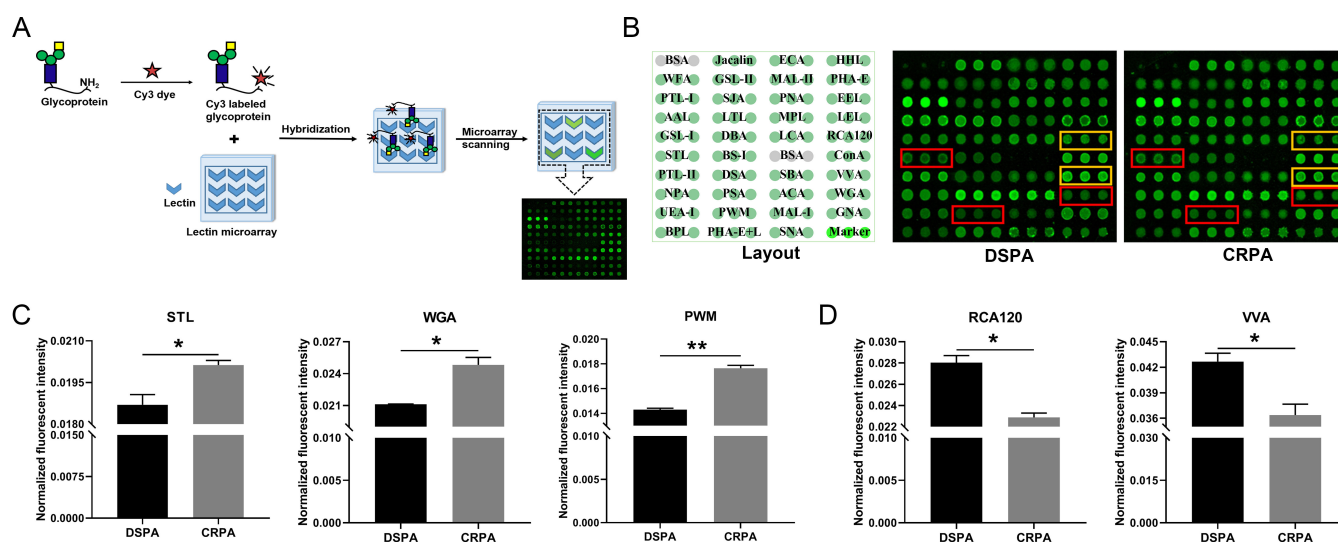


FIG 4 Lectin microarray analysis of bacterial intracellular proteins. (A) The basic procedures of glycoprotein bind to the lectin microarray. (B) Glycopatterns of lectin microarrays of intracellular protein mixtures from the DSPA and CRPA groups. Red boxes marked on the microarrays indicate increased expression and yellow boxes indicate reduced expression. (C and D) Glycan expression is up-regulated (C) and down-regulated (D) in CRPA, respectively. Data shown are means \pm SEM; * $P < 0.05$, ** $P < 0.01$ (unpaired two-tailed t -test, $n = 2$).

structures recognized by STL (Fig. 5C and D), which failed to align with the significant up-regulation detected by the lectin microarrays. This means that changes of glycan structures recognized by certain lectins (such as STL) obtained using lectin microarrays may be relative changes after normalizing fluorescence intensities. To further confirm the results, we implemented lectin blotting to analyze the binding of LCA and STL with glycan structures on whole bacterial proteins, respectively. Silver staining showed that the molecular weight and global abundance of proteins were similar in the six isolates although with several distinct bands (40–70 kDa) (Fig. 5E). Lectin blots showed the patterns of proteins loading with glycan structures bound to LCA or STL (Fig. 5F). For clarity, the results of lectin blots were quantified by the total gray values of the

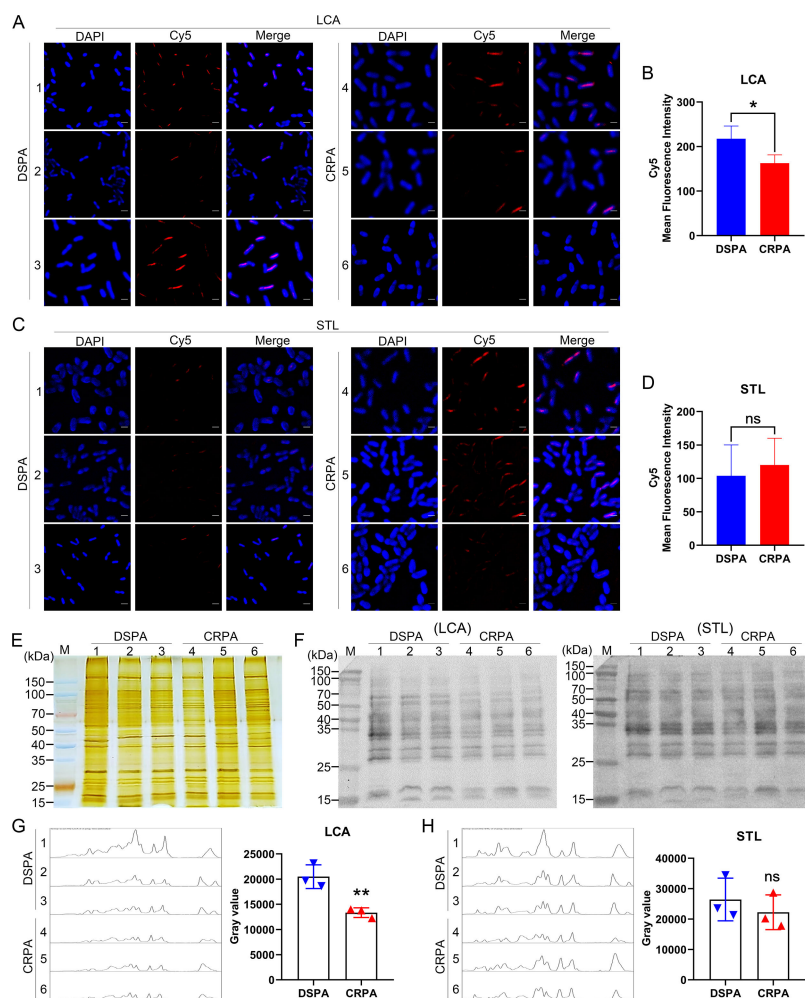


FIG 5 Validation of the differential expression of glycan structures on the cell surface. (A and C) Immunofluorescence analysis of individual strains between DSPA and CRPA. Representative images of whole bacteria stained with DAPI (blue) and hybridized with Cy5-labeled lectins [LCA (A) and STL (C)] (Cy5, red). Scale bars: 2 μ m. (B and D) Quantification of the mean fluorescence intensity of lectin-bound [LCA (B) and STL (D)] bacterial cells with Cy5 foci. Statistical significance was determined using a two-tailed unpaired *t*-test; **P* < 0.05; ns, not significant. (E) Silver-stained SDS-PAGE of whole-bacterial proteins obtained from six selected isolates. M, Protein marker. (F) The binding pattern of glycoproteins in selected samples of DSPA and CRPA isolates was analyzed by using two lectins [LCA (left) and STL (right)]. (G and H) Gray-value plots for each lane of LCA (G) and STL (H) and quantification of gray value by image. (I and J) Statistical significance was determined using a two-tailed unpaired *t*-test; ***P* < 0.01; ns, not significant. (A–H) 1–3: three DSPA strains (DSPA_1, 2, 3); 4–6: three CRPA strains (CRPA_6, 13, 20). Similar results were obtained from three independent experiments, and the data shown are from one representative experiment.

