Briefings in Bioinformatics, 2026, 27, bbag198
https://doi.org/10.1093/bib/bbag198
Published: 27 April 2026
Problem Solving Protocol

OXFORD

BACON: decoding the dynamic social networks of complex microbial communities at single-cell resolution

Wenxin Qu[1,‡], Xiaofeng Shi[2,‡], Xinxin Xu[1,‡], Chang Liu[1,3], Liguo Ding[1], Mengdi Song[1], Ziye Xu[1], Yifan Xu[1], Fangyu Mo[4], Jian Ruan[5], Michael P. Timko[6], Longjiang Fan[4], Shufa Zheng[1,3,*], Weiqin Jiang[7,*], Yongcheng Wang[1,*], Yifei Shen[1,3,*]

[1]Department of Laboratory Medicine of The First Affiliated Hospital & Liangzhu Laboratory, Zhejiang University School of Medicine, No. 79 Qingchun Road, Hangzhou, 310003, Zhejiang, China
[2]Department of Laboratory Medicine, Changxing County Hospital of Traditional Chinese Medicine, No. 99 Changlue Road, Huzhou, 313100, Zhejiang, China
[3]Key Laboratory of Clinical In Vitro Diagnostic Techniques of Zhejiang Province, No. 79 Qingchun Road, Hangzhou, 310003, Zhejiang, China
[4]Institute of Bioinformatics, No. 866 Yuhangtang Road, Zhejiang University, Hangzhou, 310058, Zhejiang, China
[5]Department of Medical Oncology of The First Affiliated Hospital, Zhejiang University School of Medicine, No. 79 Qingchun Road, Hangzhou, 310003, Zhejiang, China
[6]Departments of Biology and Public Health Sciences, University of Virginia, Charlottesville, VA 22903, United States
[7]Department of Colorectal Surgery, the First Affiliated Hospital, Zhejiang University School of Medicine, No. 79 Qingchun Road, Hangzhou, 310003, Zhejiang, China

*Corresponding authors. Shufa Zheng, Department of Laboratory Medicine of The First Affiliated Hospital & Liangzhu Laboratory, Zhejiang University School of Medicine, No. 79 Qingchun Road, Hangzhou, 310003, Zhejiang, China. E-mail: zsfzheng@zju.edu.cn; Weiqin Jiang, Department of Colorectal Surgery, the First Affiliated Hospital, Zhejiang University School of Medicine, No. 79 Qingchun Road, Hangzhou, 310003, Zhejiang, China. E-mail: WeiqinJiang@zju.edu.cn; Yongcheng Wang, Department of Laboratory Medicine of The First Affiliated Hospital & Liangzhu Laboratory, Zhejiang University School of Medicine, No. 866 Yuhangtang Road, Hangzhou, 310058, Zhejiang, China. E-mail: yongcheng@zju.edu.cn; Yifei Shen, Department of Laboratory Medicine of The First Affiliated Hospital & Liangzhu Laboratory, Zhejiang University School of Medicine, No. 866 Yuhangtang Road, Hangzhou, 310058, Zhejiang, China. E-mail: yifeishen@zju.edu.cn.
‡Wenxin Qu, Xiaofeng Shi, and Xinxin Xu contributed equally

Abstract

Microbial communities function as dynamic societies where intercellular communication governs collective behaviors. However, mapping these interaction networks has remained a fundamental challenge in microbiology. This study aims to decode the social networks of complex bacterial communities at single-cell resolution by developing BACON, a computational framework that infers quorum sensing-mediated communication from single-microbe transcriptomic data. The approach combines a curated database of signaling systems with a statistical model that quantifies communication strength through coordinated expression of signal synthesis and receptor genes. Validation in model systems demonstrated BACON's precision in reconstructing density-dependent communication trajectories in *Bacillus subtilis* and capturing rapid network reorganization in *Escherichia coli* under antibiotic stress, revealing distinct sender-receiver subpopulations. Applied to human gut microbiomes, BACON unveiled diurnal fluctuations in cross-species signaling that transcend enterotype boundaries and uncovered conserved metabolic specialization in signal-responsive bacteria. In a clinical context, analysis of an ICU patient's gut microbiome revealed how *Pseudomonas aeruginosa* establishes a self-reinforcing communication circuit that upregulates virulence pathways. This work provides a unified framework for analyzing bacterial social interactions across diverse ecosystems. It opens new avenues for understanding microbial sociology, combating antimicrobial resistance, and engineering synthetic communities.

Keywords single microbe RNA sequencing, quorum sensing, microbiome, computational framework

Introduction

Bacteria, far from existing as isolated entities, function as collective societies capable of sophisticated cooperation and coordination. This collaborative behavior is largely governed by a cell–cell communication process known as quorum sensing (QS), wherein bacteria produce, release, and detect extracellular signaling molecules called autoinducers [1, 2]. This density-dependent mechanism allows a population to synchronously regulate gene expression, thereby coordinating collective activities essential for survival and fitness, such as biofilm formation, virulence factor secretion, and antibiotic resistance [3–5].

The foundational principles of QS have been largely deciphered from studies of single species in well-defined, homogeneous laboratory environments. However, in naturally occurring microbiomes—such as the human gut—bacteria exist within immensely complex,

Received: January 9, 2026. **Revised:** March 2, 2026. **Accepted:** March 27, 2026

multispecies consortia that are subject to dynamic environmental fluctuations [6]. Within these intricate ecosystems, QS systems of different species interact, leading to cross-talk that can either antagonize or synergize interspecies communication, profoundly influencing community structure and function [7]. Traditional methodologies, including bulk RNA sequencing and reporter gene assays, have been instrumental in identifying QS components but are inherently limited in their ability to resolve the cellular heterogeneity and spatiotemporal dynamics of signaling networks within mixed communities [8–10]. The aggregate measurements provided by these approaches mask cell-to-cell variation and obscure the precise cellular origins and targets of QS signals, creating a critical gap in our understanding of how communication truly operates at the systems level.

The recent advent of single-microbe RNA sequencing (smRNA-seq) represents a transformative advance for microbial ecology, enabling the profiling of gene expression with unprecedented resolution at the level of individual microbial cells [11–13]. This technology offers a powerful lens through which to examine functional heterogeneity within isogenic populations and to map the intricate web of interactions between different species in a community [14]. For QS research, smRNA-seq provides the unique capability to simultaneously quantify the expression of key genes involved in autoinducer synthesis, signal reception, and downstream response within each individual bacterium, thus offering a direct transcriptomic readout of a cell's potential to send and receive signals [15].

To harness the potential of smRNA-seq for decoding bacterial communication, we developed BACON (BACteria COmmunicatioN), a comprehensive computational framework for the inference, visualization, and analysis of QS networks. BACON integrates a manually curated database of QS systems with a robust statistical model to quantify communication strength between microbial subpopulations or species. We rigorously validated BACON using well-characterized model organisms, demonstrating its capacity to accurately reconstruct density-dependent QS dynamics in *Bacillus subtilis* and to capture the rapid reconfiguration of signaling networks in *Escherichia coli* under antibiotic stress. Furthermore, we applied BACON to human gut microbiome samples, revealing diurnal variations in cross-species AI-2 signaling and identifying pathogen-driven QS activity in a clinical sample from an ICU patient. Our results establish BACON as a versatile and reliable tool for elucidating the social language of bacteria in health and disease, with broad applications in microbial ecology, synthetic biology, and therapeutic development.

Results

Development of BACON

First, we developed the QS database (BACON database) that contains manually curated data on both intraspecies and cross-species QS communication (Fig. 1A). The BACON database for the intraspecies QS was curated from existing databases (including Quorumpeps [8], SigMol [9], QSDB [10]) and literature mining. It contains 43 QS entries from 22 species. For cross-species QS, representative products for each QS system are listed. The synthesis group includes acyl-homoserine lactone synthase (AHLs), S-ribosylhomocysteine lyase (AI-2), accessory gene regulator protein B (AIPs), and CAI-1 autoinducer synthase (CAI-1). The receptor group includes: HTH-type transcriptional regulator LuxR (AHLs), autoinducer 2-binding protein (AI-2), accessory gene regulator A (AIPs), and CAI-1 sensor kinase (CAI-1). QS systems universally consist of two functional components: (i) signal production (synthesis group), where synthesis genes encode enzymes that catalyze the production of autoinducer molecules (e.g. LasI synthesizes AHLs, LuxS catalyzes AI-2 production), and (ii) signal detection (receptor group), where receptor genes encode proteins that bind and respond to autoinducers (e.g. LasR binds AHLs, LsrB binds AI-2). Second, we performed smRNA-seq and processed the raw sequencing data based on our previously developed sequencing and computational frames [14, 16] (Fig. 1B) (see Methods).

