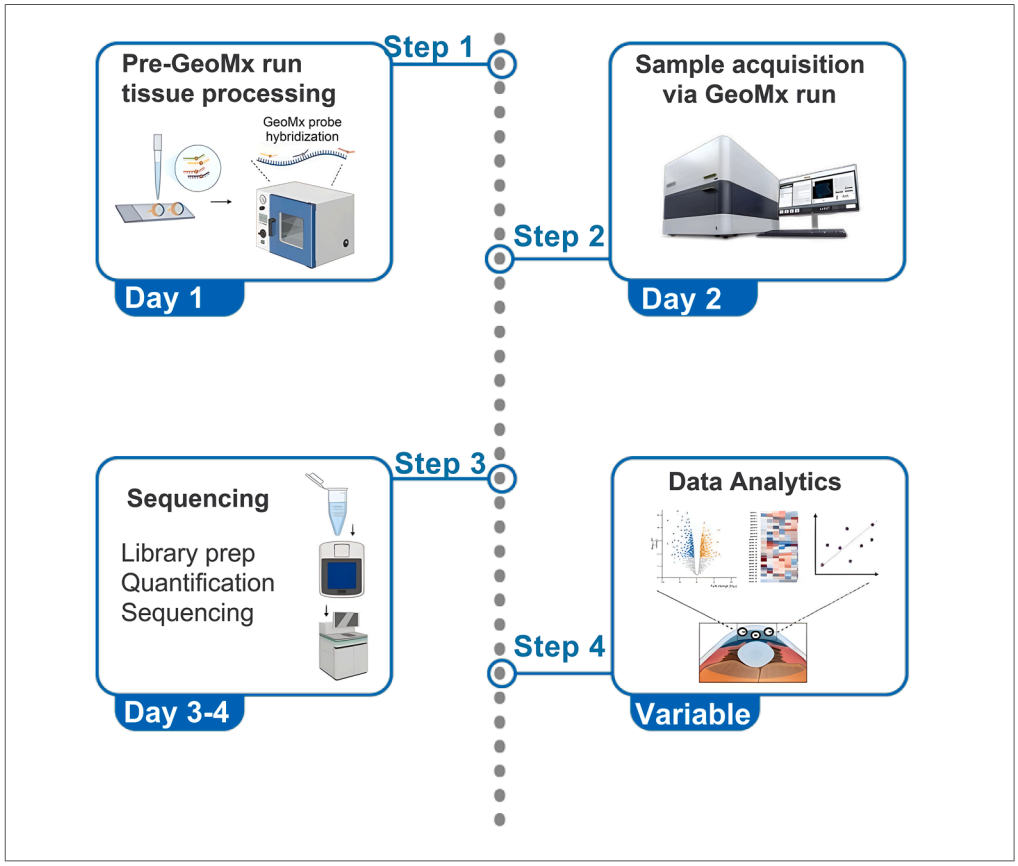
STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for dual spatial transcriptomic profiling of infected tissues



Dual spatial transcriptomics analysis offers profiling of host- and pathogen-specific transcriptional patterns in infected tissues to establish pathoadaptation signatures and predict infection outcomes. Here, we present a protocol for tissue processing, imaging, selection of regions of interest, library preparation, sequencing, and analysis pipeline. This protocol has a potential application for the analysis of any infected tissue.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Hao Zhou, Ankush Tyagi, Maxine McClain, Esperanza Anguiano, Serena Abbondante, Eric Pearlman, Mihaela Gadjeva

hao.zhou@ku.ac.ae (H.Z.)
mihaelagadjeva@gmail.com (M.G.)

Highlights

Maps host and pathogen transcripts in the same tissue sections

Works on FFPE slides with WTA and custom pathogen probes

Defines ROI by GeoMx and yields high-depth, high-quality libraries

Links host profiles to local bacterial burden using machine learning

Zhou et al., STAR Protocols 7, 104282
March 20, 2026
https://doi.org/10.1016/j.xpro.2025.104282

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for dual spatial transcriptomic profiling of infected tissues

Hao Zhou,[1,2,5,*] Ankush Tyagi,[3] Maxine McClain,[3] Esperanza Anguiano,[3] Serena Abbondante,[4] Eric Pearlman,[4] and Mihaela Gadjeva[3,6,*]

[1]Department of Biological Sciences, College of Medicine and Health Sciences, Khalifa University, Abu Dhabi, United Arab Emirates

[2]Center for Biotechnology (BTC), Khalifa University, Abu Dhabi, United Arab Emirates

[3]Bruker Spatial Biology, 3350 Monte Villa Parkway, Bothel, WA 98021, USA

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

[5]Technical contact

[6]Lead contact

*Correspondence: hao.zhou@ku.ac.ae (H.Z.), mihaelagadjeva@gmail.com (M.G.)
https://doi.org/10.1016/j.xpro.2025.104282

SUMMARY

Dual spatial transcriptomics analysis offers profiling of host- and pathogen-specific transcriptional patterns in infected tissues to establish pathoadaptation signatures and predict infection outcomes. Here, we present a protocol for tissue processing, imaging, selection of regions of interest, library preparation, sequencing, and analysis pipeline. This protocol has a potential application for the analysis of any infected tissue.

For complete details on the use and execution of this protocol, please refer to Zhou et al.[1]

BEFORE YOU BEGIN

P. aeruginosa poses a significant threat to global health, with antimicrobial resistance on the rise. The pathoadaptation mechanisms of the pathogen are understudied, and vaccines are lacking. New technologies, such as spatial transcriptomics analysis described here, offer detailed insights into the transcriptional plasticity of the pathogen during infection, thereby facilitating an improved understanding of the virulence repertoire. This protocol outlines the steps required to carry out spatial transcriptomics analysis using infected tissues.

