Spatial transcriptomics identifies novel *P. aeruginosa* virulence factors

Hao Zhou[1], Oscar Negrón[1*], Serena Abbondante[2*], Michaela Marshall[2*], Brandon Jones[1], Edison Ong[1], Nicole Chumbler[1], Christopher Tunkey[1], Groves Dixon[1], Haining Lin[1], Obadiah Plante[1**], Eric Pearlman[2**], and Mihaela Gadjeva[1]

Affiliations

[1]Moderna, Inc., 325 Binney St, Cambridge, MA 02142

[2]University of California Irvine, School of Medicine, 843 Health Sciences Rd, Irvine, CA 92697

***,**These authors have contributed equally**

Corresponding author

Mihaela Gadjeva, PhD

Moderna Inc., 325 Binney St, Cambridge, MA 02142

Email: mgadjeva@modernatx.com

Abstract

The interactions of a pathogen with the host take place within the confines of infected tissues. However, current technologies do not allow for contextual characterization of infection. We report a unified spatial transcriptomic analysis detecting pathogen- and host-transcripts simultaneously. We showcase our approach using a *P. aeruginosa*-induced ocular infection, where we observed differential enrichment of host and bacterial transcripts at specific anatomical sites. The ridge regression model, trained using a minimum of 150 host gene features, accurately predicted bacterial burden in tissue regions. The tissue-in-depth gene expression enrichment analysis identified bacterial transcript PA2590, encoding a currently uncharacterized gene as deeply penetrant. Infection experiments with the PA2590 transposon deletion mutant strain caused less disease than the WT strain. Comparative structural analysis identified the PA2590 gene product as an iron-scavenging and cobalamin transporter. Cumulatively, our data highlight coordinated spatial interplay between the host and the pathogen representing an approach to identify novel virulence traits.

36 **Introduction**

37 *Pseudomonas aeruginosa* is an opportunistic Gram-negative bacterium that belongs to the group
38 of ESKAPE pathogens, posing a growing challenge for treatment due to the rapidly increasing
39 acquisition of antimicrobial resistance (AMR) (1). Our preliminary analysis points to a 70%
40 increase in antimicrobial resistance acquisition among *P. aeruginosa* clinical isolates globally,
41 highlighting it as a major health concern. *P. aeruginosa* causes a diverse range of infections
42 including cystic fibrosis, wound and burn related skin infections, urinary tract infections, and
43 acute pneumonia (2). *P. aeruginosa* is also a major cause of contact-lens and trauma-induced
44 corneal infections worldwide. While intact corneas are typically resistant to infection, ocular
45 trauma and poor contact lens hygiene are risk factors that can allow bacterial penetration to the
46 corneal stroma (3). Infection progresses very rapidly and can lead to tissue damage that
47 manifests as severe ocular pain, visual impairment, and can cause permanent blindness if left
48 untreated (4).

49 Currently, there are no current available vaccines against *P. aeruginosa.* Most of the
50 preclinical studies leveraged individual virulence factors to generate subunit-based vaccine
51 candidates including T3SS components (e.g, PcrV), polysaccharides (e.g, alginate, psl, LPS),
52 outer membrane proteins (e.g., OprF), detoxified toxins (e.g., ExoA), or flagella (5), (6), (7).
53 Despite the promising initial preclinical data, none of the clinical trials successfully showed
54 protection, highlighting the need for new discovery pipelines to generate protective immunity in
55 susceptible populations.

56 Despite the increase in the availability of genomic data from laboratory strains and
57 clinical isolates, there are only two studies using reverse vaccinology pipelines. Bianconi et al.
58 reported the complete PAO1 proteome and predicted 718 surface-exposed and conserved

59 antigens, identifying 52 potential vaccine candidates (8). That study used only a small number of
60 clinical isolates limiting its impact given the genomic diversity of *P. aeruginosa*. In contrast,
61 Solanaki et.al. utilized 1191 predicted proteomes in their vaccine nomination pipeline and
62 generated a concatemeric vaccine candidate based on a series of *in silico* analysis including
63 location, hydrophilicity, essentiality, antigenicity, immunogenicity, and MHC class I and II
64 epitope predictions (9). However, there have been no follow-up reports addressing the efficacy
65 of the vaccine candidates.

66 Another limitation of the reverse vaccinology pipelines is the lack of consideration for
67 bacterial transcriptional plasticity, functional redundancy of metabolically adaptative
68 mechanisms, and temporal transcriptional changes that the pathogen undergoes during infection.
69 Bacterial transcriptome analysis demonstrated that *P. aeruginosa* transcriptional profiles depend
70 on the host and/or the site of infection. For example, *in vitro* bulk transcriptional responses of
71 bacteria taken from growth media were different from the *in vivo* responses in the infected lungs
72 (10,11). Exposure to Cystic Fibrosis sputum induced a distinct transcriptional profile, illustrating
73 significant transcriptional plasticity (12). Cumulatively, a detailed understanding of the
74 transcriptional behavior of the pathogen may identify new antigens for vaccine development and
75 offer an approach for antigen prioritization.

76 To achieve global profiling of host-pathogen interactions, we leveraged the resources of
77 spatial transcriptomics technology and simultaneously characterized transcriptional responses of
78 both the host and the pathogen *in situ*. We based these studies on the hypothesis that by gaining
79 insights into the host-pathogen specific temporal and spatial transcriptional gradients, new patho-
80 adaptative mechanisms can be defined.
81

82 **Results**

83 **Spatial transcriptomics reveals tissue-segment specific host transcript enrichment.**

84 To generate a spatial transcriptional profile of infection, we used a well-characterized
85 murine model of *P. aeruginosa*- corneal infection. We utilized the Nanostring GeoMx platform
86 for non-biased characterization of the host transcriptome (Fig. 1A). Sections from paraffin-
87 embedded eyes were processed and we examined the responses in the corneal epithelium, the
88 corneal stroma, and the anterior chamber as regions of interest (ROIs). Workflow is shown in
89 Fig. 1A.

90 Corneal epithelial cells were identified as keratin 6^{+}ve; neutrophils were identified as
91 $Ly6G^{+}$ve. Uninfected (contralateral eyes) and infected ocular tissues were examined (Fig. 1B).
92 Transcriptome analysis revealed common and segment-specific enrichment of host and bacterial
93 transcripts. Consistent with the rich neutrophil infiltration into the infected tissues, neutrophil-
94 specific S100a8 transcripts were highly upregulated in the epithelium and stroma (Fig. 1C). Also
95 upregulated were neutrophil and monocyte chemokines Cxcl2, and Ccl3, pro-inflammatory
96 cytokines Il-1a and Il-1b, and the neutrophil iron chelating gene Lcn2. Cxcl2, Ccl3, Csf3, Cxcr4,
97 and Il-1b, Il-1a, Il-12a cytokine transcripts were higher in the epithelium relative to the stroma
98 (Fig. 1C). These data were reflected in the Gene Ontology (GO) enrichment analysis showing
99 stronger granulocyte chemotaxis enrichment in epithelium versus stroma.

100 In contrast, transcripts associated with RNA metabolism, processing, ncRNA processing,
101 and chromatin remodeling were predominantly upregulated in the stromal layer when compared
102 to epithelium. Surprisingly, the infected stroma showed elevated anti-viral responses (Fig. 1D),
103 including Type I IFN, IFNAR, and others (Suppl. Table 1).

104 After 48 hours of infection rapid disease progression was observed, as illustrated by
105 elevated corneal opacity of the infected corneas (Suppl. Fig. 1A and B). The epithelial layer
106 displayed enrichment of transcripts associated with metabolic alterations including glucose
107 catabolism and cellular responses to glucose starvation, translation, and RNA processing (Suppl.
108 Fig. 1D and Suppl. Table 2). In the stroma, we found sustained enrichment of transcripts
109 associated with neutrophil recruitment (chemotaxis and migration), glucose catabolism, pyruvate
110 metabolism and apoptosis (Suppl. Fig. 1C and Suppl. Table 2). Cumulatively, our data provide a
111 temporal transcriptional and spatial record of tissue alterations during infection that are
112 associated with neutrophil cytokine production and metabolic changes.
113

114 **Spatial analysis of bacterial transcript abundance in infected corneas reveals distinct**
115 **enrichment profiles between the epithelium and stroma**

116 Probes were generated for 100 *P. aeruginosa* genes that were encoded outer membrane proteins,
117 extracellular proteins, reported virulence factors, and unkown, predicted genes. Probe
118 development was based on the criteria of ORF frequency in >11,000 clinical isolates,
119 evolutionary conservation, and their relation to known virulence. Three different gene-specific
120 probes were designed for each of the bacterial genes of interest and used simultaneously to
121 identify bacterial transcripts in the infected tissues (Suppl. Table 3).
122 Transcript expression was examined in PAO1 infected eyes in the positive corneal
123 epithelium compared with the stroma and anterior chamber (AC). Representative images
124 showed that infection caused loss of the cytokeratin staining in some of the surface layers of the
125 cornea, consistent with the corneal perforation (ulceration) that occurs following infection. The
126 underlying stromal layers showed rich $Ly6G^{+}$ staining, indicative of neutrophil infiltration to

127 infected cornea (Fig. 2B). Overall, we could distinguish ROIs with high and low bacterial
128 transcript presence. This was reflective of the differences in the number of bacteria at distinct
129 sites in the infected eye. While 97 transcripts were detected in almost half of all examined
130 surface segments, only 37 transcripts were detected in most of the stromal segments.
131 Interestingly, nine transcripts were consistently detected in the cytokeratin^{+} and cytokeratin^{-}
132 surface layers, Ly6G^{+} and Ly6G^{-} stromal regions, and AC segments, including AhpC, TufB,
133 PhzD1, ArcA, FpvB, PA2590, IcmP, LipA, and FemA (Fig. 2A and B).

134 Bacterial transcript enrichment analysis, normalized to the housekeeping RplL transcript
135 abundance, reflected transcript increases in the infected stromal segments when compared to the
136 infected surface layers (Fig. 2C). Nineteen transcripts showed significant upregulation in the
137 infected tissues (Suppl. Table 4). Those included transcripts with known functions such as ArcA,
138 PhzD1, IcmP, OprC, AhpC, LipA, ExoT, PhuR, OpdH, FpvB, and FemA, and transcripts
139 encoding proteins that are predicted but have not been characterized: PA2590, PA0439, PA2204,
140 PA2291, and PA4881. In contrast, transcripts for AlgE, PopD, and Piv were decreased in the
141 stroma, while ExoS and FecA showed no enrichment (Fig. 2C and Suppl. Table 4).