To characterize bacterial communication networks at single-microbe resolution, we developed BACON (BACteria COmmunicatioN), a computational framework designed to infer and quantify QS signaling in microbial populations (Fig. 1C). BACON infers communication strength between microbial subpopulations or species by analyzing the expression of QS synthesis and receptor genes from smRNA-seq data. The process involves three key steps: (i) calculation of ensemble gene expression, (ii) estimation of cell–cell QS strength, and (iii) statistical assessment of QS signal significance: (a) For each synthesis-receptor QS pair, BACON first computes the ensemble average expression of the relevant genes within each subpopulation or species. This provides a representative expression profile for each functional component of the QS system. (b) Next, BACON estimates the abundance of both the signal synthesis and receptor groups to determine the strength of communication between sender and receiver subpopulations. When a signal is encoded by a single gene, its abundance is simply the average expression of that gene. However, for more complex QS systems that involve multiple genes (e.g. synthetases, transporters, receptors, and sensor kinases), BACON aggregates the expression data from all relevant genes (selected via the curated BACON database). The synthesis abundance is modeled by integrating the expression of all involved genes, while the receptor abundance similarly reflects the expression of receptor complexes and sensor kinases. For the receptor complex, genes contributing to the receptor group are categorized into different functional groups (e.g. periplasmic receptor LuxP and sensor kinase LuxQ), and their abundance is modeled using 'AND' logic. For signal molecules that require transporters to be exported from the cell, the genes contributing to synthesis group are categorized into synthase and transporter (e.g. in *E. coli*, synthase LuxS and transporter TqsA), with their abundance modeled using AND logic. For other QS signaling molecules, abundance is calculated by the average gene expression. Communication strength between two subpopulations is calculated as the product of the synthesis and receptor abundances. (c) To assess the reliability of inferred QS interactions, BACON applies a permutation test. By randomly shuffling the cell group labels and recalculating communication strength for each permutation, BACON generates an empirical distribution of QS strengths. The significance of each QS interaction is determined by the proportion of permutations in which the recalculated communication strength exceeds the observed value. Only interactions with a P-value $<.05$ are considered significant.

To establish BACON's capability in capturing authentic QS dynamics, we applied it to two well-characterized bacterial models under distinct physiological conditions (Fig. 1D). To evaluate BACON's performance in complex, multispecies communities, we additionally profiled fecal samples from different enterotype donors collected at three diurnal timepoints, supplemented with a clinical sample from an infectious disease patient in intensive care (Fig. 1E). This multitier experimental design enabled comparative analysis of QS network

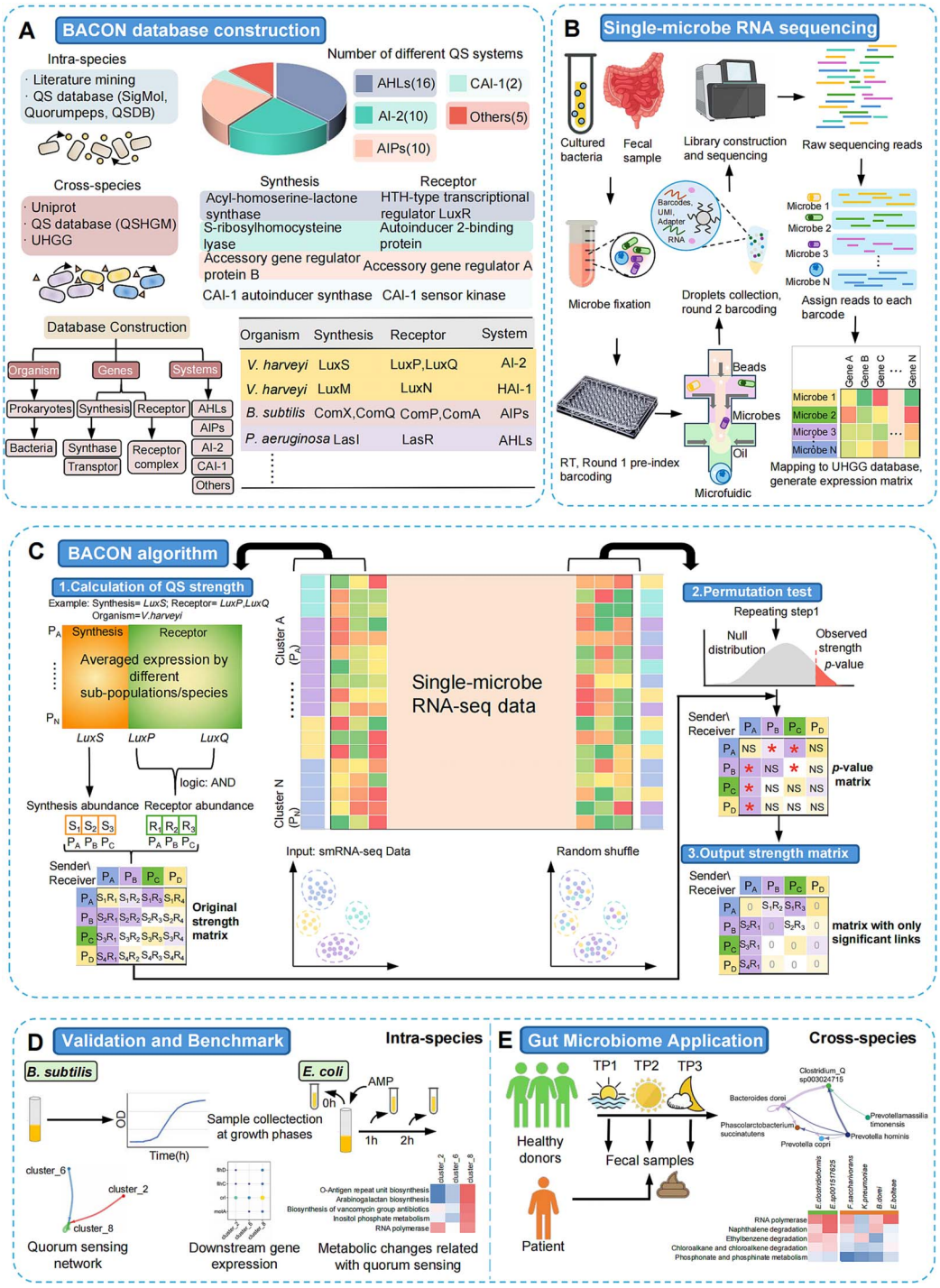

Figure 1 Overview of the BACON framework for inferring bacterial communication from single-microbe transcriptomes. (A) Construction of the BACON database. Manually curated data on intra- and cross-species QS systems were compiled from existing databases and literature. (B) Workflow for smRNA-seq. The process involves microbial fixation, permeabilization, *in situ* reverse transcription, droplet encapsulation, and library preparation for sequencing. (C) Schematic of the BACON computational model. The framework infers QS communication strength in three steps: (i) calculation of ensemble average gene expression for synthesis and receptor components within subpopulations, (ii) estimation of cell–cell communication strength based on the product of synthesis and receptor abundances, and (iii) statistical assessment of significance via permutation testing. (D) Experimental design for BACON validation. The framework was applied to single-microbe transcriptomes from model organisms (*B. subtilis* and *E. coli*) under defined physiological conditions and perturbations. (E) Application of BACON to human gut microbiome samples. Analysis included fecal samples from three healthy donors of different enterotypes collected at three diurnal timepoints, and a clinical sample from an ICU patient. The pipeline includes (i) fecal sample collection from donors representing three enterotypes at three diurnal time points, (ii) single-microbe RNA sequencing using the smRandom-seq protocol, (iii) data processing including alignment to the UHGG v2.0.1 genome database and species-level taxonomic assignment, (iv) application of BACON to infer interspecies QS communication networks; and (v) metabolic enrichment analysis of QS-responsive species to identify functional consequences of QS signaling.