Innovation

Conventional transcriptomics approaches primarily focus on characterizing host responses during infection with little emphasis on understanding pathogen-specific transcriptional adaptation mechanisms. The key innovation of this protocol is in the simultaneous capture of the transcriptional signatures of, both, the host and the pathogen. This dual profiling generates unified, spatially resolved map of infection that not only supports predictive modeling of disease dynamics and spread but also enables the discovery of previously unrecognized virulence factors.

Institutional permissions

All animal procedures and housing requirements were approved by the UC Irvine IACUC under protocol AUP-21-123. The infection studies were carried out with N=5 mice/cohort/experiment. Sample sizes were estimated based on prior experience. Exclusion criteria included cases where animals presented with wounding or corneal opacity in the eye that was not caused by infection.

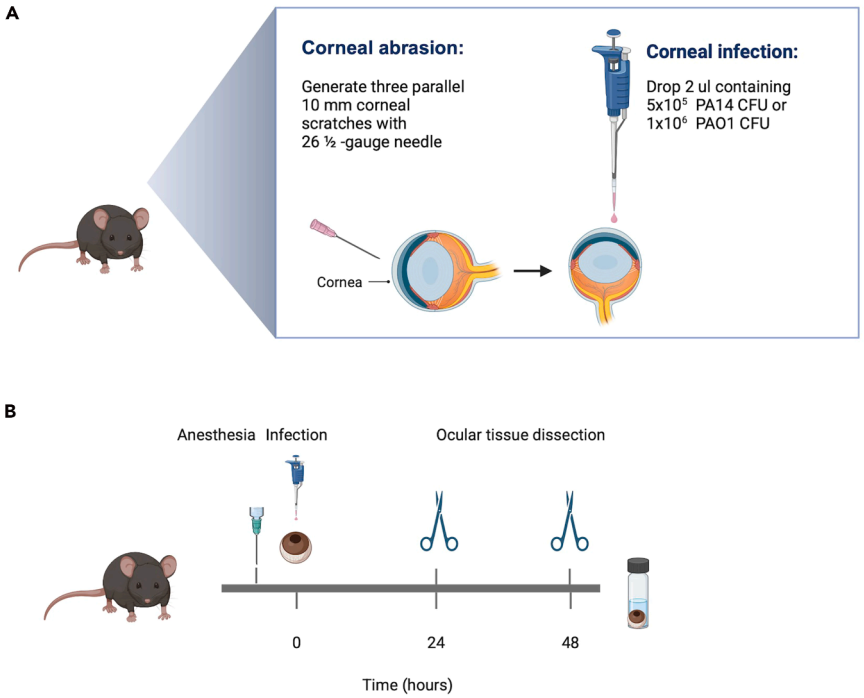


Figure 1. Ocular infection

(A) Setup and corneal abrasion technique. Mouse corneas are scratched using 26 ½ -gauge needle and 2 mL of bacterial suspension containing 5×10^5 CFU of the *P. aeruginosa* PA14 strain or 1×10^6 CFU of the *P. aeruginosa* PAO1 strain are placed onto the abrased corneas.

(B) Diagram depicting infection dissection timeline. Created in BioRender. G, M. (2025) https://BioRender.com/4fzmasq.

Infection

Timing: 24 h

This protocol is developed with tissues derived from an ocular infection model (Figure 1).

1. Anesthetize C57BL/6J mice with ketamine/xylazine.
2. Abrade corneal epithelium with three parallel 10 mm scratches with 26 ½ -gauge needle.
3. Apply 2 µL bacterial suspension topically onto the cornea.
4. Monitor infected animals and euthanize at 24 h post-pathogen challenge.

Note: Bacterial infection inoculum may vary depending on the isolate. For our experiments we used 5×10^5 CFU of the *P. aeruginosa* PA14 strain/2 µL suspension and 1×10^6 CFU of the *P. aeruginosa* PAO1 strain in 2 µL suspension dropped onto mouse corneas (Figure 1A). This protocol reflects experiments using tissues from animals infected for 24 h or 48 h (Figure 1B). The timing of the analysis can vary depending on the research objectives.

Preparation of formalin-fixed tissue blocks and slides

Timing: 48 h

This module describes tissue processing steps to generate formalin fixed tissue blocks and sections (Figure 2).

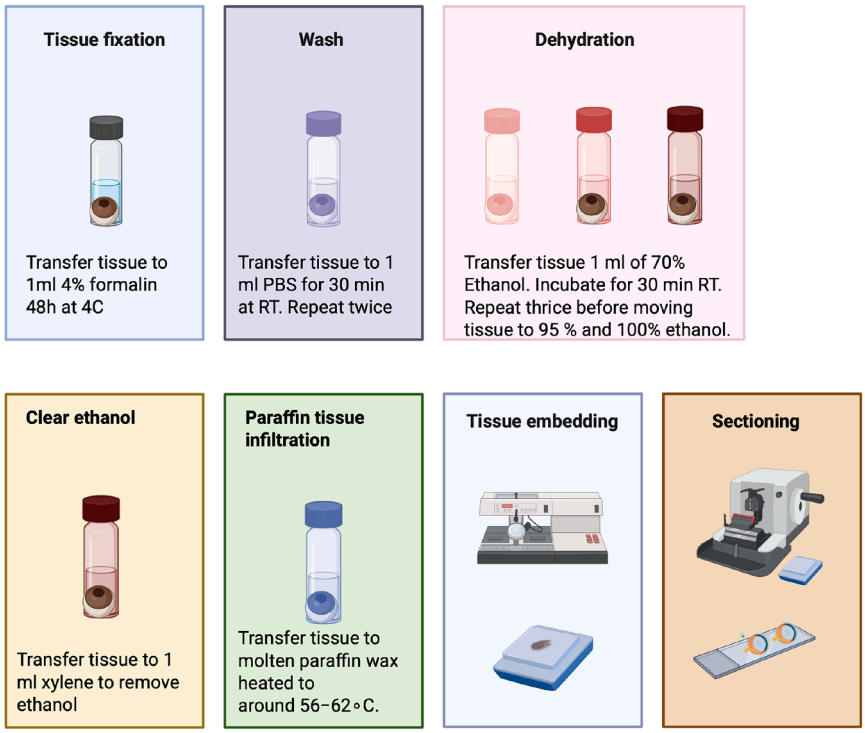


Figure 2. Tissue fixation and paraffin embedding workflow
Created in BioRender. G, M. (2025) https://BioRender.com/hgxeq50.

