Title

Population Heterogeneity and the Spatial Organization of Emergent Proteus mirabilis Swarm Collectives

ISBN

9798189272325

Author

Garling, Eliotte Eileen

Publication Date

2026-05-01

Peer reviewed|Thesis/dissertation

Population Heterogeneity and the Spatial Organization of Emergent *Proteus mirabilis* Swarm Collectives

by

Eliotte Eileen Garling

A dissertation submitted in partial fulfillment of

the requirements of the degree of

Doctor of Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Karine Gibbs, Chair
Professor Arash Komeili
Professor Kim Seed
Professor Matthew Traxler

Spring 2026

Abstract

Population Heterogeneity and the Spatial Organization of Emergent *Proteus mirabilis* Swarm Collectives

By

Eliotte Eileen Garling

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor Karine Gibbs, Chair

Collective behaviors enable individuals to aggregate, working together to accomplish tasks, and withstand adverse environments in ways that isolated organisms cannot. Collective behaviors have emerged across the tree of life, necessarily mediated by interactions among individuals. Yet, at the microbial level, the contributions of individual cells to the emergence, propagation, and maintenance of collective behaviors remain unclear. Using *Proteus mirabilis* cooperative migration (swarming) as a model system, this dissertation combines genetics, computational, mathematical, and microscopy-based methodologies to quantitatively investigate the architecture of bacterial swarm collectives.

Quantitative microscopy approaches for resolving individual cell morphometrics within dense, quickly migrating populations stymied formal descriptions of microscale contributions to collective migration. In this dissertation, I introduce Swarmetrics, a quantitative image analysis pipeline designed to tackle dynamically migrating, irregularly shaped, and densely packed bacterial populations. Application of this pipeline over a series of eight-hour time courses with two species of *P. mirabilis* cells revealed previously unrecognized heterogeneity in cell morphology throughout swarm development, but most notably during active cooperative migration. This single-cell resolution challenged traditional models of *P. mirabilis* swarm morphology (length, particularly) as a synchronized transition into swarming behaviors. A new model is presented that incorporates population-level morphological variance as a more robust indicator of a *P. mirabilis'* place in the cell cycle, comparable to established gene expression markers for swarming. Overall, this work demonstrate that some collective behaviors may be more heterogenous and dynamically structured than once believed.

Building upon these findings, I joined a collaborative team of computer scientists and mathematicians at CU Boulder to establish and characterize the population architecture of swarms at the mesoscale. Together, our efforts have yielded new tools for dissecting complex microbial communities in a quantitative manner. By comparing cell-cell contact and the clustering of cells of wild type and strains deficient in the ability to transfer the non-lethal molecularly encoded identity signal Ids, our research suggests that identity signaling may serve as a mechanism to organize individuals into collectives. Taken together, the work presented in this dissertation proposes that collective behaviors are not homogenous masses of identical cells that coordinate in perfect unison. But instead, the amalgamation of heterogenous and

dynamically structured into coordinated collective groups. The approaches developed within this body of work mark a step forward in linking individual-level behaviors to more broad population-scale behaviors leading to the emergence of collectivity.

Table of Contents

List of Figures

Chapter 2:

Chapter 3:

Dedication

This Ph.D. Dissertation is dedicated to my grandpa, Don Garling. For his unrelenting drive to endure – up until the day the gas stove arrived.

Acknowledgements

The collective behaviors explored within this dissertation illustrate instances in which individual action is insufficient to accomplish a desired goal. It is instead through collective effort that these lofty goals become tangible, maybe even, achievable.

Perhaps in the wildest dreams of my grandfather, Don -- who left school without earning a high school diploma during the Great Depression, lost his arm in a logging accident, and spent his life devising every possible logging strategy not only to keep his job, but to remain indispensable—could he have imagined building a legacy of endurance that would lead, around 100 years later, to his granddaughter completing her Ph.D. at UC Berkeley.

As I finish this chapter of my life and turn to all those who lifted me up and reminded me that I *always* existed in a collective network of support, I will never have the words to thank you.

To my parents, Tim and Kathy Garling, who never doubted, always encouraged, and always let out a hand to pick me up when I had fallen, or who embraced me in their arms when a physical reminder of their collective belief in me was needed, I owe you the world. To my friends, some of the past, and some who I know are there always, Nikki, Sarah, Rachel, Andrea, and Terra, this would not have been possible without you. And to my family, Nancy, Terry, Melissa, Dan, Kevin, Kris, Mike, Caitlin, Conor, Julia, Lauren, and Duncan – I didn't know how deep the well of Eileen's love ran until I *really* got to know and love you. To those who passed before this milestone, Don, Eileen, Joan, Rosemary, and Ken, you wove the tapestry of strength, love, and chutzpah (in equal measure) I relied on so heavily throughout this process.