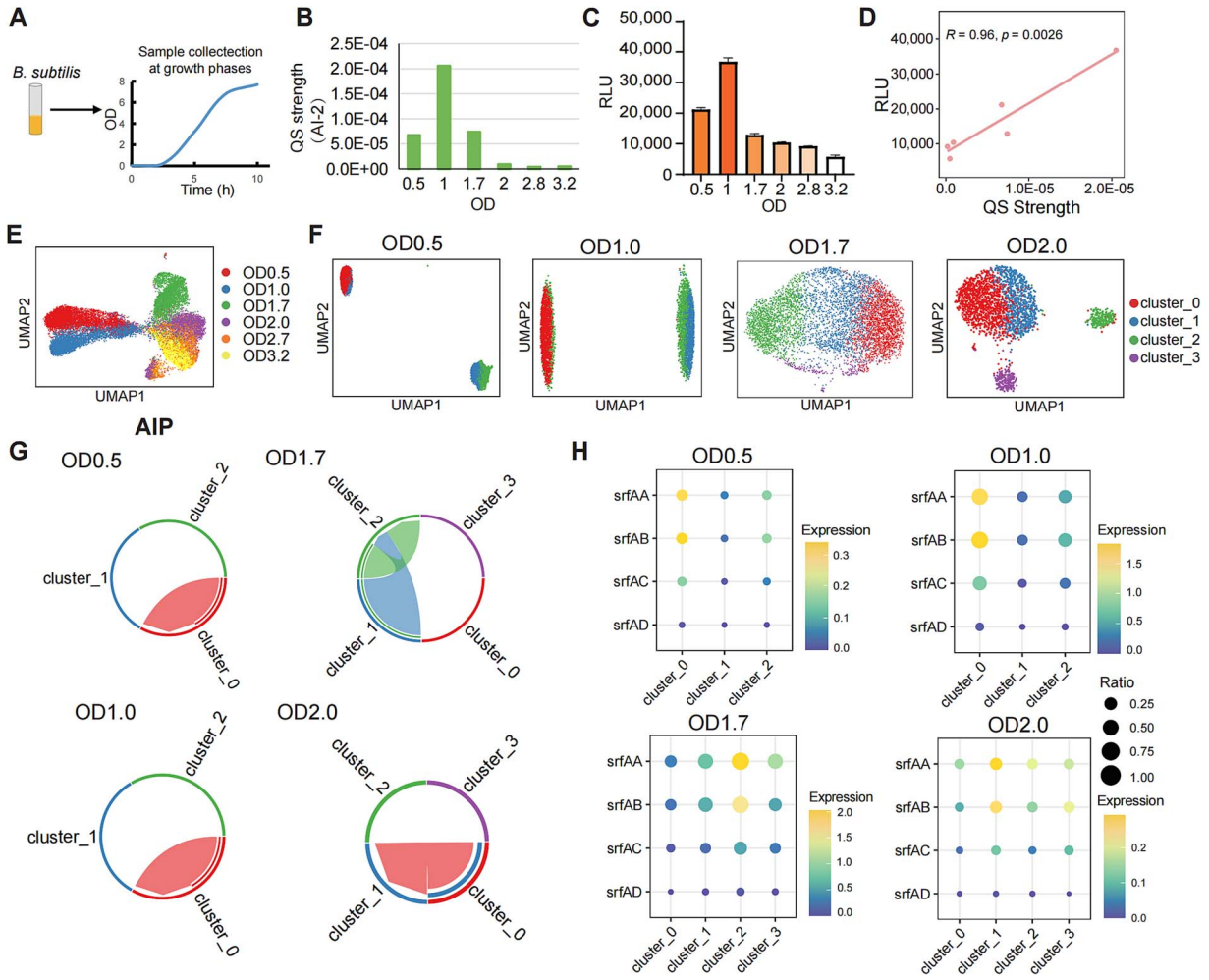

Figure 2 BACON captures density-dependent QS dynamics in *B. subtilis*. (A) Experimental design. SmRNA-seq was performed on *B. subtilis* cultures in different density. (B) BACON-inferred communication strength for the AI-2 QS system across different time points. Data are represented as mean communication strength. (C) Experimental validation of AI-2 activity measured by bioluminescence assay in *V. harveyi* BB170 reporter strain. Data points represent mean ± SD from biological replicates. (D) Correlation analysis between BACON-inferred AI-2 signaling strength and experimentally measured bioluminescence intensity. Pearson's correlation coefficient (*r*) and *P*-value are shown. (E) UMAP projection of *B. subtilis* single-microbe transcriptomes. (F) UMAP projection of *B. subtilis* single-microbe transcriptomes. (G) Inferred QS network for the ComQXP (AIP) system. Edges represent significant ($P < .05$, permutation test) communication events, with thickness proportional to communication strength. (H) Dot plot showing the expression level of downstream *srfA* operon genes (*srfAA-AD*) across different subpopulations and time points. Dot size represents the percentage of cells expressing the gene within a cluster.

architecture and associated metabolic remodeling across physiological and pathological contexts.

Validation of BACON reveals density-dependent quorum sensing dynamics

Since QS is regulated by population density [5], temporal analysis provides a critical framework for validating whether inferred communication patterns follow established physiological dynamics. To establish BACON's capability to resolve density-dependent QS dynamics, we analyzed a longitudinal single-cell RNA-sequencing dataset of *B. subtilis* across defined growth phases [12] (Fig. 2A). Application of BACON to *B. subtilis* single-microbe transcriptomes collected at optical densities of 0.5, 1.0, and 1.7 revealed dynamic QS activity in the canonical AI-2, ComQXP, and Rap-Phr systems [17]. We first focused our validation on AI-2-mediated signaling, as this conserved QS molecule—synthesized from S-adenosylmethionine by LuxS and known to regulate diverse bacterial processes including stress adaptation—provides an ideal benchmark for cross-method comparison [18]. Statistical analysis demonstrated a significant correlation ($P < .01$, Pearson's $r = 0.96$) between BACON-inferred AI-2 signaling strengths and experimentally measured bioluminescence intensities across all time points (Fig. 2B–D), establishing the method's quantitative accuracy in capturing QS dynamics. The observed reconfiguration of QS architectures across different growth phases or stress conditions mirrors the 'stage-specific' and 'dynamic modules' often identified in complex biological systems [19, 20].

Consistent with previous findings on bacterial phenotypic heterogeneity [15], our analysis revealed distinct QS variations in *B. subtilis* across growth phases (Fig. 2E). Uniform manifold approximation and projection (UMAP) clustering identified transcriptionally distinct subpopulations at different culture densities (Fig. 2F), demonstrating population-level specialization. At Optical Density (OD) 0.5, OD 1.0, and OD 1.7 in *B. subtilis*, we observed the same cluster synthesizing

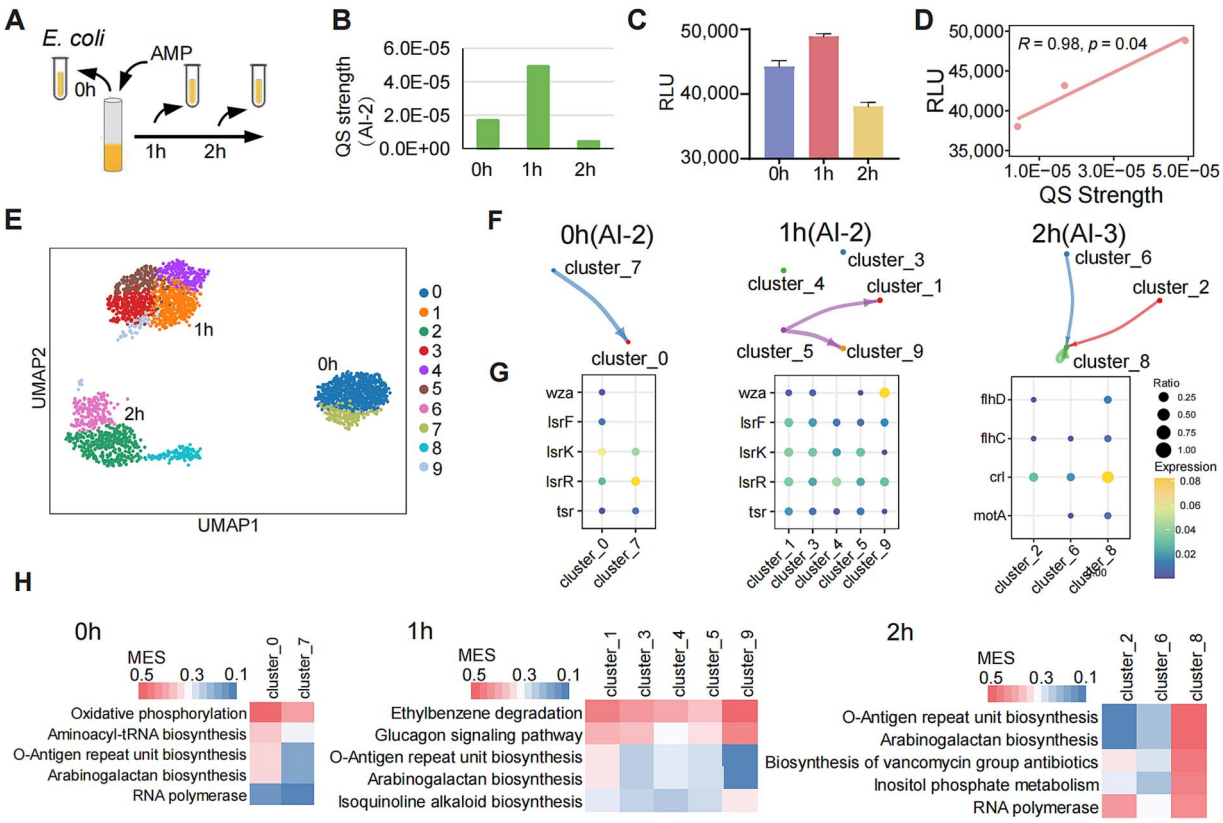

Figure 3 BACON uncovers antibiotic-perturbed QS network reconfiguration in *E. coli*. (A) Experimental design. *Escherichia coli* cultures were treated with 12.5 μg/ml ampicillin (AMP), and samples were collected at 0, 1, and 2 h post-treatment for smRNA-seq. (B) BACON-inferred communication strength for the AI-2 QS system across different time points post-antibiotic treatment. (C) Experimental validation of AI-2 activity measured by bioluminescence assay. Data points represent mean ± SD from biological replicates. (D) Correlation analysis between BACON-inferred AI-2 signaling strength and experimentally measured bioluminescence intensity. Pearson's correlation coefficient (r) and P-value are shown (E) UMAP projection of *E. coli* single-microbe transcriptomes. (F) Inferred QS networks for AI-2 and AI-3 systems at 0, 1, and 2 h post-AMP treatment. Edges represent significant ($P < .05$, permutation test) communication events, with arrow thickness proportional to communication strength. (G) Dot plot showing the expression level of key downstream QS-regulated genes across different subpopulations and time points. Dot size represents the percentage of cells expressing the gene within a cluster. (H) Heatmap showing the metabolic enrichment scores (MES, Z-score) for selected pathways in different subpopulations at 0, 1, and 2 h post-AMP treatment. Arrow thickness is normalized independently within each individual network plot.

and responding to ComX autoinducer (Fig. 2G). Functional validation through analysis of the ComA-regulated *srfA* operon confirmed pathway activation [17], with significantly elevated expression of *srfAA-AD* genes specifically in receiver clusters identified by BACON (Fig. 2H). Single-microbe resolution further revealed coexisting QS-ON and QS-OFF subpopulations, demonstrating BACON's capacity to resolve phenotypic heterogeneity in QS response. These results collectively establish BACON as a robust method for reconstructing authentic QS dynamics, capturing both the density-dependent regulation of signaling strength and cellular heterogeneity in microbial communication networks.

Benchmarking BACON with antibiotic-perturbed quorum sensing dynamics

Validating computational methods under biologically relevant perturbations is essential for establishing their robustness in capturing authentic signaling dynamics. Antibiotic stress represents a particularly suitable validation scenario, as it induces well-characterized shifts in bacterial communication while maintaining physiological relevance. To systematically evaluate BACON's performance under such conditions, we applied it to smRNA-seq data from *E. coli* cultures challenged with 12.5 μg/ml ampicillin (AMP), with samples collected at 0, 1, and 2 h post-treatment (Fig. S1A, Fig. 3A). We validated BACON's performance by comparing computationally inferred AI-2 signaling activities with parallel in vitro bioluminescence measurements. BACON-inferred AI-2 dynamics closely matched experimental measurements ($P < .05$, Pearson's $r = 0.98$), with both showing peak synthesis at 1 h followed by decline at 2 h (Fig. 3B–D), confirming the method's reliability under stress conditions.

