

RESEARCH

Open Access



ProbeST: a custom probe design pipeline for dual host–pathogen Spatial Transcriptomics

Sofia Rouot^{1†}, Ireen van Dolderen^{1,2†}, Patrick Rosendahl Andreassen^{3,4†}, Solène Frapard¹, Sybil A. Herrera-Foessel^{1,5}, Hailey Sounart¹, Sami Saarenpää¹, Julia A. Vorholt³ and Stefania Giacomello^{1*}

Abstract

Probe-based Spatial Transcriptomics profiles spatially-resolved transcriptomes using gene-specific probe pairs for both Formalin-fixed paraffin-embedded (FFPE) and Fresh Frozen samples. However, its applicability is restricted to human and mouse studies, due to commercial probe set availability. Here, we present ProbeST, an open-source computational pipeline for designing custom probe sets for genes of interest of a given organism. We validated ProbeST on FFPE mouse enteroid-derived monolayers infected with *Salmonella enterica* serovar Typhimurium, using custom pathogen probes with the available mouse probe panel. We simultaneously detected host and pathogen transcripts, with high probe specificity and low sensitivity against mCherry imaging, enabling identification of inflammatory response host genes *Mefv*, *Tnf*, and *Anxa1* colocalizing to pathogen genes. The reproducible ProbeST workflow expands probe-based Spatial Transcriptomics to studies of non-model organisms and host–pathogen interactions.

Keywords Spatial Transcriptomics, Probe Design, Pipeline, DualST, *Salmonella* Typhimurium, Mouse Enteroids, Host–pathogen interactions, Colocalization analysis, Formalin-fixed paraffin-embedded (FFPE) tissues

Background

The Spatial Transcriptomics technology (ST; 10X Genomics Visium) combines high-resolution histological imaging with sequencing-based 2D gene expression maps [1]. The original poly(A) capture approach [1, 2], compatible with Fresh Frozen (FF) tissue samples, underperforms in clinically-relevant FFPE tissue due to formalin-induced mRNA degradation and cross-linking between RNA and other biomolecules [3, 4]. To address these challenges, the Visium FFPE-compatible approach maintains sensitive transcript detection through gene-specific probe pairs [5], which consist of a Left-Hand Side (LHS) and a Right-Hand Side (RHS), each containing a hybridizing sequence and a probe handle (Fig. 1A). The LHS and RHS of a probe pair ligate only if they bind adjacent to each other on the target mRNA. After mRNA

[†]Sofia Rouot, Ireen van Dolderen and Patrick Rosendahl Andreassen first-shared authors.

*Correspondence:

Stefania Giacomello
stefania.giacomello@scilifelab.se

¹SciLifeLab, Department of Gene Technology, KTH Royal Institute of Technology, Stockholm, Sweden

²Present address: <https://www.re-peat.earth>

³Institute of Microbiology, Department of Biology, ETH Zürich, Zürich, Switzerland

⁴Present Address: Cytosurge AG, Opfikon, Switzerland

⁵Present Address: Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden



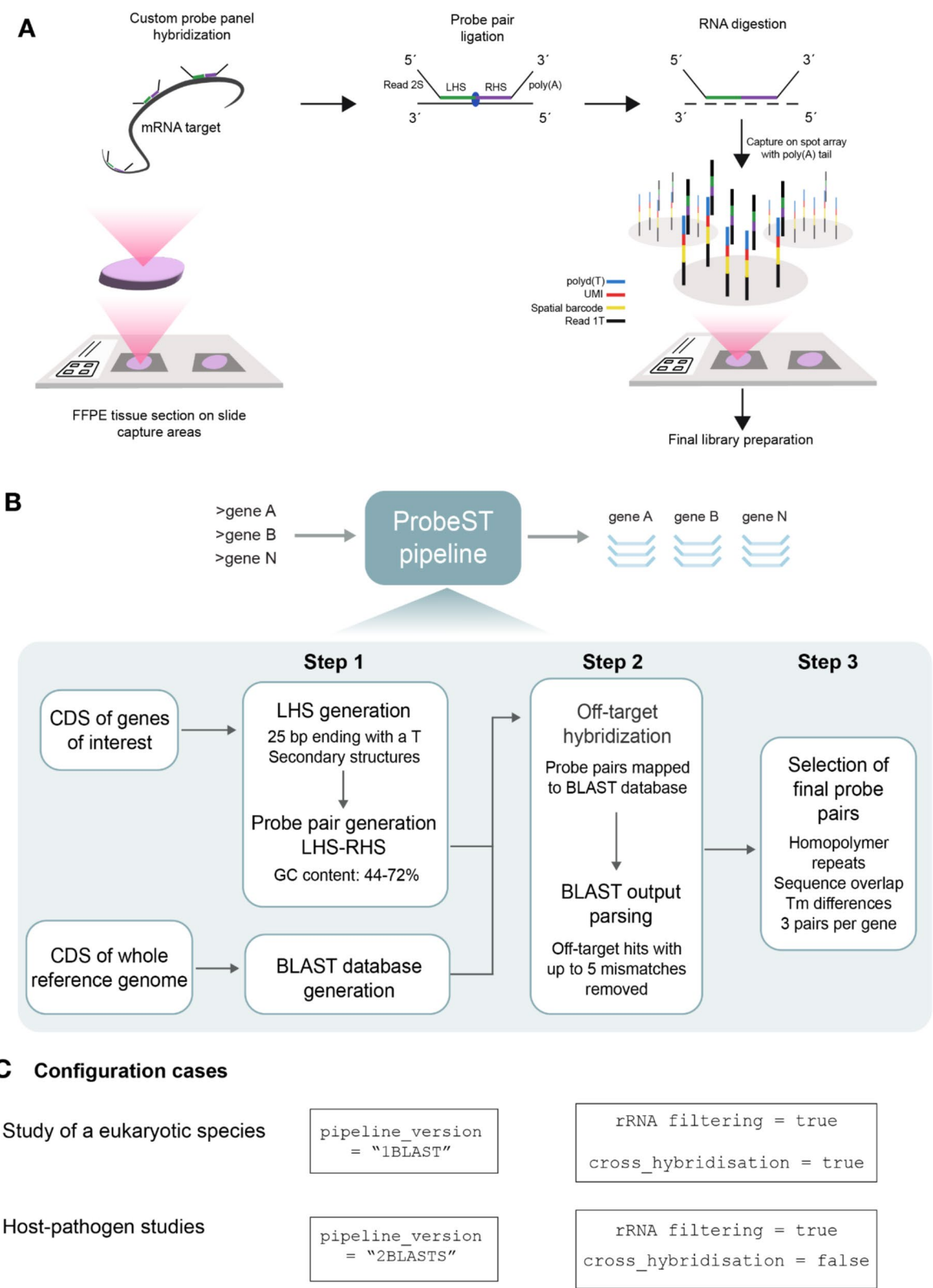


Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 An overview of the probe-based chemistry and the user-friendly ProbeST pipeline. **A** Main steps of the ST probe-based chemistry with a custom probe panel. For each mRNA target, up to three probe pairs are custom designed. Each probe consists of a 25 bp sequence reverse complementary to the sequence of the targeted transcript. The probe handle on the Left-Hand Side (LHS) probe contains the partial read 2S sequence used for downstream sequencing. The probe handle on the Right-Hand Side (RHS) probe consists of a poly(A) tail for probe capture on the Visium spot array. After probe hybridization, the probe sides are ligated to form a 50 bp probe pair. The mRNA undergoes digestion, and the probe pairs are captured by the spatial probes at specific spots of the array. Subsequent steps are performed to obtain final libraries ready for sequencing. Custom probes can be designed for any gene of interest. **B** General workflow of the ProbeST pipeline. Two input FASTA files are required. The first input file (CDS of genes of interest) is used to generate potential probe pairs, while the second input file (CDS of whole reference genome) is used to create a BLAST database. The generated potential probe pairs are then assessed for off-target hybridization against the BLAST database. The probe pairs that match to more than two sites in the genome with fewer than six mismatches are removed. If there are more than five mismatches, the match is considered insufficiently specific, resulting in a low likelihood of the probe pair binding to this target. The resulting output is parsed and up to three non-overlapping probe pairs per gene are selected for high specificity. **C** Configuration cases. The user can choose between two pipeline versions, with either 1 or 2 BLAST events. Optional steps can be included: the filtering step to remove rRNA off-target hybridization events, and the cross-hybridization step for poorly annotated genomes, to remove sequences originating from contaminants

degradation, the ligated probe pair is captured onto the Visium slide via the poly(A) tail on the RHS probe handle. This probe-based chemistry, despite requiring prior knowledge of gene sequences to design probe pairs, enables spatial profiling of whole transcriptomes and non-polyadenylated transcripts. Currently, commercially available probe sets exist only for the human and mouse whole transcriptomes, covering ~18,000 and ~19,000 genes, respectively. Therefore, custom probe design is essential to extend Visium FFPE to the study of additional organisms [6, 7].