To Nick and Jess who invited me with open arms into their growing family and let me love their kids with my entire heart, it's been the honor of a lifetime getting to be a little part of your world. To Elio and Nico, thank you for always reminding me that the world is SO wonderous and exciting, and it's imperative to take the time to slow down and *always* ask "why?"

Figgy, you remained the shining light in my life from the day I started this Ph.D. to the moment I finished it. Thank you for being my best pal.

Finally, to my mentor Dr. Karine Gibbs, I will never find the words, nor compile them in the correct order to express how monumentally you have changed me for the better. Thank you for showing me how to live according to your values and ethics regardless the context or pressure. Thank you for reminding me that science is just a lot of fallible people doing their best to understand a dynamic and beautiful world. And, finally, thank you for bearing witness to my growth at all stages, and encouraging me every step of the way.

Chapter 1: Investigating Emergent Collective Migration in *Proteus Mirabilis*

1.1 The Protean Life of *Proteus Mirabilis*

Interactions form the world we live in, from atoms, to molecules, to cells, to organisms. The expansion from the individual cell to the formation of a complex multicellular organism relies on interactions. At the same time, the transition from single cell living, to groups of cells collaborating, pushes the boundaries of the division between unicellularity and multicellularity. *Proteus mirabilis* was named after the ancient sea god Proteus who foretold the future, and morphed shapes to conceal his knowledge to those who dare attempt to capture him. The bacterium *P. mirabilis* has similarly captured the notice of scientists due to the dramatic morphological variation, and collective behavior of swarming that occurs during *P. mirabilis'* life cycle.

Grown in liquid media, *P. mirabilis* lives as a short (1-2 μm in length) rod-shaped swimmer cell expressing 6-10 flagellum (1,2) When grown on rigid surfaces, *P. mirabilis* can participate in a rapid surface-based cooperative group motility called swarming or swarm migration. There is a lag period after initial contact with a rigid surface, then swarmer cells form by elongating from short rod shape into snake-like cells that can range in length from 10-80 μm (3). Swarmer cells have multiple copies of chromosomes but lack a division plane or septum. The transition into a swarmer state is coupled with a dramatic upregulation of the flagellar master regulator *flhDC* (4), the global transcriptional regulator *rcsB*, and the direct sensing of surfaces by flagellum (5,6). Swarming cells migrate as rafts of cells connected by flagella (7), which enable groups of cells to move together much more efficiently than an individual bacterium is able. Cells in rafts are aligned in parallel as the rafts move, and populations can travel outwards as rapidly as five centimeters within 24 hours at room temperature (3). Cell rafts are transitory and the entrance and exit of the stages of consolidation and elongation results in a radial colony expansion and a stereotypical bullseye pattern when grown on nutrient agar plates in the laboratory.

The transition into a swarmer state is coupled with a dramatic upregulation of the flagellar master regulator *flhDC*, the global transcriptional regulator *rcsB*, and the direct sensing of surfaces by flagellum (5,6). However, the signals that initiate swarming motility in *P. mirabilis* remain unclear. Findings suggest that there are internal cellular components required for swarming including putrescine biosynthesis and pathways connected to amino acid metabolism (8). There are also environmental cues which initiate swarming in *P. mirabilis,* including excess L-glutamine, surface rigidity, and moisture levels (9,10). Research by Little et al. (2019) provides significant evidence that stress sensing, cell envelope structure, and the *rffG* gene encoding a sugar-modifying enzyme dTDP glucose-4,6,-dehydratase produce an LPS-linked structural component of the cell envelope. Where, under swarm-permissive conditions, pathways including the Rcs phosphorelay, *flhDC*, *rffg*, and other stress-responsive genes act as sensors to modulate *P. mirabilis'* cell cycle. Additionally, evidence also suggests that the cell envelope itself functions as a critical checkpoint, triggering the transition to cell elongation and swarming behaviors (11).

In 1946, Dienes described expanding swarm colonies of genetically identical strains growing rapidly across the petri dish from opposite ends, merging with ease (12). By contrast, a

boundary, visible by the human eye, materializes between approaching swarms of different genetic backgrounds – the result of an arrest in forward migration. This boundary is referred to as the Dienes line and was initially correlated with lethal competitive factors such as proticine and toxins due to the unoccupied regions, cell debris, and round bodies seen in the boundary at high magnification (13–16).

1.2 A Genetically Tractable Model for Dissecting Emergent Collective Behavior

The astonishing Dienes line mystified scientists due to its formidable boundary that formed when genetically different cells encountered each other, and the absence of one when genetically similar cells met at the center of a petri dish. As research continued to investigate this strange phenomenon, it became clear that lethal factors alone were not enough to fully explain the formation of this boundary. Even in the absence of proticine production, *P. mirabilis* strains can still form boundaries, highlighting that additional mechanisms beyond proticine production and sensitivity contribute to *P.mirabilis'* boundary formation. In 2008, Gibbs et al., sought to identify self-versus non-self-discrimination factors in P. mirabilis by isolating mutants that recognize their parental strain as non-self (17).

