Spatial transcriptomics identifies novel *Pseudomonas 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 authors

Mihaela Gadjeva, PhD

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

Email: mgadjeva@modernatx.com

Abstract

To holistically unravel the complexity of pathogen-host interactions within infected tissues we leverage a dual spatial transcriptomic approach that, for the first time, simultaneously captures the expression of *Pseudomonas aeruginosa* genes alongside the entire host transcriptome in a model of ocular infection. This innovative method reveals differential pathogen and host-specific gene expression patterns across specific anatomical regions generating a unified transcriptional map of infection. By integrating these data, we developed a predictive ridge regression model trained on images from infected tissues. The model achieved an R^2 score of 0.923 in predicting bacterial burden distributions by using host features thereby predicting novel biomarkers associated with disease severity. Our analysis revealed a complex interplay between *P. aeruginosa* nutritional requirements and protective host responses and identified novel interactions between bacterial metabolite transport proteins and host autophagy. Among an array of iron acquisition gene transcripts that showed significant enrichment at the host-pathogen interface, we discovered a novel virulence mediator PA2590. This study highlights the power of spatial transcriptomics, particularly in combining bacterial and host transcriptomes, to uncover novel host-pathogen interactions, advance our understanding of bacterial virulence mechanisms, and point to druggable molecules.

Introduction

Genomic and transcriptomic methodologies including next-generation sequencing and high-throughput RNA analysis open possibilities for probing the intricate interactions between pathogens and their hosts. Traditionally, these approaches produce aggregate data from entire infected tissues (1–3). Missing from these analyses is spatial resolution, which can enhance our understanding of the bacterial and host response in infected tissues. We hypothesize that transcriptional analysis combined with spatial resolution will allow us to holistically explore how pathogens adapt and survive within specific tissue microenvironments. This approach also identified novel virulence factors that are potential targets for therapeutic intervention.

Data integration to achieve a spatially resolved understanding of infection remains a formidable challenge. To address this issue, our study employs dual spatial transcriptomics profiling, a novel approach that captures the expression profiles of bacterial gene transcripts alongside the complete host transcriptome. This methodology was applied to examine the transcriptional landscape of *P. aeruginosa*-induced infection in a model of acute ocular disease. By mapping the pathogen-specific transcriptional profiles at a localized scale, we reveal temporal site-specific transcriptional adaptations and novel virulence factors. To the best of our knowledge, this is the first example of the utility of spatial data to discover novel bacterial virulence mechanisms in a non-biased manner.

Previously, machine learning methods were used to diagnose viral and bacterial infectious diseases (4),(5) and predict the risk of secondary bacterial infections (6). Most of those analyses were performed on data acquired through microarray or RNAseq approaches that lack spatial resolution. In contrast to these studies, our spatial transcriptome data analysis leverages host transcriptome features to predict infection burden and disease severity. This

63 analysis also correlates host transcript abundances with bacterial transcripts, thereby generating a
64 unique tool to predict bacterial - host interactions.
65 Cumulatively, our findings not only highlight the significance of integrating spatially
66 resolved transcriptomic data to explore the complexities of host-pathogen interactions, but also
67 suggest potential avenues for the development of targeted therapeutic interventions.
68
69 **Results**
70 **Spatial transcriptomics reveals tissue-segment specific host transcript enrichment.**
71 To generate a spatial transcriptional profile of infection, we used a well-characterized
72 murine model of *P. aeruginosa* corneal infection. We utilized the Nanostring GeoMx platform
73 for non-biased characterization of the host transcriptome (Fig. 1A). Sections from paraffin-
74 embedded eyes were processed and we examined the responses in the corneal epithelium, the
75 corneal stroma, and the anterior chamber as regions of interest (ROIs). Workflow is shown in
76 Fig. 1A.
77 Corneal epithelial cells were identified as keratin 6^{+}ve; neutrophils were identified as
78 $Ly6G^{+}$ve. Uninfected (contralateral eyes) and infected ocular tissues were examined (Fig. 1B).
79 Transcriptome analysis revealed common and segment-specific enrichment of host and bacterial
80 transcripts. Consistent with the rich neutrophil infiltration into the infected tissues, neutrophil-
81 specific S100a8 transcripts were highly upregulated in the epithelium and stroma (Fig. 1C). Also
82 upregulated were neutrophil and monocyte chemokines Cxcl2, and Ccl3, pro-inflammatory
83 cytokines Il-1a and Il-1b, and the neutrophil iron chelating gene Lcn2. Cxcl2, Ccl3, Csf3, Cxcr4,
84 and Il-1b, Il-1a, Il-12a cytokine transcripts were higher in the epithelium relative to the stroma

85 (Fig. 1C). These data were reflected in the Gene Ontology (GO) enrichment analysis showing
86 stronger granulocyte chemotaxis enrichment in epithelium versus stroma.
87 In contrast, transcripts associated with RNA metabolism, processing, ncRNA processing,
88 and chromatin remodeling were predominantly upregulated in the stromal layer when compared
89 to epithelium. Surprisingly, the infected stroma showed elevated anti-viral responses (Fig. 1D),
90 including Type I IFN, IFNAR, and others (Suppl. Table 1).
91 After 48 hours of infection rapid disease progression was observed, as illustrated by
92 elevated corneal opacity of the infected corneas (Suppl. Fig. 1A and B). The epithelial layer
93 displayed enrichment of transcripts associated with metabolic alterations including glucose
94 catabolism and cellular responses to glucose starvation, translation, and RNA processing (Suppl.
95 Fig. 1D and Suppl. Table 2). In the stroma, we found sustained enrichment of transcripts
96 associated with neutrophil recruitment (chemotaxis and migration), glucose catabolism, pyruvate
97 metabolism and apoptosis (Suppl. Fig. 1C and Suppl. Table 2). Cumulatively, our data provide a
98 temporal transcriptional and spatial record of tissue alterations during infection that are
99 associated with neutrophil cytokine production and metabolic changes.
100
101 **Spatial analysis of bacterial transcript abundance in infected corneas reveals distinct**
102 **enrichment profiles between the epithelium and stroma**
103 Probes were generated for 100 *P. aeruginosa* genes that were encoded outer membrane proteins,
104 extracellular proteins, reported virulence factors, and unknown, predicted genes. Probe
105 development was based on the criteria of ORF frequency in >11,000 clinical isolates,
106 evolutionary conservation, and their relation to known virulence. Three different gene-specific

107 probes were designed for each of the bacterial genes of interest and used simultaneously to
108 identify bacterial transcripts in the infected tissues (Suppl. Table 3).
109 Transcript expression was examined in PAO1 infected eyes in the positive corneal
110 epithelium compared with the stroma and anterior chamber (AC). Representative images
111 showed that infection caused loss of the cytokeratin staining in some of the surface layers of the
112 cornea, consistent with the corneal perforation (ulceration) that occurs following infection. The
113 underlying stromal layers showed rich Ly6G$^+$ staining, indicative of neutrophil infiltration to
114 infected cornea (Fig. 2B). Overall, we could distinguish ROIs with high and low bacterial
115 transcript presence. This was reflective of the differences in the number of bacteria at distinct
116 sites in the infected eye. While 97 transcripts were detected in almost half of all examined
117 surface segments, only 37 transcripts were detected in most of the stromal segments.
118 Interestingly, nine transcripts were consistently detected in the cytokeratin$^+$ and cytokeratin$^-$
119 surface layers, Ly6G$^+$ and Ly6G$^-$ stromal regions, and AC segments, including AhpC, TufB,
120 PhzD1, ArcA, FpvB, PA2590, IcmP, LipA, and FemA (Fig. 2A and B).
121 Bacterial transcript enrichment analysis, normalized to the housekeeping RplL transcript
122 abundance, reflected transcript increases in the infected stromal segments when compared to the
123 infected surface layers (Fig. 2C). Nineteen transcripts showed significant upregulation in the
124 infected tissues (Suppl. Table 4). Those included transcripts with known functions such as ArcA,
125 PhzD1, IcmP, OprC, AhpC, LipA, ExoT, PhuR, OpdH, FpvB, and FemA, and transcripts
126 encoding proteins that are predicted but have not been characterized: PA2590, PA0439, PA2204,
127 PA2291, and PA4881. In contrast, transcripts for AlgE, PopD, and Piv were decreased in the
128 stroma, while ExoS and FecA showed no enrichment (Fig. 2C and Suppl. Table 4).