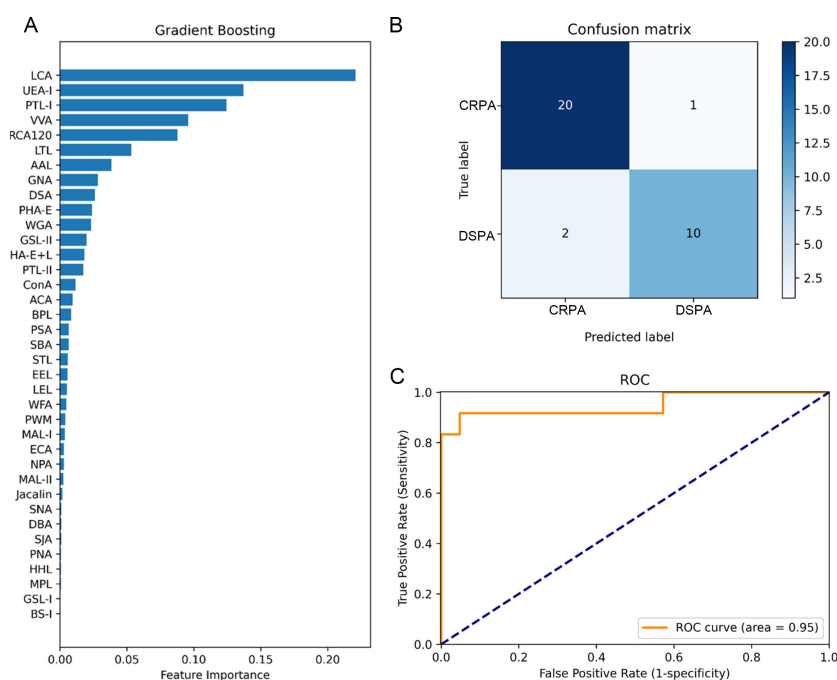


FIG 6 Performance of the GBDT model. (A) Feature importance ranking of GBDT. GBDT, gradient boosting decision tree. (B) Confusion matrix of the model. The row of the confusion matrix represents the actual value in the data set, and the column represents the predicted label of the classifier. (C) Model receiver operating characteristic (ROC) curves.

bands in each lane of the six selected strains to compare the differences between 3 DSPA and 3 CRPA, which showed that the expression of glycan structures bound by LCA was decreased in the CRPA group (Fig. 5G), whereas the expression of glycan structures bound by STL was not significantly different between the two groups (Fig. 5H), which was consistent with the results of the immunofluorescence analysis. LCA and STL were randomly selected as representatives to validate the results of the lectin microarrays; the validation for the remaining changed lectins was not shown here.

Selection of CRPA-related characteristic glycans and construction of machine-learning model using GBDT algorithm

GBDT is an integrated algorithm based on a decision tree that performs well in data analysis and prediction (45–47). To achieve data availability of the microarray assay and to identify potential features of glycan-based alterations associated with CRPA, we constructed a model for differentiating DSPA and CRPA strains through machine learning based on the GBDT algorithm. It is fruitful to screen for major glycan features associated with carbapenem resistance during the establishment of the classification model. As shown in Fig. 6A, the GBDT classifier ranked the importance of the individual variables according to their relative influence, with LCA (recognizing α -D-Man and Fuca1–6GlcNAc) as the top weighting feature. The top 5 important features also included UEA-I [recognizing Fuca1–2Gal β 1–4Glc(NAc)], PTL-I (recognizing GalNAc α 1–3Gal and GalNAc α 1–3Gal β 1–3/4Glc), VVA [recognizing GalNAc and GalNAc α -Ser/Thr(Tn)], and RCA120 (recognizing Gal β 1–4GlcNAc and Gal β 1–3GlcNAc) (Fig. 6A).

In an independent validation cohort of 33 isolates, the classifier correctly predicted 11 strains out of 12 DSPA and 21 strains out of 21 CRPA (Fig. 6B), presenting the high efficiency of the classifier. The performance of the test set was calculated with an AUROC of 0.95 (Fig. 6C), average classification precision of 89.29%, average recall of 90.9%, accuracy of 91%, and F1 score of 0.91 (Table 2). Sensitivity and specificity were 0.917 and 0.952, respectively. These exhibited the good performance of the feature-based

TABLE 2 The performance of GBDT model

	Precision	Recall	F1-score	Support
CRPA	0.91	0.95	0.93	21
DSPA	0.91	0.83	0.87	12
Accuracy	0.89	0.91	0.91	33
Macro avg	0.91	0.89	0.90	33
Weighted avg	0.91	0.91	0.91	33

classification model, which means that the glycopatterns such as α -D-Man and Fuca1-6GlcNAc recognized by LCA can be used as biomarkers for the differentiation of clinical DSPA and CRPA.

DISCUSSION

The resistance mechanisms identified in carbapenem-resistant *P. aeruginosa* are generally the loss of the expression of the outer membrane protein OprD and over-expression of the efflux pumps (MexAB-OprM and MexCD-OprJ), the expression of intracellular β -lactamases, and the acquisition of multiple carbapenemase genes, etc. (34); however, these well-defined mechanisms have not yet been effective in solving the therapeutic challenges of infections with the clinical strains. Clinical isolates often carry multiple resistance mechanisms simultaneously. It is more clinically realistic to consider the co-existence of multiple resistance mechanisms when controlling infections caused by resistant isolates. It should be perceived that the same resistance phenotype is a common feature of resistant isolates and distinguishes them from sensitive strains. In this work, we initiated a preliminary speculation as to whether clinical isolates with the same resistance phenotype would show identical glycan structures changes. Lectin microarrays were employed to compare the expression levels of glycan structures between DSPA and CRPA of clinical isolates. It was found that the glycan changes occurring on the bacterial surface were significantly abundant (Fig. 2 and 4). This is likely because the majority of bacterial carbohydrates are distributed on the bacterial surface, which acts as a direct and permeable barrier to antibiotic pressure. Of the 37 lectins we selected for fixation on microarrays, 20 lectins with alterations in NFIs contained recognition of the glycan structures of GlcNAc, Fucose, Galactose, GalNAc, LacNAc, and Mannose (Fig. 3; Table 1). The results predict the possibility of the presence of multiple glycosylation systems associated with antibiotic resistance, but the mechanisms involved have yet to be supported by sufficient and detailed evidence. Cluster analysis showed that there were differences in glycopatterns between individual isolates (Fig. 2C), and such respective differences might be related to resistance mechanisms, genotypes, or carrying resistance to other antibiotics. However, the PCA displayed the partial non-overlapping region that represents clustering separately to form DSPA and CRPA pools (Fig. 2D), suggesting that it is possible to differentiate between DSPA and CRPA based on glycopatterns. Immunofluorescence analysis and lectin blotting validated the results of lectin microarrays for two randomly selected lectins (LCA and STL) in 3 DSPA and 3 CRPA (Fig. 5). The validation results for LCA were in agreement with the results of the lectin microarrays, whereas the results for STL were inconsistent. As shown in Fig. 3A, the scatter plot for STL showed relatively discrete clustering of sample scatters. The likely reason is that the microarrays assay involves the competitive recognition and binding of 37 lectins, while the validation analysis involves the contact of individual lectins with the glycan structures on the bacterial surface without competition between lectins, indicating the occurrence of relative alterations when normalized fluorescence intensities for all lectins on glass slide. Therefore, LCA was an absolute variable involved in the differential expression of glycan structures on the bacterial surface between DSPA with CRPA, whereas the glycan structures recognized by STL may not altered. On the other hand, small sample sizes may also lead to inconsistent validation results due to the large differences in subtypes between clinical strains. To further screen for important