BACON resolved distinct QS architectures across treatment durations. UMAP analysis resolved ten transcriptionally distinct cellular clusters (Fig. 3E), revealing dynamic QS architecture reorganization over time. At baseline (0 h), AI-2 signaling occurred from cluster 7 to cluster 0, while at 1 h, the network reconfigured with cluster 5 as the primary sender and clusters 1 and 9 as receivers (Fig. 3F). Functional analysis of downstream pathway components revealed elevated expression of *lsrF* and *lsrK*—genes encoding the AI-2 phosphorylation and processing machinery—in receiver clusters (Fig. 3G) [21]. Consistent with AI-2-mediated regulation, receiver clusters exhibited upregulated expression level of *wza* (involved in colanic acid production) and *tsr* (encoding a chemotaxis receptor), indicating activation of colonization and chemotaxis programs [21].

Beyond AI-2, BACON identified autoinducer-3 (AI-3) signaling at 2 h, with clusters 2, 6, and 8 synthesizing the signal and cluster 8 functioning as receiver (Fig. 3F). This coincided with peak expression of *crl*, a curli fiber regulator promoting bacterial aggregation [21], specifically in cluster 8 (Fig. 3G). The emergence of AI-3—a signal linked to virulence regulation in *E. coli* [22]—highlights BACON's ability to detect alternative QS systems activated under prolonged antibiotic exposure.

Metabolic pathway enrichment analysis using our established pipeline [15] revealed reprogramming in QS-responsive clusters. Receiver clusters at different time points showed enhanced activity in O-antigen and arabinogalactan biosynthesis pathways (Fig. 3H), supporting recent reports of AI-2-mediated regulation of cell envelope components [23]. Notably, cluster 9 at 1 h displayed distinct activation of isoquinoline alkaloid biosynthesis, while cluster 8 at 2 h showed elevated vancomycin group antibiotic biosynthesis, suggesting AI-3 may contribute to antibiotic stress adaptation. These results demonstrate BACON's capacity to reconstruct complex QS network dynamics under antibiotic stress, revealing signaling plasticity, functional specialization, and metabolic heterogeneity in bacterial populations.

Dissecting diurnal quorum sensing activities in the human gut microbiome with BACON

The human gut hosts a diverse and complex microbial community, with dominant phyla including *Firmicutes* and *Bacteroidetes*. Recent studies have revealed that specific enterotypes, which are defined by the abundance of certain key genera, shape the microbial composition and functional activities within the gut. The *ET-P* enterotype is characterized by an overrepresentation of *Prevotella*, *ET-B* by *Bacteroides*, and *ET-F* by *Firmicutes* [24]. Understanding the microbial dynamics within these enterotypes, particularly through the lens of QS, is essential for unraveling the cross-species communication networks that influence gut microbiome function. In this study, we utilized BACON to investigate the diurnal variation in QS activities in the human gut microbiome across three healthy donors, each representing a distinct enterotype: *ET-P* (Prevotella-dominated), *ET-B* (Bacteroides-dominated), and *ET-F* (Firmicutes-dominated) (Fig. S1B, S3). Fecal samples from each donor were collected at three distinct time points during the day (morning: TP1, noon: TP2, night: TP3), providing a temporal perspective on QS dynamics [15].

AI-2 signaling, a key interspecies communication mechanism conserved across bacterial species, plays a crucial role in regulating various metabolic functions in the gut. Previous studies have demonstrated the presence of AI-2 in human stool samples, suggesting its involvement in gastrointestinal microbial communication [25]. In our analysis, we observed distinct patterns of AI-2 signaling activity across the three enterotypes. In *ET-P*, the AI-2 senders remained consistent across all time points, primarily dominated by *Prevotella* species (*Prevotella copri* (*P. copri*), *Prevotella timonensis* (*P. timonensis*), and *Prevotella hominis* (*P. hominis*)) (Fig. 4A). The AI-2 receivers varied over time: at TP1, *Phascolarctobacterium succinatutens* acted as the AI-2 receiver, while at TP2, *Clostridium_Q sp003024715* and *P. succinatutens* were the main receivers. At TP3, *Megamonas funiformis*, a *Firmicutes* species, became the AI-2 receiver. In *ET-B*, *Bacteroides dorei* and *Klebsiella pneumoniae* were the primary AI-2 senders, with *Enterocloster spp.* acting as the AI-2 receiver at all time points (Fig. 4C). In contrast, *ET-F* was characterized by *M. funiformis* as the exclusive AI-2 receiver, while the primary senders included *B. dorei* and *Phocaeicola massiliensis*, both from the *Bacteroides* genus (Fig. 4E). These results highlight significant AI-2 QS interactions between species from the *Firmicutes* and *Bacteroides* phyla, particularly across different enterotypes and time points.

The diurnal variation in AI-2 signaling indicates a dynamic, time-dependent interaction between gut microbial species. In *ET-P*, *Prevotella* species consistently served as AI-2 senders, while species from *Firmicutes* and *Bacteroides* phyla exhibited temporal shifts in receiving AI-2 signals, such as *M. funiformis* at TP3. This suggests that AI-2 signaling is involved in coordinating metabolic activities between *Firmicutes* and *Bacteroides* species, possibly to regulate nutrient acquisition and community interactions in the gut. The diversity of AI-2 signaling in these enterotypes suggests that QS could be a key player in shaping microbial interactions within the human gut microbiome.

In addition to AI-2 signaling, we further explored the metabolic activities of the AI-2 receivers. In *ET-P*, the AI-2 receivers showed enhanced activity in cobalamin (vitamin B12) transport and metabolism, which is essential for microbial survival and not synthesized by plants but exclusively produced by bacteria and archaea [26] (Fig. 4B). We also observed increased activity in the degradation of amino acids such as valine, leucine, and isoleucine, particularly at TP1 and TP2. In *ET-F*, *M. funiformis* exhibited heightened activity in Vitamin B6 metabolism and mycolic acid biosynthesis, which are key metabolic processes in bacterial survival and virulence (Fig. 4F). These findings further suggest that AI-2 plays a critical role in regulating vitamin metabolism and other biosynthetic pathways across different enterotypes.

Furthermore, the activity of RNA polymerase was elevated in *M. funiformis* and *Enterocloster spp.*, suggesting a role for AI-2 in promoting protein biosynthesis. Additionally, we observed increased activity in arabinogalactan biosynthesis and O-antigen repeat unit biosynthesis in *M. funiformis*, indicating that AI-2 may influence bacterial cell wall synthesis. These metabolic changes, particularly in *ET-B* and *ET-F*, demonstrate how AI-2 signaling may modulate fundamental processes such as fatty acid biosynthesis, which is essential for cellular integrity and growth [27]. In summary, these results provide a detailed temporal and enterotype-specific analysis of AI-2 QS in the human gut microbiome, which revealed the diurnal fluctuations in QS activities across different enterotypes, highlighting the complex interactions between *Firmicutes*, *Bacteroides*, and other gut microbial species.

Clinical application of BACON in detecting pathogen-associated quorum sensing functional alterations

The human gut microbiome plays a critical role in maintaining homeostasis, yet its dysregulation can contribute to a wide array of diseases, particularly in immunocompromised patients such as those in the intensive care unit (ICU). Recent studies have highlighted the complex interplay between pathogenic and commensal bacteria in the gut, particularly in response to infections. Understanding how microbial communities, particularly those involving pathogenic species, communicate through QS can provide insights into microbial behavior and potential therapeutic targets for infectious diseases. Here, we employed BACON to investigate the QS activities in the gut microbiome of an ICU patient diagnosed with an infectious disease. This approach allowed us to explore microbial interactions at the gene