129 To determine if bacterial transcript tissue penetrance is strain-specific, corneas were
130 infected with the cytotoxic (ExoU expressing) PA14 strain. No significant enrichment of
131 bacterial transcripts was observed at 24h post-challenge (data not shown); however, at 48h post-
132 challenge, deeper bacterial transcript penetrance and significant changes were detected. Fourteen
133 transcripts were similarly enriched between the PA14 and PAO1 data sets. These included
134 PhzD1, FpvB, PA2590, IcmP, ArcA, FemA, LipA (Fig. 3B, C, and D). The transcriptional heat
135 map comparisons between the sham and infected tissues confirmed the specificity of bacterial
136 transcript detection (Fig. 3A and Suppl. Table 5). Cumulatively, these results highlight the
137 appearance of bacterial transcriptional patterns in the infected tissues irrespective of the strain
138 types and thereby illustrate spatial organization during infection.
139
140 **Harnessing host-bacteria dynamics for accurate prediction of infection levels in tissue**
141 **regions.**
142 Next, leveraging the host transcriptional data, we examined its functional relationship
143 with the *P. aeruginosa* housekeeping gene RplL, which we used as a correlate for the number of
144 bacteria in the infected tissues. Using Ridge Regression analysis (Fig. 4A) we demonstrated
145 significant predictive accuracy, achieving an R^2 score of 0.923. The selection of 600 key host
146 gene transcript features (Fig. 4B) refined the model, enhancing our understanding of the spatial
147 distribution of bacterial burden across infection sites.
148 We found that 150 host features were sufficient to accurately predict bacterial burden
149 measured based on RplL transcript abundance, with an R^2 score of >0.8. We generated a host-
150 bacterial RplL transcript correlation network and examined it for known host protein interactions
151 (Fig. 4C). We found multiple signaling mediators associated with type I IFN, IL-1a, and TNF-

152 driven inflammatory responses, which were highly interconnected for protein-protein
153 interactions, further indicating the observation that the identified host response network reflected
154 highly coordinated immune mechanisms (Fig. 4C). Examples for individual strong correlations
155 between *P. aeruginosa* RplL and TNF-α, Trim13, and phorbol-12-myristate-13-acetate-induced
156 protein1 (Pmaip1) are shown (Fig. 4D) and demonstrate that this approach can identify known
157 (e.g., Il-1α, TNF-α) and unexpected immune determinants of infection including Trim13 and
158 Pmaip1. Employing 'leave-one-out' validation on our model with a set of mouse hosts infected
159 by PAO1 or PA14 strain, as illustrated in Suppl. Figure 2, showcases its adaptability and
160 potential for broad applicability. This approach involves training the model on data from all but
161 one host and assessing its predictive accuracy on the excluded host. The successful application of
162 this technique across biological replicates demonstrates the model's robust predictive capacity.
163 Cumulatively, our approach offers a detailed perspective on the variability of bacterial
164 load across tissue regions and demonstrated the ability to predict bacterial tissue burden by select
165 host transcript feature abundance, contributing to the field of precision medicine in diagnosing
166 bacterial infections.
167
168 **Infection with the PA2590Tn deletion mutant strain is associated with reduced pathology**
169 **and bacterial burden.**
170 Bacterial enrichment analysis identified known and unknown *P. aeruginosa* transcripts
171 that were enriched in the infected tissues. To determine the functional importance of the
172 predicted PA2590 gene that showed significant penetrance in the infected tissues from PAO1 and
173 PA14-challenged mice, corneas were infected with *P. aeruginosa* PA14 WT parental strain or the
174 PA2590 transposon mutant (PA2590Tn). Infected eyes were imaged, and corneal opacity, was

175 imaged and quantified (Fig. 5). We also assessed viable bacteria by CFU. Since we established
176 similar transcriptional enrichment for the PA2590 in the infected tissues of PAO1 and PA14-
177 induced infections, infection studies were carried out with the PA14 strain.
178 There were no significant differences in corneal opacity or bacterial count at 24h post-
179 challenge between the WT PA14 strain and PA2590Tn, consistent with the lack of enrichment for
180 PA2590 at this data point (data not shown). However, the PA2590Tn mutant showed decreased
181 virulence at 48h post-challenge as we recovered significantly less bacteria when compared to
182 WT infected tissues (Fig. 5B, C, and D). As bacterial burden at 48h post-challenge is typically
183 predictive of the disease outcomes in this model, we concluded that infection with the PA2590Tn
184 mutant causes less severe disease.
185
186 **Structural comparison analysis suggests that PA2590 transports iron or vitamin B12**.
187 To understand how PA2590 might affect *P. aeruginosa* virulence, we examined its
188 prevalence and conservation, and conducted structural studies. Analysis for PA2590 prevalence
189 in 11,447 publicly available *P. aeruginosa* genomes derived from clinical isolates revealed that
190 the PA2590 ORF is widely distributed and present in at least 97% of all clinical isolates (Fig.
191 6A). There were no significant differences in the PA2590 distribution based on geographical
192 location of the isolates. Amino acid conservation analysis was also performed using a library of
193 genomes derived from clinical isolates from different clinical presentations such as wound,
194 blood, respiratory, eye, and genitourinary tract infections. A high degree of conservation was
195 identified in all isolates at the amino acid stretches encoding the extracellular loops of the
196 predicted structure of PA2590 with AlphaFold2 (Fig. 6B) (7). Cumulatively, these data suggested
197 an important biological function for PA2590.

198 Comparative structural analysis established a high degree of conformational similarities
199 between PA2590 and *the E. coli* BtuB protein (PDB: 3m8d; R-Value Free: 0.275) present in *E.*
200 *coli* (Fig. 6C). BtuB is a vitamin B12 (VB12) transporter, suggesting that PA2590 might have a
201 similar function. Docking experiments revealed that PA2590 is likely to bind VB12 (Fig. 6E,
202 Vina docking: -9.1 kcal/mol). For comparison, the docking score for BtuB and VB12 was -8.2
203 kcal/mol. In addition to BtuB, our analysis identified structural homology between PA2590 and
204 *E. coli* FepA (PDB: 1fep; R-Value Free: 0.282) (Fig. 6D). FepA is an iron and VB12 transporter
205 (8). The acquisition of iron depends on ferric enterobactin binding to FepA where the transporter
206 preferentially recognizes the catecholate part of the complex (9). Unlike *E. coli*, our docking
207 experiments did not indicate a potential interaction between enterobactin and PA2590. However,
208 mammalian siderophores can be synthesized by eukaryotic homologues to the *E. coli* enzyme
209 generating enterobactin (10). Therefore, we examined the potential binding of PA2590 to an
210 abundant host siderophore 2,5-dihydroxybenzoic acid (2,5-DHBA). The *in silico* docking
211 experiments confirmed the predicted interaction (Suppl. Fig. 3C, Vina docking: -6.0 kcal/mol).
212 *P. aeruginosa* uses two main siderophores for iron uptake: pyoverdine and pyochelin
213 (11). As both known receptors were upregulated in the infected tissues and to determine if there
214 might be redundancy in iron uptake, the potential interaction of PA2590 with Fe(III)-pyochelin
215 was examined. The experiments showed a significant docking score of -8.2 kcal/mol, indicative
216 of a role for PA2590 in the internalization of iron complexed with *Pseudomonas*-derived
217 siderophores (Fig. 6F). Cumulatively, our data indicate that PA2590 mediates both iron and
218 VB12 transport.
219
220 **Discussion**

221 Bacterial infections remain a major global health threat, including those caused by *Pseudomonas*
222 *aeruginosa*, a highly adaptable pathogen notorious for its resistance to antibiotics and its ability
223 to cause severe infections in multiple tissues (12,13), (14). Despite its clinical significance, our
224 understanding of the molecular mechanisms that underlie *P. aeruginosa*'s ability to thrive
225 remains incomplete (15,16). The complexity of host-pathogen interactions within infected tissues
226 is a significant obstacle in developing effective therapies and vaccines (17,18).

227 Spatial transcriptomic profiling of infected tissues provides information on gene
228 expression patterns at specific infection sites. This report is the first to use dual spatial RNA-seq,
229 combining bacterial pathogen-specific and host transcriptome data with spatial resolution into a
230 unified, multidimensional dataset. The experimental approach represents a significant
231 advancement over previous studies that used either sequential analysis of eukaryotic and
232 prokaryotic features or reported bacterial analysis based only on 16S RNA seq (19,20). Our data
233 illuminate differentially enriched bacterial transcripts at the host-pathogen interface, thereby
234 identifying novel pathoadaptative mechanisms. The spatially resolved data enable machine
235 learning approaches to connect host transcriptional networks with bacterial profiles, thereby
236 identifying key nodes for targeted therapies.