142 To determine if bacterial transcript tissue penetrance is strain-specific, corneas were
143 infected with the cytotoxic (ExoU expressing) PA14 strain. No significant enrichment of
144 bacterial transcripts was observed at 24h post-challenge (data not shown); however, at 48h post-
145 challenge, deeper bacterial transcript penetrance and significant changes were detected. Fourteen
146 transcripts were similarly enriched between the PA14 and PAO1 data sets. These included
147 PhzD1, FpvB, PA2590, IcmP, ArcA, FemA, LipA (Fig. 3B, C, and D). The transcriptional heat
148 map comparisons between the sham and infected tissues confirmed the specificity of bacterial
149 transcript detection (Fig. 3A and Suppl. Table 5). Cumulatively, these results highlight the

150 appearance of bacterial transcriptional patterns in the infected tissues irrespective of the strain
151 types and thereby illustrate spatial organization during infection.
152
153 **Harnessing host-bacteria dynamics for accurate prediction of infection levels in tissue**
154 **regions.**
155 Next, leveraging the host transcriptional data, we examined its functional relationship
156 with the *P. aeruginosa* housekeeping gene RplL, which we used as a correlate for the number of
157 bacteria in the infected tissues. Using Ridge Regression analysis (Fig. 4A) we demonstrated
158 significant predictive accuracy, achieving an R^2 score of 0.923. The selection of 600 key host
159 gene transcript features (Fig. 4B) refined the model, enhancing our understanding of the spatial
160 distribution of bacterial burden across infection sites.
161 We found that 150 host features were sufficient to accurately predict bacterial burden
162 measured based on RplL transcript abundance, with an R^2 score of >0.8. We generated a host-
163 bacterial RplL transcript correlation network and examined it for known host protein interactions
164 (Fig. 4C). We found multiple signaling mediators associated with type I IFN, IL-1a, and TNF-
165 driven inflammatory responses, which were highly interconnected for protein-protein
166 interactions, further indicating the observation that the identified host response network reflected
167 highly coordinated immune mechanisms (Fig. 4C). Examples for individual strong correlations
168 between *P. aeruginosa* RplL and TNF-α, Trim13, and phorbol-12-myristate-13-acetate-induced
169 protein1 (Pmaip1) are shown (Fig. 4D), and demonstrate that this approach can identify known
170 (e.g., Il-1α, TNF-α) and unexpected immune determinants of infection including Trim13 and
171 Pmaip1. Employing 'leave-one-out' validation on our model with a set of mouse hosts infected
172 by PAO1 or PA14 strain, as illustrated in Suppl. Figure 2, showcases its adaptability and

173 potential for broad applicability. This approach involves training the model on data from all but
174 one host and assessing its predictive accuracy on the excluded host. The successful application of
175 this technique across biological replicates demonstrates the model's robust predictive capacity.
176 Cumulatively, our approach offers a detailed perspective on the variability of bacterial
177 load across tissue regions and demonstrated the ability to predict bacterial tissue burden by select
178 host transcript feature abundance, contributing to the field of precision medicine in diagnosing
179 bacterial infections.
180
181 **Infection with the PA2590Tn deletion mutant strain is associated with reduced pathology**
182 **and bacterial burden.**
183 Bacterial enrichment analysis identified known and unknown *P. aeruginosa* transcripts
184 that were enriched in the infected tissues. To determine the functional importance of the
185 predicted PA2590 gene that showed significant penetrance in the infected tissues from PAO1 and
186 PA14-challenged mice, corneas were infected with *P. aeruginosa* PA14 WT parental strain or the
187 PA2590 transposon mutant (PA2590Tn). Infected eyes were imaged, and corneal opacity, was
188 imaged and quantified (Fig. 5). We also assessed viable bacteria by CFU. Since we established
189 similar transcriptional enrichment for the PA2590 in the infected tissues of PAO1 and PA14-
190 induced infections, infection studies were carried out with the PA14 strain.
191 There were no significant differences in corneal opacity or bacterial count at 24h post-
192 challenge between the WT PA14 strain and PA2590Tn, consistent with the lack of enrichment for
193 PA2590 at this data point (data not shown). However, the PA2590Tn mutant showed decreased
194 virulence at 48h post-challenge as we recovered significantly less bacteria when compared to
195 WT infected tissues (Fig. 5B, C, and D). As bacterial burden at 48h post-challenge is typically

196 predictive of the disease outcomes in this model, we concluded that infection with the PA2590Tn
197 mutant causes less severe disease.
198
199 **Structural comparison analysis suggests that PA2590 transports iron or vitamin B12**.
200 To understand how PA2590 might affect *P. aeruginosa* virulence, we examined its
201 prevalence and conservation, and conducted structural studies. Analysis for PA2590 prevalence
202 in 11,447 publicly available *P. aeruginosa* genomes derived from clinical isolates revealed that
203 the PA2590 ORF is widely distributed and present in at least 97% of all clinical isolates (Fig.
204 6A). There were no significant differences in the PA2590 distribution based on geographical
205 location of the isolates. Amino acid conservation analysis was also performed using a library of
206 genomes derived from clinical isolates from different clinical presentations such as wound,
207 blood, respiratory, eye, and genitourinary tract infections. A high degree of conservation was
208 identified in all isolates at the amino acid stretches encoding the extracellular loops of the
209 predicted structure of PA2590 with AlphaFold2 (Fig. 6B) (13). Cumulatively, these data
210 suggested an important biological function for PA2590.
211 Comparative structural analysis established a high degree of conformational similarities
212 between PA2590 and *the E. coli* BtuB protein (PDB: 3m8d; R-Value Free: 0.275) present in *E.*
213 *coli* (Fig. 6C). BtuB is a vitamin B12 (VB12) transporter, suggesting that PA2590 might have a
214 similar function. Docking experiments revealed that PA2590 is likely to bind VB12 (Fig. 6E,
215 Vina docking: -9.1 kcal/mol). For comparison, the docking score for BtuB and VB12 was -8.2
216 kcal/mol. In addition to BtuB, our analysis identified structural homology between PA2590 and
217 *E. coli* FepA (PDB: 1fep; R-Value Free: 0.282) (Fig. 6D). FepA is an iron and VB12 transporter
218 (14). The acquisition of iron depends on ferric enterobactin binding to FepA where the

219 transporter preferentially recognizes the catecholate part of the complex (15). Unlike *E. coli*, our
220 docking experiments did not indicate a potential interaction between enterobactin and PA2590.
221 However, mammalian siderophores can be synthesized by eukaryotic homologues to the *E. coli*
222 enzyme generating enterobactin (16). Therefore, we examined the potential binding of PA2590 to
223 an abundant host siderophore 2,5-dihydroxybenzoic acid (2,5-DHBA). The *in silico* docking
224 experiments confirmed the predicted interaction (Suppl. Fig. 3C, Vina docking: -6.0 kcal/mol).
225 *P. aeruginosa* uses two main siderophores for iron uptake: pyoverdine and pyochelin
226 (17). As both known receptors were upregulated in the infected tissues and to determine if there
227 might be redundancy in iron uptake, the potential interaction of PA2590 with Fe(III)-pyochelin
228 was examined. The experiments showed a significant docking score of -8.2 kcal/mol, indicative
229 of a role for PA2590 in the internalization of iron complexed with *Pseudomonas*-derived
230 siderophores (Fig. 6F). Cumulatively, our data indicate that PA2590 mediates both iron and
231 VB12 transport.
232
233 **Discussion**
234 Spatial transcriptomic profiling of infected tissues provides specific information on gene
235 expression at specific sites of infection. Several different approaches were developed in the last
236 decade to examine spatial proteomics and transcriptomics focusing on characterizing eukaryotic
237 tissues. These tools have not been applied to study microbial pathogenesis. To the best of our
238 knowledge, this report is the first to use dual spatial RNA-seq, combining bacterial pathogen-
239 specific and host transcriptome data with spatial resolution into a unified, multidimensional
240 dataset. This experimental approach represents a significant advancement over previous studies
241 that used either sequential analysis of eukaryotic and prokaryotic features or reported bacterial

242 analysis based only on 16S RNA seq (18,19). Our data illuminate differential enriched bacterial
243 transcripts at the host-pathogen interface, thereby identifying potentially novel patho-adaptative
244 mechanisms. Acquisition of this type of data enable machine learning approaches to connect host
245 transcriptional networks with bacterial profiles, thereby identifying key nodes for targeted
246 therapies.

247 Our study uncovered a strong, predictive correlation between a network of interconnected
248 host features and the *P. aeruginosa* house-keeping transcript (RplL) during infection, offering
249 insights into the biological processes underpinning infection dynamics. The abundance of RplL
250 transcripts was used as a measure for bacterial burden during infection, and successfully
251 predicted bacterial abundance in each tissue with a significant degree of accuracy using at least
252 150 host features. Correlation analysis identified a network of IFN-inducible features: IL-1α,
253 Pmaip1, TNFα, and Trim 13. The network comprises mediators with known contributors to *P.*
254 *aeruginosa*-induced infection (e.g., IL-1α and TNF) and the previously unrecognized mediators
255 (TRIM13 and Chit1).

256 *P. aeruginosa* is a potent inducer of IL-1α and TNFα, cytokines that orchestrate
257 neutrophil infiltration to the infected site (20,21,22). Loss of IL-1α, unlike the loss of IL-1β,
258 promotes resistance to infection (23,24), suggesting that the network identifies key features
259 mediating infection outcomes. Type I IFN induced mediators inhibit IL-1α responses, and,
260 conversely, IL-1α potently antagonizes Type I IFN by regulating IFNβ transcription and
261 translation (25). In the context of infection, when inflammation onset is acute, it is likely that
262 disease outcomes are determined by IL-1α/Type I IFN balance, which are reflected in the
263 predictive analysis.

264 Type I IFN responses are induced by the cGAS/STING pathway during *P. aeruginosa*-
265 induced infection and deficiencies in either cGAS or STING promote susceptibility to infection
266 (26,27). This pathway controls autophagy, a key determinant of disease outcome (16,28).
267 Functionally, type I IFN prevents *P. aeruginosa*-induced tissue inflammation, thereby preserving
268 tissue integrity and facilitating bacterial clearance (26,29), further, reduced Type I IFN
269 responsiveness by cystic fibrosis (CF) epithelial cells promotes bacterial colonization (30). In our
270 analysis, increase in TRIM13 and Pmaip1 is consistent with known downstream events from
271 Type I IFN signaling (31). TRIM13 inhibits Type I IFNs (26), whereas Pmaip1 mediates IFN-
272 driven keratocyte and neutrophil apoptosis by mcl-1 degradation (32). Cumulatively, these data
273 identified a complex interplay between IL-1 and Type I IFN responses which predict
274 pathogenesis.

275 Pathogen biofilm gene expression is a function of bacterial location within the biofilm
276 (33). Therefore, we predicted that the bacterial transcriptional characteristics will show distinct
277 features depending on location. Consistent with this hypothesis, bacterial enrichment analysis
278 exhibited a higher tissue penetrance of transcripts associated with iron acquisition. *P. aeruginosa*
279 requires iron to detoxify reactive oxygen species (ROS), and nitric oxide (NO) and facilitates
280 pathogen survival. *P. aeruginosa* makes two major siderophores: pyoverdine and pyochelin, and
281 the latter is essential for acute infection (34). Transcripts for the pyoverdine receptor, FpvB, and
282 ferric mycobactin receptor, FemA were significantly enriched, although the pyochelin receptor
283 FptA did not show significant enrichment in our analysis and suggesting that there are other
284 mediators of pyochelin import.

285 Interestingly, we observed a significant penetrance of the PA2590 transcript encoding a
286 predicted protein. The structural analysis indicated that PA2590 belongs to a 22-stranded

287 antiparallel β-barrel domain containing Ton B-dependent transporters and the high degree of
288 structural similarity to BtuB and FemA suggested a possible role for PA2590 in mediating iron
289 intake. Our docking experiments reinforce this prediction by indicating its possible binding to
290 pyochelin. The significance of these data is highlighted by the reduction in the *P. aeruginosa*
291 virulence in the absence of PA2590. Cumulatively, our spatial transcriptome enrichment analysis
292 approach allowed us to identify novel virulence mechanisms associated with key requirements
293 for pathogen survival during infection.
294 In conclusion, we demonstrate the feasibility to spatially monitor pathogen transcriptome
295 profiles to identify interconnected host-pathogen networks and novel virulence effectors.
296
297 **Methods**
298 **Bacterial strains**
299 *P. aeruginosa* strains PAO1, PA14, PA14_30590_Tn (PA2590 transposon mutant) were grown in
300 Luria–Bertoni broth (LB) broth (BD-Difco) overnight at 37°C with agitation (200rpm). Cultures
301 were then passaged in fresh LB broth (BD-Difco) until an OD_{600} of 0.20 (midlog phase) was
302 reached. The PAO1 strain was obtained from Dr. A. Rietsch (Case Western Reserve University).
303 The PA14 and PA14_2590_Tn were generated during the PA14 transposon insertion mutant
304 library generation ((http://ausubellab.mgh.harvard.edu/cgi-bin/pa14/home.cgi)) and obtained
305 from Prof. G. Pier (BWH, HMS) (35,36). The PA2590 transposon mutant was sequence validated
306 (data not shown).
307
308 **Corneal infection model**