5. Dissect out the ocular tissues after euthanasia and place in 1 mL of iced 4% formalin.
6. Fix tissues at 4°C for 24 h.

Note: We recommend 24 h of tissue fixation for optimal performance. However, the fixation timing may vary depending on the tissue type and size. Optimal timing may need to be explored independently.

7. Place samples in 1 mL of 70% Ethanol for 30 min at 22°C–27°C, repeat for three times using fresh 70% Ethanol each time.

Pause point: Samples can be transferred to PBS at 4°C for up to a month for long-term storage.

8. Transfer tissue to 95% Ethanol for 30 min at 22°C–27°C, repeat for three times using fresh 95% Ethanol each time.
9. Transfer tissue to 100% Ethanol for 30 min at 22°C–27°C, repeat for three times using fresh 100% Ethanol each time.
10. Transfer tissue to 1 mL Xylene for 30 min at 22°C–27°C, repeat for three times using fresh Xylene each time.
11. Transfer tissue to 1 mL of molted paraffin and incubate at 56°C–60°C for 1 h, repeat four times.
12. Paraffin-embed tissues for downstream processing.

Note: For the spatial transcriptomics analysis, we recommend using 5 μm tissue sections. Each slide can harbor at least two tissue sections which maximizes data output.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
PanCK (20 mg/mL diluted 1/40)	Invitrogen	Clone:AE1 + AE3 Cat:NPB2-33200AF532
OprI (5 mg/mL diluted 1/200)	Thermo Fisher Scientific	Opr17_VHHFC
Ly6G (1 mg/mL diluted 1/100)	Novus Biologicals	Clone: 1A8 NB253131AFD
cKit (final concentration: 200 ng/mL)	Novus Biologicals	Clone: 92D NB110-93600AF532
Bacterial and virus strains		
P. aeruginosa PAO1	Stover et al.[2]	https://doi.org/10.1038/35023079
P. aeruginosa PA14	Liberati et al.[3]	https://doi.org/10.1073/pnas.0511100103
P. aeruginosa PA14 PA2590Tn	Liberati et al.	https://doi.org/10.1073/pnas.0511100103
Chemicals, peptides, and recombinant proteins		
Luria–Bertoni broth	BD-Difco	Cat# 244620
10×PBS	Sigma-Aldrich	Cat# 5368
Ketamine HCl (100 mg/mL)	MWI Animal Health	Cat# 501072
Xylazine (XylaMed)	MWI Animal Health	Cat# 119711
Luria–Bertoni Agar	BD Difco	Cat# 244520
10% formalin	Thermo Fisher Scientific	Cat# SF100-4
Proteinase K	Thermo Fisher Scientific	Cat# AM2548
100% deionized formamide	Thermo Fisher Scientific	Cat# AM9342
20×SSC	Invitrogen	Cat# AM9763
10% Tween 20	Teknova	Cat# T0710
1× EDTA pH 9.0	Sigma-Aldrich	Cat# SKU SRE0063
100% Ethanol	Fisher Chemical	Cat# CAS64-17-5
Tris Base	Fisher Bioreagents	Cat# BP150-500
Elution buffer (10 mM Tris-HCl w 0.05 Tween 20, pH 8.0	Teknova	Cat# T1485
Bond Dewax solution	Leica Biosystems	Cat# AR9222
Bond TM epitope retrieval 2	Leica Biosystems	Cat# AR9640
GeoMx DSP RNA slide FFPE prep kit (contains Buffer W, Buffer R)	NanoString Technologies	Cat# 121300313
SYTO13	Invitrogen	Cat# S11364
GeoMx morphology kit	NanoString Technologies	Cat# SH003
Invitrogen Qubit quantification assay kit	Invitrogen	Cat # Q32854
DEPC-treated water	Invitrogen	Cat # 43-879-37
Xylene	Thermo Fisher Scientific	Cat # 1330-20-7
95% Ethanol	Sigma-Aldrich	Cat # 0280
70% Ethanol	Sigma-Aldrich	Cat # 64-17-5
Paraffin	Sigma-Aldrich	Cat # 41663
Deposited data		
Transcriptomics raw data and associated images	Zhou et al.[1]	https://zenodo.org/records/14667250
Code scripts	Zhou et al.[1]	https://github.com/hz424/spatial_RNAseq_analyses/tree/main
Experimental models: Organisms/strains		
C57BL6/J; 6–8 weeks old; male and female	The Jackson Laboratory	Cat# 000664
Oligonucleotides		
Probes targeting 100 P. aeruginosa genes	Zhou et al.[1]	Table S1
GeoMx NGS RNA WTA	NanoString Technologies	Cat #12141103
Software and algorithms		
Clustal Omega	Sievers et al.[4]	https://bio.tools/clustalo
MembraneFold	Gutierrez et al.[5]	https://biolib.com/KU/MembraneFold/
Autodock Vina	Trott et al.[6]	https://vina.scripps.edu/
DESeq2	Muzellec et al.[7]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
GOATools	Klopfenstein et al.[8]	https://github.com/tanghaibao/goatools

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
GeoMxNGSPipeline v.2.3.3.10	NanoString Technologies	
mtx-COBRA	Arora et al.[9]	https://doi.org/10.1016/j.compbiomed.2024.108114
Vaxign-ML	Ong et al.[10]	https://doi.org/10.1093/bioinformatics/btaa119
BioRender		https://BioRender.com
Other		
GeoMx Seq Code Pack_ABCD (comes with Master Mix)	NanoString Technologies	Cat# 121400205
GeoMx Seq Code Pack_EFGH (comes with Master Mix)	NanoString Technologies	Cat# 121400206
GeoMx DSP Collection Plate 96	NanoString Technologies	Cat# 100473
RNase/DNase free 96-well PCR plates	Thermo Scientific	AB-1400-L
AMPure XP beads	Beckman Coulter	NC9933872

STEP-BY-STEP METHOD DETAILS

Tissue processing for spatial transcriptomics

Timing: 48 h

This module describes the steps for tissue processing including removal of paraffin, tissue hydration, and exposure of the tissue to the library probes (Figure 3).