237 Our study uncovered a strong, predictive correlation between a network of interconnected
238 host features and *P. aeruginosa* housekeeping transcript abundances (RplL) during infection.
239 Using at least 150 host features, we successfully predicted bacterial burden in each tissue
240 segment with a significant degree of accuracy. The correlative analysis identified known and
241 novel biomarkers of infection severity including IL-1α, TNF α, TRIM13, Pmaip1, and Chit1. *P.*
242 *aeruginosa* is a potent inducer of IL-1α and TNFα cytokines that orchestrate neutrophil
243 infiltration to the infected site (20,21,22). Loss of IL-1α, unlike the loss of IL-1β, promotes

244 resistance to infection (23,24). Conversely, increases in IL-1α correlates with worse disease
245 illustrating that the ML-enabled analysis correctly identified IL-1α as a biomarker for *P.*
246 *aeruginosa* disease severity. Novel biomarkers included increases in Tripartite Motif Containing
247 13 (TRIM13) and Phorbol-12-Myristate-13-Acetate-Induced Protein 1 (Pmaip1) transcripts (26).
248 These genes regulate autophagy and apoptosis (27), and while direct studies linking Trim13 and
249 Pmaip1 with *P. aeruginosa*-induced infections are lacking, their involvement in autophagy and
250 apoptosis suggest mechanistic connections worth exploring in the future.

251 We predicted that the bacterial transcriptional characteristics would show distinct features
252 depending on anatomical location. Consistently, bacterial enrichment analysis exhibited a higher
253 tissue penetrance of transcripts associated with iron acquisition. *P. aeruginosa* requires iron to
254 detoxify reactive oxygen species (ROS), and nitric oxide (NO) to ensure pathogen survival. *P.*
255 *aeruginosa* has two major siderophores - pyoverdine and pyochelin, and the latter is essential for
256 acute infection (33). Transcripts for the pyoverdine receptor, FpvB, and the ferric mycobactin
257 receptor, FemA, were significantly enriched during infection. In contrast, the pyochelin receptor,
258 FptA, did not show significant tissue enrichment, suggesting that there are other transporters for
259 pyochelin import. To this end, we observed a significant penetrance of the PA2590 transcript
260 encoding a predicted protein. Structural analysis showed that PA2590 has a 22-stranded
261 antiparallel β-barrel domain containing Ton B-dependent transporters and that 2590 has a high
262 degree of structural similarity to *P. aeruginosa* FemA transporters, thereby indicating a possible
263 role for PA2590 in mediating iron intake. Further docking analysis are indicative of pyochelin
264 binding.

265 PA2590 is also structurally similar to dual iron and vitamin B12 BtuB transporter in *E.*
266 *coli,* suggesting that PA2590 has a role in cobalbumin import (34). The dual import of iron and

267 vitamin B12 reflects the need for *P. aeruginosa'*to survive in nutritionally poor, oxygen-limited
268 environments (35). Interestingly, *P. aeruginosa* synthesizes vitamin B12 under aerobic conditions
269 although this activity is suppressed under anaerobic conditions (36). The import system that
270 facilitates anaerobic vitamin B12 intake is presently unknown (37)., and our analysis offers
271 insights into mechanisms of vitamin B12 import. In *P. aeruginosa*, vitamin B12-dependent
272 enzymes such as methylmalonyl-CoA mutase (38), methionine synthase (39), ethanolamine
273 ammonia-lyase (40), glycerol dehydratase (41), and diol dehydratase (42) play crucial roles in
274 metabolic versatility that demonstrates the ability of these bacteria to use alternative carbon
275 sources. Vitamin B12 also supports DNA replication under anaerobic conditions by serving as a
276 cofactor for essential ribonucleotide reductases (37). We also showed that PA2590 is functionally
277 important in a clinically relevant model *P. aeruginosa* corneal infection as significantly less
278 bacteria were recovered in the absence of PA2590. Cumulatively, these data demonstrate that
279 this bioinformatic approach can identify novel pathoadaptation mechanisms, including bacterial
280 virulence factors.

281 Collectively, our analyses highlight transcriptional changes associated with host
282 autophagy pathway and metabolite acquisition by the pathogen, which indicates that there may
283 be a causative link between bacterial vitamin B12, iron acquisition, and autophagy. Indeed,
284 several reports show that iron sequestration by the host to limit bacterial infection promotes
285 autophagy (43)(44)(45). Additionally, iron-siderophore complexes incubated with *C. elegans*
286 induces mitophagy, a form of damaged mitochondrial autophagy (46). Conversely, inhibition of
287 mitophagy renders *C. elegans* more susceptible to *P. aeruginosa*-induced infection (46). Vitamin
288 B12 deficiency, similar to iron starvation, promotes autophagy due to reduction in the citric acid
289 cycle and ATP production (47). Cumulatively, the spatially resolved data capture efficiently a

290 complex host-pathogen interplay that identify key metabolites such as iron and vitamin B12 that
291 regulate autophagy can also mediate resistance to infection.
292 In conclusion, the uses of spatially resolved maps of infection are widely applicable
293 towards gaining foundational insights into host-bacterial pathogen interactions that will guide
294 future mechanistic research. Our spatial transcriptome data analysis approach identified novel
295 virulence mechanisms associated with key requirements for pathogen survival during infection.
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) (48,49). 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 (50). The Python-based library
408 GOATools were used for the functional enrichment analysis (51).
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 (52). The conservation of amino acid residues within PA2590
416 was quantitatively evaluated using Shannon Entropy (53). 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 (54).
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 (55). 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 (56). 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,

we applied a grid search technique for hyperparameter tuning, ensuring that each model was fine-tuned for the best possible predictive performance.

We assessed correlations between bacterial RplL and host gene expressions using Spearman's rank correlation, with p-value adjustments made through the Benjamini-Hochberg FDR method.

References

1. Pisu D, Huang L, Grenier JK, Russell DG. Dual RNA-Seq of Mtb-Infected Macrophages In Vivo Reveals Ontologically Distinct Host-Pathogen Interactions. Cell Rep. 2020 Jan 14;30(2):335-350.e4.

2. Cornforth DM, Dees JL, Ibberson CB, Huse HK, Mathiesen IH, Kirketerp-Møller K, et al. Pseudomonas aeruginosa transcriptome during human infection. Proc Natl Acad Sci. 2018 May 29;115(22):E5125–34.

3. Harrington NE, Littler JL, Harrison F. Transcriptome Analysis of Pseudomonas aeruginosa Biofilm Infection in an Ex Vivo Pig Model of the Cystic Fibrosis Lung. Appl Environ Microbiol. 2022 Feb 8;88(3):e01789-21.

4. Mayhew MB, Buturovic L, Luethy R, Midic U, Moore AR, Roque JA, et al. A generalizable 29-mRNA neural-network classifier for acute bacterial and viral infections. Nat Commun. 2020 Mar 4;11(1):1177.

5. Xie J, Zheng X, Yan J, Li Q, Jin N, Wang S, et al. Deep learning model to discriminate diverse infection types based on pairwise analysis of host gene expression. iScience. 2024 Jun 21;27(6):109908.

6. Carney M, Pelaia TM, Chew T, Teoh S, Phu A, Kim K, et al. Host transcriptomics and machine learning for secondary bacterial infections in patients with COVID-19: a prospective, observational cohort study. Lancet Microbe. 2024 Mar 1;5(3):e272–81.

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

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

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

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

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

485 12. Reyes J, Komarow L, Chen L, Ge L, Hanson BM, Cober E, et al. Global epidemiology and
486 clinical outcomes of carbapenem-resistant Pseudomonas aeruginosa and associated
487 carbapenemases (POP): a prospective cohort study. Lancet Microbe. 2023 Mar 1;4(3):e159–
488 70.

489 13. Murray CJL, Ikuta KS, Sharara F, Swetschinski L, Aguilar GR, Gray A, et al. Global burden
490 of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet. 2022 Feb
491 12;399(10325):629–55.

492 14. Weimann A, Dinan AM, Ruis C, Bernut A, Pont S, Brown K, et al. Evolution and host-
493 specific adaptation of Pseudomonas aeruginosa. Science. 2024 Jul 5;385(6704):eadi0908.

494 15. Lee J, Zhang L. The hierarchy quorum sensing network in Pseudomonas aeruginosa.
495 Protein Cell. 2015 Jan;6(1):26–41.