309 Male C57BL/6 ("WT") mice were purchased from The Jackson Laboratory (Bar Harbor, ME)
310 and were housed in a B-level UCI vivarium room. Mice were maintained by the Pearlman Lab
311 and UCI ULAR staff. Infected mice were transferred to D-level holding rooms until the
312 experimental endpoint was reached. All protocols were approved by UCI Institutional Animal
313 Care and Use Committee under IACUC protocol AUP-21-123.
314 Overnight cultures of *P. aeruginosa* were grown to an OD_{600} of 0.20 in LB broth, and cultures
315 were then washed and resuspended in 1x sterile PBS (Corning). C57BL/6J mice were
316 anesthetized with ketamine/xylazine (MWI Animal Health) and corneal epithelium was abraded
317 with three, parallel 10mm scratches, performed with a 26 ½ -gauge needle (BD). *Pseudomonas*
318 was then applied to the abraded cornea topically with 2µl containing $5x10^5$ CFU in PBS in the
319 PA14 infection studies and with 5ml of $1x10^6$ CFU in the PAO1 infection studies. For use in the
320 infection experiments, 100µl of overnight culture was added to 10ml of fresh LB and was
321 subcultured for 2.5-3h to reach OD_{600}=0.20. 1ml of culture was spun down at 3,000 x g (rotor,
322 centrifuge) for 10 min and diluted to a desired inoculum. Experiments were carried out
323 predominantly with male C57Bl6/J mice. After 24h or 48h, mice were euthanized, and corneas
324 were imaged by brightfield microscopy to observe opacification. Bacterial CFU was quantified
325 at 24h and 48h by serial dilutions of whole eye homogenates onto LB plates and were counted
326 manually.
327
328 **Corneal opacity**
329 Corneal images were taken using a Leica MZ10 F Modular Stereo Microscope with a Leica
330 DFC450 C camera attachment. Image brightness was adjusted manually using ImageJ software
331 (NIH) and was kept consistent for all images. Images were then converted to 8-bit for

332 quantification. Using ImageJ, ROIs were drawn around the corneal area only (excluding the
333 sclera) and mean intensity was recorded for each ROI. Images were corrected by subtracting
334 intensity caused by glare of dissection lights (glare values were based on naïve corneas). Percent
335 corneal opacity was calculated by dividing corrected mean intensity by the intensity of a
336 completely white cornea.
337
338 **Bacterial quantification**
339 Once images were taken, eyes were resected and placed into a 2 ml centrifuge tube containing
340 1ml of sterile 1x PBS and a sterile steel ball bearing (5 mm diameter). Tubes were placed in a
341 Qiagen Tissue Lyser II and were homogenized at 30Hz for three minutes. After lysis,
342 homogenates were serially diluted up to 1:10,000 and 10µl of each dilution was plated onto one
343 LB agar (BD-Difco) plate (concentrations plated: undiluted, 1:10, 1:100, 1:1,000, 1:10,000).
344 Dilutions were streaked down the plate and plates were incubated at 37°C with 5% CO_2
345 overnight and colony forming units (CFU) were counted manually the next day. CFU was then
346 quantified as CFU/ml.
347
348 **Ocular tissue processing**
349 After imaging the corneas, eyes had 15µl of 4% formalin (ThermoFisher) dropped onto them and
350 were left to fix at room temperature for 15 minutes. Eyes were then carefully resected and placed
351 in 1ml of iced 4% formalin (ThermoFisher) and were fixed at 4°C for 48-72hrs. After fixation,
352 formalin was removed, and tissue was washed twice with sterile 1x PBS. Samples were then
353 placed in 2ml of 70% ethanol and were kept at 4°C until shipment to UMass Histopathology was

354 complete. Tissue embedding was done by UMass Histopathology. All tissues were processed
355 using UMass Chan Medical School Morphology Core services following standard protocols.
356
357 **Target selection and probe designs**
358 Probes targeting 100 selected genes in the PAO1 strain of *P. aeruginosa* were designed using
359 Nanostring in-house scripts (Supplementary Table 1). These probes were designed to specifically
360 bind to *Pseudomonas* genes, avoiding cross-reactivity with mouse genes and other microbes.
361
362 **Spatial transcriptomics workflow**
363 5um thick sections from formalin-fixed paraffin embedded (FFPE) tissue blocks were mounted
364 onto Superfrost Plus slides and prepped for GeoMx spatial transcriptomic profiling. Briefly,
365 slides were baked at 37° C for 1 h prior to loading on a BOND RXm (Leica) for further baking,
366 rehydration, heat-induced epitope retrieval (ER2 for 20 min at 100° C) and enzymatic digestion
367 (0.1 ug/mL proteinase K for 15 min at 37° C). Tissue sections were then hybridized with the
368 Nanostring mouse Whole Transcriptome Atlas (WTA) and custom *P. aeruginosa* probe sets at
369 37° overnight. Following hybridization, two, 25min stringency washes were performed to
370 remove off-target probes and sections were subsequently blocked and stained with the following
371 morphology marker antibodies: DNA (Syto13; 488 channel; Nanostring), PanCK (532 channel;
372 Nanostring), OprI (Opr17_VHHFC; 594 channel; Thermo Fisher Scientific) and Ly6G (1A8; 647
373 channel; Novus Biologicals). Slides were then loaded into the GeoMx DSP instrument where the
374 entire sections were scanned and imaged. Regions of interest (ROIs) were manually drawn on
375 the section images and approved for collection. The GeoMx DSP then directed UV light to each
376 ROI to cleave the barcodes (RNA ID and unique molecular identifier-containing oligonucleotide)

377 from the bound probes. Illumina i5 and i7 dual-indexing primers (Illumina) were added to the
378 collected barcodes during polymerase chain reaction (8 µl collected oligonucleotides/ROI) to
379 uniquely index each ROI. AMPure XP beads (Beckman Coulter) were used for purification of
380 PCR. Library concentration was measured using a Qubit fluorimeter (Thermo Fisher Scientific),
381 and quality was assessed using a Tape Station (Agilent Technologies).
382
383 **Illumina Sequencing and Processing**
384 The sequencing depth requirement was determined by aggregating the areas of illumination
385 (measured in μm^2) and multiplying the total area by a factor of 100, as recommended in the
386 GeoMx DSP Library and Sequencing Guide (nanostring.com). Sequencing was executed using
387 the NovaSeq 6000 System. The resultant FASTQ files were processed to Digital Count
388 Conversion (DCC) files utilizing the GeoMxNGSPipeline v.2.3.3.10, a pipeline publicly
389 available through NanoString Technologies. This pipeline facilitated individual probe alignment
390 by identifying read tag sequences (RTS). Additionally, it eliminated duplicates in probe counts
391 generated during the library preparation's polymerase chain reaction, through a meticulous
392 analysis of the unique molecular identifiers (UMIs) associated with each probe.
393
394 **Quantification and Statistical Analysis**
395 The DCC files underwent analysis using NanoString Technologies' GeoMx™ Digital Spatial
396 Profiler. Key metrics such as the number of raw reads, the alignment rate of these reads, and the
397 sequencing saturation for each area of illumination were scrutinized to evaluate the quality of
398 library preparation. Criteria for inclusion in the analysis were set: more than 10^6 raw reads per
399 ROI, an alignment rate exceeding 80%, and a sequencing saturation greater than 80%. A "no

400 template control" sample, devoid of target probes, was integrated into the library preparation and
401 subsequent sequencing to gauge potential contamination. To estimate background noise and
402 determine the limit of quantification, negative control probes employing External RNA Controls
403 Consortium (ERCC) sequences were included in each GeoMx™ RNA assay. Target probe data
404 were filtered based on frequency within each ROI, followed by normalization against the highest
405 negative control counts using background subtraction. Differential gene expression analysis
406 across various groups was performed using DESeq2, a statistical method designed to identify
407 significant changes in normalized gene expression levels (37). The Python-based library
408 GOATools were used for the functional enrichment analysis (38).
409
410 **Analysis of PA2590**
411 To identify the PA2590 gene, we analyzed more than 11,000 genome sequences from clinical
412 isolates collected between 2000 and 2023. These sequences, sourced from the NCBI database,
413 were subjected to BlastP analysis with the criteria of >85% identity and >70% query and >70%
414 subject coverage, to identify homologous genes of PA2590. Following this, Clustal Omega was
415 utilized for sequence alignment (39). The conservation of amino acid residues within PA2590
416 was quantitatively evaluated using Shannon Entropy (40). This approach effectively highlights
417 conserved regions by examining the variability of amino acid positions across the aligned
418 sequences, offering insights into the evolutionary stability of the gene. MembraneFold was
419 employed to predict membrane topology of PA2590 (41).
420
421 **Structural homology analysis and molecular docking**

422 For structural insights, we employed Foldseek to align PA2590
423 (www.uniprot.org/uniprotkb/Q9I0P7/entry#structure) with existing structures in the Protein Data
424 Bank (PDB) in 01/2024 (42). This alignment was instrumental in hypothesizing the functional
425 structure of PA2590, based on similarities with known protein structures. Structural
426 superimposition was conducted using the cmd.align function in PyMOL v2.5.4, involving 1000
427 alignment cycles with the transform parameter set to 1. *In silico* docking experiments were
428 conducted using Autodock Vina (43). The pose with the lowest energy score (kcal/mol) was
429 chosen for representation. Pymol scripts was used to highlight residues potentially interacting
430 with PA2590, focusing on polar contacts within a 3.5Å range.
431
432 **Predictive Models and Correlative Analysis**
433 We employed Random Forest (RF), Ridge Regression (RR), and Support Vector Machine (SVM)
434 algorithms for predictive analysis, with these models trained on a complete set of host features
435 (as detailed in Suppl. Fig. 2A). Among these, Ridge Regression emerged as the most accurate
436 model and was selected for further in-depth analysis and optimization.
437 The Ridge Regression model was trained using host response profiles and subsequently used to
438 predict the expression data of the PA RplL gene. For normalization of data values, a log
439 transformation ($\log_2$(values + 1)) was applied. To assess the model's predictive accuracy, we
440 implemented a five-fold cross-validation strategy, utilizing Regions of Interest (ROIs) from mice
441 that had been infected. Additionally, cross-validation was carried out across different hosts to
442 validate the model's capacity to predict outcomes in hosts at similar infection stages, as detailed
443 in Supplemental Figure 2. The effectiveness of the model was thoroughly evaluated based on the
444 average Mean Absolute Error (MAE), Explained Variance Score (EVS), and R^2 scores.

445 Further, we assessed the model's performance when trained with varying numbers of mouse gene
446 features. For this, we selected the top-performing features based on their coefficient values in the
447 model trained with the complete dataset. The primary metric for assessing these subset-trained
448 models was the average R^2 score. To optimize these models, based on different sets of features,
449 we applied a grid search technique for hyperparameter tuning, ensuring that each model was
450 fine-tuned for the best possible predictive performance.
451 We assessed correlations between bacterial RplL and host gene expressions using Spearman's
452 rank correlation, with p-value adjustments made through the Benjamini-Hochberg FDR method.
453

454 **References**

455 1. Mladenovic-Antic S, Kocic B, Velickovic-Radovanovic R, Dinic M, Petrovic J, Randjelovic
456 G, et al. Correlation between antimicrobial consumption and antimicrobial resistance of
457 Pseudomonas aeruginosa in a hospital setting: a 10-year study. J Clin Pharm Ther.
458 2016;41(5):532–7.

459 2. Gellatly SL, Hancock REW. Pseudomonas aeruginosa : new insights into pathogenesis and
460 host defenses. Pathog Dis. 2013 Apr 1;67(3):159–73.

461 3. Fleiszig SMJ, Kroken AR, Nieto V, Grosser MR, Wan SJ, Metruccio MME, et al. Contact
462 lens-related corneal infection: Intrinsic resistance and its compromise. Prog Retin Eye Res.
463 2020 May 1;76:100804.

464 4. Sy A, Srinivasan M, Mascarenhas J, Lalitha P, Rajaraman R, Ravindran M, et al.
465 Pseudomonas aeruginosa Keratitis: Outcomes and Response to Corticosteroid Treatment.
466 Invest Ophthalmol Vis Sci. 2012 Jan 25;53(1):267–72.