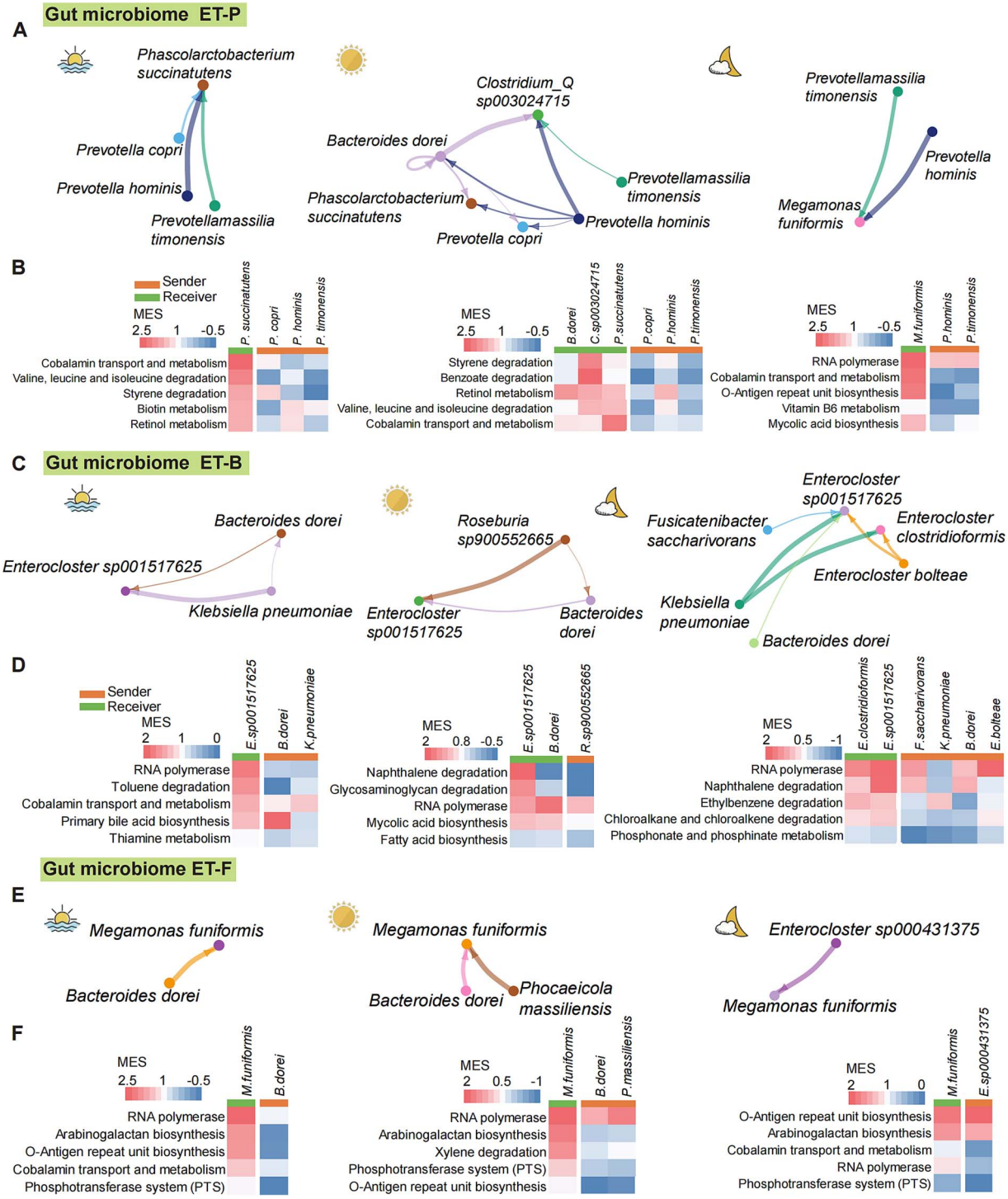

Figure 4 Diurnal fluctuations in AI-2-mediated QS across human gut enterotypes. (A, C, E) Inferred AI-2 QS networks for Donor *ET-P* (Prevotella-dominated) (A), Donor *ET-B* (Bacteroides-dominated) (C), and Donor *ET-F* (Firmicutes-dominated) (E) at three diurnal time points (TP1: morning, TP2: noon, TP3: night). Edges represent significant ($P < .05$, permutation test) communication events, with arrow thickness proportional to communication strength. (B, D, F) Heatmaps showing the metabolic enrichment scores (MES, *Z*-score) for selected pathways in the primary AI-2 sender and receiver species identified in Donor *ET-P* (B), Donor *ET-B* (D), and Donor *ET-F* (F) across the three time points. Arrow thickness is normalized independently within each individual network plot.

expression level, shedding light on the potential roles of QS in microbial pathogenicity and its relationship with clinical outcomes.

The fecal sample from the ICU patient was analyzed, capturing over 6058 genes, with a total of 8234 barcodes detected (Fig. S1C). The microbiota in this sample was dominated by *Pseudomonas aeruginosa*, a well-known opportunistic pathogen that is commonly associated with nosocomial infections [28]. In addition to *P. aeruginosa*, the gut microbiota also contained *K. pneumoniae*, *Enterococcus spp.*, and *E.*

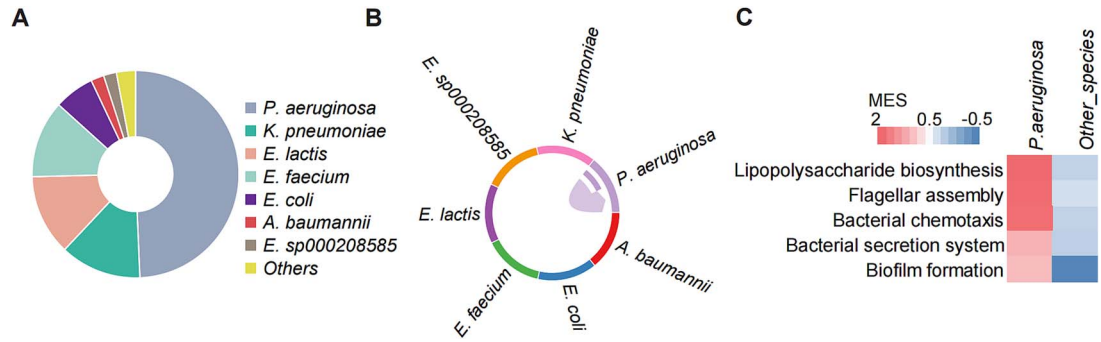

Figure 5 BACON detects pathogen-associated QS functional alterations in a clinical ICU microbiome sample. (A) Species composition of the ICU patient's gut microbiome. Bar plot shows the relative abundance (percentage of barcodes) of different bacterial species. (B) Inferred AHL (N-acyl homoserine lactone) QS network. *Pseudomonas aeruginosa* acts as both the primary sender and receiver in a self-reinforcing LasI/LasR circuit. The edge represents a significant ($P < .05$, permutation test) communication event, with thickness proportional to communication strength. (C) Heatmap showing the metabolic enrichment scores (MES, *Z*-score) for virulence-associated pathways in *P. aeruginosa* and other species present in the patient's microbiome.

coli, all of which are potential pathogens frequently found in ICU patients (Fig. 5A). The analysis revealed that *P. aeruginosa* was not only the most abundant organism in the microbiome but also acted as both the sender and receiver of N-acyl homoserine lactones (AHLs), a key signaling molecule in QS systems such as LasI/LasR (Fig. 5B). The LasI/LasR system in *P. aeruginosa* is known to control a variety of virulence factors and is crucial for the early stages of biofilm formation, which is a hallmark of chronic infections [7].

Our BACON analysis further indicated that the LasI/LasR QS system was highly active in this ICU patient's microbiome. Specifically, *P. aeruginosa* exhibited upregulation of genes related to several pathogenic processes, including lipopolysaccharide biosynthesis, flagellar assembly, bacterial chemotaxis, and biofilm formation (Fig. 5C). These pathways are critical for the virulence of *P. aeruginosa* and are tightly regulated by QS mechanisms, which allow the bacteria to coordinate their behavior in response to environmental cues and population density [7]. The upregulation of these pathways suggests that QS plays a central role in the pathogenicity of *P. aeruginosa* within the ICU patient's gut microbiome, potentially facilitating its survival and persistence in the host. Additionally, the ability of *P. aeruginosa* to both send and receive AHLs highlights its capacity to engage in intra- and interspecies communication, which may contribute to its ability to adapt to the stressful and competitive environment of the gut, particularly during periods of infection.

Discussion

BACON represents a computational framework specifically designed to decipher bacterial communication networks from smRNA-seq data. By constructing a comprehensive, manually curated database of synthesis-receptor pairs for both intra- and cross-species QS systems and developing a robust statistical model to quantify communication strength, we have established a method that reliably infers QS dynamics at single-microbe resolution. Our validation in model organisms—capturing the density-dependent biphasic trajectory of ComX signaling in *B. subtilis* and the antibiotic-induced reconfiguration of AI-2 and AI-3 networks in *E. coli*—demonstrates that BACON recapitulates established physiological dynamics and uncovers nuanced cellular heterogeneity within clonal populations.

A key finding enabled by BACON's resolution was the identification of functional specialization in QS tasks among cellular subpopulations (Fig. 6). In *B. subtilis*, we observed that the synthesis and reception of the ComX autoinducer occurred within the same cellular cluster across growth phases. This finding aligns with the concept of 'self-sensing' in QS, where autoinducer-secreting cells exhibit a stronger QS response, a phenomenon previously reported in *Bacillus* populations [17]. BACON now provides the toolset to observe such functional stratification directly from transcriptomic data, moving beyond bulk measurements that average out this critical heterogeneity. Furthermore, in the complex gut microbiome, BACON revealed intricate cross-species AI-2 interactions, such as *Bacteroides* species consistently acting as AI-2 senders while *M. funiformis* (a *Firmicutes*) served as a receiver in multiple donors. This suggests a potential metabolic division of labor, where different species assume specialized roles in the production and perception of a universal signal to coordinate community-level functions.

The application of BACON to a clinical sample from an ICU patient underscores its utility in uncovering pathogen-centric communication networks with potential therapeutic implications. We found that *P. aeruginosa*, a dominant pathogen in the patient's gut (Fig. 5A), engaged in intense AHL-based communication, primarily with itself, which was associated with the upregulation of virulence pathways, including lipopolysaccharide biosynthesis, flagellar assembly, and biofilm formation. This is consistent with extensive literature establishing that the LasI/LasR system controls a regulon of virulence factors in *P. aeruginosa* and that self-produced AHLs can activate a positive feedback loop enhancing pathogenicity [29, 30]. The ability of BACON to pinpoint such active, pathogen-driven QS hubs *in situ* highlights its potential for identifying patients where targeting QS, rather than bacterial viability, could be a viable strategy to disarm pathogens and mitigate infection severity.

While AI-2 was a focus of this study due to its role as a universal cross-species signal [18], the BACON database incorporates other major QS systems (AHLs, AIPs, CAI-1), allowing researchers to tailor analyses to their specific interests. Our analysis of gut microbiomes also yielded findings with ecological and potential applied relevance. For instance, the association of AI-2 reception in *Enterocloster spp.* with elevated naphthalene degradation activity hints at unexplored roles for gut microbes in the bioremediation of environmental pollutants [31], meriting further investigation.