496 16. Moradali MF, Ghods S, Rehm BHA. Pseudomonas aeruginosa Lifestyle: A Paradigm for
497 Adaptation, Survival, and Persistence. Front Cell Infect Microbiol [Internet]. 2017 Feb 15
498 [cited 2024 Aug 20];7. Available from: https://www.frontiersin.org/journals/cellular-and-
499 infection-microbiology/articles/10.3389/fcimb.2017.00039/full

500 17. Bjarnsholt T, Alhede M, Alhede M, Eickhardt-Sørensen SR, Moser C, Kühl M, et al. The *in*
501 *vivo* biofilm. Trends Microbiol. 2013 Sep 1;21(9):466–74.

502 18. Rossi E, La Rosa R, Bartell JA, Marvig RL, Haagensen JAJ, Sommer LM, et al.
503 Pseudomonas aeruginosa adaptation and evolution in patients with cystic fibrosis. Nat Rev
504 Microbiol. 2021 May;19(5):331–42.

505 19. Galeano Niño JL, Wu H, LaCourse KD, Kempchinsky AG, Baryiames A, Barber B, et al.
506 Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. Nature.
507 2022;611(7937):810–7.

508 20. Lyu L, Li X, Feng R, Zhou X, Guha TK, Yu X, et al. Simultaneous profiling of host expression
509 and microbial abundance by spatial metatranscriptome sequencing. Genome Res. 2023 Mar
510 1;33(3):401–11.

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

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

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

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

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

526 26. Narayan K, Waggoner L, Pham ST, Hendricks GL, Waggoner SN, Conlon J, et al. TRIM13 is
527 a negative regulator of MDA5-mediated type I interferon production. J Virol. 2014
528 Sep;88(18):10748–57.

529 27. Ahsan N, Shariq M, Surolia A, Raj R, Khan MF, Kumar P. Multipronged regulation of
530 autophagy and apoptosis: emerging role of TRIM proteins. Cell Mol Biol Lett. 2024 Jan
531 16;29:13.

532 28. Tata M, Wolfinger MT, Amman F, Roschanski N, Dötsch A, Sonnleitner E, et al. RNASeq
533 Based Transcriptional Profiling of Pseudomonas aeruginosa PA14 after Short- and Long-Term
534 Anoxic Cultivation in Synthetic Cystic Fibrosis Sputum Medium. PLOS ONE. 2016 Jan
535 28;11(1):e0147811.

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

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

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

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

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

551 34. Pieńko T, Trylska J. Extracellular loops of BtuB facilitate transport of vitamin B12 through
552 the outer membrane of E. coli. PLoS Comput Biol. 2020 Jul;16(7):e1008024.

553 35. Hood MI, Skaar EP. Nutritional immunity: transition metals at the pathogen-host
554 interface. Nat Rev Microbiol. 2012 Jul 16;10(8):525–37.

555 36. Lee KM, Go J, Yoon MY, Park Y, Kim SC, Yong DE, et al. Vitamin B12-Mediated Restoration
556 of Defective Anaerobic Growth Leads to Reduced Biofilm Formation in Pseudomonas
557 aeruginosa. Infect Immun. 2012 Apr 12;80(5):1639–49.

558 37. Crespo A, Blanco-Cabra N, Torrents E. Aerobic Vitamin B12 Biosynthesis Is Essential for
559 Pseudomonas aeruginosa Class II Ribonucleotide Reductase Activity During Planktonic and
560 Biofilm Growth. Front Microbiol. 2018 May 15;9:986.

561 38. Gross F, Ring MW, Perlova O, Fu J, Schneider S, Gerth K, et al. Metabolic engineering of
562 Pseudomonas putida for methylmalonyl-CoA biosynthesis to enable complex heterologous
563 secondary metabolite formation. Chem Biol. 2006 Dec;13(12):1253–64.

564 39. Foglino M, Borne F, Bally M, Ball G, Patte JC. A direct sulfhydrylation pathway is used for
565 methionine biosynthesis in Pseudomonas aeruginosa. Microbiol Read Engl. 1995 Feb;141 (Pt
566 2):431–9.

567 40. Lundgren BR, Sarwar Z, Pinto A, Ganley JG, Nomura CT. Ethanolamine Catabolism in
568 Pseudomonas aeruginosa PAO1 Is Regulated by the Enhancer-Binding Protein EatR (PA4021)
569 and the Alternative Sigma Factor RpoN. J Bacteriol. 2016 Sep 1;198(17):2318–29.

570 41. Liu Y, Sun W, Ma L, Xu R, Yang C, Xu P, et al. Metabolic Mechanism and Physiological Role
571 of Glycerol 3-Phosphate in Pseudomonas aeruginosa PAO1. mBio. 2022 Dec
572 20;13(6):e0262422.

573 42. Shibata N, Masuda J, Tobimatsu T, Toraya T, Suto K, Morimoto Y, et al. A new mode of
574 B12 binding and the direct participation of a potassium ion in enzyme catalysis: X-ray
575 structure of diol dehydratase. Structure. 1999 Aug 15;7(8):997–1008.

576 43. Chadha J, Harjai K, Chhibber S. Revisiting the virulence hallmarks of Pseudomonas
577 aeruginosa: a chronicle through the perspective of quorum sensing. Environ Microbiol.
578 2022;24(6):2630–56.

579 44. Mukhopadhyay S, Encarnación-Rosado J, Lin EY, Sohn ASW, Zhang H, Mancias JD, et al.
580 Autophagy supports mitochondrial metabolism through the regulation of iron homeostasis in
581 pancreatic cancer. Sci Adv. 2023 Apr 19;9(16):eadf9284.

582 45. Tam E, Reno C, Nguyen K, Cho S, Sweeney G. Importance of Autophagy in Mediating
583 Cellular Responses to Iron Overload in Cardiomyocytes. Rev Cardiovasc Med. 2022 May
584 7;23(5):167.

585 46. Wilson BR, Bogdan AR, Miyazawa M, Hashimoto K, Tsuji Y. Siderophores in Iron
586 Metabolism: From Mechanism to Therapy Potential. Trends Mol Med. 2016
587 Dec;22(12):1077–90.

588 47. Tsalamandris G, Hadjivassiliou M, Zis P. The Role of Nutrition in Neurological Disorders.
589 Nutrients. 2023 Nov 7;15(22):4713.

590 48. Kulasakara H, Lee V, Brencic A, Liberati N, Urbach J, Miyata S, et al. Analysis of
591 Pseudomonas aeruginosa diguanylate cyclases and phosphodiesterases reveals a role for bis-
592 (3'-5')-cyclic-GMP in virulence. Proc Natl Acad Sci U S A. 2006 Feb 21;103(8):2839–44.

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

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

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

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

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

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

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

609 56. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a
610 new scoring function, efficient optimization, and multithreading. J Comput Chem.
611 2010;31(2):455–61.

612

613

614

615 **Supplementary Tables**

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

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

618

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

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

621

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

623 experiments and corresponding Nanostring detection probes.

624

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

626 fold changes at 24h post-challenge.

627

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

629 fold changes at 48h post-challenge.

630

631 **Data and Code Availability**

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

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

634 for training predictive models and conducting correlative analysis can be found at

635 https://github.com/modernatx/spatial_RNAseq_analyses.