467 5. Inoue K, Kinoshita M, Muranishi K, Ohara J, Sudo K, Kawaguchi K, et al. Effect of a Novel
468 Trivalent Vaccine Formulation against Acute Lung Injury Caused by Pseudomonas
469 aeruginosa. Vaccines. 2023 Jun;11(6):1088.

470 6. Killough M, Rodgers AM, Ingram RJ. Pseudomonas aeruginosa: Recent Advances in Vaccine
471 Development. Vaccines. 2022 Jul 8;10(7):1100.

472 7. Döring G, Meisner C, Stern M, Flagella Vaccine Trial Study Group. A double-blind
473 randomized placebo-controlled phase III study of a Pseudomonas aeruginosa flagella vaccine
474 in cystic fibrosis patients. Proc Natl Acad Sci U S A. 2007 Jun 26;104(26):11020–5.

475 8. Bianconi I, Alcalá-Franco B, Scarselli M, Dalsass M, Buccato S, Colaprico A, et al. Genome-
476 Based Approach Delivers Vaccine Candidates Against Pseudomonas aeruginosa. Front
477 Immunol. 2018;9:3021.

478 9. Solanki V, Tiwari M, Tiwari V. Prioritization of potential vaccine targets using comparative
479 proteomics and designing of the chimeric multi-epitope vaccine against Pseudomonas
480 aeruginosa. Sci Rep. 2019 Mar 27;9(1):5240.

481 10. Rossi E, Falcone M, Molin S, Johansen HK. High-resolution in situ transcriptomics of
482 Pseudomonas aeruginosa unveils genotype independent patho-phenotypes in cystic fibrosis
483 lungs. Nat Commun. 2018 Aug 27;9(1):3459.

484 11. Damron FH, Oglesby-Sherrouse AG, Wilks A, Barbier M. Dual-seq transcriptomics
485 reveals the battle for iron during Pseudomonas aeruginosa acute murine pneumonia. Sci Rep.
486 2016 Dec 16;6(1):39172.

487 12. Tata M, Wolfinger MT, Amman F, Roschanski N, Dötsch A, Sonnleitner E, et al.
488 RNASeq Based Transcriptional Profiling of Pseudomonas aeruginosa PA14 after Short- and
489 Long-Term Anoxic Cultivation in Synthetic Cystic Fibrosis Sputum Medium. PLOS ONE.
490 2016 Jan 28;11(1):e0147811.

491 13. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly
492 accurate protein structure prediction with AlphaFold. Nature. 2021 Aug;596(7873):583–9.

493 14. Barnard TJ, Watson ME, McIntosh MA. Mutations in the Escherichia coli receptor FepA
494 reveal residues involved in ligand binding and transport. Mol Microbiol. 2001 Aug;41(3):527–
495 36.

496 15. Buchanan SK, Smith BS, Venkatramani L, Xia D, Esser L, Palnitkar M, et al. Crystal
497 structure of the outer membrane active transporter FepA from Escherichia coli. Nat Struct
498 Biol. 1999 Jan;6(1):56–63.

499 16. Moretti J, Roy S, Bozec D, Martinez J, Chapman JR, Ueberheide B, et al. STING Senses
500 Microbial Viability to Orchestrate Stress-Mediated Autophagy of the Endoplasmic Reticulum.
501 Cell. 2017 Nov 2;171(4):809-823.e13.

502 17. Ghssein G, Ezzeddine Z. A Review of Pseudomonas aeruginosa Metallophores:
503 Pyoverdine, Pyochelin and Pseudopaline. Biology. 2022 Dec;11(12):1711.

504 18. Galeano Niño JL, Wu H, LaCourse KD, Kempchinsky AG, Baryiames A, Barber B, et al.
505 Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. Nature.
506 2022;611(7937):810–7.

507 19. Lyu L, Li X, Feng R, Zhou X, Guha TK, Yu X, et al. Simultaneous profiling of host
508 expression and microbial abundance by spatial metatranscriptome sequencing. Genome Res.
509 2023 Mar 1;33(3):401–11.

510 20. Mizgerd JP. Competing Benefits of Tumor Necrosis Factor-α for Bacteria and for Host
511 Defense. Am J Respir Crit Care Med. 2003 Dec 15;168(12):1410–1.

512 21. Minns MS, Liboro K, Lima TS, Abbondante S, Miller BA, Marshall ME, et al. NLRP3
513 selectively drives IL-1β secretion by Pseudomonas aeruginosa infected neutrophils and
514 regulates corneal disease severity. Nat Commun. 2023 Sep 20;14(1):5832.

515 22. Lagoumintzis G, Xaplanteri P, Dimitracopoulos G, Paliogianni F. TNF-alpha induction
516 by Pseudomonas aeruginosa lipopolysaccharide or slime-glycolipoprotein in human
517 monocytes is regulated at the level of Mitogen-activated Protein Kinase activity: a distinct
518 role of Toll-like receptor 2 and 4. Scand J Immunol. 2008 Feb;67(2):193–203.

519 23. Ratitong B, Marshall ME, Dragan MA, Anunciado CM, Abbondante S, Pearlman E.
520 Differential Roles for IL-1α and IL-1β in Pseudomonas aeruginosa Corneal Infection. J
521 Immunol Baltim Md 1950. 2022 Aug 1;209(3):548–58.

522 24. Kugadas A, Christiansen SH, Sankaranarayanan S, Surana NK, Gauguet S, Kunz R, et al.
523 Impact of Microbiota on Resistance to Ocular Pseudomonas aeruginosa-Induced Keratitis.
524 PLOS Pathog. 2016 Sep 22;12(9):e1005855.

525 25. Mayer-Barber KD, Yan B. Clash of the Cytokine Titans: counter-regulation of
526 interleukin-1 and type I interferon-mediated inflammatory responses. Cell Mol Immunol.
527 2017 Jan;14(1):22–35.

528 26. Chen K, Fu Q, Liang S, Liu Y, Qu W, Wu Y, et al. Stimulator of Interferon Genes
529 Promotes Host Resistance Against Pseudomonas aeruginosa Keratitis. Front Immunol. 2018
530 Jun 5;9:1225.

531 27. Woznica A, Kumar A, Sturge CR, Xing C, King N, Pfeiffer JK. eLife. eLife Sciences
532 Publications Limited; 2021 [cited 2024 Apr 2]. STING mediates immune responses in the
533 closest living relatives of animals. Available from:
534 https://elifesciences.org/articles/70436/figures

535 28. Yuan K, Huang C, Fox J, Laturnus D, Carlson E, Zhang B, et al. Autophagy plays an
536 essential role in the clearance of Pseudomonas aeruginosa by alveolar macrophages. J Cell
537 Sci. 2012 Jan 15;125(2):507–15.

538 29. Zhou C min, Wang B, Wu Q, Lin P, Qin S gang, Pu Q qin, et al. Identification of cGAS as
539 an innate immune sensor of extracellular bacterium Pseudomonas aeruginosa. iScience. 2020
540 Dec 10;24(1):101928.

541 30. Parker D, Cohen TS, Alhede M, Harfenist BS, Martin FJ, Prince A. Induction of Type I
542 Interferon Signaling by Pseudomonas aeruginosa Is Diminished in Cystic Fibrosis Epithelial
543 Cells. Am J Respir Cell Mol Biol. 2012 Jan;46(1):6–13.

544 31. Knowlton JJ, Dermody TS, Holm GH. Apoptosis Induced by Mammalian Reovirus Is
545 Beta Interferon (IFN) Independent and Enhanced by IFN Regulatory Factor 3- and NF-κB-
546 Dependent Expression of Noxa. J Virol. 2012 Feb;86(3):1650–60.

547 32. Kirschnek S, Vier J, Gautam S, Frankenberg T, Rangelova S, Eitz-Ferrer P, et al.
548 Molecular analysis of neutrophil spontaneous apoptosis reveals a strong role for the pro-
549 apoptotic BH3-only protein Noxa. Cell Death Differ. 2011 Nov;18(11):1805–14.

550 33. Amieva M, Peek RM. Pathobiology of Helicobacter pylori-Induced Gastric Cancer.
551 Gastroenterology. 2016 Jan;150(1):64–78.

552 34. Takase H, Nitanai H, Hoshino K, Otani T. Impact of siderophore production on
553 Pseudomonas aeruginosa infections in immunosuppressed mice. Infect Immun. 2000
554 Apr;68(4):1834–9.

555 35. Kulasakara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, et al. Analysis of
556 Pseudomonas aeruginosa diguanylate cyclases and phosphodiesterases reveals a role for bis-
557 (3'-5')-cyclic-GMP in virulence. Proc Natl Acad Sci U S A. 2006 Feb 21;103(8):2839–44.

558 36. Liberati NT, Urbach JM, Miyata S, Lee DG, Drenkard E, Wu G, et al. An ordered,
559 nonredundant library of Pseudomonas aeruginosa strain PA14 transposon insertion mutants.
560 Proc Natl Acad Sci U S A. 2006 Feb 21;103(8):2833–8.

561 37. Muzellec B, Teleńczuk M, Cabeli V, Andreux M. PyDESeq2: a python package for bulk
562 RNA-seq differential expression analysis. Bioinformatics. 2023 Sep 1;39(9):btad547.

563 38. Klopfenstein DV, Zhang L, Pedersen BS, Ramírez F, Warwick Vesztrocy A, Naldi A, et
564 al. GOATOOLS: A Python library for Gene Ontology analyses. Sci Rep. 2018 Jul
565 18;8(1):10872.

566 39. Sievers F, Higgins DG. Clustal Omega. Curr Protoc Bioinforma. 2014;48(1):3.13.1-
567 3.13.16.

568 40. Capra JA, Singh M. Predicting functionally important residues from sequence
569 conservation. Bioinformatics. 2007 Aug 1;23(15):1875–82.

570 41. Gutierrez S, Tyczynski WG, Boomsma W, Teufel F, Winther O. MembraneFold:
571 Visualising transmembrane protein structure and topology [Internet]. bioRxiv; 2022 [cited
572 2024 Apr 12]. p. 2022.12.06.518085. Available from:
573 https://www.biorxiv.org/content/10.1101/2022.12.06.518085v1

574 42. van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, et al. Fast and
575 accurate protein structure search with Foldseek. Nat Biotechnol. 2024 Feb;42(2):243–6.

576 43. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a
577 new scoring function, efficient optimization, and multithreading. J Comput Chem.
578 2010;31(2):455–61.

579

580

581

582 **Supplementary Tables**

583 **Supplementary Table 1.** Host-specific transcriptional changes induced by infection at 24 hours

584 post-challenge. This includes GO enrichment analysis results (Supplementary data to Figure 1D).

585

586 **Supplementary Table 2.** List of genes differentially expressed in infected epithelial and stromal

587 ROIs when comparing 48-hour infection to 24-hour infection.

588

589 **Supplementary Table 3.** List of 100 *P. aeruginosa* transcripts profiled in the GeoMx

590 experiments and corresponding Nanostring detection probes.

591

592 **Supplementary Table 4**. Enrichment analysis reflecting *P. aeruginosa* PAO1 bacterial transcript

593 fold changes at 24h post-challenge.

594

595 **Supplementary Table 5.** Enrichment analysis reflecting *P. aeruginosa* PA14 bacterial transcript

596 fold changes at 48h post-challenge.

597

598 **Data and Code Availability**

599 Upon acceptance of the manuscript, the raw sequencing data from our genomics and spatial

600 transcriptomics studies will be made available in the NCBI database. Additionally, the code

601 employed for training predictive models and performing correlative analysis will be shared

602 openly on our Github page.

603

Acknowledgments

We would like to acknowledge the outstanding contributions of the NanoString team: Maxine Macclain, Espy Anguiano, Tim Riordan, Ozge Cagsal-Getkin, Mary Hentschel, and Ankush Tyagi who worked tirelessly to help us conduct the feasibility studies and establish the technology. We would like to thank University of Massachusetts Histopathology core, Yu Liu and Jayme Haywosz, for embedding and processing services. We would like to thank Prof. G. Pier (MGB, HMS) for providing the PA14 and PA14 PA2590 strains.