1. Process slides using BondRxV.
   a. Bake 5 μm thick sections from formalin-fixed paraffin embedded (FFPE) tissue blocks at 60°C for 1 h prior to loading on Bond RXm.
   b. Once loaded, bake the slides for 30 min at 60°C.
   c. Dewax with 150 μL of Bond Dewax Solution for 30 s 5 times at 72°C.
   d. Rehydrate with 150 μL of reagent-grade ethanol 3 times at 22°C–27°C for 15 min.
   e. Perform heat-induced epitope retrieval with Bond Tm epitope retrieval 2 Solution for 20 min at 100°C.
   f. Enzymatic digest with 0.1 mg/mL Proteinase K for 15 min at 37°C.
   g. Stop the digest with 10% Neutral Buffered Formalin for 5 min at RT.
   h. Wash with NBF Stop Buffer (0.1M tris, 0.1M glycine in DEPC-treated water) for 5 min at RT.
   i. Remove the slides from the Bond RXm.

   Note: A fully automated protocol for FFPE slide processing is available utilizing the BondRx instrument.

   Note: Non-FFPE samples like fresh frozen tissue blocks can be used in the workflow following an alternative protocol (https://university.nanostring.com/geomx-dsp-automated-slide-preparation-user-manual).

2. Hybridize with library probes.
   a. Hybridize the slides with 25 μL Nanostring mouse Whole Transcriptome Atlas and 12.5 μL custom *pathogen-specific* probes in 200 μL Buffer R per slide at 37° for 16–18h.
   b. Wash the slides with a 1:1 mix of 100% deionized formamide and 4× SSC for 25 min at 25°C.
   c. Discard the wash.
   d. Repeat the washing step by adding 250 μL 100% deionized formamide and 4× SSC per slide for 25 min at 25°C.

   Note: Bruker Spatial Biology offers, both, mouse and human whole transcriptome atlases containing probes to detect approximately 22,000 host transcripts. The pathogen-specific probes were custom-designed using *P. aeruginosa* PA14 genomic information[1–3] with a computational *in silico* preselection filters.[1,10]

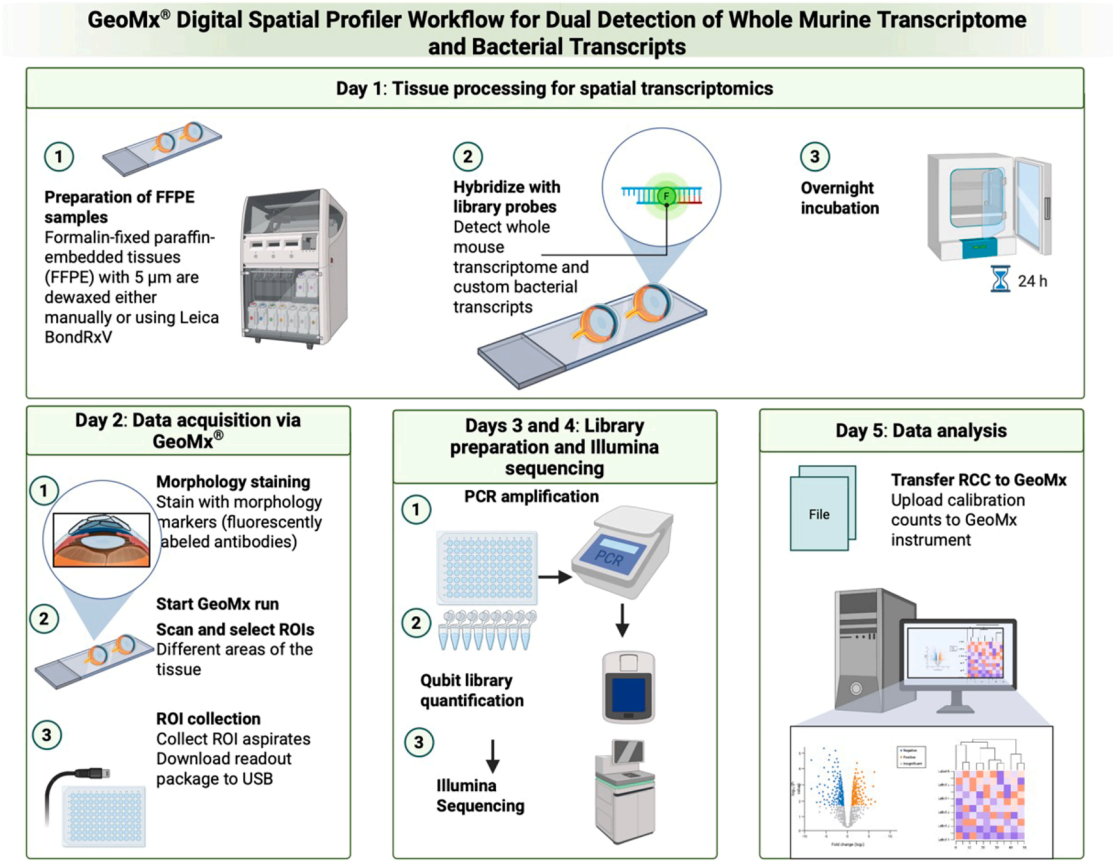


Figure 3. Spatial transcriptomics workflow using GeoMx DSP instrument

The procedure covers the following steps: Day1 - Slide processing using Bond RxV pipeline and probe library hybridization, Day 2 - GeoMx run consisting of staining of slides with morphology markers, GeoMx imaging, ROI selection, and sample acquisition, Day 3 - PCR-based library preparation, quantification, and sequencing, Day 4 - Data analytics. Created in BioRender. G, M. (2025) https://BioRender.com/p7q0gem.