636

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

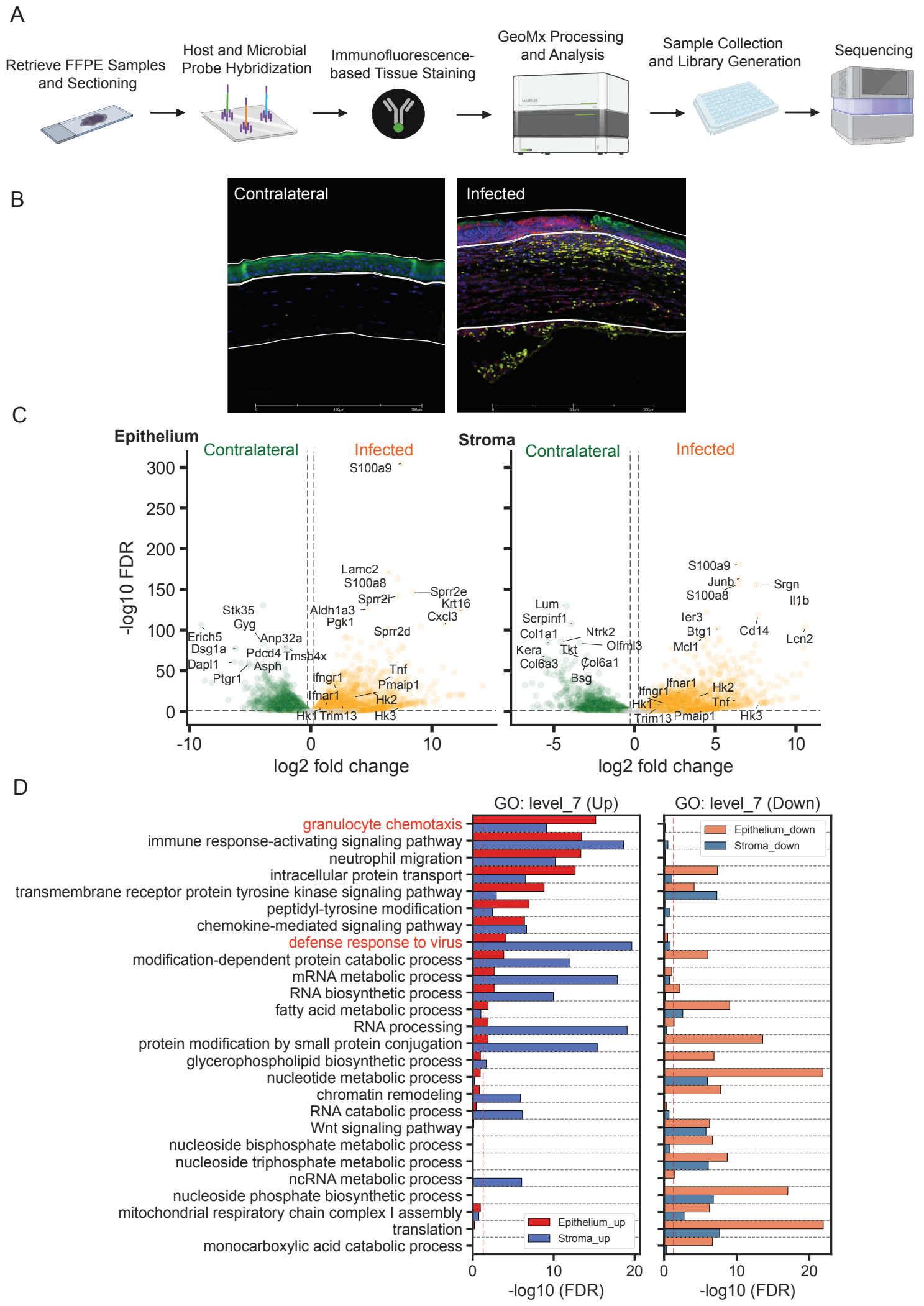


Figure 1. Non-biased whole mouse transcriptome analysis of P. aeruginosa-infected tissues reveal differential enrichment for host transcripts depending on the anatomical site.

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

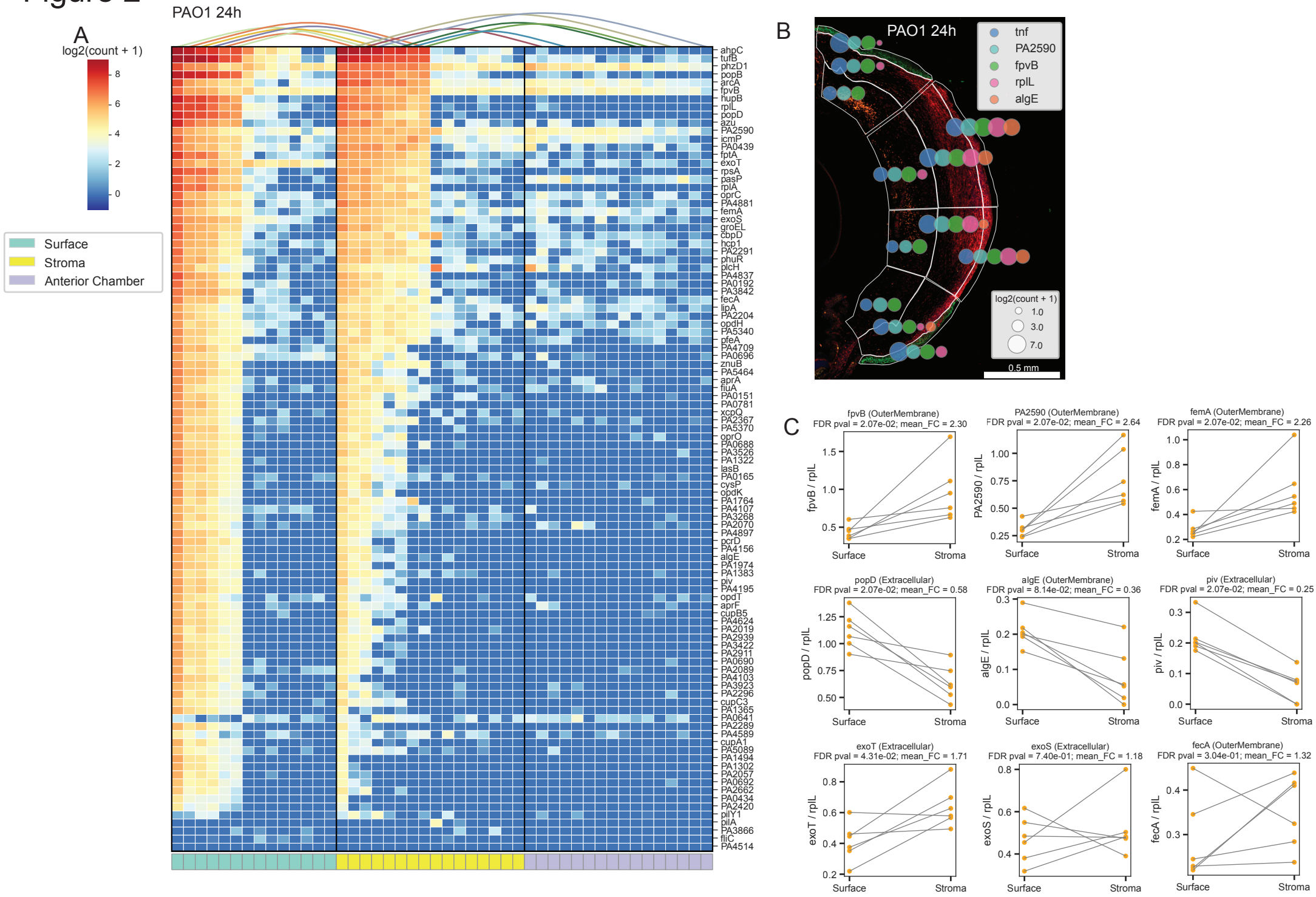


Figure 2. Bacterial transcript profiling of infected with P. aeruginosa PAO1 corneas reveal differential tissue enrichment.