glycopolymers capable of distinguishing carbapenem resistance, machine learning was then implemented to construct a classifier model. Our results illustrate a series of good performance parameters for the predicted model with lectins such as LCA and UEA-I as the important features (Fig. 6; Table 2). In turn, the model performance reflects the fact that glycan structures recognized by the lectins can be used as biomarkers to distinguish clinical CRPA from DSPA.

Several studies have found that *P. aeruginosa* isolated from hospital environments exhibits high genotypic diversity and broad antimicrobial resistance spectra. Similarly, the isolates we collected also exhibited MDR to various antibiotics, with the highest ratios of resistance to carbapenems (Fig. 1A and B). Therefore, analyzing only a single isolate may be confounded by genotypic and multiple resistance, resulting in reduced clinical significance. For this reason, we analyzed a total of 110 clinical isolates from a hospital, divided into a DSPA group and a CRPA group, and we believe that the analysis of multiple samples in each group is more relevant for clinical guidance. This is also inspired by the widespread application of lectin microarray to assess tumor characteristics and screen new biomarkers for diagnostic cancer (48–52). When a large number of samples are collected within each group, the changes in expression levels of glycan structures sought are more representative of the various pathological and physiological states, as a clinical disease is often accompanied by other conditions, such as different complications between individuals with diabetes. Data from the microarray showed that glycopolymers differed between strains of different subtypes (Fig. 2C), but that the CRPA group exactly possessed significant glycan alterations that differed from the DSPA group (Fig. 2C and D), an exciting starting point for discovering the glycopolymers as biomarkers associated with antibiotic resistance.

The development of artificial intelligence as machine learning has shown prospects in the evaluation of disease-based diagnostic efficacy (53, 54). We follow the idea of differentiating clinical disease by identifying biomarkers through models that look for potential markers associated with CRPA, an approach that has shown credible accuracy and specificity when evaluating diagnostic efficacy through model construction (55). In this study, we constructed a GBDT-based machine-learning model to differentiate between DSPA and CRPA. During the construction process, α -D-Man and Fuc α 1–6GlcNAc recognized by LCA were found to be the most important variables that can differentiate between clinical DSPA and CRPA with good performance (Fig. 6). The representational glycan structures identified through this method are extremely reliable as biomarkers for CRPA.

Analysis of glycan structures on the bacterial surface realistically reflects the transformation in glycan signaling associated with clinical infectious *P. aeruginosa* under different selective pressures of antibiotics. The known monosaccharides structure on *P. aeruginosa* surface includes GlcNAc, GalNAc, ManNAc, FucNAc, and other unique monosaccharides (e.g., Xyl, Kdo, Rha, etc.) (8, 56). The association between these monosaccharide-related glycan structures and antimicrobial resistance in *P. aeruginosa* has not been clearly described. *C. jejuni* is a well-studied model organism for bacterial glycosylation. The *N*-oligosaccharyltransferase PglB is the key enzyme that catalyzes *N*-glycosylation of, at least, 53 proteins in *C. jejuni* (57, 58). Disruption of *pglB* ablates the glycosylation of CmeABC, resulting in impairment of the efflux activity of CmeABC, which reduces resistance to four different antibiotic classes (ampicillin, erythromycin, tetracycline, and ciprofloxacin) (59). Abouelhadid et al. showed that the loss of the *N*-linked glycans in CmeABC is the sole reason for the multidrug efflux pump impairment phenotype and not a pleiotropic effect caused by the disruption of the *N*-oligosaccharyltransferase PglB (37). Interestingly, *P. aeruginosa* also possesses the gene *pgl*, which was characterized as 6-phosphogluconolactonase. In most cases, the genes responsible for glycosylation are located in close vicinity of the gene encoding the target protein (60). In *P. aeruginosa* PAO1, the *pgl* is located adjacent to the glucose/carbohydrate outer membrane porin OprB (<https://www.pseudomonas.com/>). The Pgl responsible for the glycosylation of OprB remains unclear. Mutant strains lacking OprB have not

been reported to be associated with carbapenem resistance and the porin protein associated with carbapenem resistance in *P. aeruginosa* is usually OprD (34). It remains to be elucidated which glycosylases regulate the glycosylation of OprD and whether this results in changes in antimicrobial resistance. Furthermore, if the function of Pgl in *P. aeruginosa* also affects the activity of the multidrug efflux pump, it is also possible that resistance to carbapenems could be affected. In contrast to the above thinking, the significant reduction in the expression level of Fucose-associated glycan structures recognized by LCA occurred on the bacterial surface, while the presence of Fucose in *P. aeruginosa* has not been reported. Fucosyltransferases (FucTs) are important tools for the synthesis of fucosylated glycoconjugates. The understanding of the fucosylation process and related FucTs in bacteria is also quite limited (61, 62) (<http://www.cazy.org/bP.html>). The presence of FucTs in clinical *P. aeruginosa* isolates is highly questionable because many pathogens can evolve to manufacture host-like carbohydrates on their cell surfaces to mimic host-promoting survival and infection behaviors (40, 63, 64). *P. aeruginosa* is highly adapted to its environment. In a host microenvironment receiving antibiotic treatment, bacteria may also be able to perform molecular mimicry to harness host glycosylation resources in response to antibiotic pressure. Therefore, the significant changes in Fucose detected by microarrays may represent a characteristic simulation advantage of clinical strains isolated from human hosts, compared to laboratory strains. At the same time, it also illustrates that the contribution of glycans in biological processes is complex. Fortunately, the specific glycan structures recognized by lectins are well defined, and it is, therefore, promising that we used 37 lectins recognizing different glycan structures immobilized on microarrays to detect resistance-related changes in bacterial surface glycans.