Data acquisition via GeoMx

Timing: 24 h

Tissue sections were subjected to immunofluorescent staining to visualize tissue architecture and assist in the selection of regions of interest (ROIs). Following ROI selection, the GeoMx DSP instrument photocleaves the hybridized probes from each ROI, and the released oligonucleotides are subsequently collected for downstream analysis. See Figure 3 for a pipeline overview.

3. Tissue section morphology staining.
   a. Block the sections with 200 µL Buffer W for 30 min in a humidified chamber at 25°C.
   b. Stain the sections with the morphology marker antibodies: Syto13, anti-PanCK, anti-OprI, and anti-Ly6G in a total volume of 200 µL in Buffer W media for 1h in a humidified chamber at 25°C.
   c. Wash the slides with 200 µL in Buffer W media.
4. Process the slides using the GeoMx DSP instrument.
   a. Scan slides and define regions of interest (ROI) manually per each slide. Select three corneal epithelial ROIs, three stromal ROIs, and three anterior chamber ROIs per ocular tissue section. See Figure 4 for an example of a tissue scan with ROI selections.
   b. Collect individual ROI post UV-cleavage into a 96-well collection plate.

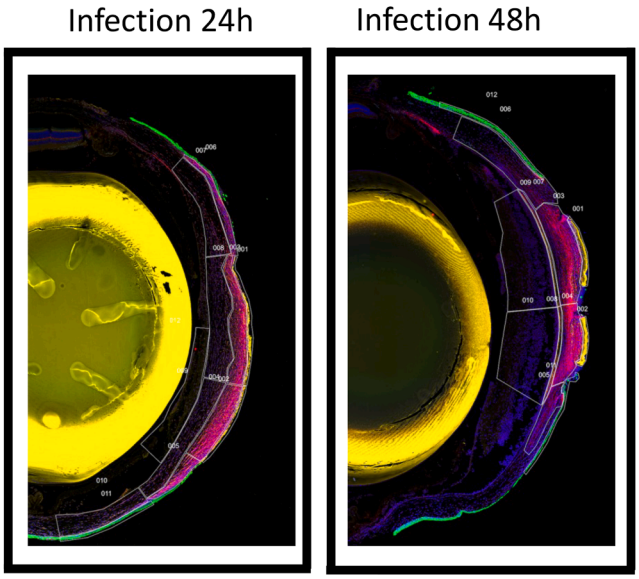


Figure 4. Representative images of GeoMx ROI selection on infected corneas

Cross-sections of murine eyes at 24 h (left) and 48 h (right) post-infection. Fluorescent morphology markers visualize tissue architecture and immune cell infiltration (pink), guiding the placement of geometric Regions of Interest (ROIs, outlined in white) for spatial transcriptomic analysis.

c. Seal the collection plate with a permeable membrane and dry down on a thermocycler at 65° C for 1h.

Note: At this step you can store the plate at −20°C for up to 30 days.

d. After the dry down, replace the permeable seal with a new permeable membrane.
e. Spin the plate down to ensure lack of remaining liquid.
f. Rehydrate the plate with 10 μL DEPC water for 10 min at 25°C.

Library preparation and Illumina sequencing

Timing: 48 h

The collected probes are used to generate libraries that are sequenced.

5. PCR amplification.
   a. In a 96-well PCR plate, add 2 μL of master mix, 4 μL of seq-code indexes and 4 μL of re-suspended barcodes.
   b. Mix and place in a thermocycler.

PCR cycling conditions

Steps	Temperature	Time	Cycles
Initial Incubation	37°C	30 min	1
Initial deactivation	50°C	10 min	1
Denaturation	95°C	15 s	18 cycles
Annealing	65°C	60 s	
Extension	68°C	30 s	
Final extension	68°C	5 min	1
Hold	4°C	forever	

6. PCR clean-up.
   a. Add 1.2× volume of AMPure XP beads to non-template Control (NTC) sample and pool tubes. Mix and incubate for 5 min at 25°C.
   b. Place on DynaMag-2 magnetic stand for additional 5 min at 25°C.
   c. While keeping the tube on the stand, discard the liquid.
   d. Washed the beads twice with 50 μL 80% Ethanol for 30 s each time.
   e. Allow to air dry for 5 min at 25°C.
   f. Elute the library from the beads by adding 54 μL of elution buffer to each tube followed by pelleting beads on the magnetic stand for 5 min.
   g. Transfer 50 μL of each elution to a new tube and add 60 μL of AMPure XP beads. Incubate for 5 min at 25°C.
   h. Pellet on the magnetic stand for 5 min, remove liquid, wash the beads twice with 80% Ethanol for 30 s each time. Air dry for 5 min.
   i. Add 5 μL of elution buffer to the NTC sample and 12 μL added to the library tube(s). Incubate at 25°C for 5 min.
   j. Measure library concentration with Qubit fluorimeter and assess quality by Tape Station.

   Note: The library concentration will depend on tissue type, tissue state, ROI size and should be taken into consideration when proceeding with sequencing to best fit experimental objectives.

7. Illumina sequencing.
   a. Adjust the final DSP library concentration to 1.25 nM in 12 μL.
   b. Spike with 3.8 μL 2.5 nM PhiX in 10 mM Tris HCl pH8.0 into the denatured library pool.
   c. 77 μL of 0.2 N NaOH is added to 310 μL of the pooled, undenatured library for S4 flow cell.
   d. Vortex the mixture, centrifuge at 280*g* for 1 min, incubate for 8 min at RT.
   e. Add 78 μL of 400 mM Tris-HCl (pH 8.0) to neutralize.
   f. Transfer the denatured library to the designated tube on NovaSeg6000 reagent cartridge.