We present ProbeST, an open-source and scalable pipeline to custom design probe pairs compatible with the probe-based Visium assays [8]. We experimentally validated ProbeST by designing probe pairs specific to coding regions of the *Salmonella enterica* subsp. *enterica* serovar Typhimurium strain SL1344 (*S. Tm*) and combined them with the mouse whole transcriptome probe set to analyze adult derived mouse intestinal organoids (enteroids) infected with *S. Tm* with high specificity, thus enabling the study of specific host–pathogen interactions. ProbeST is publicly available on GitHub as an automated Snakemake workflow, allowing reproducible and ease-of-use custom probe design for ST studies beyond human and mouse to any eukaryotic and prokaryotic organism [9].

Results

ProbeST custom probe design pipeline

We developed ProbeST to address the need for an automated custom probe design pipeline for Visium studies of genes of interest of an organism. ProbeST is designed to output a final probe set optimized for specific hybridization to target genes by meeting the following guidelines (as suggested by 10X Genomics probe design instructions [8]): the probes should target coding regions, the 25th nucleotide of the LHS probe (at the 3' end) must be a T, and the GC content should be between 44–72% for each probe. The probe design is compatible to all probe-based Visium assays, namely the Visium CytAssist Spatial Gene Expression (FFPE, Fresh Frozen, Fixed Frozen),

Visium Spatial Gene Expression for FFPE and the Visium HD Spatial Gene Expression.

ProbeST consists of three main steps integrated into a user-friendly Snakemake pipeline (Fig. 1B). In step 1, the pipeline first utilizes Primer3, an open-source program for designing hybridization probes [10], to generate 25 bp LHS probe sequences complementary to the Coding DNA Sequences (CDS) of the genes of interest, and then extends these sequences into 50 bp probe pairs (LHS + RHS). The LHS sequences are filtered based on secondary structure formation (see Methods). In step 2, ProbeST uses the CDS reference genome to include a BLAST step that removes probe pairs with off-target hybridization on the reference genome [11], corresponding to probe pairs with at least two BLAST hits of less than six mismatches [8]. In step 3, the pipeline selects the final probe pair set by keeping up to three non-overlapping probe pairs per gene, while excluding probes with homopolymeric regions longer than 5 bp. Finally, for each gene, the pipeline calculates and reports the melting temperature (T_m) of each probe as well as the T_m difference (ΔT_m) between them. ProbeST generates three files as final outputs, including a file formatted for use as input for the Space Ranger software, the bioinformatics tool used to analyze and process Visium data [12].

Configuration options of ProbeST are study-dependent and include off-target filtering against a second reference genome in host–pathogen cases, cross-hybridization filtering, and rRNA filtering (Fig. 1C). The cross-hybridization filter can be included to remove sequences in the CDS references that arise from cross-hybridization, defined as contamination from the environment. Indeed, poorly annotated genomes often contain sequences from environmental, viral, or prokaryotic DNA [13, 14]. The cross-hybridization step keeps probe pairs of higher specificity to the organism of interest and removes pairs with off-target hybridization to environmental, viral, or prokaryotic references. This step is strongly recommended for probe design from poorly annotated eukaryotic genomes. The rRNA filtering removes probe pairs binding to any rRNA sequences in the current SILVA

database, which includes all three kingdoms [15] (see [Methods](#)). Overall, ProbeST designs highly specific custom probe pairs that exclusively target the gene transcripts, including non-polyadenylated transcripts, and allows flexibility in its configuration.

Dual Spatial Transcriptomics study of *S. Tm* infection in mouse enteroid-derived monolayers as a demonstration of ProbeST custom probes

To assess the ProbeST pipeline, we applied it to study the *Salmonella enterica* subsp. *enterica* serovar Typhimurium strain SL1344 (*S. Tm*) infection process and the intestinal host response to pathogen invasion leveraging mouse enteroid-derived monolayers. We selected 19 genes from a previous study that investigated the transcriptional adaptations of *Salmonella enterica* during infection (Additional file 1) [16]. We used the *S. Tm* CDS of the selected genes as input for ProbeST, with the header format modified (see [Methods](#)), and generated a total of 51 *S. Tm* probe pairs (Additional files 2–5).

The gene panel included genes preferentially induced in *S. Tm*'s cytosolic and vacuolar life cycles, representing different stages of *S. Tm* infection (Additional file 1). Indeed, during *S. Tm* infection of the small intestine, the pathogen invades intestinal epithelial cells (IECs), adopts a vacuolar phase, and, depending on host cell conditions, can potentially transition to a cytosolic phase. The host IEC layer senses the invasion through the NAIP/NLRC4 inflammasome, leading to Caspase-1 activation, Gasdermin D-mediated pore formation, and cytokine IL-18 release [17]. The panel also included *S. Tm* housekeeping genes to capture highly abundant transcripts, as *S. Tm* transcripts are less abundant than host transcripts in the tissue.

To validate our custom probes, we applied dual Spatial Transcriptomics (DualST), an approach that uses two probe sets to simultaneously detect spatially resolved transcripts from both host and pathogen cells [6]. Specifically, we targeted 19,405 mouse genes with the commercially available 55,538 mouse probes and 19 *S. Tm* genes with our custom-designed 51 *S. Tm* probes, corresponding to a 1089-fold excess of host probes. We applied DualST to four enteroid-derived monolayers, representing different experimental conditions: wild-type non-infected (WT), WT infected (WT+), *Nlrc4*^{-/-} infected, and *GsdmD*^{-/-} infected, with an *S. Tm* infection load equivalent to about 500 bacteria per condition. NLRC4 is an inflammasome sensor that detects intracellular bacterial components, triggering epithelial defense responses including interleukin maturation, cell expulsion from the monolayer and GSDMD cleavage. Cleaved GSDMD N-terminal forms pores in the membrane that allows mature interleukins to be secreted and thus contributes to

pyroptotic cell death by amplifying host inflammation [18]. These samples were designed to assess the effect of *S. Tm* infection on host transcriptional responses and the contribution of *Nlrc4* and *GsdmD* for those.

ProbeST designs custom probes with high specificity

First, we tested the specificity of the custom probe set to detect *S. Tm* transcripts by comparing *S. Tm* signals across infected conditions (WT+, *Nlrc4*^{-/-}, *GsdmD*^{-/-}) and the control condition (WT). We observed a high number of *S. Tm* UMIs in the infected conditions (482 UMIs for WT+, 562 for *GsdmD*^{-/-} and 766 for *Nlrc4*^{-/-}), compared to only 24 UMIs in WT (Fig. 2A). The low background signal observed in the latter can be explained by unspecific binding. The dual capture of host and low-abundance *S. Tm* transcripts in the infected samples verified probe sensitivity, and the detection of *S. Tm* transcripts demonstrates that the ProbeST custom probes hybridize to the target genes, undergo ligation, and are detected with host probes, despite the 1089-fold difference in probe abundance. We observed that the total UMI counts per *S. Tm* gene varied considerably, from 764 UMIs for *sipC* and 340 for *lpp*, to 1 UMI for *uhpT* and *zinT* (Fig. 2B, Additional file 6: Fig. S1–3). The UMI variation within this gene set has a strong Spearman correlation (0.683) with the Transcript per Million (TPM) observed in an earlier study [16] (Additional file 7: Table S4), which supports this variation may reflect differences in transcript abundance.