As a preliminary investigation of the relevance of CRPA to glycopatterns, our study also has some limitations. For example, we used Cy3 to label the samples to make the assay consistent between the bacterial surface and intracellular space, and the final detection was mainly the changes of glycan structures on each part of the glycoprotein. To detect changes in LPS structure, enzymatic hydrolysis may be required before labeling to expose the reactive amino groups for binding to Cy3. Moreover, our microarray contained only 37 lectins, of which did not print CSL that bind to a Rha structure, and Rha was the main backbone structure of LPS of *P. aeruginosa*. Therefore, we lacked a complete analysis of the LPS structure.

A series of detailed analysis and the implementation of machine learning have made the data obtained by microarray available and finally approached important glycopatterns for CRPA. Future studies of a clear mechanism will focus on these important glycan structures to find new targets for therapeutic intervention and novel glycol-antimicrobial agents to alleviate the fatal threat posed by clinical MDR *P. aeruginosa*. In conclusion, the application of lectin microarrays to explore glycan biomarkers associated with clinical bacterial resistance is feasible and it is promising to employ this technology to further understand the mechanism of the effects of antibiotic resistance and glycosylation.

MATERIALS AND METHODS

Bacterial strains and susceptibility assays

The strains used in this study were collected from clinical *P. aeruginosa* isolates isolated from the Department of clinical laboratory, No. 215 Hospital of Nuclear Industry in Shaanxi Province from December 2019 to June 2021. A total of 140 isolates were obtained from clinical samples, including sputum (75%), skin secretions (13%), and others of peritoneal drainage, whole blood culture, bile, urine, pleural effusion, cerebrospinal fluid, and catheters (12% in total). The clinical laboratory performed an identification of the strains using a fully automated bacterial analyzer such as VITEK 2 System (BbioMérieux), and the antibiotic susceptibility of the strains with 14 antimicrobial drugs of Carbapenems (Imipenem, Meropenem), Cephalosporins (Cefoperazone/Sulbactam, Ceftazidime, Cefepime), Penicillins (Piperacillin, Piperacillin/tazobactam),

Aminoglycosides (Amikacin, Gentamicin, Tobramycin), Quinolones (Ciprofloxacin, Levofloxacin), Monocyclic β -lactams (Aminotransomide), Peptides (Polymyxin B) was presented by MIC values, with the VITEK 2 AST-GN09 card for the identification of Gram-negative bacilli. The interpretation of MICs was divided into three categories according to the Clinical Laboratory Standard Institute (CLSI) guidelines: antibiotic sensitivity, intermediate, and resistance. One hundred and ten isolates (DSPA = 53 and CRPA = 57) were included in a series of analyses of this study (Fig. 1B).

Bacterial sample preparation and fluorescent labeling

Cells were harvested from bacterial cultures during the mid-exponential phase of growth at an OD₆₀₀ of 1.0. The collected bacterial precipitate was washed 3 times with an equal volume of phosphate-buffered saline (PBS) by centrifugation at $6,000 \times g$ for 5 min at 4°C.

For the preparation of intracellular protein samples, the bacterial cells were first resuspended in 2 mL PBS and added with 1% phenylmethylsulfonyl fluoride. Following sonication, cells were ruptured and then centrifuged at $8,000 \times g$ for 10 min at 4°C to separate unbroken cells and any insoluble debris. The supernatant was filtered twice through a 0.22- μ m filter to completely remove cell debris to obtain a clarified protein extract. One hundred microliters of prepared protein samples was extracted into 100 μ L of 0.1M Na₂CO₃ (pH 9.3) solution and added 5 μ L of 100 mg/mL activated Cy3 fluorescent dye. The mixture was incubated at 100 rpm for 0.5 h at 4°C, protected from the light. Next, 20 μ L of 4 M hydroxylamine hydrochloride solution was added and reacted on ice in the dark for 20 min to terminate the covalent binding between the Cy3 and the amino group on the protein. The fluorescent-labeled protein was separated and purified using the Sephadex G-25 column (GE Healthcare Life Sciences). The protein concentration was quantified using BCA assay, and samples were stored at -20°C until used.

Differently, for the fluorescent labeling of the whole bacteria, we resuspended the bacteria with 200 μ L of 0.1 M Na₂CO₃ (pH 9.3) solution, added 20 μ L of 100 mg/mL activated Cy3 fluorescent dye, and incubated at 80 rpm for 1 h at room temperature protected from light. After the Na₂CO₃ solution containing the free Cy3 dye was removed by centrifugation at $3,000 \times g$ for 5 min, the fluorescently labeled bacterial precipitate was collected. The bacteria were washed five times with 800 μ L PBS to remove the unbound Cy3 dye wrapped between the bacterial precipitates. The fluorescently labeled bacterial cells were then resuspended in 100 μ L PBS and temporarily stored at 4°C protected from the light.

Lectin microarray analysis

Lectin microarray analysis was performed as previously described (65) with minor modifications. Briefly, 37 different lectins were immobilized on slides modified by epoxidation. Each lectin was printed in triplicate spots in four replicate subarrays on each microarray slide. After ensuring that the lectin was firmly adhered to and stable on the chip surface, the lectin microarray was submerged into PBST buffer (10 mM PBS containing 0.2% Tween 20) and washed twice at 80 rpm for 5 min, and twice with PBS. Six hundred microliters of blocking buffer (10 mM PBS; 2% BSA; 500 mM glycine; 0.05% Tween 20) was added to the microarray hybridization cassette, covered with a cleaned lectin microarray and incubated in the microarray hybridization chamber for 1 h at 37°C protected from light. After the blocking step, the slides were washed twice with PBST and PBS, respectively.

Subsequently, for the analysis of bacterial intracellular protein samples, two micrograms of fluorescently labeled protein samples was mixed with incubation solution (10 mM PBS, 2% BSA, 500 mM glycine, 0.05% Tween 20, 10% 4 M hydroxylamine) and then added to a microarray hybridization cassette containing four subarrays (mixture system of 120 μ L for each of microarray sectors), covered with a cleaned lectin microarray, and incubated in the microarray hybridization chamber for 1 h at 37°C protected

from light. After incubation, the hybridized microarrays were washed, spin-dried, and scanned using an AXON GenePix (4100A) fluorescence microarray scanner to read the raw intensity values. For the detection of the whole bacteria, 40 μL of fluorescently labeled bacterial suspension was directly mixed with 90 μL of incubation solution. Then, 120 μL of the mixture was transferred to the microarray hybridization cassette according to the subarray region. The cleaned lectin microarray was covered similarly and incubated in the microarray hybridization chamber at 37°C for 3 h protected from light. After incubation, the hybridized microarrays were gently washed and air-dried in the dark. The final lectin microarray was scanned by AXON GenePix (4100A) scanner and the fluorescence signal value was obtained.