   Note: The tube is kept on ice and loaded onto the instrument within 30 min to minimize library degradation and the risk of increased duplicate reads.

   Note: GeoMx libraries can be sequenced using other instruments and platforms like *Illumina NextSeq1000, NextSeq2000, or NovaSeq X Plus Platform.*

Computational analysis

Timing: 48 h

The computational analysis includes data conversion, normalization, and custom-generated pipeline. We explored differential gene expression, functional enrichment, predictive modeling and correlative analytics (Figure 5).

8. Sequencing data processing and quality control.
   The initial and most critical phase of the analysis is to convert raw sequencing data into a quantifiable format and ensure its quality.
   a. Initial data conversion: Process the raw FASTQ files from the Illumina NovaSeq 6000 system into Digital Count Conversion (DCC) files using the GeoMxNGSPipeline (v.2.3.3.10) from NanoString Technologies.

   Note: The pipeline performs probe alignment by matching each sequencing read to its unique probe sequence (Read Tag Sequence or RTS).

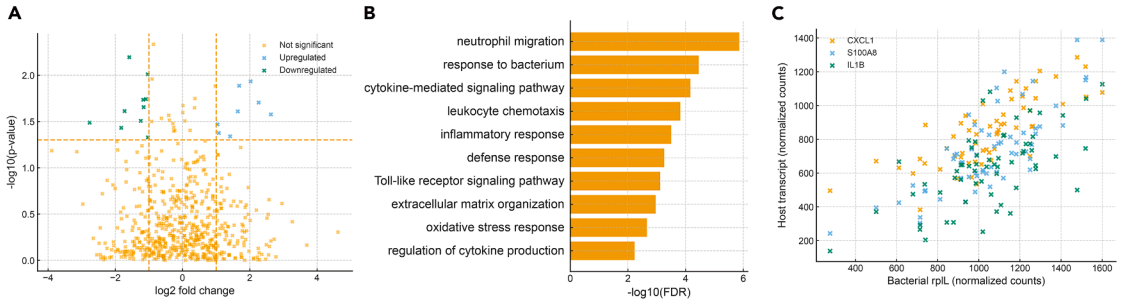


Figure 5. Illustrative data analysis outputs (non-data example)

(A) Volcano plot of differentially expressed host genes from a simulated analysis, highlighting significantly upregulated and downregulated genes using canonical thresholds ($|\log_2 FC| \geq 1$, $p \leq 0.05$).

(B) Bar chart of top enriched Gene Ontology (GO) terms for the Biological Process domain derived from a simulated list of upregulated genes, displayed as $-\log_{10}(FDR)$.

(C) Example simulated correlations between the bacterial transcript rplL and selected host transcripts (CXCL1, S100A8, IL1B). These panels are illustrative placeholders only; they are not generated from the study's experimental data; and are provided to clarify expected analytical outputs and figure formats.

b. UMI-based deduplication: Collapse PCR duplicates by identifying reads with the same Unique Molecular Identifier (UMI).

Note: This ensures the final count accurately reflects the original number of RNA molecules, removing amplification bias.

c. Quality control (QC): Load the generated DCC files into the GeoMx Digital Spatial Profiler software and assess each Region of Interest (ROI) based on the following metrics:
   i. Raw Reads: Ensure > 10^6 reads per ROI for sufficient statistical power.
   ii. Alignment Rate: Verify an alignment rate > 80% to confirm experimental specificity.
   iii. Sequencing Saturation: Confirm a sequencing saturation > 80% to ensure sequencing depth was adequate.

d. Background estimation and filtering:
   i. Estimate technical noise and the limit of quantification (LOQ) using External RNA Controls Consortium (ERCC) spike-in controls.
   ii. Assess potential contamination by analyzing the No Template Control (NTC) sample that was processed alongside experimental samples.

9. Data normalization and differential expression.
   After QC, the data is normalized to allow for accurate comparisons between different ROIs and conditions.
   a. Adjust for technical variations (e.g., ROI size, sequencing depth) within the GeoMx software. This process involves background subtraction using negative control probe counts and normalization based on library size (total counts) within each ROI.
   b. Differential gene expression (DGE) Analysis: Identify genes with significant expression changes between experimental groups using the DESeq2 package in R.

Note: This package models the count data using a negative binomial distribution.

c. Analyze DGE output: From the DESeq2 analysis, obtain a list of genes with their associated log2 fold change, p-value, and False Discovery Rate (FDR).

Note: Genes with an FDR < 0.05 are typically considered significantly differentially expressed.[7]

10. Functional Enrichment and Pathway Analysis.

Identifying a list of differentially expressed genes is only the first step; the next is to understand their collective biological meaning.

a. Gene ontology (GO) Analysis: Determine if specific biological themes are over-represented in the list of differentially expressed genes using the Python library GOATools.[8]
b. Review enriched pathways: Analyze the enriched GO terms across the three domains: Biological Process (e.g., "neutrophil migration"), Molecular Function (e.g., "cytokine receptor activity"), and Cellular Component (e.g., "extracellular space").

Note: The analysis uses a statistical test to identify pathways where the number of significant genes is greater than expected by random chance.