Genetic analysis revealed a mutation in the operon termed *ids* (identification of self). Additionally, a complete *idsA-F* cluster across several different *P. mirabilis* strains revealed that the amino acid sequences of IdsD and IdsE are very similar besides a variable region within each protein (17). Further biochemical investigation demonstrated that the IdsD and IdsE proteins interact in an allele-specific manner (18). The IdsD protein binds primarily to the IdsE protein encoded adjacently; variants of IdsE that are encoded elsewhere within the genome -- or from foreign strains -- will not bind. Modifications to the variable region amino acid sequences within IdsD and IdsE, can either disrupt or generate binding capabilities between the two proteins.

Transfer of IdsD from a donor to a recipient cell is likely mediated by the type VI secretion system (T6S) (18–20). T6S systems are translocation apparatuses that span the cell envelope and are found in Gram-negative bacteria, resembling an inverted bacteriophage-puncturing device (21,22). Incidentally, IdsC contains a DUF4123 domain (Pfam family PF13503) encoded upstream of the T6S effectors and was previously used as a marker to identify unknown T6S effectors (19,23). Biochemical, molecular, and genetic evidence has demonstrated that the DUF4123 domain of IdsC is essential for the binding interaction between IdsC and IdsD (19). IdsC is also necessary for the formation of subcellular clusters of IdsD along the cell periphery and is necessary for the delivery of IdsD from a donor to a recipient cell through T6S machinery. IdsC's role in the localization and transport of IdsD can be disentangled through the introduction of mutations to the amino acid sequence of IdsC (24). Interestingly, a single cell can produce mismatched versions of IdsD and IdsE without disrupting swarm expansion or Dienes line formation, as long as IdsD transfer is inhibited (25). Together, these findings highlight the genetic tractability of *P. mirabilis* in which precise genetic perturbations can be harnessed to dissect the molecular mechanisms underlying self/non-self-recognition and collective migration.

1.3 Social Identity in a Migrating Collective

As *P. mirabilis* enters the snake-like phenotype of a swarmer cell and migrates in tandem with neighboring cells, T6S machinery is visible via TssB-sfGFP reporter strains as elongated sheaths (25). However, during consolidation when swarming arrestsT6S machinery is diffuse, suggesting that IdsD is exchanged between cells during active collective migration (26,27). Within the context of swarming this implies that each cell can exchange IdsD with the ever-changing cast of neighbors it encounters.

A cell is considered "self" if it produces an IdsE protein that is cognate to the IdsD it receives from a neighboring cell (18). A non-cognate IdsD/E interaction, either from the deletion of *idsE* or from a non-self-strain encounter, results in the attenuation of swarm expansion without the loss of cell viability (14,18). Non-self-recognition and the resultant exclusion from the leading swarm edge is termed "Ids mismatch." Cells that are excluded from the leading swarm edge due to Ids mismatch can grow and divide at a rate comparable to non-excluded cells (25). Additionally, further examination of an active swarm environment revealed that *P. mirabilis* exchange IdsD via T6S during migration (26). However, the signal transduction and recognition response is delayed until the cells have stopped moving and enter the consolidation phase of their life cycle (28). This underscores that Ids-mediated self-recognition functions not by lethal means, but by coordinating participation with non-lethal identity signaling within the swarm collective.

The story does not end here, however. When *P. mirabilis* cells have halted migration and transitioned into the non-migratory, dividing, consolidation phase of their life cycle, Ids mismatch results in a striking transcriptional shift (28). Specifically, the upregulation of guanosine tetraphosphate (p)ppGpp, and subsequent activation of the stringent response. Initially described in *E. coli* under starvation conditions (29,30), a rise in (p)ppGpp levels and entrance into the stringent response is due to the selective downregulation of rRNA, and protection from DNA from damage. Research since the initial discovery of (p)ppGpp in 1969 has demonstrated that (p)ppGpp regulatory pathways are highly conserved and function as cellular alarm systems. Activation of (p)ppGpp regulatory pathways can result in widespread transcriptional, physiological, and metabolic adaptations to environmental conditions, thereby increasing the probability of survival (31–33). In *P. mirabilis* cells, entrance into the stringent response is marked by the transition into a non-swarming state, where cells are unable to elongate into swarmer cells and participate in collective swarm motility and winnowing from the leading swarm edge (28). While the fate of these cells in a swarm remains unknown, our current hypothesis suggests that Ids mismatch results in kin cells slowly dominating the outer edges of a swarm. Thus, Ids proteins act to non-lethally maintain a dominant clonality at the leading edge of the swarm front, biasing access to potentially nutrient-rich and environmentally favorable conditions. It is worth noting that the quantity or degree of Ids mismatch required for a *P. mirabilis* cell to receive before entering the stringent response remains unknown.