Lectin-based immunofluorescence staining of whole bacteria and confocal microscopy

One hundred microliters of 1 mg/mL lectin solution was mixed with 100 μL of 0.1 M Na_2CO_3 solution (pH 9.3), and 5 μL of 100 mg/mL activated Cy5 fluorescent dye (GE Healthcare) was added to the mixed solution. The mixture was incubated in the dark at room temperature for 2 h. The incubated reaction was terminated by the addition of 20 μL of 4 M hydroxylamine hydrochloride on ice for 20 min. The fluorescent-labeled lectin was purified by Sephadex G-25 column. The concentration was determined by BCA assay, and the labeled lectins were stored at -20°C in darkness until use.

Bacteria were cultured as described above. One milliliter culture (OD_{600} of 1.0) was spun down for 5 min; 4°C; 8,000 rpm. Pellets were washed three times with precooled PBS at equal volume and then fixed in 750 μL of 4% precooled paraformaldehyde for 20 min at 4°C. Subsequently, bacteria were washed 3 times with 800 μL PBS and blocked with 1 mL Carbo-Free blocking solution for 1 h at room temperature. The blocked bacteria were then incubated with Cy5-labeled lectins (10 $\mu\text{g}/\text{mL}$ in Carbo-Free blocking solution) with gentle shaking overnight at 4°C in the dark. The bacteria were washed 3 times with 800 μL PBS and then stained with 4,6-diamidino-2-phenylindol [DAPI (10 $\mu\text{g}/\text{mL}$), Sigma] for 5 min. Subsequently, bacteria were washed 2 times and resuspended with 200 μL PBS. The bacterial cell suspension (80 μL) was plated into a 15-mM glass bottom cell culture dish (NEST, 801002) and dried in the dark.

The Leica SP8 confocal microscope was used to capture the fluorescence signal of bacterial cells with a 63 \times oil objective, and the 638 nm (the Cy5 channel) and 405 nm (the DAPI channel) lasers were opened for fluorescence imaging. The bacterial fluorescence signal was quantitatively analyzed by Image J software.

Silver staining and lectin blotting

Overnight cultures of bacteria were sub-cultured with 1% dilution in fresh LB medium and grown at 37°C with continuous shaking (220 rpm) to an OD_{600} of 1.0. The pelleted bacteria collected by centrifugation were resuspended in 100 μL of 1 \times SDS-loading buffer and incubated at 100°C for 20 min. Allow the solution to cool at room temperature for 15 min. Ten microliters of the prepared bacterial protein sample was separated by 12% SDS-PAGE gels and visualized using silver staining according to the standard protocol. In brief, the SDS gel was immersed in a fixative solution (50% ethanol, 10% glacial acetic acid) at room temperature for 1 h. Afterward, the gel was sensitized with sensitizing solution containing 0.2% sodium-thiosulfate for 30 min with gentle shaking and then washed four times with deionized water (DI water). The gel was stained with 1 g/L AgNO_3 for 20 min and washed with DI water for 10–20 s. Carry out the color reaction for 5 min using a chromogen solution containing 35 g/L Na_2CO_3 and 0.1% formaldehyde. The reaction was terminated by a stop solution (10% glacial acetic acid). The gel was finally washed with DI water and imaged using a gel imaging system (Tanon, Shanghai, China).

For lectin blotting, the bacterial proteins in gels were transferred onto a polyvinylidene difluoride membrane and blocked with Carbo-Free blocking solution for 1 h at room temperature. The blocked membrane was then incubated with Cy5-labeled lectins

(10 µg/mL in Carbo-Free blocking solution) with gentle shaking overnight at 4°C in the dark. The membranes were then washed twice each for 5 min with TBST and scanned by red fluorescence channel (635 nm excitation/650LP emission) with the voltage of 800 PMT using a phosphorimager (Storm 840, Molecular Dynamics). The gray values were calculated by Image J software.

Construction of machine-learning model

In this study, 110 clinical *P. aeruginosa* isolates were divided into the DSPA group (53 strains) and CRPA group (57 strains). The data from the lectin microarray analysis of 110 strains were randomly split into a 70% training set and a 30% test set (66–69). We utilized the open-source Python machine-learning library sci-kit-learn to train and build a machine-learning model based on a GBDT algorithm. We applied fivefold cross-validation to determine the optimal parameters for the training set. With the optimal parameters tuned, 70% of lectin microarray data were used to train machine-learning algorithms to construct predictive models. Then, the performance of the established model was verified using 30% of the lectin microarray data. Calculation of the model performance measurement parameters was achieved in the test cohort with the help of feature importance analysis, confusion matrix, and receiver operating characteristic (ROC) curve analysis. The classification performance of the model was evaluated by calculating the F1 score, precision, recall, and support value (70).

Data pre-processing and statistical analysis

After extracting the accurate microarray analysis data from the scanned images by GenePix Pro 6.0 Microarray Image Analysis Software, the raw data were subjected to normalization. The signal intensity value corresponding to each lectin spot is subtracted from the background signal intensity value of the spot as the effective signal value for each spot. Each lectin corresponds to three replicate spots, and the median of the effective signal values corresponding to the three replicate spots is taken so that each lectin in a subarray region corresponds to a median value. The proportion of the median value of each lectin to the sum of the median values of all the lectins was then calculated to perform the global normalization of the data, i.e., the normalized fluorescent intensities for each lectin. Tables were visualized in Microsoft Excel, and Cluster analysis was performed using Expander 8.0 software. Graphs were generated using GraphPad Prism 8.0 software and statistically analyzed by the *t*-test parametric test, with differences considered statistically significant at *P* value <0.05. In addition, the data were further analyzed by PCA.

ACKNOWLEDGMENTS

This study was supported by grants from the National Natural Science Foundation of China (32101030) and the Natural Science Foundation of Shaanxi Province (2021JQ-446).

AUTHOR AFFILIATIONS

¹Laboratory of Functional Glycomics, College of Life Sciences, Northwest University, Xi'an, Shaanxi, China

²Hospital of Shaanxi Nuclear Industry, Xianyang, Shaanxi, China

AUTHOR ORCIDs

Jing Dang  <http://orcid.org/0009-0000-7608-021X>

Jian Shu  <http://orcid.org/0000-0002-2938-4218>

Huizheng Hu  <http://orcid.org/0000-0002-8850-0864>

Zheng Li  <http://orcid.org/0000-0002-1839-3450>

FUNDING

Funder	Grant(s)	Author(s)
MOST National Natural Science Foundation of China (NSFC)	32101030	Jian Shu
陕西省科学技术厅 Natural Science Foundation of Shaanxi Province (Shaanxi Natural Science Foundation)	2021JQ-446	Jian Shu

AUTHOR CONTRIBUTIONS

Jing Dang, Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft, Writing – review and editing | Ruiying Wang, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation | Hanjie Yu, Formal analysis, Investigation, Resources, Software, Supervision | Zhuo Chen, Conceptualization, Formal analysis, Methodology | Wenbo Yan, Investigation, Resources | Bingxiang Zhao, Investigation | Yuzi Wang, Investigation | Huizheng Hu, Conceptualization, Data curation, Investigation, Project administration, Resources, Supervision, Validation.