We acknowledge that BACON operates on several assumptions. It infers communication potential from the coordinated expression of synthesis and receptor genes, and while we validated these inferences with orthogonal experimental assays, the actual concentration

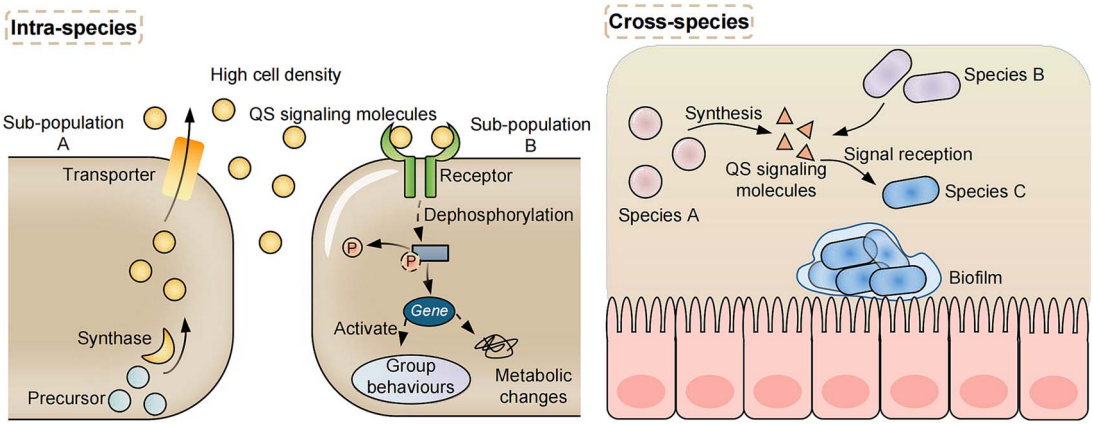

Figure 6 Model of functional specialization in QS tasks among bacterial subpopulations. Schematic summarizing the key finding that bacterial populations exhibit functional heterogeneity in QS tasks. Within a population, distinct subpopulations can specialize as primary signal senders (synthesizing autoinducers), receivers (expressing receptor complexes), or can be nonparticipating. This specialization enables complex communication architectures and coordinated community behaviors, as observed in *B. subtilis*, *E. coli*, and the human gut microbiome.

of signaling molecules in the microenvironment and the presence of unknown inhibitors or synergists can affect functional outcomes. Furthermore, the current database will benefit from continuous expansion as new QS systems and components are characterized. Because BACON relies on known genomic annotations, uncharacterized or highly divergent QS systems in nonmodel taxa may be missed, leading to a potential underestimation of communication networks. As the database expands to include newly discovered signaling molecules and receptors, we anticipate that the inferred network topologies will become significantly denser. Additionally, smRNA-seq can be affected by low mRNA content, dropout, and operon-related coverage biases. In the current version of BACON, we mitigate these effects by (i) using ensemble average expression within the subpopulation or species level, (ii) using an AND-logic formulation for multicomponent modules, and (iii) employs a rigorous permutation test to ensure statistical robustness. Moreover, larger cohort studies will be necessary to establish the generalizability of enterotype-specific QS patterns and to perform robust statistical comparisons across enterotypes. Future iterations of BACON could leverage machine learning (ML) to learn complex, nonlinear interaction patterns directly from data. While our current framework relies on a knowledge-based statistical model and curated annotations, ML architectures could identify latent communication motifs and uncharacterized signaling systems that may be missed by current assumptions.

In conclusion, BACON provides a powerful and flexible framework for transitioning the study of bacterial communication from simplified models to complex, real-world communities. By leveraging the resolution of single-microbe transcriptomics, it opens a new window into the social lives of bacteria, revealing the cellular heterogeneity, dynamic network reconfiguration, and functional specialization that underpin collective behaviors (Fig. 6). We anticipate that BACON will become a valuable tool for fundamental microbial ecology research, for uncovering novel therapeutic targets in infectious diseases, and for guiding the rational design of synthetic microbial consortia.

Materials and methods

BACON database construction

The BACON database was constructed through manual curation of known QS systems. For intraspecies QS, data were compiled from existing databases—Quorumpeps [8], SigMol [9], and QSDB [10]—and supplemented by extensive literature mining in PubMed Central. The resulting intraspecies database comprises 43 distinct QS entries across 22 bacterial species. For cross-species QS, the database was built using information from the QSHGM database [32] and the Unified Human Gastrointestinal Genome (UHGG) collection v2.0.1 [33], with gene information validated via UniProt. We focused on four major QS systems: acyl-homoserine lactones (AHLs), autoinducer-2 (AI-2), autoinducer-3 (AI-3), and the accessory gene regulator (AIP) system. The synthesis group contains 899 entries from 663 species, and the receptor group contains 701 entries from 538 species. Given the current state of cross-species QS research and the significant variability in gut microbial compositions, we adopted a permissive approach for the cross-species database. For instance, in *E. coli*, all genes required for AI-2 reception (*lsrB*, *lsrA*, *lsrC*, *lsrD*) are included, whereas for intraspecies QS, only the crucial ABC transporter gene *lsrB* is required [34]. The database is designed to be flexible; users can subset the provided data to construct a custom dataset relevant to their species of interest for input into the BACON framework. The QS database was constructed to maximize sensitivity. While this is reasonable for discovery purposes, such a permissive strategy may increase the risk of false positives. To counterbalance this permissive strategy and minimize false positives arising from partial pathway expression, BACON strictly relies on the permutation test and AND logic for multigene QS systems (described below) as statistical negative control to establish a baseline null distribution for interaction strengths.

Inference of quorum sensing communication networks with BACON

The BACON framework infers QS communication strength from smRNA-seq data in three sequential steps.

1. ***Calculation of ensemble average expression.***

For each gene involved in a synthesis-receptor QS pair, the ensemble average expression within a defined cell group (e.g. a cluster or species) was calculated as the arithmetic mean of the normalized gene expression across all cells in that group.

2. ***Estimation of cell–cell QS strength.***

BACON estimates the abundance of synthesis and receptors for each group, and computes cell–cell QS strength. The abundance of the synthesis group (S_i) and receptor group (R_u) for cell groups i

(sender) and u (receiver) was estimated based on the expression of their constituent genes. When the synthesis or receptor group is a signal that corresponds to a single gene, the abundance of synthesis or receptor group is set as the ensemble average expression defined in step1. For the other signals, the abundance of the synthesis or receptor group depends on expression levels of corresponding synthetase and transporter (e.g. in *E. coli*, the synthetase LuxS and transporter TqsA), receptor and sensor kinase (e.g. in *Vibrio harveyi* receptor LuxP and sensor kinase LuxQ), respectively. Assume that the synthetase requires r_1 catalyzing steps, for the n-th catalyzing step ($n = 1,2,\ldots,r_1$), let g_n denote the number of synthetase complex that catalyze the same chemical reaction (e.g. in *P. aeruginosa*, the AmbBCDE complex encoding IQS signal, thus there were four catalyzing steps). $E_{i,n,l}$ ($l = 1,2,\ldots g_n$) denote the ensemble average expression of l-th synthetase complex for step n in cell group i. Likewise, let q denote the number of transporter for the signals, and let Y_i denote the ensemble average expression of transporter. Then, the abundance of the synthesis group is modeled by the $1+ r_1$ functional groups of genes including one group for transporter and r_1 groups for the r_1 steps of synthesis. Because a high abundance of synthesis group requires high expressions of all the $1+ r_1$ groups of genes, so the AND logic (i.e., geometric mean) is applied among different functional group genes. Thus, the abundance of the synthesis group for the i-th subpopulations/species is modeled as:

$$S_i = \sqrt[1+r_1]{\frac{\sum_1^q \mathrm{Y}_i}{q} \cdot \frac{\sum_{l=1}^{g_1} E_{i,1,l}}{g_1} \cdot \ldots \cdot \frac{\sum_{l=1}^{g_{r_1}} E_{i,r_1,l}}{g_{r_1}}}$$

For the receptor group, assume that the receptor of the receptor group requires r_2 steps; for the n-th reception step ($n = 1,2,\ldots, r_2$), let j_n denote the number of receptor complex (e.g. in *E. coli*, the LsrBACD complex transport AI-2 signal). $W_{u,n,l}$ ($l = 1,2,\ldots j_n$) denote the ensemble average expression of l-th receptor complex for step n in cell group u. Likewise, let z denote the number of sensor kinase for the signals, and let K_u denote the ensemble average expression of sensor kinase. Then, the abundance of the receptor group is modeled by the $1+ r_2$ functional groups of genes including one group for sensor kinase and r_2 groups for the r_2 steps of reception. The AND logic is applied among different functional group genes. Thus, the abundance of the receptor group for the i-th subpopulation/species is modeled as follows:

$$R_u = \sqrt[1+r_2]{\frac{\sum_1^z \mathrm{K}_u}{z} \cdot \frac{\sum_{l=1}^{j_1} W_{u,1,l}}{j_1} \cdot \ldots \cdot \frac{\sum_{l=1}^{j_{r_2}} W_{u,r_2,l}}{j_{r_2}}}$$

Then the signal strength from subpopulation/species i to subpopulation/species u is defined as follows:

$$S_iR_u = S_i \cdot \ R_u$$

3. *Statistical significance of QS signals.*

The statistical significance of each inferred QS interaction was assessed using a permutation test. Cell group labels were randomly shuffled, and the communication strength ${S_iR_u}^{(m)}$ was recalculated for each permutation m. The P-value was calculated as follows:

$$p_{i,u} = \frac{1}{M}\sum_{m=1}^{M} I_{\left\{{S_iR_u}^{(m)} > S_iR_u\right\}}$$

Where ${S_iR_u}^{(m)}$ is recalculated synthesis-receptor signaling score for the m-th permutation and M is the total number of permutations (set $M = 100$ as default); $I_{\left\{{S_iR_u}^{(m)} > L_iT_u\right\}}$ is an indicator function of m, and equals 1 if ${S_iR_u}^{(m)} > L_iT_u$ and 0 otherwise. Only the communications with P-value $<.05$ are considered significant; for nonsignificant communication interactions, the communication strength values are set to be zeros.