1.3 The Emergence of Collective Behaviors from Single Cell States

Multicellular group formation of unicellular organisms has been observed across the tree of life, and includes the bacterium *P. mirabilis* and *Myxococcus xanthus*, the slime mold *Dictyostelim*

discoideum, and the green alga *Cholerra vulgaris.* These organisms blur the boundary between unicellularity and multicellularity through their participation in group behaviors (34,35). *P. mirabilis* strikes an intricate balance between single-cell living and the multicellular migration termed swarming. Swarming, wherein thousands of cells move cooperatively in a unified mass, allows cells to migrate greater distances than a single cell could alone (1,9). Coordinated and cooperative behaviors are hypothesized to have facilitated the evolutionary transition from group behaviors to multicellularity, a trait that has arisen more than 25 times across the tree of life (36). While the evolution of multicellularity is still not well understood, analysis of group behaviors can inform the constraints and advantages of multicellular existence.

Individual and group behaviors have been studied across many disciplines and scales, from economics-based game theory to social psychology, and developmental biology to ecology (26,37–40). Indeed, the range of disciplines that have encountered and attempted to understand group behaviors is vast. Given the range of organisms that exhibit group behaviors, it's important to define the terms of group behaviors. Particularly due to the undefined or "gray zones" existing outside of said clearly defined terms. Group behaviors emerge from social interactions that may be influenced by genetic relatedness (clonality), proximity, or other features (41). Within the realm of possible group behaviors are collective behaviors. Collective behaviors can be considered as behaviors that arise without centralized control, where the properties and functionality of the group emerge through interactions among individuals and the expansive networks they form (34,42). Population dynamics emerge out of interactions and describe the ways the groups and individuals communicate, sense, and respond to their surroundings. Individuals participating in group behaviors may perform different tasks, and the division of labor may be distributed based on several different factors including interactions, the shapes of colonies and swarms, and stages of development (43). Additionally, the division of labor and interactions between individuals may shape the overall composition of a population. For example, female African lions (*Panthera leo*) engage in many group behaviors including group hunting, communal cub rearing, and group territoriality(44,45). However, during group-territorial conflict simulating an intruder, some members of the pack would lead the approach to protect the territory whereas other females would stay behind thereby avoiding danger (46). The differences in the community structure and complexity in behaviors in response to approaching danger demonstrate that within cooperative group behaviors, individuals cannot be expected to behave identically.

Within the bacterial realm of group behaviors, there are several different stereotypically observed behaviors. Biofilms and inoculation colonies are examples of micro-movement and sessile behaviors. Biofilms are surface-adhered or suspended aggregates of bacterial communities that are often multilayered and experience cell-to-cell contact within extracellular polymeric substances (EPS) that provides a defined architecture to the biofilm (47). These sessile communities can be attached to solid surfaces, form pellicles at the liquid-air interface, or free float in liquid (48,49). Recent research has suggested that social cooperativity plays a role in the early stages of biofilm development during the period of transient surface attachment, called reversible attachment, a process which initiates the formation of *Pseudomonas aeruginosa* biofilms (50,51). Additionally, biofilms are often polymicrobial communities that participate in dynamic intercellular signaling through the exchange of metabolites, genetic material, and other compounds (48,52,53). Continued research has highlighted the heterogeneity of biofilms,

bringing forth the idea that macro-scale behaviors and physiological structures formed by individuals within a collective may be influenced by microenvironmental conditions that regulate signaling, metabolism, and gene expression (54–56).

1.4 Navigating the Continuum of Colony Structure, Local Environments and Cellular Interactions

Highly motile macro-movement behaviors, like swarming, are another example of stereotypical bacterial group behaviors. Swarming describes the coordinated movement of bacteria (several mm s^{-1}) across rigid surfaces and is hypothesized to require a quorum of cells to initiate swarming. The transition to swarming is coupled with drastic morphological changes during elongation, *P. mirabilis* dramatically upregulates their flagellum and orient along the lengths of neighboring cells to form cell "rafts". These rafts of cells can migrate up to 100 times in surface distance (in centimeters) than individual cells are able to migrate (1,2). The decision to move between sessile and macro movement group behaviors requires the integration of many environmental cues: nutrient accessibility, oxygen levels, temperature, pH, and surface tension, all of which are abiotic factors known to influence the initiation and coordination of group behaviors (57). However, the exact cue(s) for *P. mirabilis* swarm initiation remains unknown.

Other microbial systems have uncovered the mechanisms that demarcate the transition from individual cells to a sessile collective state. Carbon catabolite control mechanisms provide a scaffolding by which microbes can preferentially catabolize the most energy-efficient carbon source available in an environment to achieve a stable growth rate (58,59). Selecting an optimal carbon source, however, can constrain the expression of genes involved in the catabolism of other carbon sources. This regulatory carbon catabolite control is known as CCR and enables microbes to survive and compete in dynamic environments through controlled metabolic flexibility (60). In cases of nutrient depletion, or other stressors, cells can secrete signaling molecules into the environment to trigger the shift from a unicellular to a multicellular lifestyle.