REFERENCES

- Kothari A, Kumar SK, Singh V, Kumar P, Kaushal K, Pandey A, Jain N, Omar BJ. 2022. Association of multidrug resistance behavior of clinical *Pseudomonas aeruginosa* to pigment coloration. *Eur J Med Res* 27:120. <https://doi.org/10.1186/s40001-022-00752-6>
- Hirsch EB, Tam VH. 2010. Impact of multidrug-resistant *Pseudomonas aeruginosa* infection on patient outcomes. *Expert Rev Pharmacoecon Outcomes Res* 10:441–451. <https://doi.org/10.1586/erp.10.49>
- Kaye KS, Pogue JM. 2015. Infections caused by resistant gram-negative bacteria. *Epidemiol Management Pharmacother* 35:949–962. <https://doi.org/10.1002/phar.1636>
- Theuretzbacher U. 2017. Global antimicrobial resistance in gram-negative pathogens and clinical need. *Curr Opin Microbiol* 39:106–112. <https://doi.org/10.1016/j.mib.2017.10.028>
- Hassuna NA, Darwish MK, Sayed M, Ibrahim RA. 2020. Molecular epidemiology and mechanisms of high-level resistance to meropenem and imipenem in *Pseudomonas aeruginosa* Infect Drug Resist 13:285–293. <https://doi.org/10.2147/IDR.S233808>
- Sleiman A, Fayad AGA, Banna H, Matar GM. 2021. Prevalence and molecular epidemiology of carbapenem-resistant gram-negative Bacilli and their resistance determinants in the Eastern mediterranean region over the last decade. *J Glob Antimicrob Resist* 25:209–221. <https://doi.org/10.1016/j.jgar.2021.02.033>
- Shrivastava S, Shrivastava P, Ramasamy J. 2018. World health organization releases global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. *J Med Soc* 32:76. https://doi.org/10.4103/jms.jms_25_17
- Yakovlieva L, Fülleborn JA, Walvoort MTC. 2021. Opportunities and challenges of bacterial glycosylation for the development of novel antibacterial strategies. *Front Microbiol* 12:803203. <https://doi.org/10.3389/fmicb.2021.803203>
- Couvigny B, Lapaque N, Rigottier-Gois L, Guillot A, Chat S, Meylheuc T, Kulakauskas S, Rohde M, Mistou M-Y, Renault P, Doré J, Briandet R, Serron P, Guédon E. 2017. Three glycosylated serine-rich repeat proteins play a pivotal role in adhesion and colonization of the pioneer commensal bacterium, *Streptococcus salivarius*. *Environ Microbiol* 19:3579–3594. <https://doi.org/10.1111/1462-2920.13853>
- Tan FYY, Tang CM, Exley RM. 2015. Sugar coating: bacterial protein glycosylation and host-microbe interactions. *Trends Biochem Sci* 40:342–350. <https://doi.org/10.1016/j.tibs.2015.03.016>
- Pellic V. 2008. Type IV Pili: *e pluribus unum* Mol Microbiol 68:827–837. <https://doi.org/10.1111/j.1365-2958.2008.06197.x>
- Smedley JG III, Jewell E, Roguskie J, Horzempa J, Syboldt A, Stolz DB, Castric P. 2005. Influence of pilin glycosylation on *Pseudomonas aeruginosa* 1244 pilus function. *Infect Immun* 73:7922–7931. <https://doi.org/10.1128/IAI.73.12.7922-7931.2005>
- Latousakis D, Juge N. 2018. How sweet are our gut beneficial bacteria? A focus on protein glycosylation in lactobacillus. *Int J Mol Sci* 19:136. <https://doi.org/10.3390/ijms19010136>
- Frese SA, Mackenzie DA, Peterson DA, Schmaltz R, Fangman T, Zhou Y, Zhang C, Benson AK, Cody LA, Mulholland F, Juge N, Walter J. 2013. Molecular characterization of host-specific biofilm formation in a vertebrate gut symbiont. *PLoS Genet* 9:e1004057. <https://doi.org/10.1371/journal.pgen.1004057>
- Aas FE, Vik A, Vedde J, Koomey M, Egge-Jacobsen W. 2007. *Neisseria gonorrhoeae* O-linked pilin glycosylation: functional analyses define both the biosynthetic pathway and glycan structure. *Mol Microbiol* 65:607–624. <https://doi.org/10.1111/j.1365-2958.2007.05806.x>
- Charbonneau ME, Mourez M. 2008. The *Escherichia coli* AIDA-I autotransporter undergoes cytoplasmic glycosylation independently of export. *Res Microbiol* 159:537–544. <https://doi.org/10.1016/j.resmic.2008.06.009>
- Sherlock O, Dobrindt U, Jensen JB, Munk Vejborg R, Klemm P. 2006. Glycosylation of the self-recognizing *Escherichia coli* AG43 autotransporter protein. *J Bacteriol* 188:1798–1807. <https://doi.org/10.1128/JB.188.5.1798-1807.2006>
- Sherlock O, Vejborg RM, Klemm P. 2005. The TibA adhesin/invasin from enterotoxigenic *Escherichia coli* is self recognizing and induces bacterial aggregation and biofilm formation. *Infect Immun* 73:1954–1963. <https://doi.org/10.1128/IAI.73.4.1954-1963.2005>
- Avilés-Reyes A, Miller JH, Simpson-Haidaris PJ, Hagen FK, Abranches J, Lemos JA. 2014. Modification of *Streptococcus mutans* Cnm by PgfS contributes to adhesion, endothelial cell invasion, and virulence. *J Bacteriol* 196:2789–2797. <https://doi.org/10.1128/JB.01783-14>
- Bu S, Li Y, Zhou M, Azadin P, Zeng M, Fives-Taylor P, Wu H. 2008. Interaction between two putative glycosyltransferases is required for glycosylation of a serine-rich streptococcal adhesin. *J Bacteriol* 190:1256–1266. <https://doi.org/10.1128/JB.01078-07>
- Bensing BA, Gibson BW, Sullam PM. 2004. The *Streptococcus gordonii* platelet binding protein GspB undergoes glycosylation independently of export. *J Bacteriol* 186:638–645. <https://doi.org/10.1128/JB.186.3.638-645.2004>
- Zeng L, Burne RA. 2009. Transcriptional regulation of the cellobiose operon of *Streptococcus mutans*. *J Bacteriol* 191:2153–2162. <https://doi.org/10.1128/JB.01641-08>
- Jennings MP, Jen FE-C, Roddam LF, Apicella MA, Edwards JL. 2011. *Neisseria gonorrhoeae* pilin glycan contributes to CR3 activation during challenge of primary cervical epithelial cells. *Cell Microbiol* 13:885–896. <https://doi.org/10.1111/j.1462-5822.2011.01586.x>
- Li S, Zhang L, Yao Q, Li L, Dong N, Rong J, Gao W, Ding X, Sun L, Chen X, Chen S, Shao F. 2013. Pathogen blocks host death receptor signalling by arginine glcnacylation of death domains. *Nature* 501:242–246. <https://doi.org/10.1038/nature12436>
- Newton HJ, Pearson JS, Badea L, Kelly M, Lucas M, Holloway G, Wagstaff KM, Dunstone MA, Sloan J, Whistock JC, Kaper JB, Robins-Browne RM, Jans DA, Frankel G, Phillips AD, Coulson BS, Hartland EL. 2010. The type III effectors NleE and NleB from enteropathogenic *E. coli* and OspZ from