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

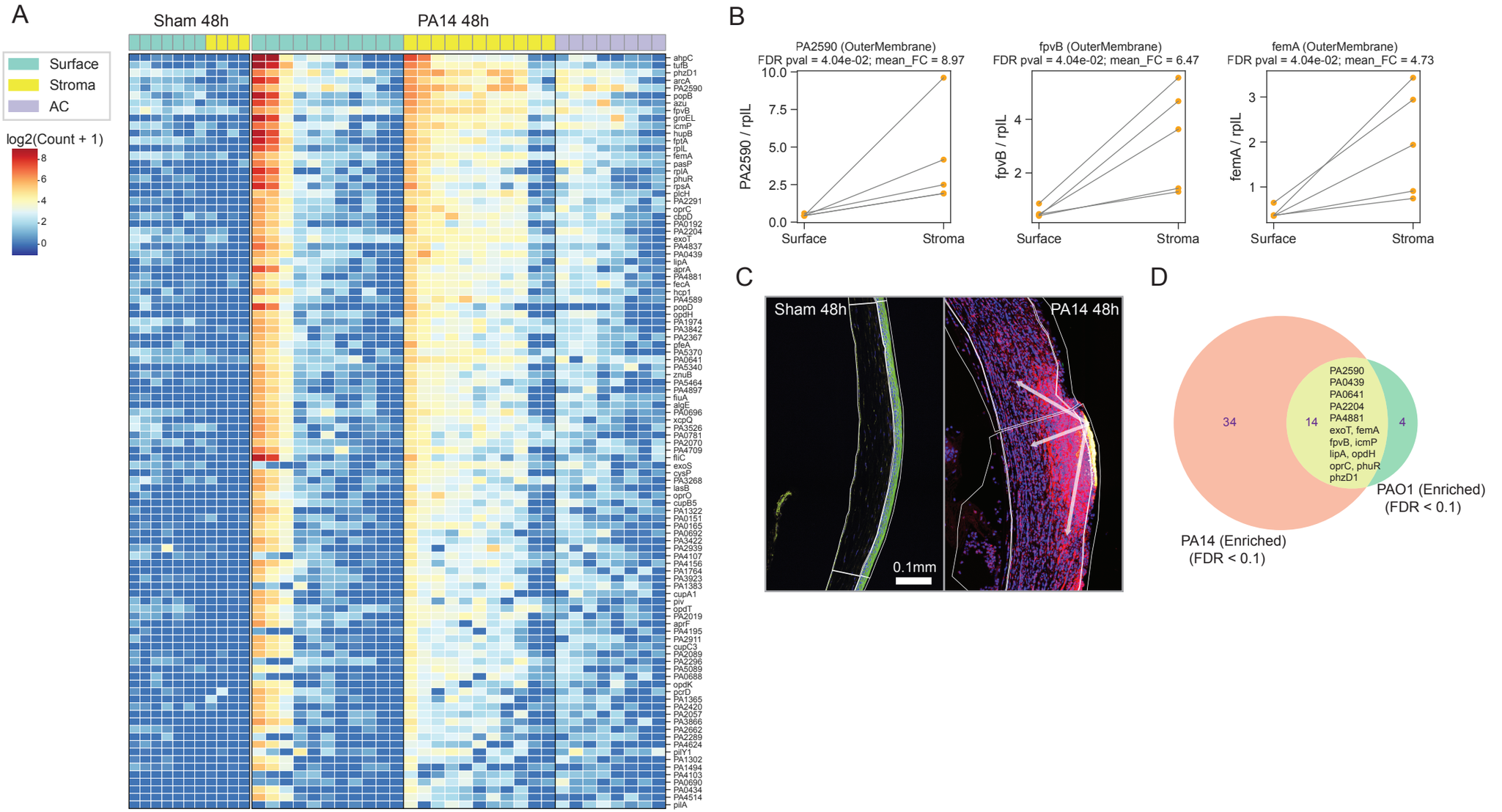


Figure 3. Bacterial transcript profiling of PA14-infected tissues exhibits similar tendencies as the PAO1-induced infections.

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

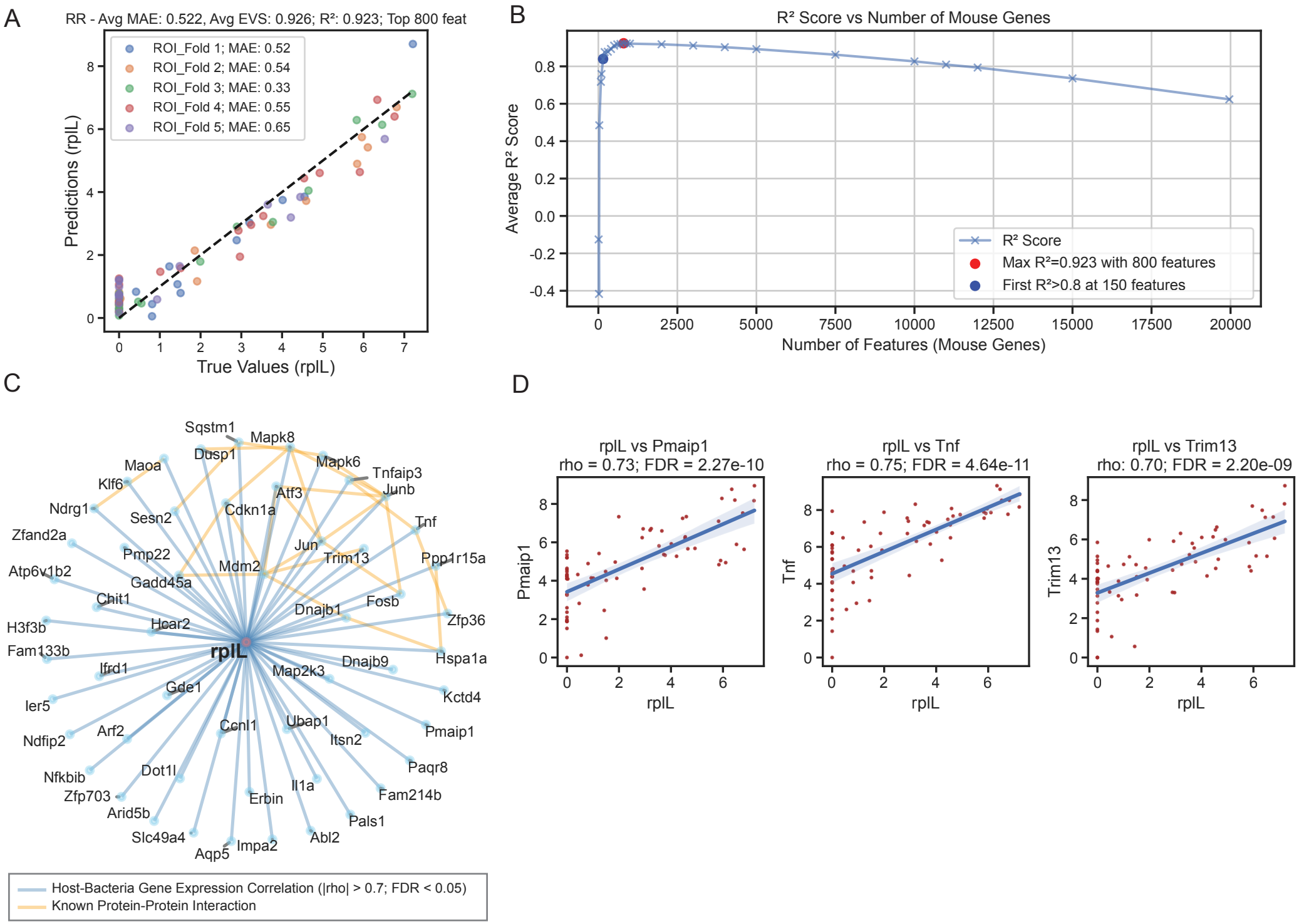


Figure 4. Harnessing host responses for highly predictive analysis of infection levels in tissue regions.

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

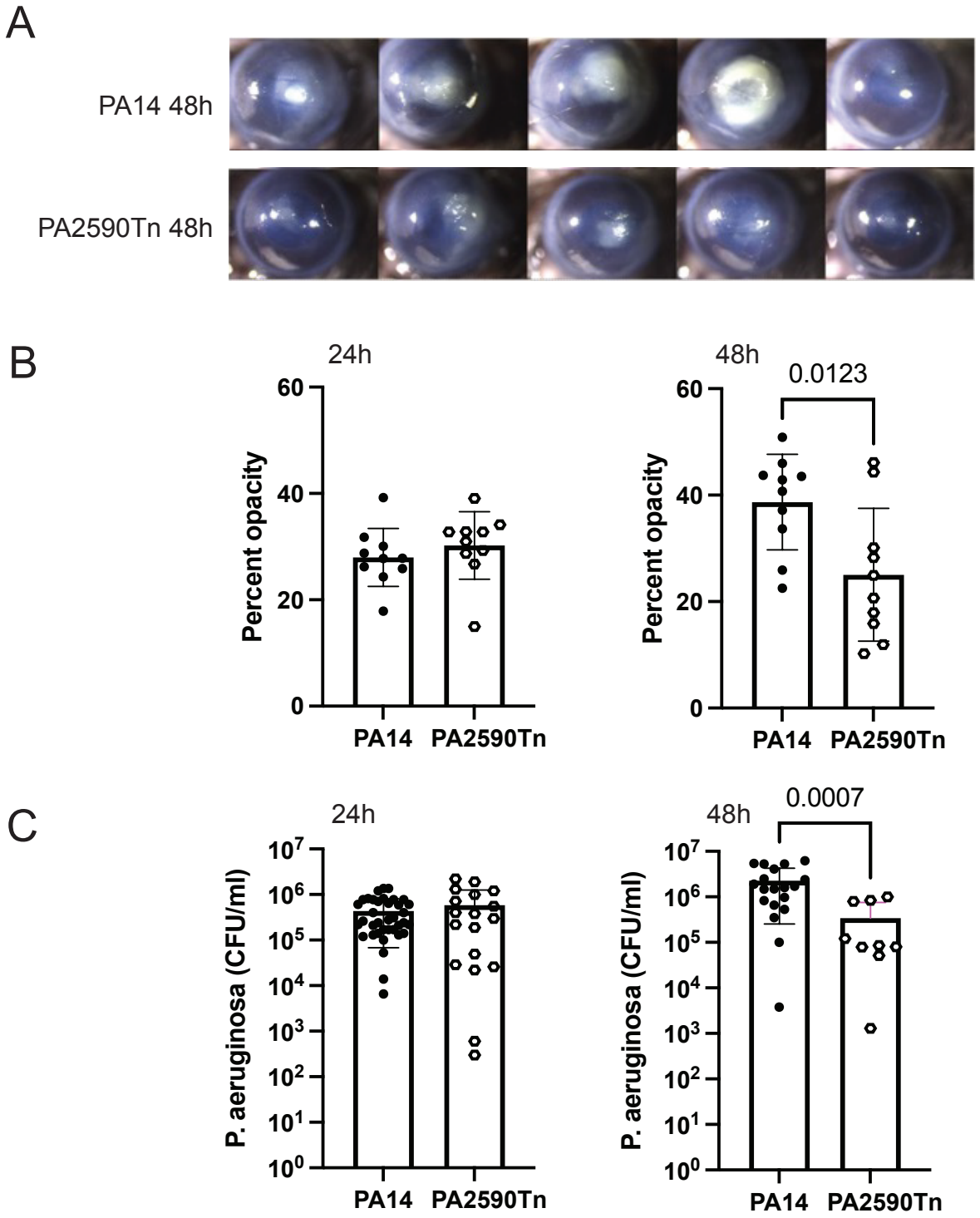


Figure 5. Infection with the PA2590Tn deletion mutant strain is associated with reduced pathology and bacterial burden.