Self/non-self-recognition is intimately tied with participation in the collective migration of *P. mirabilis* – though, this underscores a larger theme of microbial interactions: interactions between cells can drive behavioral shifts, lifestyle adaptations, and dramatic morphological changes (61,62) These changes can exert profound effects on the structure and function of microbial communities as a whole. Unicellular organisms possess sophisticated mechanisms for self-recognition that can govern a wide variety of behaviors, such colony formation, mating, and immune defense. For example, *Staphylococcus aureus* uses the accessory gene regulator (*agr*) quorum-sensing system to initiate pathogenesis (63). Marking the switch from colonization to virulence, *S.* aureus cells secrete autoinducing peptides (AIP) into the extracellular environment. The concentration of AIPs allows cells to communicate relative population density. Once a threshold AIP level is reached in the local environment, *S. aruerus* can upregulate genes involved in pathogenesis, allowing for the establishment of infection (64).

Unlike quorum sensing, which may indicate how *many* others are around, it may convey little about *who* those others are. The social amoeba *D. discoideum* survives by weaving together indiscriminate quorum sensing with identity signaling (65). Under starvation conditions, *D. discoideum* cells release the diffuse secondary signaling molecule, cyclic adenosine

monophosphate (cAMP). This signaling molecule travels like a wave through the local environment, calling other *D. discoideum* cells to aggregate. Once assembled, the proximity of cells allows for self/non-self-recognition to occur. This process is mediated by the polymorphic proteins, TgrB1 and TgrC1(66). These proteins function as a ligand-receptor pair in a sequence specific manner. When neighboring cells are recognized as "self" through cell-cell adhesion, the cells will merge into clonal fruiting bodies, while those who recognize each other as "non-self" will not. This mix of signaling systems demonstrates alternative modes of social behavior that are not mutually exclusive, (1) soluble cAMP aggregates cells into a local environment at proximity to (2) undergo cell-cell contact mediated self/non-self-recognition. These different signaling systems reveal how the transition from a unicellular to a multicellular life phase can arise from environmental signaling, and identity-dependent social selection.

For some organisms, the transition from unicellular to multicellular life stages could be described in shades of gray, as opposed to a strict shift from one form to the other. Interestingly, both swarms and biofilm structures display heterogeneity in cellular biomass, cellular signaling, metabolism, and gene expression within one microbial collective (58,67). Not only this, but the spatial organization of collectives has shown to be heterogenous in biofilms as well. Evans *et al.* (2023) demonstrated a tradeoff in biofilm structure, where some cells may be in more direct contact with redox-active compounds naturally toxic to *P. aeruginosa* closer to the air interface (68). While cells deeper in the biofilm in hypoxic/anoxic conditions are protected from the toxicity of redox-active compounds closer to the air-interface of the biofilm. These observations suggest that microbial collectives can exhibit hallmark features of multicellular organization, in which cellular specialization and spatial structuring enable the division of physiological burdens across the community.

Overall, the ability to sense and respond to environmental stressors is fundamental for the survival of organisms and multicellular networks, whether sessile or macro movement, may act as survival mechanisms. Temporal and spatial organization of individual cells can lead to the establishment of resilient bacterial communities. *P. mirabilis* represents a paradigm for organisms at the interface of unicellularity and multicellularity, particularly exemplified by its macro-scale cooperative migration. In its single-cell state, *P. mirabilis* combines the experimental advantages of single-cell systems with the added complexity of cooperative migration, opening the door to address broader questions about the relationship between group behaviors, interactions, the role of non-lethal identity signaling in colony building, and broader evolutionary questions about the origins of multicellularity.

Chapter 2: Population heterogeneity revealed in morphometric analysis of densely populated microbial swarm collectives

This work has been published in mBio
Garling EE, Li S, Ho K, Gibbs KA. 2026. Population heterogeneity revealed in morphometric analysis of densely populated microbial swarm collectives. *mBio* **17**:e03342-25.
https://doi.org/10.1128/mbio.03342-25

I. Chapter 2 Preface

Bacteria and animals utilize group migration to survive and propagate as a population. A subset engages in collective migration, which is an emergent behavior, organized spatiotemporally at two different levels: the individual and the collective. While much is known about collective behaviors, research into the roles that individuals play has been stymied. In microbial systems, this was particularly due to fundamental unknowns, such as the role of individual cell development within collectives. The following manuscript focused our investigation on individual *P. mirabilis* cells as they participated in collective migration over an eight-hour life cycle. The central goal of this manuscript was to gain insight into approaches individual cells may use to access cooperative behaviors in highly competitive environments.