We observed that the fraction of bacterial-positive spots across all four biological conditions was 4.76%, with 0.132 average *S. Tm* genes per spot, 0.171 average *S. Tm* UMIs per spot, and an average of 1.875 *S. Tm* UMIs per bacterial-positive spot (Additional file 8). Additionally, *S. Tm* genes that were more abundant showed a higher number of spots with only that bacterial gene detected, termed singleton spots (Fig. 2C). The distribution of *S. Tm* UMIs per spot, or the amount of spots detecting a given amount of *S. Tm* UMIs, revealed that most spots which captured the pathogen transcripts have up to 3 UMIs, while only one spot captured 27 UMIs across all pathogen genes, highlighting the sparse nature of the bacterial information (Additional file 6: Fig. S4).

mCherry fluorescence validates the high specificity of ProbeST custom probes

Next, we validated *S. Tm* custom probe detection by comparing the DualST signal with mCherry fluorescence, a constitutively expressed red fluorescent protein in the *S. Tm* used for infection studies [17]. The mCherry imaging approach was ideal for enteroid-derived monolayers grown directly on slides, as it allowed bacterial imaging on the same monolayer prior to the Visium workflow. We developed a computational approach to

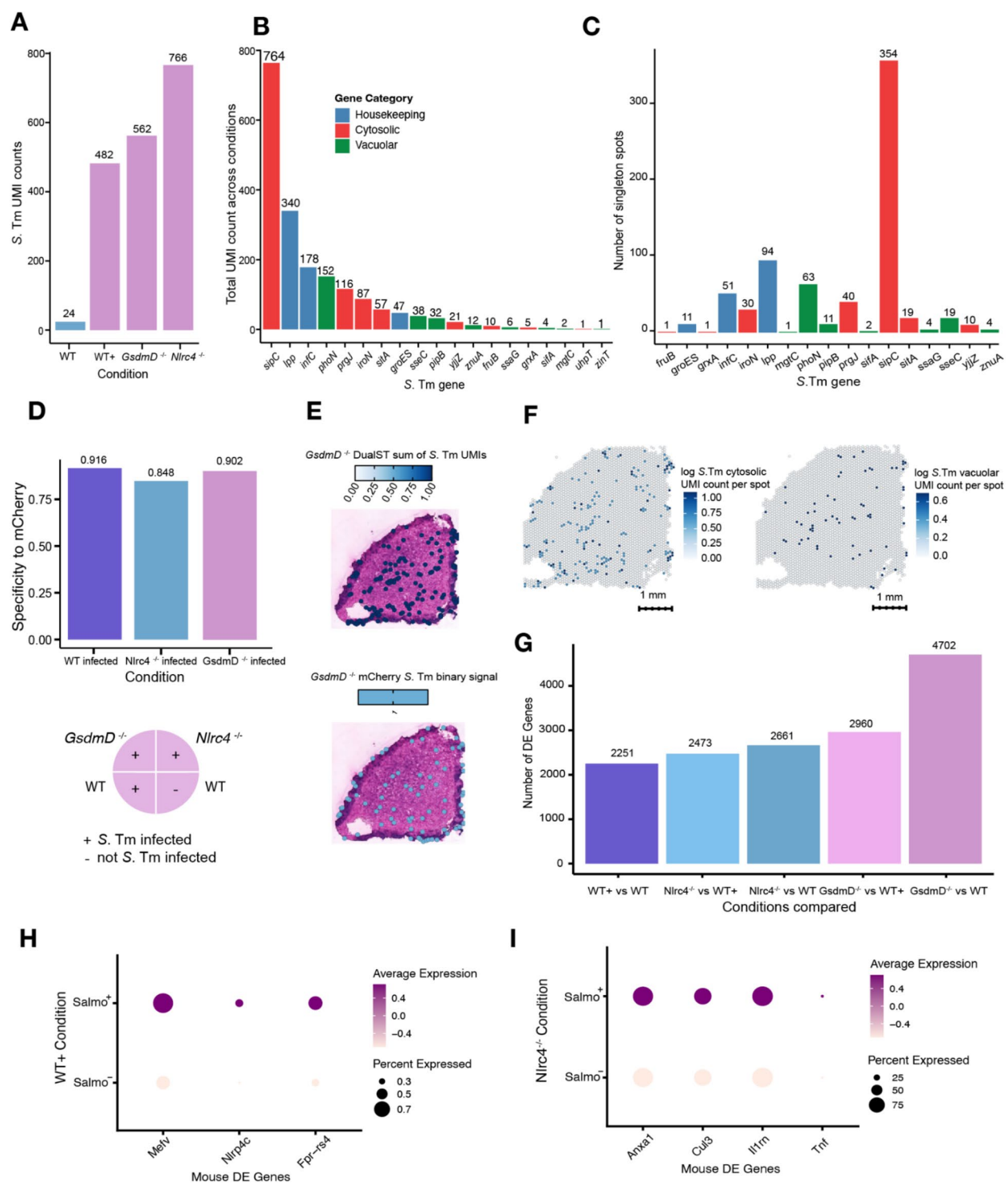


Fig. 2 (See legend on next page.)

compare spot coordinates from DualST data with pixel coordinates from mCherry fluorescence imaging (see [Methods](#)). Across all three infected enteroid conditions, we observed a spatial overlap with high specificity (WT+: 0.916, *GsdmD*^{-/-}: 0.848, and *Nlrc4*^{-/-}: 0.902) and lower sensitivity (WT+: 0.167, *GsdmD*^{-/-}: 0.163, and

Nlrc4^{-/-}: 0.117) (Fig. 2D). The high specificity confirmed that the spatial distributions of the DualST and the mCherry signals generally match. Indeed, we observed a similar spatial distribution of the *S. Tm* UMI counts and the mCherry signal across all three infected conditions (Fig. 2E, Additional file 6: Fig S5). The lower sensitivity

(See figure on previous page.)

Fig. 2 Application of ProbeST for DualST of *S. Tm* infection in mouse enteroid-derived monolayers. **A** *S. Tm* detection by the ProbeST custom probe panel in the three infected enteroid-derived monolayer conditions WT+, *Nlrc4*^{-/-} infected, and *GsdmD*^{-/-} infected (pink), compared to the control non-infected enteroid-derived monolayer WT (blue). The sum of UMI counts across all *S. Tm* unique gene transcripts was calculated for each condition. **B** Plot of the total UMI counts for each *S. Tm* gene across all spots in all four conditions. **C** Number of singleton spots for each bacterial gene, defined as spots detecting only UMIs from a single *S. Tm* gene. The genes *uhpT* and *zinT* have no singleton spots and are not shown. **D** Top: Validation of the *S. Tm* DualST signal with the fluorescent mCherry signal. The mCherry spatial coordinates were aligned to the DualST coordinates with a computational approach. Specificity was calculated for each infected condition through a confusion matrix, with the mCherry signal considered as ground-truth. Bottom: schematic overview of the shapes of the enteroid-derived monolayer conditions, indicating which conditions were infected by *S. Tm*. **E** The custom probe *S. Tm* DualST coordinates (top) and the mCherry coordinates (bottom) mapped onto the *GsdmD*^{-/-} infected condition. The DualST UMI counts shown are log₁₀-normalized with parameter max_cutoff = 0.99. **F** Spatial mapping of the targeted cytosolic-specific and vacuolar-specific *S. Tm* genes in the *GsdmD*^{-/-} infected condition. The UMI counts are log₁₀-normalized. **G** Total number of mouse DE genes between different enteroid conditions. **H** Colocalization analysis. Dot plot of the mouse DE genes between spots with *S. Tm* probe capture (Salmo⁺) and spots without *S. Tm* capture (Salmo⁻) in the WT+ condition. **I** Colocalization analysis. Dot plot of the mouse DE genes between Salmo⁺ and Salmo⁻ spots in the *Nlrc4*^{-/-} infected condition. UMI = Unique Molecular Identifier; DE = Differentially Expressed