11. Predictive Modeling and Correlative Analysis.
This step leverages machine learning to integrate the host and pathogen data, aiming to build predictive models of the infection state.
    a. Model training: Train a machine learning model (e.g., Ridge regression) on the host transcriptomic profiles to predict the abundance of the *Pseudomonas aeruginosa* housekeeping gene rplL,[4] which serves as a proxy for local bacterial burden.
    b. Model validation: Assess the model's accuracy and generalizability using two methods:
        i. 5-fold cross-validation: Split the data into five parts, train on four, and test on the fifth, rotating through all parts.
        ii. Leave-one-out validation: Train the model on data from all but one host animal and test its predictive accuracy on the excluded host.
    c. Evaluate performance: Quantify the model's performance using Mean Absolute Error (MAE), Explained Variance Score (EVS), and the R^2 score.
    d. Identify predictive features: Use the trained model's coefficients to identify the most important host genes (biomarkers) for predicting bacterial load.
    e. Assess the direct relationship between individual host gene transcripts and the bacterial transcript using Spearman's rank correlation.
12. Pathogen-specific sequence and structural analysis.
For promising bacterial genes identified as important during infection (like PA2590), a deep bioinformatic investigation is performed to hypothesize their function.
    a. Homology and conservation analysis:
        i. Use BlastP to find homologous protein sequences across the NCBI database.
        ii. Align these sequences with Clustal Omega.[5]
        iii. Calculate Shannon Entropy for each position in the alignment to identify highly conserved amino acids critical for protein function.
    b. Membrane topology prediction: Use MembraneFold[6] to predict the protein's orientation within the bacterial membrane and identify potential extracellular loops that may interact with the host.
    c. Structural homology modeling: Predict the 3D structure of the target protein. Use Foldseek to identify a known protein structure with a similar sequence in the Protein Data Bank (PDB) to serve as a template.
    d. Molecular docking simulation: Perform in silico docking experiments with Autodock Vina[11] to predict how small molecules (ligands like nutrients or drugs) bind to the predicted 3D structure.

Note: A highly negative docking score suggests a strong, favorable interaction, providing hypotheses about the protein's function.

EXPECTED OUTCOMES

The spatial transcriptomic analysis generates tissue region specific transcriptional profiles with high reproducibility and sensitivity. Typically, a successful experiment should offer more than 10^6 raw

reads per ROI, an alignment rate exceeding 80%, and a sequencing saturation greater than 80%. Downstream computational analysis offers information on differentially expressed transcripts and, or transcripts enrichment across different ROIs thereby supporting biomarker discovery and validation pipelines, and mechanistic analysis.

LIMITATIONS

This approach offers tissue region-specific transcriptional signatures with the minimum number of 200 cells per region of interest. The analysis may require independent protein-based validation.

TROUBLESHOOTING

Problem 1

Tissue section detaches from slide during processing (related to "tissue processing for spatial transcriptomics" step).

Tissue detachment commonly occurs when slides lack sufficient adhesion, when tissue sections are not fully dried or baked prior to dewaxing, or when harsh buffer conditions/agitation dislodge the section during epitope retrieval and washing. These issues compromise downstream staining, hybridization, and imaging quality.

Potential solution

- **Insufficient slide adhesion**: Use positively charged slides (e.g., Superfrost Plus) to ensure strong electrostatic binding between the tissue and the slide surface.
- **Incomplete drying or baking**: Follow baking protocols strictly (e.g., 60°C for 1 h) to thoroughly melt paraffin and promote proper tissue adhesion before dewaxing.
- **Harsh retrieval or washing conditions**: Handle slides gently during manual washes and ensure Bond RXm liquid dispensing is calibrated to prevent high-pressure streams that may detach tissue sections.

Problem 2

Low RNA quality or weak RNA signal from FFPE samples (related to "preparation of infected tissues" step).

Low RNA integrity in FFPE tissues can result from prolonged fixation, ineffective deparaffinization or digestion, or degradation prior to fixation. Poor RNA quality reduces probe performance and overall assay sensitivity.

Potential solution

- **Over-fixation**: Standardize fixation to 24–72 h at 4°C. For long-term storage, transfer tissues to PBS after 70% ethanol to prevent additional cross-linking.
- **Sub-optimal digestion**: Optimize Proteinase K incubation based on tissue density. Dense/fibrotic tissues may require longer digestion, while delicate tissues may require shorter times; perform a pilot test if possible.
- **Pre-fixation degradation**: Minimize the time from dissection to fixation and keep tissues on ice to reduce RNase-mediated degradation.

Problem 3

Poor ROI segmentation or low signal-to-noise ratio during imaging (related to "data acquisition via GeoMx" step).

Suboptimal image quality often results from non-specific binding of morphology antibodies, high intrinsic tissue autofluorescence, or inadequate blocking. These issues can compromise ROI selection and downstream quantification.

Potential solution

- **Non-specific antibody binding**: Titrate morphology marker antibodies to determine the optimal concentration that maximizes signal while minimizing background.
- **High autofluorescence**: Verify filter compatibility on the GeoMx instrument and image using wavelengths that minimize intrinsic autofluorescence, which is pronounced in certain tissues (e.g., cornea).
- **Insufficient blocking**: Ensure thorough blocking (related to Step 3a). Perform antibody incubations in a humidified chamber to prevent evaporation and reduce non-specific staining.

Problem 4

Low sequencing quality or insufficient sequencing yield (related to Library Prep and illumina sequencing step).

Low read quality or under-clustering may occur due to inaccurate library quantification, presence of adapter dimers, or degradation of libraries prior to loading.

Potential solution

- **Inaccurate quantification**: Validate library concentration using at least two methods, such as Qubit fluorometry for concentration and TapeStation/Bioanalyzer for size distribution and purity.
- **Adapter-dimer contamination**: If a 120-bp peak is observed on TapeStation, perform an additional 0.2× AMPure cleanup to remove adapter dimers.
- **Library degradation**: Keep final libraries on ice and load onto the sequencer within 30 min of denaturation to prevent degradation.

Problem 5

High signal detected in the No-Template Control (NTC) (related to Library Prep and illumina sequencing step).

Elevated NTC counts suggest contamination of water, primers, master mix, or cross-contamination during pipetting. This can compromise data interpretation and sensitivity.

Potential solution

- **Reagent contamination**: Maintain physically separate pre-PCR and post-PCR areas with dedicated pipettes, filter tips, and reagents.
- **Cross-contamination during handling**: Aliquot water, master mix, and other reagents into single-use volumes to avoid repeated exposure of stock materials.
- **Persistent NTC signal**: Discard any suspect reagents and repeat library preparation with fresh materials.

Problem 6

Suspected batch effects in the final dataset (related to computational analysis steps).