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

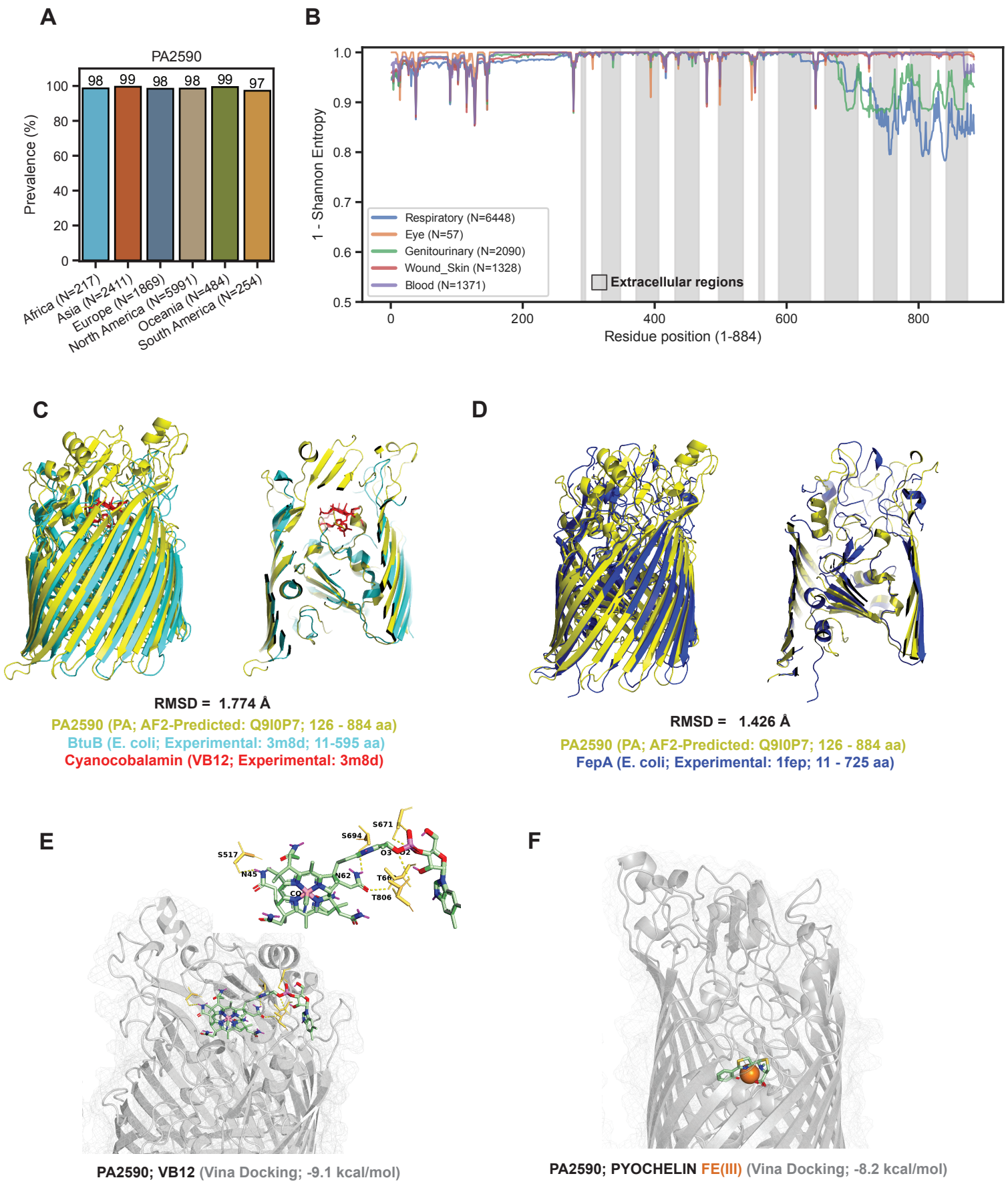


Figure 6. Structural homology-analysis predicts PA2590 gene to function as a vitamin B12 and iron transporter.

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.

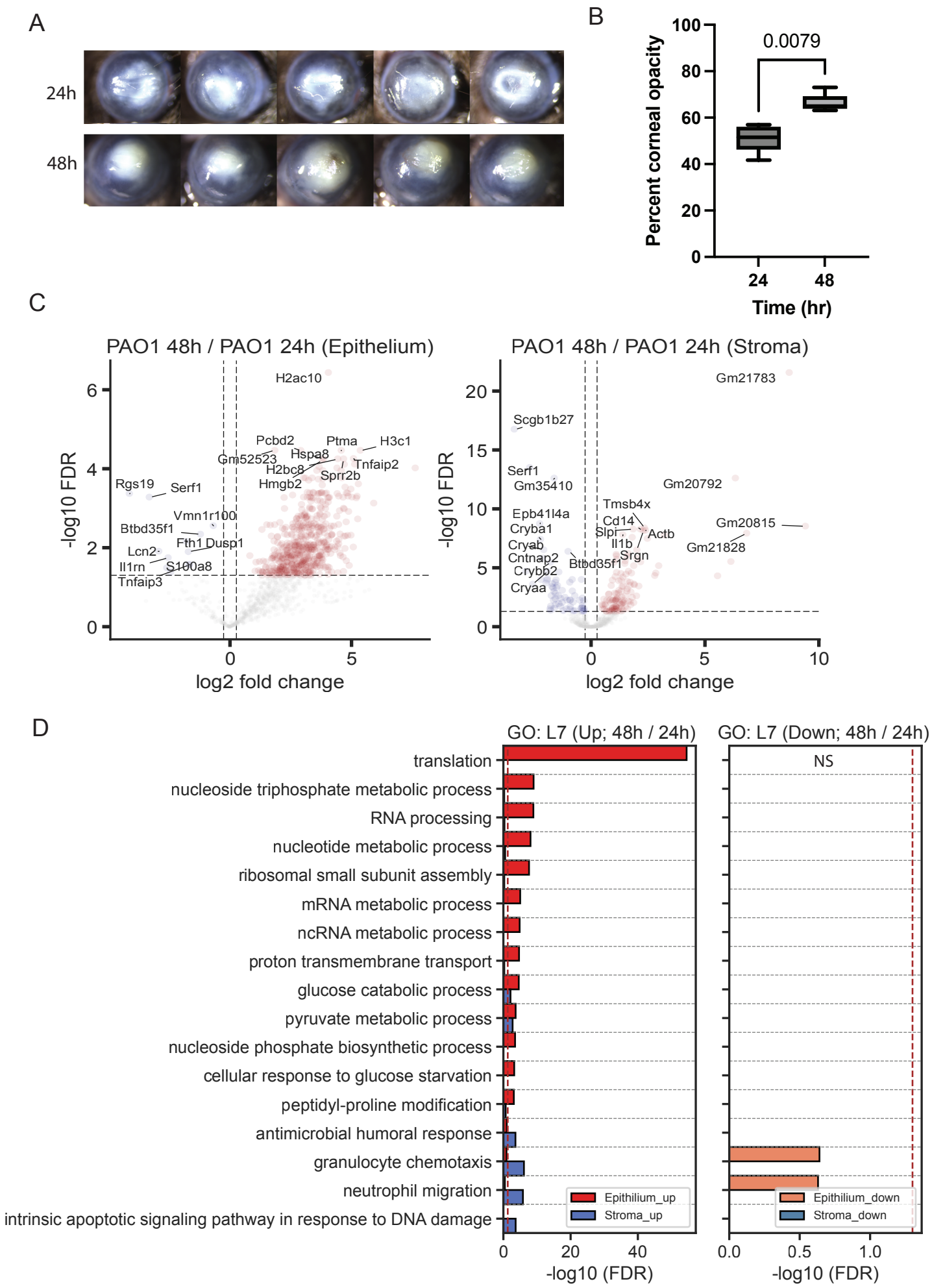
D. Structural homology modeling highlights parallels between the PA2590 structure predicted by AlphaFold2 and the experimentally determined structure of the E. coli FepA protein. The figure depicts the predicted model (light green) with the FepA structure (dark blue). The RMSD is 1.426 Å over amino acid 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.