- Shigella* block nuclear translocation of NF-kappaB p65. PLoS Pathog 6:e1000898. <https://doi.org/10.1371/journal.ppat.1000898>
26. Nadler C, Baruch K, Kobi S, Mills E, Haviv G, Farago M, Alkalay I, Bartfeld S, Meyer TF, Ben-Neriah Y, Rosenshine I. 2010. The type III secretion effector NleE inhibits NF-kappaB activation. PLoS Pathog 6:e1000743. <https://doi.org/10.1371/journal.ppat.1000743>
 27. Bastos PAD, da Costa JP, Vitorino R. 2017. A glimpse into the modulation of post-translational modifications of human-colonizing bacteria. J Proteomics 152:254–275. <https://doi.org/10.1016/j.jprot.2016.11.005>
 28. Schmidt MA, Riley LW, Benz I. 2003. Sweet new world: glycoproteins in bacterial pathogens. Trends Microbiol 11:554–561. <https://doi.org/10.1016/j.tim.2003.10.004>
 29. Lu Q, Yao Q, Xu Y, Li L, Li S, Liu Y, Gao W, Niu M, Sharon M, Ben-Nissan G, Zamyatina A, Liu X, Chen S, Shao F. 2014. An iron-containing dodecameric heptosyltransferase family modifies bacterial autotransporters in pathogenesis. Cell Host Microbe 16:351–363. <https://doi.org/10.1016/j.chom.2014.08.008>
 30. Lu Q, Li S, Shao F. 2015. Sweet talk: protein glycosylation in bacterial interaction with the host. Trends Microbiol 23:630–641. <https://doi.org/10.1016/j.tim.2015.07.003>
 31. Wu H, Zeng M, Fives-Taylor P. 2007. The glycan moieties and the N-terminal polypeptide backbone of a fimbria-associated adhesin, Fap1, play distinct roles in the biofilm development of *Streptococcus parasanguinis*. Infect Immun 75:2181–2188. <https://doi.org/10.1128/IAI.01544-06>
 32. Lucchese A. 2017. *Streptococcus mutans* antigen I/II and autoimmunity in cardiovascular diseases. Autoimmun Rev 16:456–460. <https://doi.org/10.1016/j.autrev.2017.03.009>
 33. Alemka A, Nothaft H, Zheng J, Szymanski CM. 2013. N-Glycosylation of *Campylobacter jejuni* surface proteins promotes bacterial fitness. Infect Immun 81:1674–1682. <https://doi.org/10.1128/IAI.01370-12>
 34. Pang Z, Raudonis R, Glick BR, Lin TJ, Cheng Z. 2019. Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and alternative therapeutic strategies. Biotechnol Adv 37:177–192. <https://doi.org/10.1016/j.biotechadv.2018.11.013>
 35. Flemming HC, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. 2016. Biofilms: an emergent form of bacterial life. Nat Rev Microbiol 14:563–575. <https://doi.org/10.1038/nrmicro.2016.94>
 36. Becker K, Haldemann K, Selchow P, Reinau LM, Dal Molin M, Sander P. 2017. Lipoprotein glycosylation by protein-O-mannosyltransferase (MAB_1122C) contributes to low cell envelope permeability and antibiotic resistance of *Mycobacterium abscessus* Front Microbiol 8:2123. <https://doi.org/10.3389/fmicb.2017.02123>
 37. Abouelhadid S, Raynes J, Bui T, Cuccui J, Wren BW. 2020. Characterization of posttranslationally modified multidrug efflux pumps reveals an unexpected link between glycosylation and antimicrobial resistance. mBio 11:e02604-20. <https://doi.org/10.1128/mBio.02604-20>
 38. Pilobello KT, Krishnamoorthy L, Slawek D, Mahal LK. 2005. Development of a lectin microarray for the rapid analysis of protein glycopatterns. ChemBiochem 6:985–989. <https://doi.org/10.1002/cbic.200400403>
 39. Rosenfeld R, Bangio H, Gerwig GJ, Rosenberg R, Aloni R, Cohen Y, Amor Y, Plaschkes I, Kamerling JP, Maya RB-Y. 2007. A lectin array-based methodology for the analysis of protein glycosylation. J Biochem Biophys Methods 70:415–426. <https://doi.org/10.1016/j.jbbm.2006.09.008>
 40. Campanero-Rhodes MA, Palma AS, Menéndez M, Solís D. 2019. Microarray strategies for exploring bacterial surface glycans and their interactions with glycan-binding proteins. Front Microbiol 10:2909. <https://doi.org/10.3389/fmicb.2019.02909>
 41. Hsu KL, Pilobello KT, Mahal LK. 2006. Analyzing the dynamic bacterial glycome with a lectin microarray approach. Nat Chem Biol 2:153–157. <https://doi.org/10.1038/nchembio767>
 42. Yasuda E, Sako T, Tateno H, Hirabayashi J. 2014. Application of lectin microarray to bacteria including *Lactobacillus casei/paracasei* strains. Methods Mol Biol 1200:295–311. https://doi.org/10.1007/978-1-4939-1292-6_25
 43. Kilcoyne M, Twomey ME, Gerlach JQ, Kane M, Moran AP, Joshi L. 2014. *Campylobacter jejuni* strain discrimination and temperature-dependent glycome expression profiling by lectin microarray. Carbohydr Res 389:123–133. <https://doi.org/10.1016/j.carres.2014.02.005>
 44. Morita Y, Tomida J, Kawamura Y. 2014. Responses of *Pseudomonas aeruginosa* to antimicrobials. Front Microbiol 4:422. <https://doi.org/10.3389/fmicb.2013.00422>
 45. Kurian B, Jyothi VL. 2022. Comparative analysis of machine learning methods for breast cancer classification in genetic sequences. J Environ Public Health 2022:7199290. <https://doi.org/10.1155/2022/7199290>
 46. Li K, Shi Q, Liu S, Xie Y, Liu J. 2021. Predicting in-hospital mortality in ICU patients with sepsis using gradient boosting decision tree. Medicine 100:e25813. <https://doi.org/10.1097/MD.00000000000025813>
 47. Zhang Y, Zhang R, Ma Q, Wang Y, Wang Q, Huang Z, Huang L. 2020. A feature selection and multi-model fusion-based approach of predicting air quality. ISA Trans 100:210–220. <https://doi.org/10.1016/j.isatra.2019.11.023>
 48. Hashim OH, Jayapalan JJ, Lee CS. 2017. Lectins: an effective tool for screening of potential cancer biomarkers. PeerJ 5:e3784. <https://doi.org/10.7717/peerj.3784>
 49. Syed P, Gidwani K, Kekki H, Leivo J, Pettersson K, Lamminmäki U. 2016. Role of lectin microarrays in cancer diagnosis. Proteomics 16:1257–1265. <https://doi.org/10.1002/pmic.201500404>
 50. Shu J, Ma J, Ren X, Wang J, Wang Y, Zhang K, Yu H, Guo X, Li Z. 2021. The abnormal glycopatterns of salivary glycoproteins in esophageal squamous cell carcinoma patients. Front Chem 9:637730. <https://doi.org/10.3389/fchem.2021.637730>
 51. Liu L, Li D, Shu J, Wang L, Zhang F, Zhang C, Yu H, Chen M, Li Z, Guo X. 2020. Protein glycopatterns in bronchoalveolar lavage fluid as novel potential biomarkers for diagnosis of lung cancer. Front Oncol 10:568433. <https://doi.org/10.3389/fonc.2020.568433>
 52. Liu X, Yu H, Qiao Y, Yang J, Shu J, Zhang J, Zhang Z, He J, Li Z. 2018. Salivary glycopatterns as potential biomarkers for screening of early-stage breast cancer. EBioMedicine 28:70–79. <https://doi.org/10.1016/j.ebiom.2018.01.026>
 53. Ren X, Shu J, Wang J, Guo Y, Zhang Y, Yue L, Yu H, Chen W, Zhang C, Ma J, Li Z. 2022. Machine learning reveals salivary glycopatterns as potential biomarkers for the diagnosis and prognosis of papillary thyroid cancer. Int J Biol Macromol 215:280–289. <https://doi.org/10.1016/j.ijbiomac.2022.05.194>
 54. Tang Z, Zhang F, Wang Y, Zhang C, Li X, Yin M, Shu J, Yu H, Liu X, Guo Y, Li Z. 2022. Diagnosis of hepatocellular carcinoma based on salivary protein glycopatterns and machine learning algorithms. Clin Chem Lab Med 60:1963–1973. <https://doi.org/10.1515/cclm-2022-0715>
 55. Dang K, Zhang W, Jiang S, Lin X, Qian A. 2020. Application of lectin microarrays for biomarker discovery. ChemistryOpen 9:285–300. <https://doi.org/10.1002/open.201900326>
 56. Lam JS, Taylor VL, Islam ST, Hao Y, Kocincová D. 2011. Genetic and functional diversity of *Pseudomonas aeruginosa* lipopolysaccharide. Front Microbiol 2:118. <https://doi.org/10.3389/fmicb.2011.00118>
 57. Scott NE, Parker BL, Connolly AM, Paulech J, Edwards AVG, Crossett B, Falconer L, Kolarich D, Djordjevic SP, Højrup P, Packer NH, Larsen MR, Cordwell SJ. 2011. Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*. Mol Cell Proteomics 10:M000031-MCP201. <https://doi.org/10.1074/mcp.M000031-MCP201>
 58. Jervis AJ, Wood AG, Cain JA, Butler JA, Frost H, Lord E, Langdon R, Cordwell SJ, Wren BW, Linton D. 2018. Functional analysis of the *Helicobacter pullorum* N-linked protein glycosylation system. Glycobiology 28:233–244. <https://doi.org/10.1093/glycob/cwx110>
 59. Abouelhadid S, North SJ, Hitchen P, Vohra P, Chintoan-Uta C, Stevens M, Dell A, Cuccui J, Wren BW. 2019. Quantitative analyses reveal novel roles for N-glycosylation in a major enteric bacterial pathogen. mBio 10:e00297-19. <https://doi.org/10.1128/mBio.00297-19>
 60. Verma A, Schirm M, Arora SK, Thibault P, Logan SM, Ramphal R. 2006. Glycosylation of B-type flagellin of *Pseudomonas aeruginosa*: structural and genetic basis. J Bacteriol 188:4395–4403. <https://doi.org/10.1128/JB.01642-05>
 61. Ma B, Simala-Grant JL, Taylor DE. 2006. Fucosylation in prokaryotes and eukaryotes. Glycobiology 16:158R–184R. <https://doi.org/10.1093/glycob/cwl040>
 62. Brzezinski K, Dauter Z, Jaskolski M. 2012. Structures of NodZ alpha1,6-fucosyltransferase in complex with GDP and GDP-fucose. Acta