Author contributions

HZ conducted all computational analyses. SB and MM carried out all in vivo experiments and analyzed their results. ON was responsible for data acquisition within the GeoMx workflow. BJ primarily handled the sequencing analysis portion of the GeoMx workflow. EO and HZ selected the PA genes. CT advised on the sequencing aspects of the GeoMx workflow. GD and HL validated the PA2590 Tn mutant. HZ, NC, EP, and SB played key roles in study design and data interpretation. MG devised the study concept, and contributed to data collection, analysis, and interpretation. HZ, ON, NC, EP, and MG drafted the initial manuscript. All authors discussed the data, reviewed, and revised the manuscript.

Ethics declarations

HZ, ON, NC, BJ, EO, CT, GD, HL, OP, and MG are employee of Moderna Inc. and may hold stock/stock options in Moderna, Inc. SA, MM, and EP have no conflicts to declare.

Figure 1

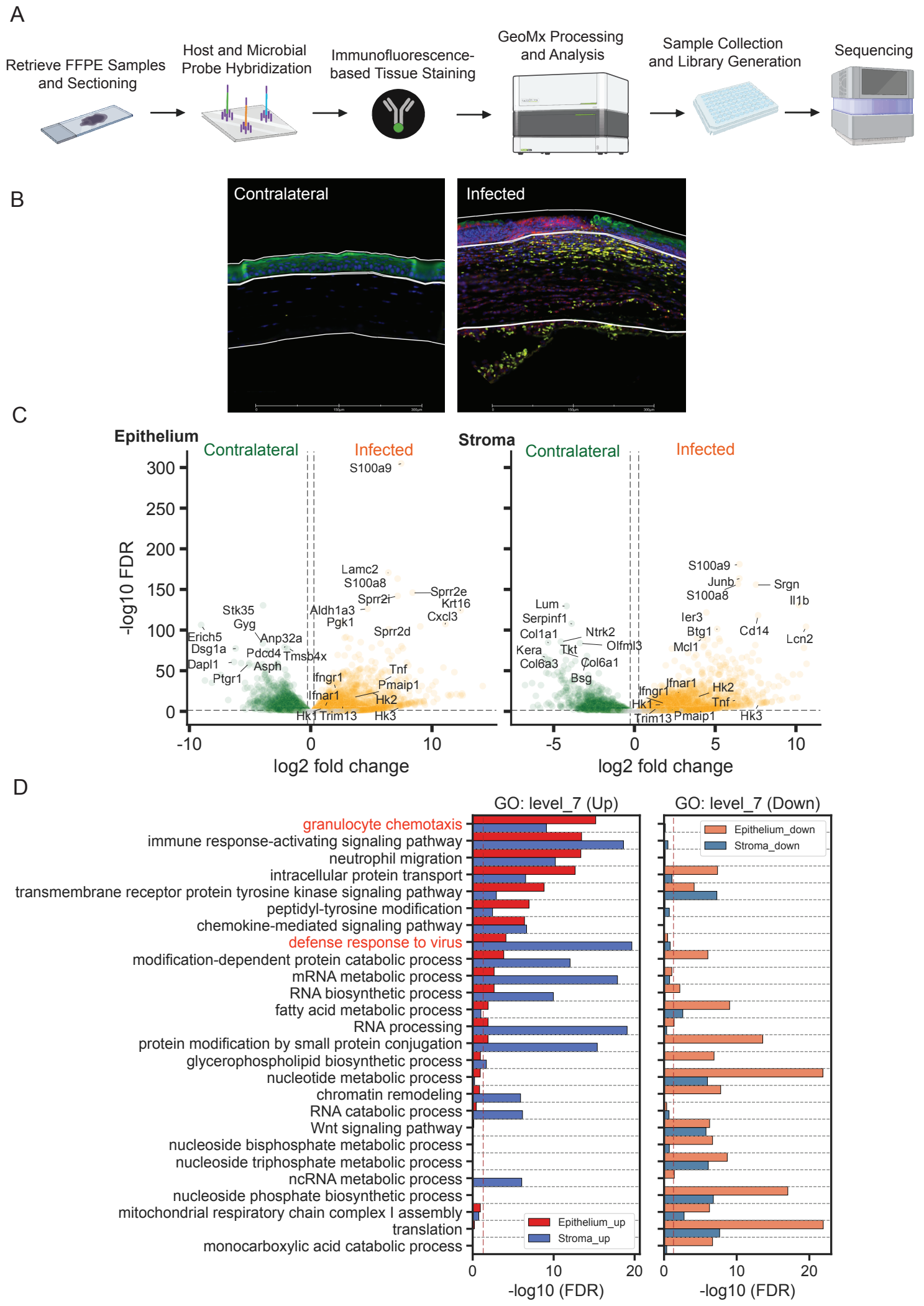


Figure 1. Non-biased whole mouse transcriptome analysis of P. aeruginosa-infected tissues reveal differential enrichment for host transcripts depending on the anatomical site.

A. A schematic diagram of the workflow. Ocular tissues were harvested at 24h post-challenge with $1x10^6$ CFU PAO1. Tissues were paraffin-embedded, sectioned, stained for morphological markers and hybridized with whole host transcriptome probe library. Regions of interest were selected to encompass epithelial layer, stromal layer, and anterior chamber of the eye. Probes were cleaved post-hybridization, collected with GeoMx, amplified, and generated libraries sequenced. Image generated with BioRender.

B. Representative immunohistochemical images of noninfected (contralateral, control eyes) and P. aeruginosa-infected ocular tissues stained for cytokeratin (green), Ly6G (red), DNA (blue), and c-KIT (yellow) were visualized with GeoMx (300 µm size bar, N=2 biological replicas). Robust neutrophil infiltration (red) in the corneal stroma and epithelium is readily observed in the infected site.

C. Volcano plots illustrating differentially expressed transcripts in epithelium and stroma tissue layers of the eye following infection, compared to tissues of the non-infected contralateral eye at 24 hours post-challenge. (N=2, biological replicas). The x-axis indicates fold change (log_2 scale), and the y-axis shows the statistical significance ($-log_{10}$ FDR-adjusted p-values). Sample metadata included.

D. Bar plot of GO (Gene Ontology) functional enrichment analysis, highlighting tissue site-specific gene transcript clustering at level 7. The x-axis displays the FDR-adjusted p-values ($-log_{10}$ scale), indicating the statistical significance of the enrichment. The y-axis lists the level 7 GO categories with enriched transcripts. This analysis incorporates data from ROI profiles comprising 24 epithelial and 24 stromal samples from ocular tissues harvested at 24 hours post-challenge and 8 epithelial and 8 stromal samples from control samples. (N=2, biological replicas). GO enrichment analysis listed in Suppl. Table 2.

Cumulatively these data offer a comprehensive tissue map of infection indicating region-specific alterations.

Figure 2

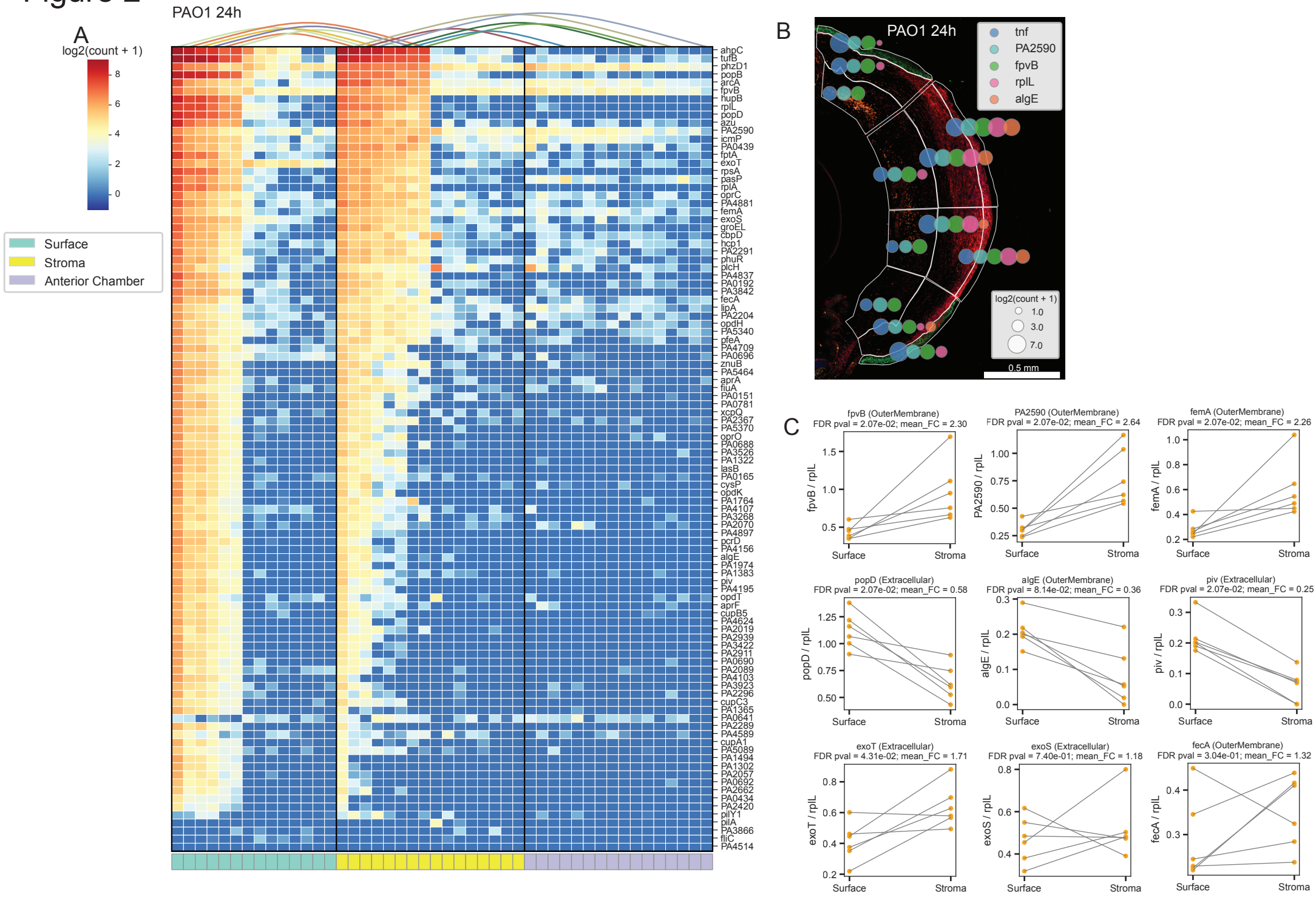


Figure 2. Bacterial transcript profiling of infected with P. aeruginosa PAO1 corneas reveal differential tissue enrichment.

A. Heatmap showing the distribution of 100 bacterial transcripts in infected eyes. Each column represents data from individual ROIs. Surface epithelial ROIs (N=14) (green), stromal ROIs (16) (yellow), and Anterior Chamber (AC) ROIs (N=16) (purple) are grouped together. The curved lines on the top of the heatmap connect juxtaposing ROIs. Data are presented cumulatively from ROI selections derived from two biological replicas.

B. Immunohistochemistry analysis of infected corneal section stained for cytokeratin (green) and Ly6G+ neutrophils (red). Size bar: 0.5 mm. Overlay plot of host and bacterial transcript levels in PAO1-infected tissues. Logarithmic transcript abundances ($\log_2$(count + 1)) of host TNFα, and bacterial PA2590, FpvB, RplL, and AlgE are represented as circles, whose diameter reflects transcript burden in the individual ROIs.