likely resulted from several factors. First, mCherry fluorescence is not at single-molecule resolution and is inherently less sensitive than probe-based detection. Moreover, a likely explanation for the low sensitivity in this study is potential competition of probe sequences on the target mRNA. Indeed, we found that for genes with more than 1 probe pair, 0–90% of the 50 bp sequences overlapped (Additional file 9). We addressed this issue in the ProbeST pipeline, reducing the overlap to 0% between probe pairs targeting the same gene. A higher number of genes targeted could also help to increase bacterial detection and sensitivity, as well as an extra permeabilization step for the bacterial cells. Overall, the spatial overlap between fluorescent and probe *S. Tm* signals validated that custom probes captured *S. Tm* transcript distribution with high specificity and low sensitivity.

Gene-level analyses of the custom *S. Tm* probes allow gene-pair colocalization analysis of the cytosolic and vacuolar lifestyles

While the 55 µm spot resolution and the selected bacterial genes did not enable a distinction of cytosolic *S. Tm* and vacuolar *S. Tm*, the latter could infer whether genes associated with the cytosolic or vacuolar lifestyle exhibited pairwise spatial co-expression. We observed the detection of both cytosolic-associated and vacuolar-associated genes in the infected conditions (Fig. 2F), suggesting that we captured *S. Tm* from both stages of the infection process. We also observed a higher abundance of cytosolic-associated genes, however this may be driven by *sipC*, the gene with the highest UMI abundance. Lifestyle-specific gene-pair colocalization analysis using Fisher's exact test revealed several gene pairs that significantly colocalize (*p*-value < 0.05), including *sitA-iroN*, *sipC-prgJ*, *sitA-fruB*, and *sitA-yjz* among the cytosolic lifestyle genes, and *sseC-sifA*, *znuA-phoN* and *sseC-pipB* among the vacuolar lifestyle genes (Additional File 6: Fig. S6–7). These lifestyle-specific pairs have previously been described as being co-regulated, and indeed yielded a high Pearson correlation from the RNA-seq data of a previous study [16] (Additional file 7: Tables S7–8). For

instance, colocalizing cytosolic gene pair *sitA-iroN* (Pearson correlation of 0.58) is among the genes shown to be up-regulated by low Fe²⁺ + shock in vitro, and the vacuolar pair *sseC-pipB* (Pearson correlation of 0.98) is positively co-regulated when *Salmonella* is within the *Salmonella*-containing vacuole (SCV). Spatial gene-pair colocalization analysis identified previously reported genes whose expression is co-regulated in different stages of the *Salmonella* infection process.

To further investigate host–pathogen interactions at the spatial level, we performed a colocalization analysis to examine localized host transcriptional changes at *S. Tm* infection sites, enabled by our ProbeST custom probe set.

ProbeST allows host–pathogen colocalization exploratory analysis and highlights the role of the innate immune system in *S. Tm*-infected IECs

To investigate condition-specific host responses to *S. Tm* infection, we performed Differential Expression Analysis (DEA) between WT and each of the infected conditions (WT+, *GsdmD*^{-/-}, and *Nlrc4*^{-/-}).

The comparison between infected *GsdmD*^{-/-} and uninfected WT conditions yielded 4702 significantly differentially expressed (DE) genes (adjusted *p*-value < 0.05), the highest among all comparisons, followed by 2960 DE genes observed between infected *GsdmD*^{-/-} and infected WT+enteroids (Fig. 2G, Additional file 10). While DE gene counts do not directly reflect infection severity, these differences suggest distinct transcriptional responses in the absence of *GsdmD* and *Nlrc4*. *GsdmD* encodes a pore-forming effector involved in pyroptotic cell death downstream of inflammasome activation, and its deletion may alter IEC death and inflammatory signaling dynamics [19]. NLRC4, in contrast, acts upstream as a pattern recognition receptor sensing bacterial flagellin and type III secretion system components, triggering inflammasome activation and downstream responses, including cytokine release and, in IECs, epithelial cell extrusion.

While only one biological sample per condition was included, the spatial resolution of the data allows for an exploratory comparison of transcriptional trends across conditions. We observed significant downregulation of the transcription factors *Stat3*, *Fos*, *Fosl1*, *Egr1*, and *Atf3*, in both *GsdmD*^{-/-} and *Nlrc4*^{-/-} monolayers compared to WT and WT+ monolayers (Additional file 10: Tables S10–13). These transcription factors are known to be induced during early *S. Tm* infection and to regulate inflammatory and stress response pathways in epithelial cells [18, 20, 21]. Consistent with these findings, the GO term “STAT Family Protein Binding” (GO:0097677) showed significant downregulation in the *Nlrc4*^{-/-} and WT+ comparison (Additional file 11), suggesting attenuation of STAT-mediated signaling in the absence of NLRC4.

In contrast, the transcriptional differences between WT+ and WT enteroids were relatively modest. A likely explanation is the immature state of IECs in monolayer culture, which has been reported to dampen inflammasome activity following NLRC4 activation [22]. Nevertheless, several biological processes previously associated with host epithelial defense including regulation of nitric oxide metabolism, lysozyme production, and peroxisomal transport were upregulated in WT+ monolayer relative to WT monolayer (Additional file 12) [23–26]. Taken together, these spatially resolved transcriptional patterns are consistent with an altered epithelial response to *S. Tm* infection in the absence of NLRC4 or GSDMD. While biological replicates and further functional insights are needed to confirm these findings, the observed differences suggest that inflammasome signaling contributes to shaping the early transcriptional landscape of infected enteroids and demonstrates that DE genes can be detected between conditions.

To further delve into the local effect of *S. Tm* infection on the mouse IEC transcriptional profile, we performed a colocalization analysis comparing gene expression profiles between spots that contained *S. Tm* signal (Salmo⁺, at least 1 UMI of any *S. Tm* gene) and spots with zero *S. Tm* signal (Salmo⁻).

In the WT+ condition, DEA and pathway analysis [27] revealed upregulation of *Mefv*, *Nlrp4c*, and *Fpr-rs4* (adjusted p-value < 0.05), genes involved in inflammatory responses (Fig. 2H, Additional file 13: Table S14, Additional file 6: Fig. S8). *Mefv* encodes pyrin, a key regulator of caspase-1 and IL-18 expression in innate immunity [28, 29], while *Nlrp4c* is a gene of unknown function that is part of the Nlrp4 family [30], some members of which have been suggested to be part of autoimmune responses.