By performing the three steps for each synthesis-receptor interaction pair, we obtain the communication strength for any interaction pair x from any subpopulation/species i to subpopulation/species u, $Q^x_{i,u}$ ($i = 1,2,\ldots,T$, $u = 1,2,\ldots,T$, $x = 1,2,\ldots X$), which can be written into a three dimensional array $\mathbf{Q}$ (T^*T^*X), where T is the number of subpopulations/species and X is the number of synthesis-receptor groups pairs.

Aggregation of quorum sensing networks

For each interaction pair x, denote $\mathbf{Q^{(x)}} = Q^x_{i,u}$ the communication strength matrix for the significant signals among all subpopulations/species. The aggregated network can then be obtained by summarizing all $\mathbf{Q^{(x)}}$ over all synthesis-receptor pairs, using the following aggregation method:

$$Weight = \sum_x \mathbf{Q^{(x)}}$$

Given the sender and receiver, this aggregation method sums QS strength values over all synthesis-receptor pairs, denoted as 'weight.'

Bacterial culture

Escherichia coli BW25113 was obtained from Sir Run Run Shaw Hospital, Zhejiang University School of Medicine. Cultures were grown in LB medium at 37°C with shaking until the mid-logarithmic phase ($OD_{600} = 0.5$), at which point the T0 sample was collected. Ampicillin (AMP) was added to the remaining culture to a final concentration of 12.5 μg/ml. Subsequent samples were collected at 1 and 2 h post-treatment. All samples were centrifuged at $2000 \times g$ for 2 min at 4°C, washed twice with phosphate-buffered saline, and processed for smRNA-seq using the smRandom-seq protocol [16].

Autoinducer-2 bioassay

Autoinducer-2 culture supernatant preparation. *Bacillus subtilis* was grown at 37°C in Luria-Bertani (LB) medium. The culture supernatant at different status was collected by centrifuging at 5000 g for 10 min and passed through the 0.22 μm pore size filter. *Escherichia coli* was cultured in LB medium until the mid-logarithmic growth phase ($OD_{600} = 0.5$), then 0 h sample was taken. Ampicillin (AMP) was added to the rest culture to a final concentration of 12.5 μg/ml. The following samples were taken from the culture at 1 and 2 h time points, respectively. Subsequently, the supernatant of the *E. coli* culture was collected by using the method for collecting the supernatant of *B. subtilis*. Filtered culture supernatants were stored at −80°C until used. AI-2 reporter strain *V. harveyi BB170*, which responds only to the AI-2 molecule, was used in the AI-2 bioassay. *V. harveyi BB170* was grown in AB medium at 30°C for 16 h until the stationary phase ($OD_{600} = 1.0$–1.1). The overnight culture diluted 1:5000 in AB medium was mixed 1:9 with the medium supernatant of the tested bacteria and added to a dark 96-well plate. The culture was incubated at 30°C with shaking at a rotate speed of 180 rpm for 4 h. The luminescent intensity that

corresponds to the activity level of AI-2 was measured by using the luminescent module of the multifunctional microplate reader.

Human fecal sample collection

This study was approved by the Ethics Committee of the First Affiliated Hospital, Zhejiang University School of Medicine (Approval No. 2021IIT A0239). All participants provided written informed consent. Fecal samples were collected from three healthy male donors (age range 20–30) at three time points: TP1 (6–9 a.m., after breakfast), TP2 (11 a.m.-1 p.m., after lunch), and TP3 (6–9 p.m., after dinner). One additional fecal sample was collected from a male ICU patient (age 59) with an infectious disease (COVID-19) at a stage of disease progression prior to antibiotic administration. All samples were processed immediately after collection. Impurities were removed by centrifugation at $500 \times g$ for 3 min at 4°C. The purified samples were then centrifuged at $3900 \times g$ for 5 min at 4°C to pellet microbial cells for subsequent smRNA-seq.

Single-microbe RNA sequencing and data analysis

Library preparation and sequencing

Single-microbe RNA sequencing libraries were prepared as previously described [14, 15]using the VITApilote-PFT1200 kit (M20 Genomics). Briefly, fixed and permeabilized microbes were subjected to in situ reverse transcription and dA-tailing. Single microbes were then co-encapsulated with barcoded hydrogel beads in droplets using a microfluidic chip (VITAcruizer DP400). On-bead cDNA synthesis and amplification were performed within the droplets. The resulting cDNA was purified, quantified (Qubit 3.0, Thermo Fisher Scientific), and qualified (DNA Fragment Analyzer, Bioptic). Sequencing libraries were constructed using the DNA Library Prep Kit for Illumina V3 (Vazyme) and sequenced on an Illumina NovaSeq 6000 platform (S4 Reagent Kit, 150 bp paired-end).

Data processing and analysis

Raw sequencing data were processed using our established pipeline [14, 15]. Primer and dA-tailed sequences were trimmed from the reads. Cell barcodes and UMIs were extracted from the R1 file, allowing for a Hamming distance of $\leq$2 bp. The R2 file was aligned to the UHGG v2.0.1 genome database [31] using STAR (v2.7.10a) [35]. Gene-level counts were generated with featureCounts (v2.0.3) [36], and UMI counts were deduplicated using umi_tools (v1.1.2) [37]. Downstream analysis, including dimensionality reduction (UMAP) and clustering, was performed using Seurat (v4.1.3) [38]. Microbes with nCount_RNA > 30 and nFeature_RNA > 15 were retained for further study. Dimensionality reduction was performed using principal component analysis (PCA) and clustering was performed using UMAP. Clusters were identified with the 'FindClusters' function (resolution = 0.5) in Seurat. Prevalent bacterial species, defined as those with >100 total barcodes, were selected for subpopulation-level analysis.

Single-microbe metabolic enrichment analysis

Metabolic activity was quantified using a rank-based metabolic enrichment score [15]. KEGG Orthology (KO) annotations for gut microbiome species were obtained from the UHGG database [33], and pathway gene sets were compiled using the KEGGREST package (v1.38.0). The analysis involved three steps: (i) raw MES generation: single-sample Gene Set Enrichment Analysis (ssGSEA) [39] was applied to the gene expression matrix to calculate an initial enrichment score for each pathway in each microbe. (ii) Permutation-based inference: for each pathway, genes were permuted 1000 times to generate a null distribution of random enrichment scores. (iii) MES normalization: the raw MES was compared to the null distribution, and a final normalized MES (Z-score) was calculated using the scale function in R.

Key Points

- A computational framework that infers quorum sensing from single-microbe RNA sequencing data was developed.
- BACON accurately reconstructs density-dependent and stress-responsive quorum sensing dynamics in model bacteria at single-cell resolution. Application to human gut microbiomes reveals diurnal and enterotype-specific cross-species communication networks.

Acknowledgements

We thank the core facilities of Zhejiang University Medical Center and Liangzhu Laboratory for their excellent technical support. We acknowledge the contributions of the clinical and technical staff of the First Affiliated Hospital, Zhejiang University School of Medicine. The authors gratefully acknowledge the support of the high-performance computing cluster and the help of Enhui Shen from Center for Bioinformatics and Big data Technology, Zhejiang University.

Author contributions

Wenxin Qu (Investigation, Data curation, Formal analysis, Writing—original draft), Xiaofeng Shi (Investigation, Data curation, Formal analysis, Writing—original draft), Xinxin Xu (Formal analysis, Writing—original draft), Chang Liu (Investigation), Liguo Ding (Investigation), Mengdi Song (Investigation), Ziye Xu (Investigation), Yifan Xu (Investigation), Fangyu Mo (Investigation), Jian Ruan (Investigation), Michael P. Timko (Formal analysis, Writing—original draft), Longjiang Fan (Formal analysis, Writing—original draft), Shufa Zheng (Investigation, Formal analysis, Writing—original draft, Supervision), Weiqin Jiang (Conceptualization, Investigation, Formal analysis, Writing—original draft, Supervision), Yongcheng Wang (Conceptualization, Investigation, Formal analysis, Writing—original draft, Supervision), and Yifei Shen (Conceptualization, Investigation, Formal analysis, Writing—original draft, Supervision).

Supplementary material

Supplementary material is available at *Briefings in Bioinformatics* online.

Funding

The project was supported by the National Natural Science Foundation of China (No. 82402729, Y.S. and 32200073, 32250710678, Y.W.), Zhejiang Provincial Natural Science Foundation of China (No. LQ23H200003, Y.S.), Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (No. 2021R01012, Y.W.), 'Pioneer' R&D programs of Zhejiang Province (No. 2024C03005, Y.W.), Key R&D

Program of Zhejiang (No.2024SSYS0022, Y.W.), and Zhejiang Provincial Natural Science Foundation for Distinguished Young Scholar (No. LR23H200002, S.Z.).

Data availability

The single-microbe RNA-seq data of *E.coli* generated in this study have been deposited in the Genome Sequence Archive database under accession code CRA011274. The public *B. subtilis* data used in this study are available in the Gene Expression Omnibus database under accession code GSE151940. The single-microbe RNA sequencing datasets of gut microbiome produced in this research are available in Genome Sequence Archive database under accession code CRA034559 (https://ngdc.cncb.ac.cn/gsa/browse/CRA034559).