After much trial and error, and incredible collaborative conversations, the Swarmetrics pipeline emerged as a methodology to collect morphometric data on individual *P. mirabilis* cells as they participated in collective migration. This pipeline addressed two major limitations to collecting morphometric data of individual cells. First, during *P. mirabilis* collective migration, the cells are densely packed and often irregularly shaped. This led to tremendous difficulty in the image segmentation process. A step that also happens to be essential (currently) to gain morphometric insights of individual cells. Through a combination of manual annotation, and machine learning, the hurdle was overcome. Next, the Swarmetrics pipeline allows for the morphometrics of cells to be measured in their native environment. Meaning, a swarm front could be imaged with phase contrast microscopy, and individual cell morphometrics could be obtained without having to sub-sample or use other labeling techniques.

Overall, this manuscript revealed that during collective migration – once thought to be *only* carried forward by long *P. mirabilis'* cells (>4 μm), that there was an unexpected amount of heterogeneity in individual cell morphology, particularly with regards to cell length. This manuscript opens interesting questions about what *is* required for *P. mirabilis* collective migration to occur and suggests that swarming may not be as synchronized (or synchronized differently) than once thought.

2.1 Abstract

Uncovering cell morphology within communities is crucial to understanding how collective groups of organisms can function and adapt to their environments. Key questions remain regarding how cell morphology influences population behaviors and whether uniformity is required for coordinated actions such as collective migration. These gaps hinder the ability to describe how individuals contribute to collective adaptability and coordinated behaviors. We developed an image capture and analysis pipeline, named Swarmetrics, for studying individual cells that are within dense bacterial communities, can participate in collective behaviors, and may have irregular cell morphologies. We used *Proteus mirabilis* swarms as a proof-of-concept. Swarmetrics achieved an unbiased analysis of most cells, even those that were unlabeled and in a densely populated environment. In *P. mirabilis* swarms, we found unexpected heterogeneity in cell length throughout the swarm development cycle, particularly during active collective migration. Variance in cell length was revealed to be a reliable indicator of swarm development stage, comparable to a gene expression marker. These findings questioned the traditional view of uniform or synchronized transitions in bacterial swarming for the *Proteus* species. This study introduces new tools and insights for studying cellular variation in complex microbial environments, with broad applications to other dense communities, and sheds light on the physiology of individuals during collective behavior.

2.2 Introduction

Individual cells and their interactions can cause wide-scale changes to emergent behaviors and community structure. Many emergent behaviors can be organized at two different levels: the individual and the collective. While much is studied about the collective population, there remain many unanswered questions about the individual, such as its development. Some mechanisms that contribute to emergent collectivity in animals include cell-cell communication, physical contact with neighbors, nutrient availability, and cell development. For instance, ants, birds, and termites use collective behavior to collect, process, and distribute resources to build communities and defend their territories (69–72). Similarly, the bacterium *Myxococcus xanthus* and the eukaryotic slime mold *Dictyostelium discoideum* use collective behaviors to either prey on competitors or respond to physical changes in their environment (73,74). A common thread between these systems is that they involve densely packed communities, where an individual organism's movements contribute to population or colony expansion. Major challenges in characterizing individual organisms have included technical and methodological limitations. Here, we overcome these barriers for a model microbial system and reduce the gap between collective migration and cell development.

Studies of group dynamics in microbes and eukaryotic systems have provided clues for mechanisms of intercellular communication in the development of complex multicellular structures and group behaviors. For example, while individual *M. xanthus* cells induce prey cell lysis (75), the consumption of a whole prey community requires transition of these *M. xanthus* cells into collective multicellular aggregates capable of swarming (reviewed in Thiery and Kaimer (76)). The dynamics of cell-cell interactions drive multicellular *M. xanthus* groups like aggregates and swarms (73,74). Indeed, physical interactions between cells influence

multicellular aggregates that contribute to collective migration for other organisms, including the social amoeba *D. discoideum*, *Pseudomonas aeruginosa*, and *Escherichia coli* (77–80). Studies on the multicellular lifestyle of many microorganisms, particularly those in biofilms, reveal striking phenotypic and genetic heterogeneity, highlighting its potential role in microbial adaptation and resilience (81). While much research on collective migration focus on populations uniform in cell size and shape, many microbes that collectively migrate on firmer surfaces like catheters and glass are composed of individuals with varied lengths, for example, the pathogen *Vibrio parahaemolyticus*, the social microbe *Paenibacillus* spp., the pathogen *Serratia marcescens*, and the social opportunistic pathogen *Proteus mirabilis* (82–85). To capture single-cell dynamics in robust swarms with elongated cells (85), prior studies used isolation of cells from swarms before imaging, analysis of low-density regions, and speckling 10% of the population via expression of a fluorescent protein (resulting in just 1 in every 10 cells being analyzed) (1,11). These tools lack the ability to isolate and analyze most individual cell dynamics directly in the context of an active swarm.