In the *Nlrc4*^{-/-} condition, the genes *Anxa1*, *Cul3*, *Il1rn*, and *Tnf* were upregulated in Salmo⁺ spots (Fig. 2I, Additional file 13: Table S15, Additional file 6: Fig. S8). Notably, *Anxa1* plays a regulatory role during intestinal

mucosal inflammation [31], and *Tnf* is known to be induced by *S. Tm* flagellin upon intestinal invasion, acting as a proinflammatory signaling molecule for neighboring cells [32]. These findings suggest that inflammatory response genes are consistently activated in response to *S. Tm* infection in both WT+ and *Nlrc4*^{-/-} Salmo⁺ spots. Finally, the *GsdmD*^{-/-} Salmo⁺ spots revealed upregulation of *Tunar*, *Gsd3*, *Aadacl4*, and *Hsf2*, genes with predictive transmembrane helix domains (Additional file 13: Table S16, Additional file 6: Fig. S8). Given the involvement of *GsdmD* in forming pores in the infected IEC membrane, the four genes *Tunar*, *Gsd3*, *Aadacl4*, and *Hsf2* might be upregulated as a result of the lack of *GsdmD* expression or pore-formation. Together, the exploratory analysis of differential gene expressions of innate immune regulators in Salmo⁺ spots suggest condition-specific local host responses to *S. Tm* infection, and highlight the use of custom probes to perform Spatial Transcriptomics host–pathogen colocalization studies.

Discussion

Overall, we present ProbeST, a Snakemake pipeline that automates and standardizes the design of custom gene-specific probe sets for ST studies of any organism in a reproducible manner. ProbeST minimizes manual input and enables the spatial capture of both polyadenylated and non-polyadenylated transcripts, supporting studies of non-model organisms and host–pathogen interactions. While the current ProbeST version does not yet account for parameters such as common SNPs, polymorphisms, bacterial horizontal gene transfer events or conserved gene families, future updates aim to improve design precision by incorporating such features.

ProbeST generates probe pairs with high specificity by filtering out probes with off-target binding events against the reference genome, and, when applicable, the host genome, rRNA sequences and against contaminant sequences in the reference genome. The last option, called cross-hybridization, accounts for environmental, prokaryotic and viral sequences and aims to address poorly annotated eukaryotic reference genomes. However, an external limitation of in silico bacterial probe design stems from potential common database contamination and taxonomic errors in metagenomic sequencing [13, 33, 34]. Therefore, using the latest bacterial genome annotations of interest as input for ProbeST is recommended.

Alternative pipelines have been developed for imaging-based ST, including the pipeline Spapros that covers both gene set selection and probe design [35]. Though similar in structure, ProbeST is tailored for Visium, the most widely used spatially-resolved transcriptomics technology [36], and enables multi-organism transcript detection. ProbeST's flexible workflow allows for potential

adaptation to other ST platforms, including imaging-based methods, by modifying parameters such as handle sequence and probe length.

We demonstrate the utility of custom probes generated by ProbeST in a DualST experiment, where we detect non-polyadenylated transcripts of the bacteria *S. Tm* in infected mouse enteroid-derived monolayers, despite a lower number of both prokaryotic transcripts and probes compared to the host. The DualST pathogen data shows high specificity and lower sensitivity to the mCherry bacterial signal from the same monolayers. The results corroborate the key role of the innate immunity in the host transcriptional responses to *S. Tm* infection, although biological inference remains limited due to only one biological replicate per condition and to data sparsity. Indeed, low-abundance bacterial signals may arise from host-dominated transcript environments and impact data processing and biological interpretation. The sparse bacterial data may be mediated with a higher number of selected genes for a strong statistical power and robust gene-level interpretations. Probe sensitivity is not constrained by host-dominated transcript environments, as ProbeST designs probes of high specificity to the genes of interest. We summarize what is supported by our data, what remains exploratory, and what requires future validation in Additional file 14: Table S17. Our study validates the high specificity of the ProbeST custom probes, demonstrating the pipeline's potential to detect transcripts from multiple organisms in a single tissue section, and to advance ST probe-based studies beyond mouse and human organisms.

Conclusions

ProbeST is a custom design Snakemake pipeline for generating specific probe panels for the probe-based Visium chemistry that can be applied on both FF and FFPE tissue sections, aimed at maintaining spatial transcript retrieval. ProbeST facilitates performing Visium experiments with gene-specific binding probes for organisms other than human or mouse, making it highly applicable for the study of other eukaryotes, microorganisms, and any non-model organism.

The DualST dataset generated to study host–pathogen interactions during *S. Tm* infection of a mouse enteroid-derived monolayer validated the specificity of the binding probes by targeting specific gene transcripts of interest. ProbeST has the potential to achieve scalability to the whole transcriptome of any organism, providing a novel way of applying probe-based Visium experiments to study spatial transcriptomics across kingdoms.

Methods

ProbeST pipeline

Input files

ProbeST takes as input the Coding DNA Sequence (CDS) in FASTA file format, trimmed for the genes of interest. The input file consists of a header and its corresponding sequence for each gene transcript, with each header manually modified to contain at least the gene ID, the gene name and the transcript ID in the following order and format:

```
>key:value key:value key:value
>gene_ID:CBW18961.1 gene_name:sipC transcript_ID:CBW18961.1_2869
```

As the CDS file formatting may vary between databases and species, manual editing of the file prior to running ProbeST is necessary. The CDS headers of the 19 *S. Tm* selected genes were modified accordingly and used for the ProbeST pipeline. In total, 51 probe pairs were obtained.

Pipeline parameters

BLAST parameters and parsing

The E-value is set to 30 000, based on primer-blast [37], and the word size is set to 7, a small word size chosen for better sensitivity with short probes, also based on primer-blast. The cost to open a gap is 3, the penalty for mismatches to enforce specificity –1, and the task is "blastn-short".

The BLAST output is parsed to identify probes with a high likelihood of off-target hybridization. The metric "n_corrected_mismatches" is calculated and the probes are removed if it has corrected mismatches up to 5. Above 6 mismatches, the probe hybridisation to that hit is considered not specific enough and therefore not considered an off-target. The filtering step depends on the pipeline configuration: for BLASTing against the genome from which the genes of interest are from, then a probe pair is removed if it has at least two significant hits, while for BLASTing against the host genome, any probe pair with at least one significant off-target hit is removed.

Secondary structure parameters from Primer3

The filter to control the formation of secondary structures is implemented with Primer3 to address specifically self-complementary binding and hairpin formations [10], increasing probe hybridization efficiency and specificity. The threshold PRIMER_INTERNAL_SELF_ANY = 47ΔG is set for self-complementarity anywhere in the oligo, while PRIMER_INTERNAL_SELF_END = 47ΔG is set for self-complementarity at the 3' end of the oligo, critical to avoid as the ligation happens on that end, and finally PRIMER_INTERNAL_HAIRPIN_TH = 47ΔG is set to

control hairpin stability. The thresholds may be modified by the user.

Configuration options

Pipeline versions

The pipeline version “1BLAST” is designed for cases where a single organism is of interest. Probe pairs with off-target hits to sequences in the organism’s genome are filtered out. The pipeline version “2BLASTS” is designed for cases where two organisms are of interest, such as in host–pathogen studies. Here, probe pairs with significant off-target hits to sequences in the first organism’s genome are filtered out, as well as probes pairs with any significant hit against sequences in the second organism. A hit is considered significant if there are less than 6 mismatches.

Ribosomal RNA binding

The final probe pairs from the ProbeST pipeline are taken as input for the optional rRNA filtering step. The 50 bp probe pair binding sequences are BLASTed against the SILVA rRNA database, containing sequences belonging to all three domains of life [15]. Any probe with at least one off-target hit of less than 6 mismatches is removed. The rRNA filtering step is strongly recommended for any robust custom probe design, as rRNA sequences are abundant in any organism and are highly diverse, increasing the risk for off-target binding.

Cross-hybridization

The final probe pairs from the ProbeST pipeline or from the rRNA filtering step are taken as input for the optional cross-hybridization steps. The 50 bp probe pair binding sequences are BLASTed against the environmental, prokaryotic and virus databases downloaded from NCBI. Any probe with at least one off-target hit of less than 6 mismatches is removed from the final probe pair selection, resulting in a probe set with higher specificity. The cross-hybridization step is recommended for poorly annotated eukaryotic genomes.