Batch effects may arise when samples are processed on different days, with different reagent lots, or sequenced across multiple lanes. These can obscure true biological variation and compromise downstream analyses.

Potential solution

- **Experimental design**: Randomize biological replicates across batches, reagent lots, and sequencing runs whenever possible.
- **Data visualization**: Before differential expression analysis, perform PCA or UMAP on normalized counts and visualize clustering by batch date, operator, or sequencing run to assess potential batch structure.
- **Statistical correction**: If batch effects are detected, include batch as a covariate in your statistical model (e.g., in DESeq2, use design = ~ batch + condition).

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Mihaela Gadjeva (mihaelagadjeva@gmail.com).

Technical contact

Technical questions on executing this protocol should be directed to and will be answered by the technical contact, Dr. Hao Zhou (mihaelagadjeva@gmail.com).

Materials availability

This study did not generate unique materials.

Data and code availability

Original source data are available on Zenodo (Zenodo: https://zenodo.org/records/14667250).[12] Raw sequencing data have been deposited in NCBI GEO under accession number NCBI: PRJNA1202130. The complete computational workflow, including executable Jupyter notebooks, is available at Github: https://github.com/hz424/spatial_RNAseq_analyses/tree/main. The notebooks include Bacterial_target_expression_plot.ipynb: visualization of pathogen transcript abundance across ROIs; GO_enrichment_host_response.ipynb: Gene Ontology enrichment of host responses; Predictive_models.ipynb: Ridge regression-based modeling of bacterial load from host transcriptomes.

ACKNOWLEDGMENTS

Protocol deployment was funded by Moderna Therapeutics and Bruker Spatial Biology.

We would like to acknowledge the key roles of Moderna colleagues Oscar Negron, Nicole Chumbler, Brandon Jones, Edison Ong, and Christopher Tunkey and UC Irvine colleagues Serena Abbondante and Michaela Marshall in protocol deployment. We would like to thank University of Massachusetts Histopathology core, Yu Liu and Jayme Haywosz, for embedding and processing services.

AUTHOR CONTRIBUTIONS

H.Z. established and performed the computational analysis. A.T. and M.M. supported and advised the GeoMx workflow. E.P. and S.A. supported the underlying animal studies. H.Z., A.T., M.M., E.P., and M.G. played key roles in study design, troubleshooting, and data interpretation. M.G. coordinated and conceptualized the study. All authors discussed the protocol description, reviewed, and revised the manuscript.

DECLARATION OF INTERESTS

M.G., E.A., A.T., and M.M. are employee of Bruker Spatial Biology and may hold stock/stock options in Bruker Spatial Biology.

REFERENCES

1. Zhou, H., Negrón, O., Abbondante, S., Marshall, M., Jones, B., Ong, E., Chumbler, N., Tunkey, C., Dixon, G., Lin, H., et al. (2025). Spatial transcriptomics identifies novel Pseudomonas aeruginosa virulence factors. Cell Genom. *5*, 100805. https://doi.org/10.1016/j.xgen.2025.100805.

2. Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., et al. (2000). Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature *406*, 959–964. https://doi.org/10.1038/35023079.

3. Liberati, N.T., Urbach, J.M., Miyata, S., Lee, D.G., Drenkard, E., Wu, G., Villanueva, J., Wei, T., and Ausubel, F.M. (2006). An ordered, nonredundant library of Pseudomonas aeruginosa strain PA14 transposon insertion mutants. Proc. Natl. Acad. Sci. USA *103*, 2833–2838. https://doi.org/10.1073/pnas.0511100103.

4. Abildgaard, L., Schramm, A., Rudi, K., and Højberg, O. (2009). Dynamics of plc gene transcription and alpha toxin production during growth of Clostridium perfringens strains with contrasting alpha toxin production. Vet. Microbiol. *139*, 202–206. https://doi.org/10.1016/j.vetmic.2009.05.014.

5. Sievers, F., and Higgins, D.G. (2014). Clustal Omega. Curr. Protoc. Bioinform. *48*, 13.1–3.13.16. https://doi.org/10.1002/0471250953.bi0313s48.

6. Gutierrez, S., Tyczynski, W.G., Boomsma, W., Teufel, F., and Winther, O. (2022). MembraneFold: visualising transmembrane protein structure and topology. Preprint at bioRxiv. https://doi.org/10.1101/2022.12.06.518085.

7. Muzellec, B., Teleńczuk, M., Cabeli, V., and Andreux, M. (2023). PyDESeq2: a python package for bulk RNA seq differential expression analysis. Bioinformatics *39*, btad547. https://doi.org/10.1093/bioinformatics/btad547.

8. Klopfenstein, D.V., Zhang, L., Pedersen, B.S., Ramírez, F., Warwick Vesztrocy, A., Naldi, A., Mungall, C.J., Yunes, J.M., Botvinnik, O., Weigel, M., et al. (2018). GOATOOLS: a Python library for gene ontology analyses. Sci. Rep. *8*, 10872. https://doi.org/10.1038/s41598018289482.

9. Arora, I., Kummer, A., Zhou, H., Gadjeva, M., Ma, E., Chuang, G.Y., and Ong, E. (2024). mtx COBRA: subcellular localization prediction for bacterial proteins. Comput. Biol. Med. *171*, 108114. https://doi.org/10.1016/j.compbiomed.2024.108114.

10. Ong, E., Wang, H., Wong, M.U., Seetharaman, M., Valdez, N., and He, Y. (2020). Vaxign ML: supervised machine learning reverse vaccinology model for improved prediction of bacterial protective antigens. Bioinformatics *36*, 3185–3191. https://doi.org/10.1093/bioinformatics/btaa119.

11. Trott, O., and Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. *31*, 455–461. https://doi.org/10.1002/jcc.21334.

12. Zhou, H., Negrón, O., Chumbler, N., Gadjeva, M., and Pearlman, E. (2024). Spatial Transcriptomics Datasets: P. aeruginosa Corneal Infection (Zenodo). https://doi.org/10.5281/zenodo.14497751.