To address this knowledge gap about the impacts of individual development and morphology on collective migration, we turned to a microbial system that shows significant changes in morphology over the course of its lifecycle and engages in the collective behavior of swarming. *P. mirabilis* is a major cause of recurrent complicated urinary tract infections. Part of its ability to induce disease is linked to its ability to engage in swarming, where cells move cooperatively, allowing for colony expansion on the centimeter scale (7,86). *P. mirabilis* cell length changes during their swarm development life cycle, from short ~2-μm long cells to elongated cells, reportedly up to 100 μm in length (2). During elongation to form swarmer cells, *P. mirabilis* cells dramatically increase the number of flagella per cell via upregulation of the *flhDC* regulon, which directly controls expression of the *fliA* gene (87–89). Swarming cells are thought to use their flagella to align along each other; and this state is transcriptionally tied to virulence (90). These rafts of cells can migrate further (perhaps as much as 100-fold) than a single-cell can. Cell development is varied within a swarm, and the visual periodicity of swarm expansion and stopping is reportedly correlated with changes in cell morphology and development (91,92). Although there are images of individual cells at nanometer to micrometer resolutions, studies of individual *P. mirabilis* cells within swarms are underexplored.

A significant hurdle to studying individual cell morphology and behavior during swarming is due to limitations in image analysis, including resolution and irregular shape detection. Prior challenges include constraints in computing power for image processing and data management in high-throughput quantitative analysis for individuals, particularly in the context of swarm development or in animal hosts. The lack of methodologies restricts the establishment of a clear, robust description of cell development during active swarms. However, analysis of single-cell morphology within a swarm over time requires high resolution and high-throughput analysis techniques. Over the past two decades, an increasing number of software packages were developed for segmenting and analyzing cells with different shapes and densities. Some software packages can track the interactions between individual cells and how these changes can induce changes in population structure (50,93–97). However, none of these analysis software packages adequately segment cells in *P. mirabilis* swarms. This result was perhaps due to the density of swarming populations and the irregular morphology of elongated cells; these organisms can

often fold in on themselves or are twisted during swarming. A new and different image analysis pipeline was needed.

Therefore, we integrated multiple software packages that, when used together, form a pipeline, named "Swarmetrics," which enables the analysis of irregular, unlabeled cells within densely packed swarms. We utilized this pipeline, which integrates Trainable Weka Segmentation, MicrobeJ, Fiji, and R (98–104), to generate a quantitative dataset characterizing the morphological changes of *P. mirabilis* cells during the swarm development cycle within the native context of an active swarm. We found significant heterogeneity in cell morphology during each stage of swarm development and discovered that cell development is unlikely to be synchronized during swarming. Only a fraction of the population was elongated during active collective migration, and cell length was weakly correlated with expression of a *fliA* chromosomal reporter, suggesting that absolute cell length may not reflect individual participation in active swarms. Variance in cell length appeared to indicate stages in the swarm development cycle, similarly to that of *fliA* gene expression. Altogether, Swarmetrics permitted the examination of single-cell dynamics during active collective migration and the initial integration of individual cell development with collective behaviors, all while using unlabeled cells in a native, dense environment.

2.3 Results

2.3.1 An image capture and analysis pipeline identifies and characterizes individual cells with various morphologies in densely packed communities.

Major challenges to image analysis of microbial swarms are the density of cells and the variation in cell lengths and orientations. These hurdles are especially acute for robust swarmers like *P. mirabilis*, whose shape can range from linear to spiral. We sought to capitalize on recently advanced machine-learning tools to develop an expanded imaging and analysis pipeline for densely packed bacterial communities.

We aimed to acquire information about individual cell morphology within the native context of a dense swarm and without bias due to cell labeling. *P. mirabilis* is a sufficient model, because its macroscale swarm behaviors are correlated with microscale cell physiology (2,91). We integrated two open-source software packages within the Fiji ecosystem; all classifier models and code are publicly posted as outlined in Materials and Methods. Phase contrast images were segmented after capture of in-focus, phase-contrast images using light microscopy (Figure 2.1). We next applied Trainable Weka Segmentation, which is a collection of machine-learning algorithms that uses manual annotations to train a classifier to segment stacks of images (104). We optimized classifier models to separate between cells, background, and intercellular spaces by manually tracing cells from two to four images per condition. The intracellular spaces were traced during classifier training due to the presence of light-gray extracellular flagella-like materials, which required additional classification and, in doing so, resulted in a clearer delineation between cells and background in the masks. Trainable Weka Segmentation uses the trained classifiers to generate probability masks of the images; these then produce binary masks. We used MicrobeJ, a tool for high-throughput quantitative analysis of individual cells from binary masks (38), to generate quantitative measurements for cells identified in the binary mask (Figure 2.1). To reduce mislabeling of background noise, such as debris and lysed cells, in an

automated manner, we excluded cropped cells at the image edge, and optimized the area, length, and circularity filters. We also manually removed the few remaining mislabeled components within each image. Inspection of the images revealed that particles of 1 µm or less in length were debris (Supplemental Figure S2.1), and so, we also applied a filter of greater than 1 µm in length for cell analysis. By completion, most cells within each frame have a corresponding binary outline and can be analyzed for morphological and spatial properties.