C. Bacterial enrichment was calculated as relative transcript abundance in the surface segments compared to the stromal segments. Bacterial transcript levels were normalized to the housekeeping bacterial RplL transcripts. FDR values and fold change (FC) are annotated per each gene. Representative bacterial enrichment data depicting: 1) transcripts that were enriched in the stroma FpvB, PA2590, FemA; 2) transcripts showing no enrichment, and 3) transcripts showing decreased abundances such as AlgE and Piv. Complete enrichment analysis is provided in Suppl Table 4. Data are representative of two experiments (Suppl. Fig. 2)

Figure 3

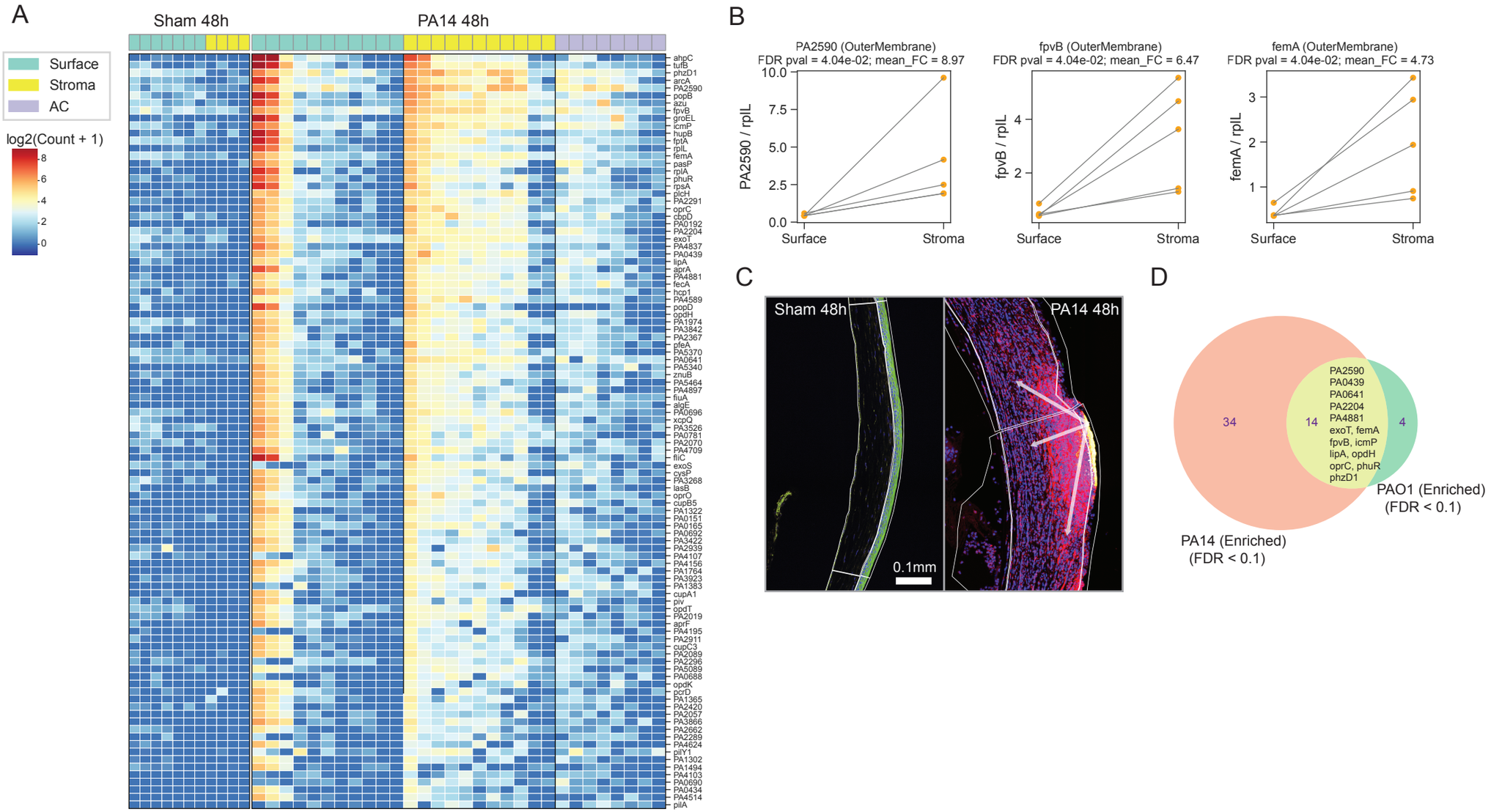


Figure 3. Bacterial transcript profiling of PA14-infected tissues exhibits similar tendencies as the PAO1-induced infections.

A. Heat map showing the distribution of 100 bacterial transcripts in shams and PA14-infected eyes. The shams and the infected tissues were harvested and processed 48h post-challenge. Each column represents data from individual ROIs. Transcript abundances were compared between surface epithelial sham ROIs (N=7) (green), stromal sham ROIs (N=4) (yellow), and infected surface ROIs (N=11) (green), infected stromal ROIs (N=11) (yellow), infected AC ROIs (N=8) (purple) are grouped together. Data are presented cumulatively from ROI selections derived from three biological replicates.

B. Bacterial enrichment was calculated as relative transcript abundances in the surface segments compared to the stromal segments. Bacterial transcript levels were normalized to the housekeeping bacterial RplL transcripts. FDR values and fold change (FC) are annotated for FpvB, PA2590, and FemA. Complete enrichment analysis is provided in Suppl Table 5.

C. Immunohistochemistry analysis of shams and PA14-infected corneal section stained for OprI (yellow), cytokeratin (green), Ly6G+ neutrophils (red), and DNA (blue). Size bar: 0.1 mm. Arrows indicate bacterial spread from the infected surface lesion into the stroma. Shams, N=3, infected tissues N=11 biological replicates.

D. Venn diagram depicting differentially present and shared bacterial transcripts that are enriched within the PAO1 and PA14-infected tissues.

Cumulatively, data point to commonalities in infection architecture that are not strain-specific.

Figure 4

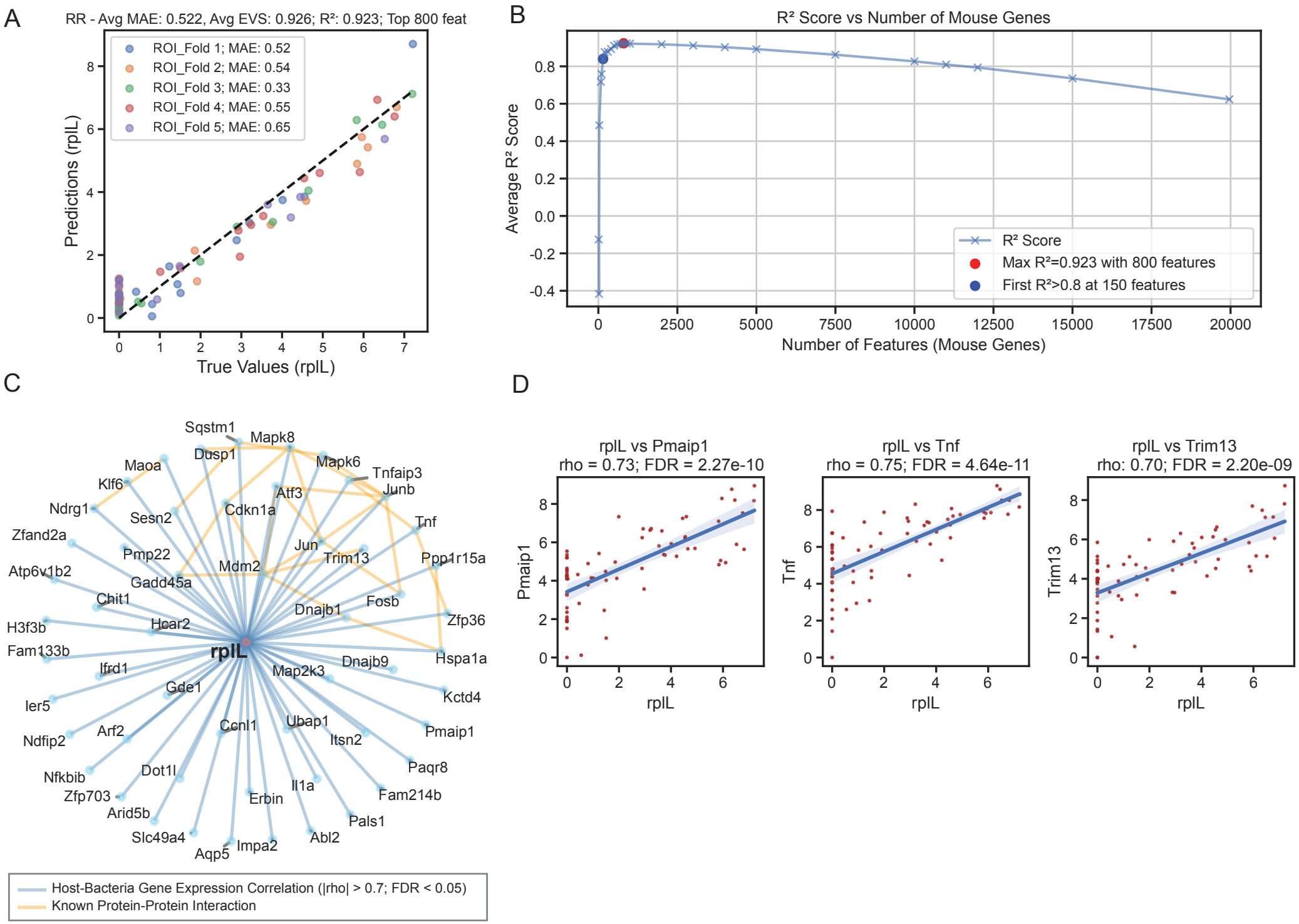


Figure 4. Harnessing host responses for highly predictive analysis of infection levels in tissue regions.

A. Residual plot of a Ridge Regression model trained on host expression profiles to predict bacterial RplL gene expression levels. Data values are transformed using $\log_2$(values + 1). The model underwent five-fold cross-validation using ROIs from tissues derived from 24h post-challenge. Model performance was evaluated based on average MAE, EVS, and R2 scores.

B. Evaluation of model performance when trained with varying numbers of mouse gene features. Top features were chosen for subset training based on the coefficient values from the model trained using the complete dataset. Average R^2 scores provided the metric of assessment. Models, based on different feature sets, underwent optimization using grid search for hyperparameter tuning.

C. Host-bacteria gene expression correlation network, including features with absolute rho values greater than 0.7 and an FDR less than 0.05, determined using Spearman's correlation.

Cumulatively, these data offer a path towards predicting pathogen burden through the surrogate measurements of RplL abundance.