Code availability

BACON is an R package available at GitHub (https://github.com/qu-wx/BACON/). The package dependencies include data.table v1.14.2, dplyr v1.0.9, Seurat v4.1.0, SeuratObject v4.1.0, circlize v0.4.14, ComplexHeatmap v2.8.0, igraph v1.3.4, and ggplot2 v3.3.6. The code used to reproduce the analysis in this study is available at (https://github.com/qu-wx/BACON-database/tree/main).

Ethics declarations

The study protocol was approved by the Ethics Committee of the First Affiliated Hospital, Zhejiang University School of Medicine, China (2021IIT A0239).

Consent to participate, and consent to publish declarations

Not applicable.

Consent to Publish declaration

Not applicable.

References

1. Keller L, Surette MG. Communication in bacteria: an ecological and evolutionary perspective. *Nat Rev Microbiol* 2006;**4**:249–58. https://doi.org/10.1038/nrmicro1383
2. Hawver LA, Jung SA, Ng WL. Specificity and complexity in bacterial quorum-sensing systems. *FEMS Microbiol Rev* 2016;**40**:738–52. https://doi.org/10.1093/femsre/fuw014
3. Cheng XL, Xu Q, Sun JD *et al.* Quorum sensing signals improve the power performance and chlortetracycline degradation efficiency of mixed-culture electroactive biofilms. *iScience* 2022;**25**:104299. https://doi.org/10.1016/j.isci.2022.104299
4. Rattray JB, Thomas SA, Wang Y *et al.* Bacterial quorum sensing allows graded and bimodal cellular responses to variations in population density. *mBio* 2022;**13**:e0074522. https://doi.org/10.1128/mbio.00745-22
5. Oña L, Shreekar SK, Kost C. Disentangling microbial interaction networks. *Trends Microbiol* 2025;**33**:619–34. https://doi.org/10.1016/j.tim.2025.01.013
6. Mukherjee S, Bassler BL. Bacterial quorum sensing in complex and dynamically changing environments. *Nat Rev Microbiol* 2019;**17**:371–82. https://doi.org/10.1038/s41579-019-0186-5
7. Oliveira RA, Cabral V, Torcato I *et al.* Deciphering the quorum-sensing lexicon of the gut microbiota. *Cell Host Microbe* 2023;**31**:500–12. https://doi.org/10.1016/j.chom.2023.03.015
8. Wynendaele E, Bronselaer A, Nielandt J *et al.* Quorumpeps database: chemical space, microbial origin and functionality of quorum sensing peptides. *Nucleic Acids Res* 2013;**41**:D655–9. https://doi.org/10.1093/nar/gks1137
9. Rajput A, Kaur K, Kumar M. SigMol: repertoire of quorum sensing signaling molecules in prokaryotes. *Nucleic Acids Res* 2016;**44**:D634–9. https://doi.org/10.1093/nar/gkv1076
10. Klein K, Garkov D, Rütschlin S *et al.* QSDB-a graphical quorum sensing database. *Database (Oxford)* 2021;baab058, **2021**. https://doi.org/10.1093/database/baab058
11. Ma P, Amemiya HM, He LL *et al.* Bacterial droplet-based single-cell RNA-seq reveals antibiotic-associated heterogeneous cellular states. *Cell* 2023;**186**:877–891.e14. https://doi.org/10.1016/j.cell.2023.01.002
12. Kuchina A, Brettner LM, Paleologu L *et al.* Microbial single-cell RNA sequencing by split-pool barcoding. *Science* 2021;**371**:eaba5257. https://doi.org/10.1126/science.aba5257
13. Pountain AW, Yanai I. Dissecting microbial communities with single-cell transcriptome analysis. *Science* 2025;**389**:eadp6252. https://doi.org/10.1126/science.adp6252
14. Shen Y, Qian Q, Ding L *et al.* High-throughput single-microbe RNA sequencing reveals adaptive state heterogeneity and host-phage activity associations in human gut microbiome. *Protein Cell* 2024;**16**:211–26. https://doi.org/10.1093/procel/pwae027
15. Shen Y, Qu W, Song M *et al.* Single-microbe RNA sequencing uncovers unexplored specialized metabolic functions of keystone species in the human gut. *Imeta* 2025;**4**:e70035. https://doi.org/10.1002/imt2.70035
16. Xu Z, Wang Y, Sheng K *et al.* Droplet-based high-throughput single microbe RNA sequencing by smRandom-seq. *Nat Commun* 2023;**14**:5130. https://doi.org/10.1038/s41467-023-40137-9
17. Bareia T, Pollak S, Eldar A. Self-sensing in Bacillus subtilis quorum-sensing systems. *Nat Microbiol* 2018;**3**:83–9. https://doi.org/10.1038/s41564-017-0044-z
18. Pereira CS, Thompson JA, Xavier KB. AI-2-mediated signalling in bacteria. *FEMS Microbiol Rev* 2013;**37**:156–81. https://doi.org/10.1111/j.1574-6976.2012.00345.x
19. Ma X, Tang W, Wang P *et al.* Extracting stage-specific and dynamic modules through analyzing multiple networks associated with cancer progression. *IEEE/ACM Trans Comput Biol Bioinform* 2018;**15**:647–58. https://doi.org/10.1109/tcbb.2016.2625791
20. Ma X, Yu L, Wang P *et al.* Discovering DNA methylation patterns for long non-coding RNAs associated with cancer subtypes. *Comput Biol Chem* 2017;**69**:164–70. https://doi.org/10.1016/j.compbiolchem.2017.03.014
21. Mayer C, Borges A, Flament-Simon SC *et al.* Quorum sensing architecture network in Escherichia coli virulence and pathogenesis. *FEMS Microbiol Rev* 2023;**47**:fuad031. https://doi.org/10.1093/femsre/fuad031
22. Witsø IL, Valen Rukke H, Benneche T *et al.* Thiophenone attenuates enteropathogenic Escherichia coli O103:H2 virulence by interfering with AI-2 signaling. *PloS One* 2016;**11**:e0157334. https://doi.org/10.1371/journal.pone.0157334
23. Hu X, Wang Z, Wang W *et al.* Irisin as an agent for protecting against osteoporosis: a review of the current mechanisms and

pathways. *J Adv Res* 2024;**62**:175–86. https://doi.org/10.1016/j.jare.2023.09.001
24. Ley RE, Hamady M, Lozupone C *et al.* Evolution of mammals and their gut microbes. *Science* 2008;**320**:1647–51. https://doi.org/10.1126/science.1155725
25. Sperandio V, Torres AG, Jarvis B *et al.* Bacteria-host communication: the language of hormones. *Proc Natl Acad Sci U S A* 2003;**100**:8951–6. https://doi.org/10.1073/pnas.1537100100
26. Degnan PH, Taga ME, Goodman AL. Vitamin B12 as a modulator of gut microbial ecology. *Cell Metab* 2014;**20**:769–78. https://doi.org/10.1016/j.cmet.2014.10.002
27. Li J, Liu H, Zhao C *et al.* Autoinducer-2 quorum sensing regulates biofilm formation and chain elongation metabolic pathways to enhance caproate synthesis in microbial electrochemical system. *Chemosphere* 2023;**344**:140384. https://doi.org/10.1016/j.chemosphere.2023.140384
28. Curran CS, Bolig T, Torabi-Parizi P. Mechanisms and targeted therapies for Pseudomonas aeruginosa lung infection. *Am J Respir Crit Care Med* 2018;**197**:708–27. https://doi.org/10.1164/rccm.201705-1043SO
29. O'Loughlin CT, Miller LC, Siryaporn A *et al.* A quorum-sensing inhibitor blocks Pseudomonas aeruginosa virulence and biofilm formation. *Proc Natl Acad Sci U S A* 2013;**110**:17981–6. https://doi.org/10.1073/pnas.1316981110
30. Azam MW, Khan AU. Updates on the pathogenicity status of Pseudomonas aeruginosa. *Drug Discov Today* 2019;**24**:350–9. https://doi.org/10.1016/j.drudis.2018.07.003
31. Parolini M. Toxicity of the non-steroidal anti-inflammatory drugs (NSAIDs) acetylsalicylic acid, paracetamol, diclofenac, ibuprofen and naproxen towards freshwater invertebrates: a review. *Sci Total Environ* 2020;**740**:140043. https://doi.org/10.1016/j.scitotenv.2020.140043
32. Wu S, Feng J, Liu C *et al.* Machine learning aided construction of the quorum sensing communication network for human gut microbiota. *Nat Commun* 2022;**13**:3079. https://doi.org/10.1038/s41467-022-30741-6
33. Almeida A, Nayfach S, Boland M *et al.* A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat Biotechnol* 2021;**39**:105–14. https://doi.org/10.1038/s41587-020-0603-3
34. Li X, Chen L, Zhou H *et al.* LsrB, the hub of ABC transporters involved in the membrane damage mechanisms of heavy ion irradiation in Escherichia coli. *Int J Radiat Biol* 2021;**97**:1731–40. https://doi.org/10.1080/09553002.2021.1987565
35. Dobin A, Davis CA, Schlesinger F *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;**29**:15–21. https://doi.org/10.1093/bioinformatics/bts635
36. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014;**30**:923–30. https://doi.org/10.1093/bioinformatics/btt656
37. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res* 2017;**27**:491–9. https://doi.org/10.1101/gr.209601.116
38. Hao Y, Hao S, Andersen-Nissen E *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021;**184**:3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048
39. Barbie DA, Tamayo P, Boehm JS *et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 2009;**462**:108–12. https://doi.org/10.1038/nature08460

Briefings in Bioinformatics, 2026, **27**, bbag198
https://doi.org/10.1093/bib/bbag198
Problem Solving Protocol