DualST

Sample generation

Enteroid culture and 2D monolayer preparation

Enteroids were derived from adult C57BL/6 mice isolated in previous studies [22]. Cryo-stocked enteroids were revived by thawing in a water bath at 37 °C and centrifuged in 5 mL DMEM/F-12 at 200 RCF for 5 min. The supernatant was discarded, and the crypts were resuspended in Matrigel [38]. Approximately 50 µL of the Matrigel-crypt suspension was plated per well in pre-warmed 24-well plates, followed by incubation at 37 °C for 10 min to solidify the Matrigel. The embedded crypts were overlaid with 700 µL of Mouse IntestiCult™

Organoid Growth Medium and cultured at 37 °C in a 5% CO₂ incubator. The medium was replaced every 2–3 days and passaged every 5–7 days as described previously [39]. At least 2 passages were done before revived enteroids were used for sample preparations as described below.

For the preparation of the enteroid 2D monolayer, enteroids were first enriched for proliferative cells by growing them in Human IntestiCult™ Organoid Growth Medium for 4 days after passage. To prepare the growth surface, an Ibidi 4-well removable inserts was placed on a microscope slide. The area in the wells were then coated with 5 µL/cm² Matrigel and allowed to solidify at 37 °C for 1 h. Enteroids enriched for proliferative cells were harvested in ice-cold DMEM/F-12 and spun at 200 rcf to remove the Matrigel and dissociated into single cells by resuspending and incubating in TrypLE Express for 5 min at 37 °C, followed by gentle pipetting. Dissociated cells were centrifuged at 300 rcf for 5 min, and the pellet was resuspended in 100 µL of Human IntestiCult™ Organoid Growth Medium supplemented with 10 µM Y-27632. Approximately 10⁵ cells were seeded into each well of the Ibidi 4-well removable insert, on top of the Matrigel-coated microscope slides. Cells were incubated at 37 °C in a 5% CO₂ incubator for 4 days with a medium exchange at day 3. Monolayers were monitored under a microscope to ensure confluency and cell health.

Bacterial Infection and FFPE preservation

The enteroid monolayers were washed with DMEM/F-12 medium. *Salmonella enterica* serovar Typhimurium strain SL1344/pFPV25.5 (constitutive mCherry) was grown overnight in LB medium [40]. The bacterial culture was then diluted 1:100 in fresh LB medium and incubated on a rolling wheel for 4 h at 37 °C. Optical density (OD) was measured to determine bacterial concentration, and 500 bacteria in 100 µL of Human IntestiCult™ Organoid Growth Medium (verified by plating on LB agar plates) were used to infect each cell layer. An infection of 500 bacteria per condition was chosen to better spatially resolve each individual infection site and ensure they would not overlap, enabling observation of transcriptional responses in regions with and without nearby infections. The infection was carried out for 2 h at 37 °C, where approximately half of all identified bacteria was observed to be associated with an expelled pyroptotic cell, an indication of NAIP/NLRC4 inflammasome activation, to capture both early responses to pyroptotic cells as well as responses to cells that have not yet undergone full pyroptosis. Following infection, the medium was removed, and samples were fixed in 4% paraformaldehyde (PFA). The samples were imaged (see Visual detection section below) and dehydrated by adding a series of 70%, 95% and 100% ethanol for 5 min. Dehydrated samples were embedded in paraffin by heating the

microscopy slide to 90 °C, adding 100 µL molten paraffin and placing a 90 °C clean microscopy slide on top of the paraffin to spread the paraffin thin. The top slide was removed and the sample slide with a thin layer of paraffin was cooled down and solidified. Samples were stored at −80 °C until used.

Visual detection of bacteria in infected monolayers

Imaging of the infected monolayers was performed on fixed samples in PBS with 1 U/mL RNase inhibitor (Takara) using a Zeiss Celldiscoverer 7 platform and imaged with a PlanApochromat 5x/0.35/WD: 5.1 mm objective captured on an AxioCam 712 mono camera. Entire monolayers were imaged in both brightfield and red fluorescence channels. Red fluorescence was excited using Calibri.2 LED Illumination Diodelaser 561 nm 10 mW, and emission was filtered with a 515/30 nm band-pass filter. The imaging was conducted across 7 focal planes to capture the full depth of the monolayers. Image analysis was performed in FIJI [41]. Potential *S. Tm* sites were automatically assigned by subtracting a Z-projection of the median intensity from a Z-projection of max intensity, followed by particle analysis filtering for circularity and size. Automatically assigned *S. Tm* sites were manually curated in the original images containing all 7 focal planes and filtered for actual *S. Tm* sites.

Spatial Transcriptomics

The experimental workflow consisted of the CytAssist-enabled FFPE Visium protocol [42]. The main steps in the workflow are: tissue preparation, imaging, probe mix addition, CytAssist-enabled RNA digestion, probe release, reverse transcription, library preparation and Illumina sequencing [43].

Prior to the probe hybridization step 1.1 of the User Guide [43], the *S. Tm* custom probes, ordered as oPool Oligos (50 pmol/oligo scale), were first resuspended in 62.5 µL IDTE (pH 8.0) to a concentration of 800 nM/oligo. A spike-in pool working stock was then prepared with 3 µL of each of the LHS and RHS resuspended stocks and completed to 100 µL with nuclease-free water for a final concentration of 24 nM/probe. The probe mix was prepared with 10 µL of the working stock, 10 µL of RHS Mouse Probes, 10 µL of LHS Mouse Probes, and 70 µL of FFPE Hybridization buffer [8].

Spatial Transcriptomics data generation

The gene expression count matrices were generated by Space Ranger ‘count’ (version 3.0.0) [12].

A custom reference file was made from Space Ranger ‘mkref’ using the Mouse reference dataset (GRCm38 Reference—2020-A) and the *S. Tm* CDS file (GenBank Assembly GCA000210855.2). A custom probe set was used, by appending the *S. Tm* custom probes to the

Mouse Probe Set from 10X Genomics (Visium Mouse Transcriptome Probe Set v1.0).

From this step, data processing and analysis were performed in R (version 4.3.2). The package Semla (version 1.1.6) was used for its tools developed specifically for Spatially Resolved Transcriptomics data analysis and visualization [44].

Validation of the *S. Tm* probes with mCherry fluorescence

The *S. Tm* spatial distribution between the DualST data and the mCherry signal were compared. Firstly, brightfield images were acquired prior to the ST workflow to obtain visual data on *S. Tm* localization in the mouse enteroid monolayers. Following imaging, potential *S. Tm* locations were automatically detected based on mCherry signal intensity and size. Each site was then manually assessed to remove false positives. Pixel coordinates with mCherry red signal were recorded, and the images were vertically flipped to match the DualST images.

Then, the brightfield mCherry images were aligned to the DualST image with the Loupe Browser tool v7.0.1 [45], and the pixel coordinates with mCherry signal were transformed with the resulting transformation matrix. As the fluorescent and DualST data are from the same samples, it was not necessary to adjust for the use of consecutive sections. After alignment and transformation, the pixel coordinates were read as a Seurat object and subsequently merged with the DualST Seurat object. Each pixel coordinate was attributed to the closest Visium spot coordinate within a radius of 100 µm, based on the k-nearest neighbors approach. The DualST data was split into Salmo⁺ and Salmo[−] spots, with Salmo⁺ corresponding to spots with at least 1 UMI from any *S. Tm* gene. The coordinates of the DualST Salmo⁺ spots were compared with the mCherry spot coordinates.

A confusion matrix was generated, taking the mCherry data as reference. The specificity of the DualST to detect *S. Tm* was calculated with the following formula:

$$\text{Specificity} = \#TN / (\#TN + \#FP)$$

With #TN the spot coordinates with neither mCherry nor DualST signal, and #FP the number of spots with DualST signal but no fluorescent signal.