- Crystallogr D Biol Crystallogr 68:160–168. <https://doi.org/10.1107/S0907444911053157>
63. Rabinovich GA, van Kooyk Y, Cobb BA. 2012. Glycobiology of immune responses. *Ann N Y Acad Sci* 1253:1–15. <https://doi.org/10.1111/j.1749-6632.2012.06492.x>
 64. Carlin AF, Lewis AL, Varki A, Nizet V. 2007. Group B streptococcal capsular sialic acids interact with siglecs (immunoglobulin-like lectins) on human leukocytes. *J Bacteriol* 189:1231–1237. <https://doi.org/10.1128/JB.01155-06>
 65. Du H, Yu H, Yang F, Li Z. 2021. Comprehensive analysis of glycosphingolipid glycans by lectin microarrays and MALDI-TOF mass spectrometry. *Nat Protoc* 16:3470–3491. <https://doi.org/10.1038/s41596-021-00544-y>
 66. Noorbakhsh J, Farahmand S, Foroughi Pour A, Namburi S, Caruana D, Rimm D, Soltanieh-Ha M, Zarringhalam K, Chuang JH. 2020. Deep learning-based cross-classifications reveal conserved spatial behaviors within tumor histological images. *Nat Commun* 11:6367. <https://doi.org/10.1038/s41467-020-20030-5>
 67. Li Z, Kaiser L, Holzgreve A, Ruf VC, Suchorska B, Wenter V, Quach S, Herms J, Bartenstein P, Tonn JC, Unterrainer M, Albert NL. 2021. Prediction of TERTp-mutation status in IDH-wildtype high-grade gliomas using pre-treatment dynamic [¹⁸F]FET PET radiomics. *Eur J Nucl Med Mol Imaging* 48:4415–4425. <https://doi.org/10.1007/s00259-021-05526-6>
 68. Singhanian A, Verma R, Graham CM, Lee J, Tran T, Richardson M, Lecine P, Leissner P, Berry MPR, Wilkinson RJ, Kaiser K, Rodrigue M, Woltmann G, Haldar P, O'Garra A. 2018. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. *Nat Commun* 9:2308. <https://doi.org/10.1038/s41467-018-04579-w>
 69. Hu Y, Zhong J, Xiao Y, Xing Z, Sheu K, Fan S, An Q, Qiu Y, Zheng Y, Liu X, Fan G, Liu Y. 2020. Single-cell RNA cap and tail sequencing (scRCAT-seq) reveals subtype-specific isoforms differing in transcript demarcation. *Nat Commun* 11:5148. <https://doi.org/10.1038/s41467-020-18976-7>
 70. Zhang Q, Yang L, Zhou F. 2021. Attention enhanced long short-term memory network with multi-source heterogeneous information fusion: an application to BGI genomics. *Inf Sci (N Y)* 553:305–330. <https://doi.org/10.1016/j.ins.2020.10.023>