Figure 5

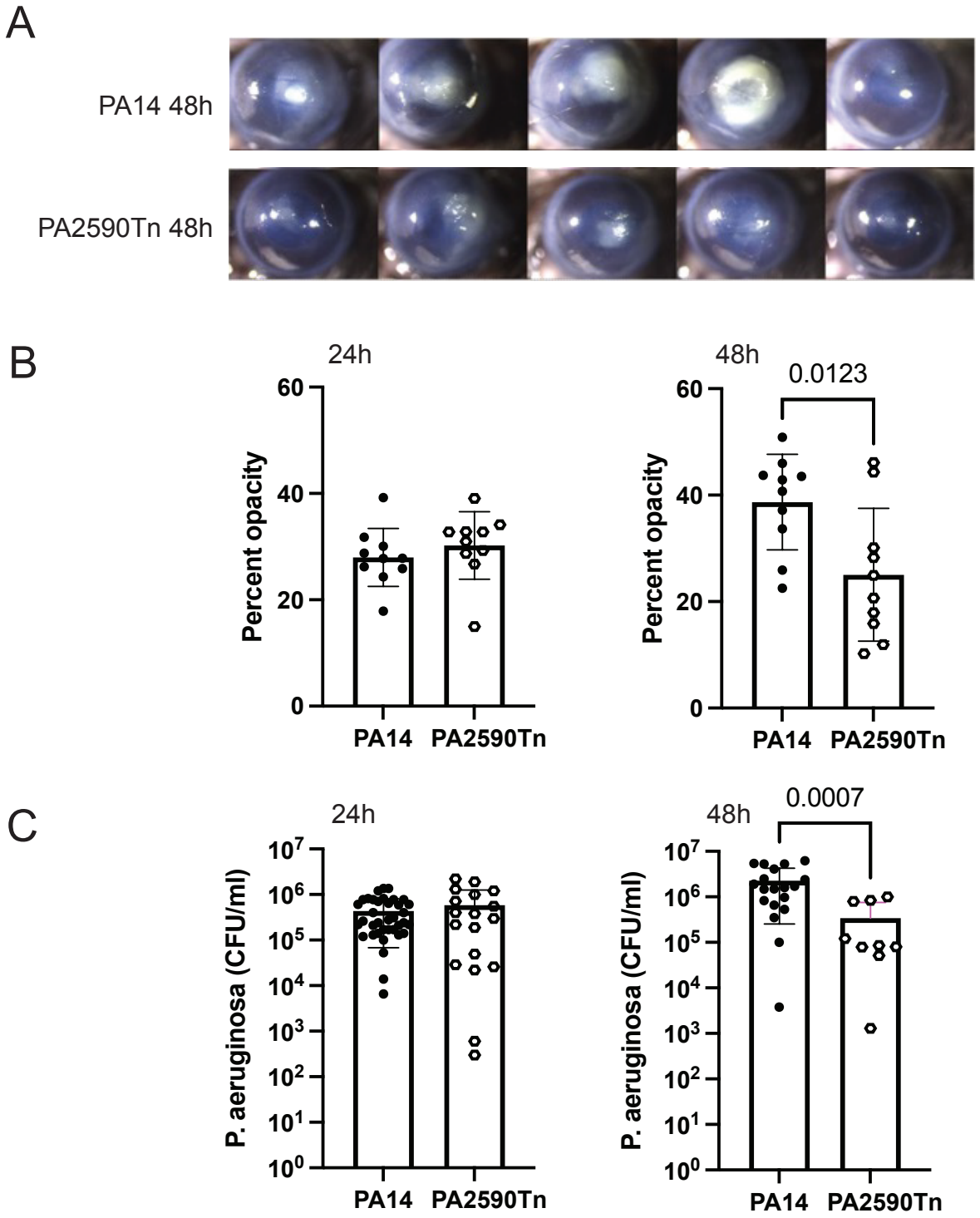


Figure 5. Infection with the PA2590Tn deletion mutant strain is associated with reduced pathology and bacterial burden.

A. Representative images from infected with PA14 and PA14 PA2590Tn deletion mutant mice from two independent experiments. Images were collected at 48h post-challenge.

B. Corneal opacity scores reflective of disease severity were quantified by image analysis at 24h and 48h post-challenge. Data are presented cumulatively from two independent experiments. The individual symbols present a single infected cornea. The filled symbols represent infected C57Bl6/J mice with the PA14 strain (N=10) and the open circles represent C57Bl6/J mice infected with the transposon deletion mutant for PA2590 (N=10). Unpaired Student's t-test with Welch's correction (P=0.01).

C. C57Bl6/J mice were infected with $5x10^5$ CFU of PA14 (N=19) (filled circles) or the PA14 PA2590Tn deletion mutant (N=9) (open circles), and bacterial burden was quantified at 24h and 48h post-challenge. Data are presented cumulatively from two independent experiments. The individual symbols present individual infected corneas. Unpaired Student's t-test with Welch's correction (P=0.0007).

Cumulatively, data point to decreased disease severity in the absence of the PA2590 gene product.

Figure 6

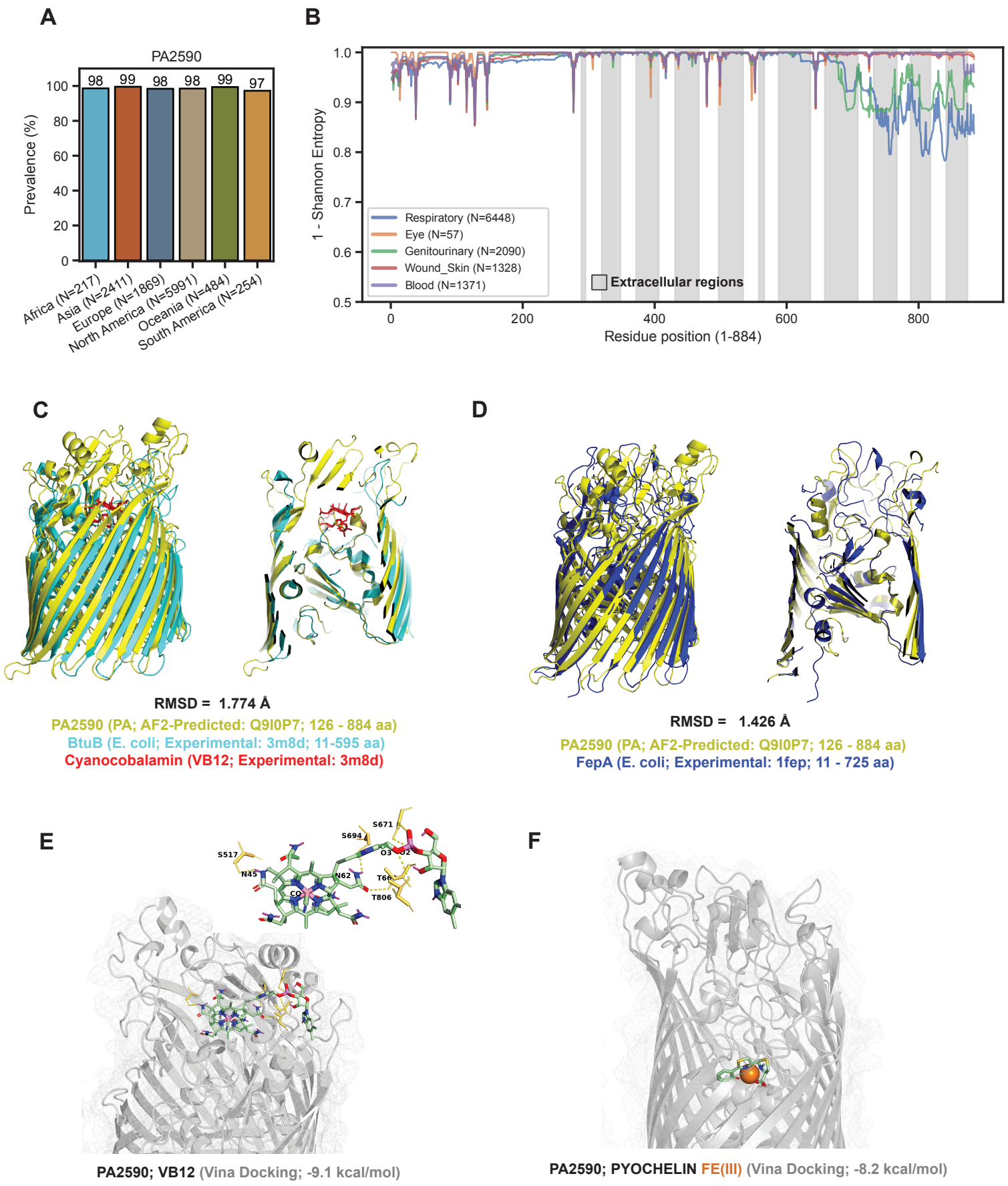


Figure 6. Structural homology-analysis predicts PA2590 gene to function as a vitamin B12 and iron transporter.

A. Prevalence of the PA2590 gene in the genomes of P. aeruginosa clinical isolates sourced from various regions: Africa (N=216), Asia (N=2278), Europe (N=1680), North America (N=5763), Oceania (N=366), and South America (N=253). Here, 'N' denotes the total number of clinical isolates analyzed from each region.

B. Amino acid conservation analysis for the PA2590 gene, illustrating data from clinical isolates of P. aeruginosa obtained from different infection types: genitourinary (blue line, N=1991), respiratory (orange line, N=6132), blood (green line, N=1202), and wound/skin infections (red line, N=1302). The x-axis displays the residue position, while the y-axis shows the 1-Shannon entropy values. Regions shaded indicate the extracellular loops of the protein.

C. Structural homology modeling highlights parallels between the PA2590 structure predicted by AlphaFold2 and the experimentally determined structure of the E. coli BtuB protein. The figure contrasts the predicted model (light green) with the Btub structure (dark green), with VB12 depicted in red. Structural superimposition was conducted using the cmd.align function in PyMOL, involving 1000 alignment cycles with the transform parameter set to 1. This analysis resulted in a Root Mean Square Deviation (RMSD) of 1.774 Å for residues 126-884 of PA2590 and 11-595 of BtuB, indicating a significant level of structural similarity.

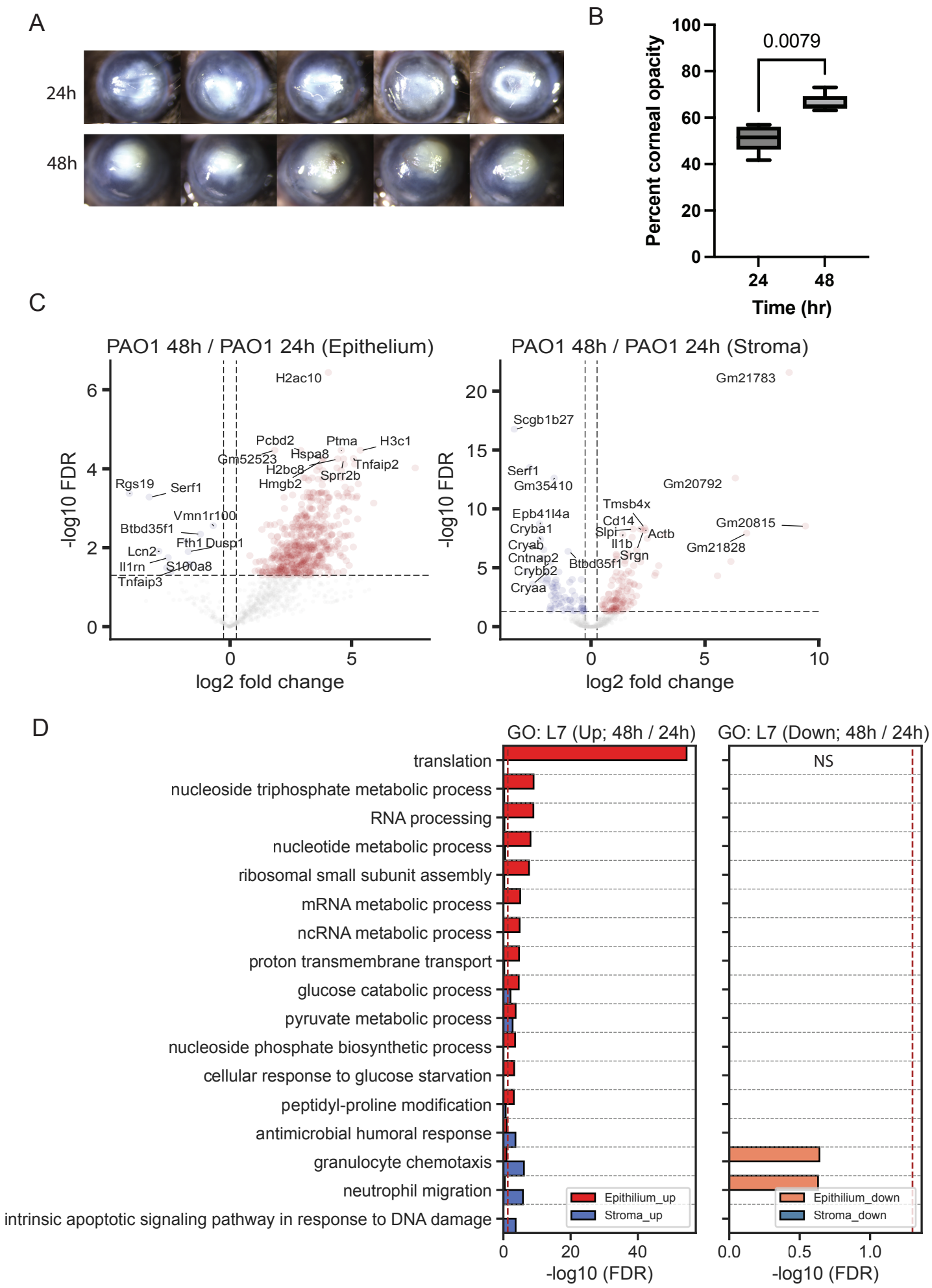
D. Structural homology modeling highlights parallels between the PA2590 structure predicted by AlphaFold2 and the experimentally determined structure of the E. coli FepA protein. The figure depicts the predicted model (light green) with the FepA structure (dark blue). The RMSD is 1.426 Å over amino acid residues 126-884 of PA2590 and 11-725 of FepA, indicating a significant degree of structural similarity.