Spot-level and gene-pair colocalization analysis

The mouse gene expression data was first filtered and normalized with the Seurat (version 5.0.2) module SCTransform v2 [46], then Differential Expression Analysis (DEA) was performed on mouse genes with the Seurat ‘FindMarkers’ function and the Wilcoxon rank-sum test. More specifically, the conditions WT+, *GsdmD*^{−/−}, *Nlrc4*^{−/−} were each compared to WT to analyze the different effects of *S. Tm* infection in relation to

the non-infected control. Then, *GsdmD*^{-/-} and *Nlrc4*^{-/-} were each compared to WT + to analyze the effects of the gene knockouts on the enteroid defense response to *S. Tm* infection.

The GO terms for each list of DE genes were obtained with the EnrichR package for the libraries “GO_Biological_Processes_2023”, “KEGG_2019_Mouse”, and “GO_Molecular_Functions_2023” (version 3.4) [47].

The colocalization analysis was performed with a DEA between Salmo⁺ and Salmo⁻ spots, for each enteroid condition WT, WT +, *GsdmD*^{-/-}, and *Nlrc4*^{-/-}, where Salmo⁺ is attributed to spots with at least 1 UMI from any *S. Tm* gene. The same parameters as above were applied for the DEA. Pathway analysis was performed with the DAVID functional annotation tool (v2023q4) [27].

The gene-pair colocalization analysis was performed using Fisher's exact test on all spots with *S. Tm* UMIs. Gene pairs with a *p*-value < 0.05 were considered to be significantly colocalizing.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-026-13077-z>.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.
Supplementary Material 5.
Supplementary Material 6.
Supplementary Material 7.
Supplementary Material 8.
Supplementary Material 9.
Supplementary Material 10.
Supplementary Material 11.
Supplementary Material 12.
Supplementary Material 13.
Supplementary Material 14.

Acknowledgements

We would like to thank the Scientific Center for Optical and Electron Microscopy (ScopeM) of ETH Zurich for providing helpful technical advice. The authors thank Michael Stebler (ScopeM, Department of Health Sciences and Technology, ETH Zurich, Switzerland) for his support and assistance with the ZEISS Celldiscoverer 7, Lovisa Franzén for her guidance in the data analysis, and Jeremie Charbord for helpful feedback on the manuscript.

Authors' contributions

S.G., P.R.A. and J.A.V. were responsible for the conceptualization of the study, and secured the necessary funding for the project. S.R., I.v.D., and S.H.F. developed the ProbeST computational pipeline, and S.S. and H.S. helped in its design. P.R.A., S.F., and S.R. conducted the research and investigation process. P.R.A. designed and provided the mouse intestinal organoid samples, collected and analyzed the mCherry-fluorescent imaging data of the pathogen *S*. Tm. S.R. designed the custom probes for *S*. Tm. S.F. and S.R.

conducted the CytAssist-enabled DualST experiments. S.R. curated the DualST sequencing data and performed the formal analysis of the data. S.R. and P.R.A. validated the results. S.R. created the visualizations and wrote the original draft of the manuscript. S.G., H.S., S.S., P.R.A. and J.A.V. contributed to the review and editing of the manuscript. S.F. and P.R.A. carried out the project administration and responsibility for the research planning and execution. Study supervision was carried out by S.G.

Funding

Open access funding provided by Royal Institute of Technology. This work was supported by the Swedish Research Council (VR, 2020–04864) and the European Research Council (ERC) under the Horizon Europe research and innovation programme (Synergy Grant, CartoHostBug, no. 101118531). P.R.A. was supported by a Novo Nordisk Foundation “Postdoc Fellowship for Research Abroad 2020—Bioscience and Basic Biomedicine” grant (NNF20OC0059485).

Data availability

ProbeST is implemented as a Snakemake pipeline and is available on our GitHub repository with its related workflow (<https://github.com/giacomellolab/ProbeST>) [9]. The scripts for the data analysis can be accessed on our GitHub repository. Sequencing data have been deposited at NCBI-SRA under the BioProject PRJNA1218431. The gene count matrices, the seurat object, as well as the high-resolution ST H&E images are available on Figshare [50, 51].

Declarations

Ethics approval and consent to participate

No mice were sacrificed for this publication, as all enteroids are derived from another study [19, 22, 48].

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 16 February 2026 / Accepted: 10 June 2026

Published online: 25 June 2026

References

1. Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*. 2016;353(6294):78–82.
2. 10x Genomics. Visium Spatial Gene Expression Reagent Kits User Guide. Available from: <https://www.10xgenomics.com/support/spatial-gene-expression-fresh-frozen/documentation/steps/library-construction/visium-spatial-gene-expression-reagent-kits-user-guide>. Cited 17 Jul 2025.
3. Gracia Villacampa E, Larsson L, Mirzazadeh R, Kvastad L, Andersson A, Mollbrink A, et al. Genome-wide spatial expression profiling in formalin-fixed tissues. *Cell Genomics*. 2021;1(3):100065.
4. Hoffman EA, Frey BL, Smith LM, Auble DT. Formaldehyde crosslinking: a tool for the study of chromatin complexes. *J Biol Chem*. 2015;290(44):26404–11.
5. 10x Genomics. Visium Spatial Gene Expression Reagent Kits for FFPE User Guide. Available from: <https://www.10xgenomics.com/support/spatial-gene-expression-ffpe/documentation/steps/library-construction/visium-spatial-gene-expression-reagent-kits-for-ffpe-user-guide>. Cited 16 Apr 2025.
6. Sounart H, Lázár E, Masarapu Y, Wu J, Várkonyi T, Glasz T, et al. Dual spatially resolved transcriptomics for human host–pathogen colocalization studies in FFPE tissue sections. *Genome Biol*. 2023;24(1):237.
7. Di Marco F, Nicola F, Giannese F, Saliu F, Tonon G, de Pretis S, et al. Dual spatial host-bacterial gene expression in *Mycobacterium abscessus* respiratory infections. *Commun Biol*. 2024. <https://doi.org/10.1038/s42003-024-06929-5>.
8. 10x Genomics. Custom Probe Design for Visium Spatial Gene Expression and Chromium Single Cell Gene Expression Flex. Available from: <https://www.10xgenomics.com/support/spatial-gene-expression-ffpe/documentation/steps/experimental-design-and-planning/custom-probe-design-for-visium-spatial-gene-expression-and-chromium-single-cell-gene-expression-flex>