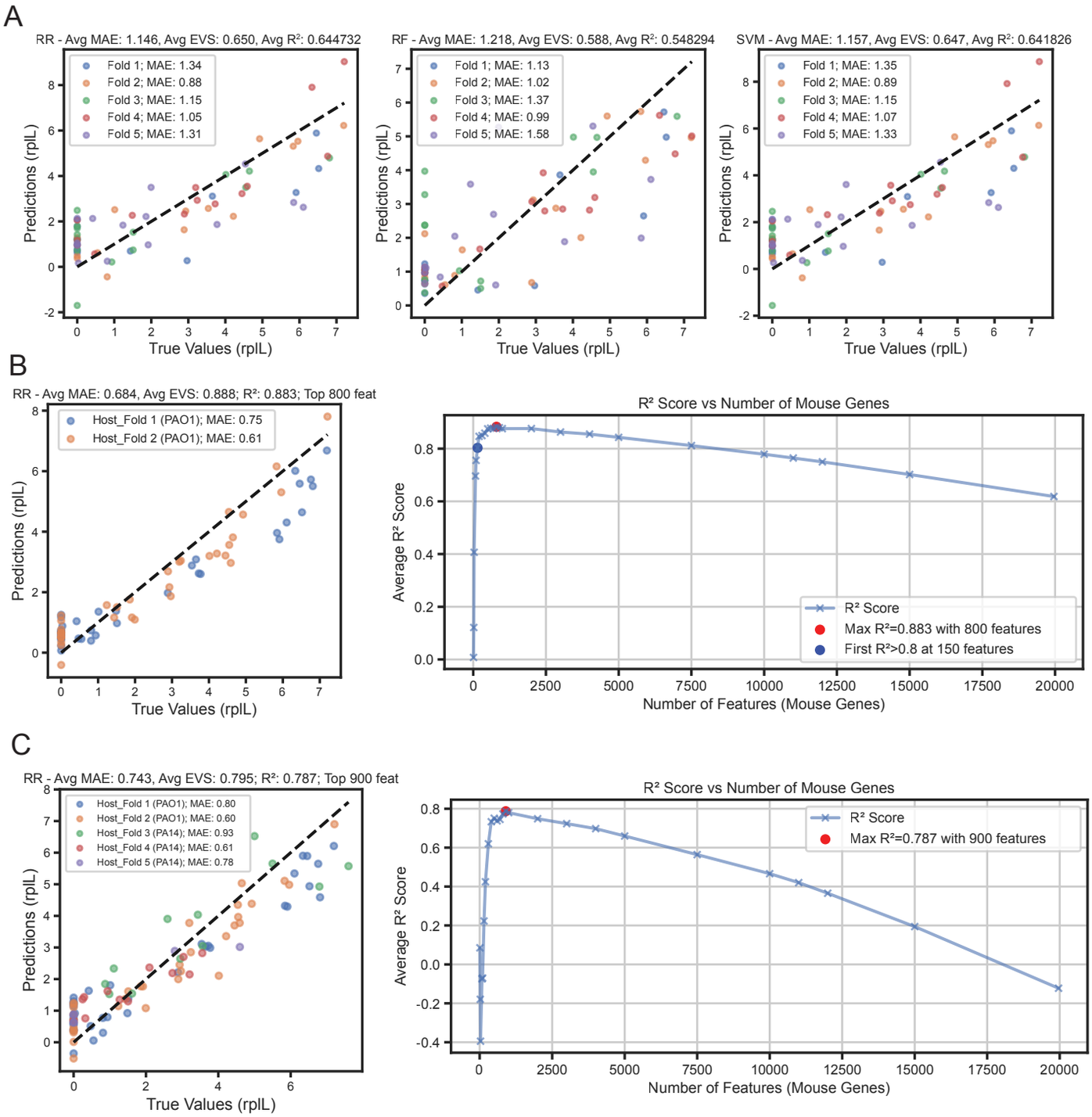
A. Representative images from infected with 5x105 CFU P. aeruginosa PAO1 strain C57Bl6/J mice (N=5 biological replicas). Images are collected at 24h and 48h post-challenge.

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



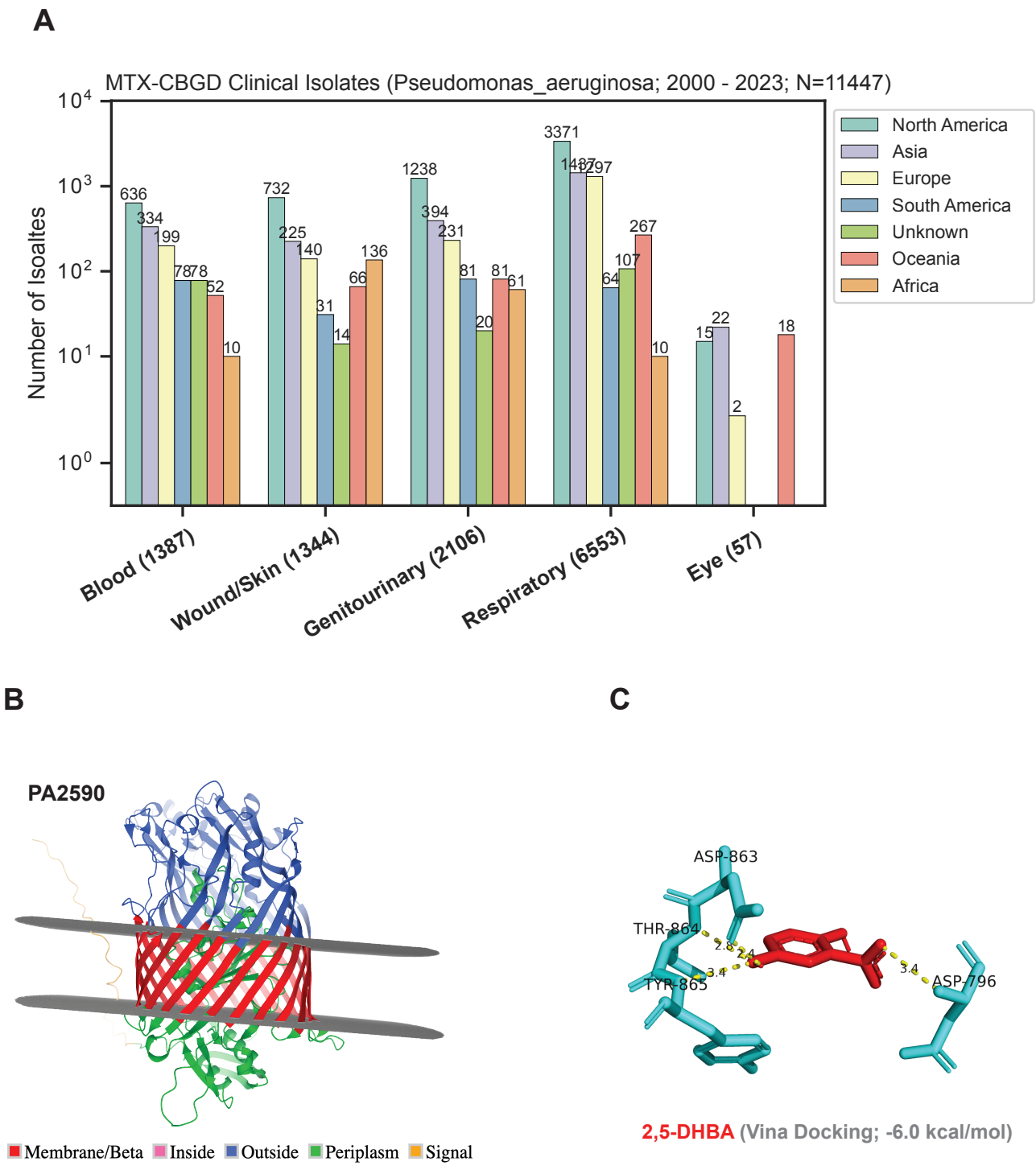
Supplementary Figure 2. Evaluation of model performance in predicting pathogen-specific gene expression based on host response profiles from multiple biological replicas

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.