E. In silico docking experiments were conducted between PA2590 and VB12 using Autodock Vina. The pose with the lowest energy (kcal/mol) was chosen for representation. Pymol scripts were used to highlight residues potentially interacting with PA2590, focusing on polar contacts within a 3.5 Å range.

F. In silico docking experiments were conducted between PA2590 and pyochelin using Autodock Vina. The pose with the lowest energy (kcal/mol) was chosen for representation, indicating a docking score of -8.2 kcal/mol.

Cumulatively, data indicate that the PA2590 might function as either an iron imported utilizing a host siderophore or VB12 importer.

Suppl. Fig. 1



Supplementary Figure 1. Temporal profiling of host response confirms sustained transcriptional remodeling and tissue-specific gene enrichment.

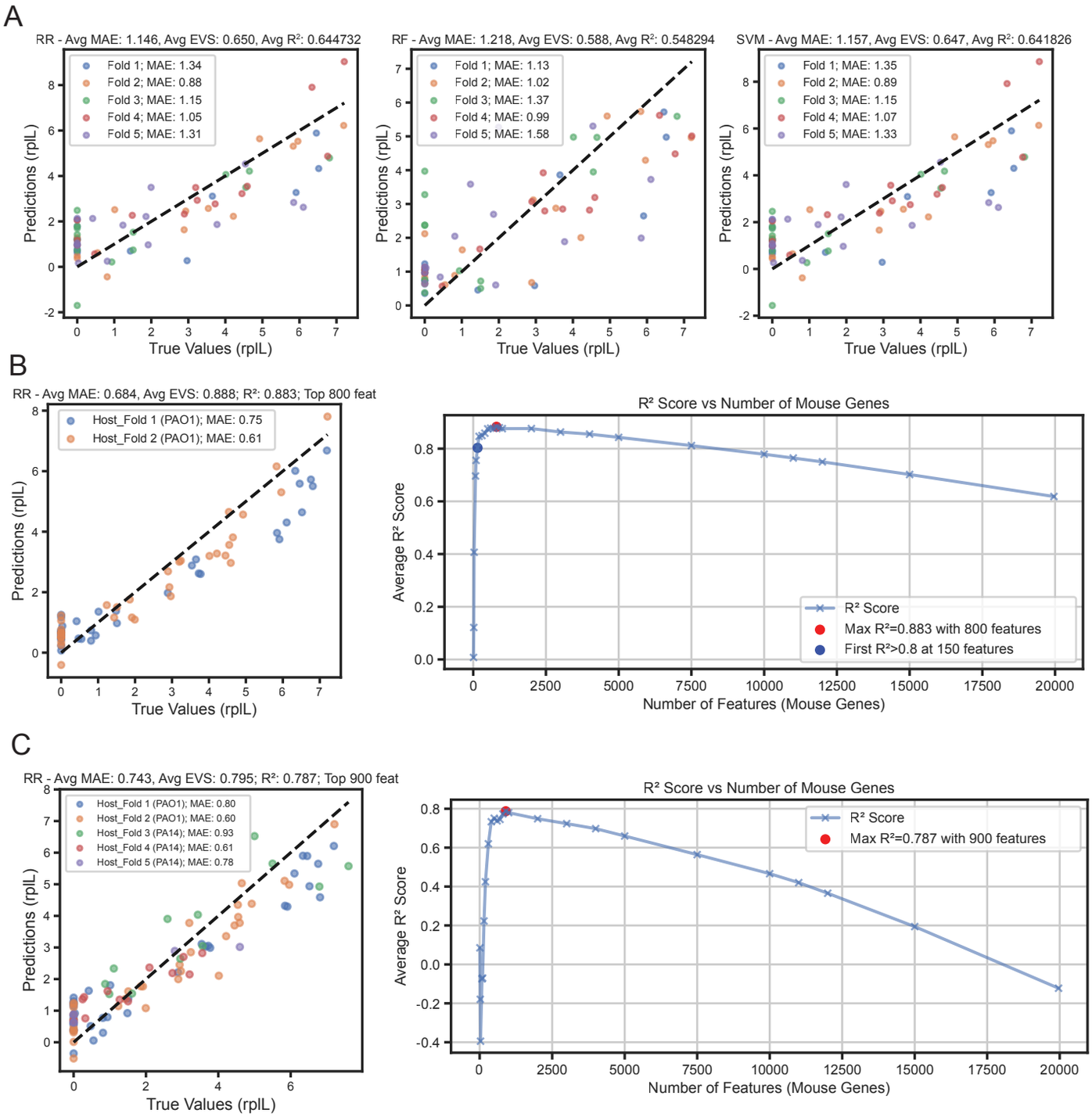
A. Representative images from infected with 5x105 CFU P. aeruginosa PAO1 strain C57Bl6/J mice (N=5 biological replicas). Images are collected at 24h and 48h post-challenge.

B. Corneal opacity scores reflective of disease severity were estimated at 24h and 48h post-challenge. (N=5 biological replicas).

C. Volcano plots illustrating differentially expressed transcripts in epithelium and stroma tissue layers of the eye following infection, compared to tissues of the non-infected control eyes at 48h post-challenge. Mice were challenged with 5x105 CFU P. aeruginosa PAO1 and tissues were harvested at 24h (N=1) and 48h post-infection (N=2). Data are presented cumulatively. The x-axis indicates fold change (log2 scale), and the y-axis shows the statistical significance (-log10 FDR-adjusted p-values). Detailed sample metadata is available in Suppl. Table 2.

D. Bar plot of GO (Gene Ontology) functional enrichment analysis, highlighting tissue site-specific gene transcript clustering at level 7. The x-axis displays the FDR-adjusted p-values (-log10 scale), indicating the statistical significance of the enrichment. The y-axis lists the top level 7 GO categories with enriched transcripts. This analysis incorporates data from ROI profiles comprising 24 epithelial and 24 stromal samples from ocular tissues harvested at 24 hours post-challenge and eight epithelial and eight stromal samples from non-infected eyes, control samples.

Suppl. Fig. 2



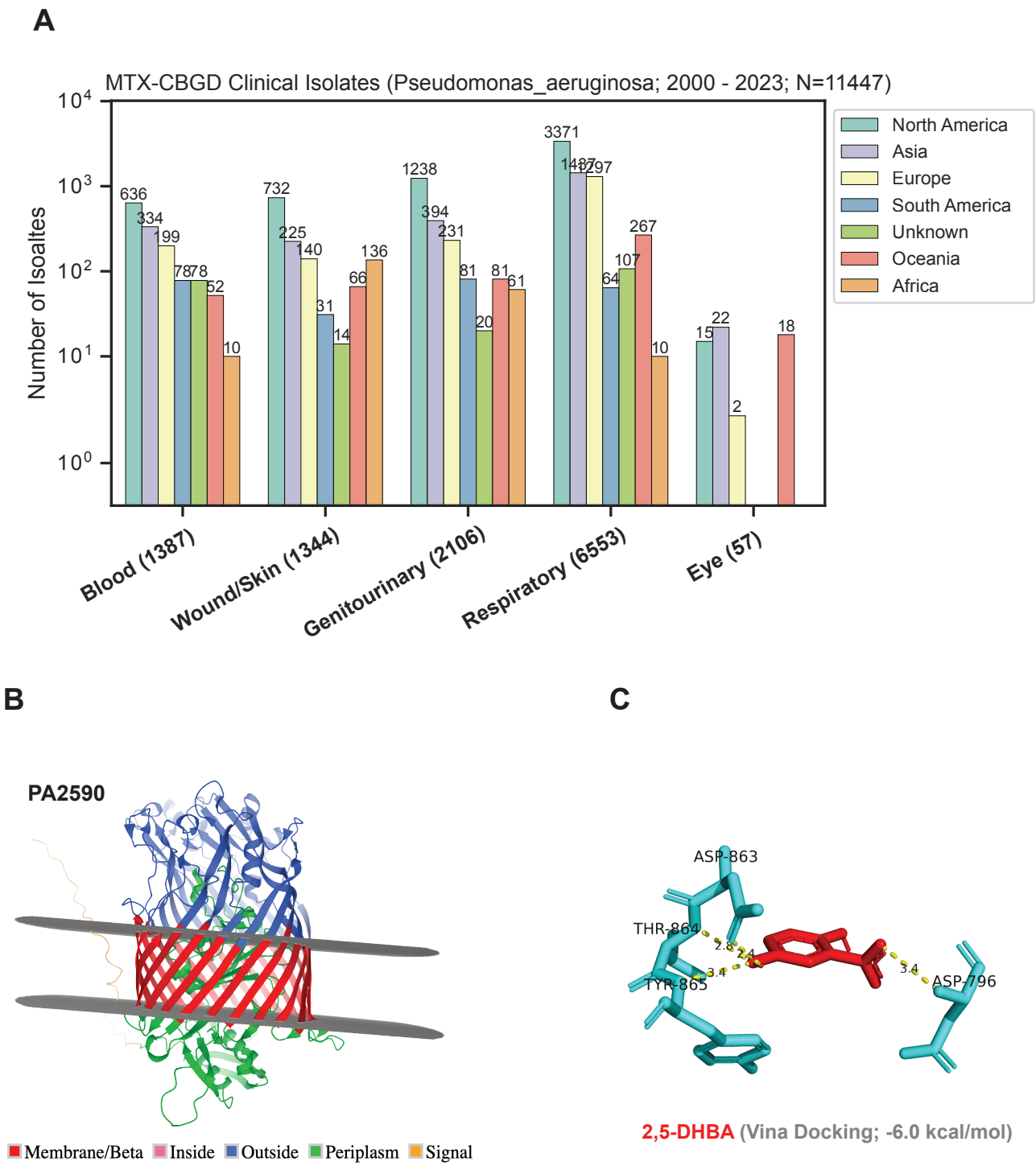
Supplementary Figure 2. Evaluation of model performance in predicting pathogen-specific gene expression based on host response profiles from multiple biological replicas

A. Residual plots of models RF, RR, and SVM generated from host response profiles to predict pathogen-specific RplL gene expression levels. Data values are $\log_2$-transformed (values + 1). Models were trained using ROIs selected from tissues collected 24h post-challenge and cross-validated against ROI datasets from different biological replicas. Model performance was assessed using average MAE, EVS, and R^2 scores.

B. Evaluation of model performance when trained on a single biological data set and validated against a second data set derived from a biological replica. Average R2 scores provided the metric of assessment. Models, based on different feature sets, underwent optimization using grid search for hyperparameter tuning.

C. The assessment of model performance involved training on multiple biological replicas and validating on a single excluded replica, which included samples infected with PA14 and PAO1. Average R^2 scores served as the evaluation metric. The models, based on various feature sets, underwent optimization through grid search for hyperparameter tuning.

Suppl. Fig. 3



Supplementary Figure 3. Structural and functional insights into PA2590.

A. Genomic data from 11,447 *P. aeruginosa* clinical isolates was sourced from NCBI reflecting strain diversity for the past 23 years. The bar graph depicts the global distribution of the number of isolates plotted on the y-axis, and the types of associated infections plotted on the x-axis.

B. Membrane topology of PA2590 structure predicted by MembraneFold.

C. In silico docking experiments were conducted between PA2590 and 2,5-DHBA. The pose with the lowest energy (kcal/mol) was chosen for representation.