9. Rouot S, van Dolderen I, Rosendahl Andreassen P, Frapard S, A. Herrera-Foessel S, Sounart H, et al. giacomellolab/ProbeST. Available from: <https://github.com/giacomellolab/ProbeST>
10. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 2012;40(15):e115.
11. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403–10.
12. 10x Genomics. Space Ranger 3.0. Available from: <https://www.10xgenomics.com/support/software/space-ranger/3.0>. Cited 16 Apr 2025.
13. Steinegger M, Salzberg SL. Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in GenBank. *Genome Biol.* 2020;21(1):115.
14. Cornet L, Meunier L, Van Vlierberghe M, Léonard RR, Durieu B, Lara Y, et al. Consensus assessment of the contamination level of publicly available cyanobacterial genomes. *PLoS ONE.* 2018;13(7):e0200323.
15. Chuvpochina M, Gerken J, Frentrop M, Sandikci Y, Goldmann R, Freese HM, et al. SILVA in 2026: a global core biodata resource for rRNA within the DSMZ digital diversity. *Nucleic Acids Res.* 2026. <https://doi.org/10.1093/nar/gkaf1247>.
16. Powers TR, Haeberle AL, Predeus AV, Hammarlöf DL, Cundiff JA, Saldaña-Ahuactzi Z, et al. Intracellular niche-specific profiling reveals transcriptional adaptations required for the cytosolic lifestyle of *Salmonella enterica*. *PLoS Pathog.* 2021;17(8):e1009280.
17. Samperio Ventayol P, Geiser P, Di Martino ML, Florbrant A, Fattinger SA, Walder N, et al. Bacterial detection by NAIP/NLRC4 elicits prompt contractions of intestinal epithelial cell layers. *Proc Natl Acad Sci U S A.* 2021;118(16):e2013963118.
18. Zhao H, Zhang X, Zhang N, Zhu L, Lian H. The interplay between *Salmonella* and host: mechanisms and strategies for bacterial survival. *Cell Insight.* 2025;4(2):100237.
19. Fattinger SA, Maurer L, Geiser P, Bernard EM, Enz U, Ganguillet S, et al. Gasdermin D is the only Gasdermin that provides protection against acute *Salmonella* gut infection in mice. *Proc Natl Acad Sci U S A.* 2023;120(48):e2315503120.
20. Hannemann S, Gao B, Galán JE. Salmonella Modulation of Host Cell Gene Expression Promotes Its Intracellular Growth. Valdivia RH, editor. *PLoS Pathog.* 2013;9(10):e1003668.
21. Bruno VM, Hannemann S, Lara-Tejero M, Flavell RA, Kleinstein SH, Galán JE. Salmonella Typhimurium Type III Secretion Effectors Stimulate Innate Immune Responses in Cultured Epithelial Cells. Isberg RR, editor. *PLoS Pathog.* 2009;5(8):e1000538.
22. Andreassen PR, Ernst C, Pimentel Mendes M, Bill Andersen T, Guillaume-Gentil O, Fattinger SA, et al. NAIP/NLRC4 inflammasome dynamics in murine enteroids are tuned by NAIP ligand concentration and epithelial cell differentiation. *Cell Rep.* 2025. <https://doi.org/10.1016/j.celrep.2025.116143>.
23. Jiang L, Li W, Hou X, Ma S, Wang X, Yan X, et al. Nitric oxide is a host cue for *Salmonella* Typhimurium systemic infection in mice. *Commun Biol.* 2023;6(1):501.
24. Bumann D, Schothorst J. Intracellular *Salmonella* metabolism. *Cell Microbiol.* 2017;19(10):e12766.
25. Raj D, Nair AV, Singh A, Basu S, Sarkar K, Sharma J, et al. Salmonella Typhimurium effector Ssel regulates host peroxisomal dynamics to acquire lysosomal cholesterol. *EMBO Rep.* 2024;26(3):656–89.
26. Han J, Balasubramanian I, Flores JA, Bandyopadhyay S, Yang J, Liu Y, et al. Intestinal lysozyme engagement of *Salmonella* Typhimurium stimulates the release of barrier-impairing InvE and Lpp1. *J Biol Chem.* 2024;300(7):107424.
27. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* 2022;50(W1):W216–21.
28. Papin S, Cuenin S, Agostini L, Martinon F, Werner S, Beer HD, et al. The SPRY domain of Pypin, mutated in familial Mediterranean fever patients, interacts with inflammasome components and inhibits proIL-1 β processing. *Cell Death Differ.* 2007;14(8):1457–66.
29. Chae JJ, Wood G, Masters SL, Richard K, Park G, Smith BJ, et al. The B30.2 domain of pypin, the familial Mediterranean fever protein, interacts directly with caspase-1 to modulate IL-1 β production. *Proc Natl Acad Sci U S A.* 2006;103(26):9982–7.
30. Eibl C, Grigoriu S, Hessenberger M, Wenger J, Puehringer S, Pinheiro AS, et al. Structural and functional analysis of the NLRP4 pypin domain. *Biochemistry.* 2012;51(37):7330–41.
31. Babbitt BA, Laukoetter MG, Nava P, Koch S, Lee WY, Capaldo CT, et al. Annexin A1 regulates intestinal mucosal injury, inflammation, and repair. *J Immunol.* 2008;181(7):5035–44.
32. Ciacci-Woolwine F, Blomfield IC, Richardson SH, Mizel SB. *Salmonella* flagellin induces tumor necrosis factor alpha in a human promonocytic cell line. *Infect Immun.* 1998;66(3):1127–34.
33. Chorlton SD. Ten common issues with reference sequence databases and how to mitigate them. *Front Bioinforma.* 2024;4. Available from: <https://www.frontiersin.org/journals/bioinformatics/articles/https://doi.org/10.3389/fbinf.2024.1278228/full>
34. Cornet L, Baurain D. Contamination detection in genomic data: more is not enough. *Genome Biol.* 2022;23(1):60.
35. Kuemmerle LB, Luecken MD, Firsova AB, Barros de Andrade e Sousa L, Straßer L, Mekki II, et al. Probe set selection for targeted spatial transcriptomics. *Nat Methods.* 2024;21(12):2260–70.
36. Moses L, Pachter L. Museum of spatial transcriptomics. *Nat Methods.* 2022;19(5):534–46.
37. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics.* 2012;13(1):134.
38. 354230 | Corning® Matrigel® Growth Factor Reduced (GFR) Basement Membrane Matrix, LDEV-free, 10 mL | Corning. Available from: <https://ecatalog.corning.com/life-sciences/b2c/US/en/Surfaces/Extracellular-Matrices-ECMs/Corning%C2%AE-Matrigel%C2%AE-Matrix/p/354230>. Cited 16 Apr 2025.
39. Ernst C, Andreassen PR, Giger GH, Nguyen BD, Gäbelein CG, Guillaume-Gentil O, et al. Direct *Salmonella* injection into enteroid cells allows the study of host–pathogen interactions in the cytosol with high spatiotemporal resolution. *PLoS Biol.* 2024;22(4):e3002597.
40. Furter M, Sellin ME, Hansson GC, Hardt WD. Mucus architecture and near-surface swimming affect distinct *Salmonella typhimurium* infection patterns along the murine intestinal tract. *Cell Rep.* 2019;27(9):2665–2678.e3.
41. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods.* 2012;9(7):676–82.
42. 10x Genomics. Visium CytAssist Spatial Gene Expression for FFPE – Deparaffinization, H&E Staining, Imaging & Decrosslinking. Available from: <https://www.10xgenomics.com/support/cytassist-spatial-gene-expression/documentation/steps/tissue-staining/visium-cyt-assist-spatial-gene-expression-for-ffpe-deparaffinization-hand-e-staining-imaging-and-decrosslinking>
43. 10x Genomics. Visium CytAssist Spatial Gene Expression Reagent Kits User Guide. Available from: <https://www.10xgenomics.com/support/cytassist-spatial-gene-expression/documentation/steps/library-construction/visium-cyt-assist-spatial-gene-expression-reagent-kits-for-ffpe>
44. Larsson L, Franzén L, Ståhl PL, Lundeberg J. Sema: a versatile toolkit for spatially resolved transcriptomics analysis and visualization. *Bioinformatics.* 2023;39(10):btad626.
45. 10x Genomics. Loupe Browser. Available from: <https://www.10xgenomics.com/support/software/loupe-browser/latest>
46. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol.* 2024;42(2):293–304.
47. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics.* 2013;14(1):128.
48. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics.* 2013;14(1):128.
49. Fattinger SA, Geiser P, Samperio Ventayol P, Di Martino ML, Furter M, Felmy B, et al. Epithelium-autonomous NAIP/NLRC4 prevents TNF-driven inflammatory destruction of the gut epithelial barrier in *Salmonella*-infected mice. *Mucosal Immunol.* 2021;14(3):615–29.
50. Rouot S, van Dolderen I, Rosendahl Andreassen P, Frapard S, A. Herrera-Foessel S, Sounart H, et al. Spatial Transcriptomics Gene count matrices. Figshare. Available from: <https://doi.org/10.17044/scilifelab.28350424>.

51. Rouot S, van Dolderen I, Rosendahl Andreassen P, Frapard S, A. Herrera-Foessel S, Sounart H, et al. Microscope images of H&E stained samples. Figshare. Available from: <https://doi.org/10.17044/sciifelab.28350185>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.