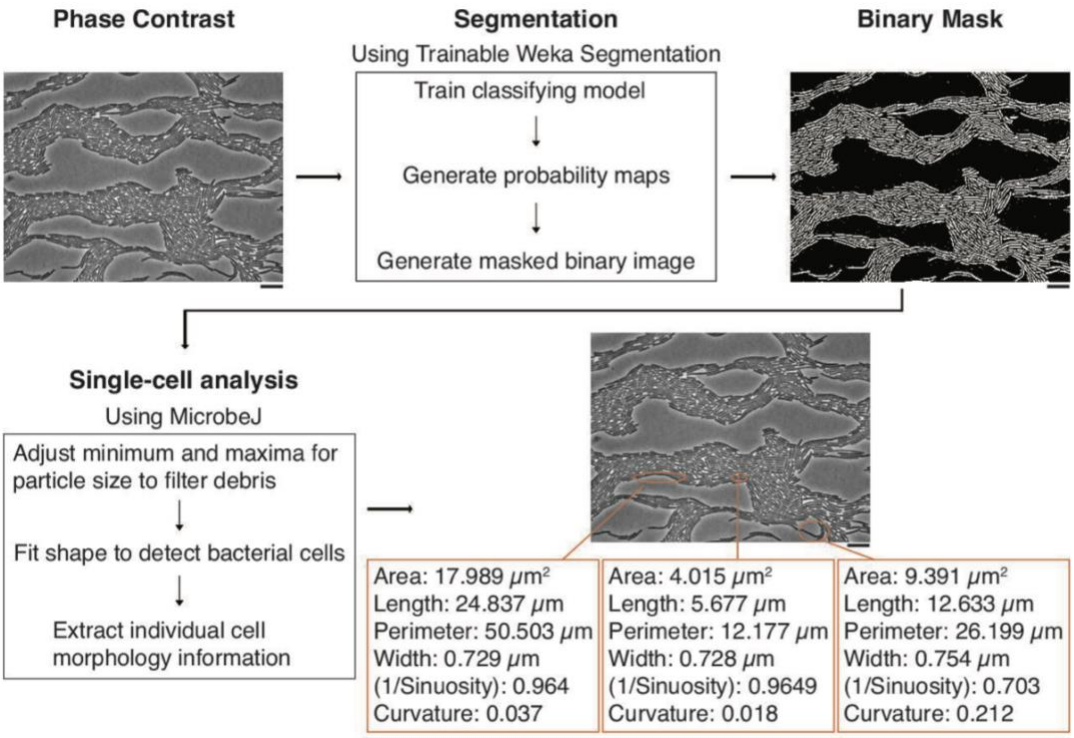


Figure 2.1 Swarmetrics pipeline for single-cell analysis in swarms. Upper left, a representative phase contrast image of a swarm at four hours. Cells (black, rod-like shapes) are densely packed in the population; the grey is the rich nutrient-agar on top of which cells move. Scale bar, 10 µm. Using Weka Segmentation, we generated a binary mask (see Materials and Methods), where cells are black, and all other features (background and the space in-between cells) are white. Once the mask (upper right image) was imported into Microbe J, we extracted morphological information about individual cells. A representative three cells have been circled in red to show the recorded morphological information (bottom right image). Cell measurements were taken for cells in the population except those along the perimeter.

Individual cell data is accessible once an object is defined in the mask. Using MicrobeJ, we focused on six features: width, area, length, perimeter, and sinuosity. We used R to analyze the quantitative measurements. As an example, three cells are highlighted in Figure 1; their quantitative values correspond to their visible differences in morphology. From this sample image in Figure 2.1, we quantified 657 cells, representing > 89% of the visible cells and particles when counted by hand (Supplement Figure 1.2), which is similar to other cell detection algorithms (105–108). In reviewing images, missed particles are present at all time points, with no apparent bias for length. Further, the absence of ~10% of particles from detection appears to be due to many factors such as (i) inaccuracies in cell detection, (ii) out-of-focus cell capture, (iii) low-level cell death and debris, and (iv) cells at the edges of the micrograph image or

partially in the image. Given that the sizes of the final datasets were over 7,000 cells (Figure 2.2), we considered this data to be robust, rigorous, and representative of the populations. Altogether, the implementation of this integrated pipeline for analyzing individual cells can lead to a deeper understanding of single-cell morphologies within a dense population.

2.3.2 Swarm development is a gradient of single-cell morphologies.

For a first application of the Swarmetrics pipeline, we turned to the *P. mirabilis* swarm development cycle. The ability to capture correlated developmental time points in this cycle using photography and microscopy has seeded several outstanding questions, including the extent of cell morphologies in a swarm and the potential synchronization and periodicity of the swarm developmental cycle (11,86,91,93,109). Swarm development connects visual migration with distinct cellular morphologies; however, historically, higher resolution of single-cell dynamics within a broader population was difficult to achieve. The literature speaks of a defined set of stages during each round of the swarm development cycle (110). Specifically, short cells appear to double more than twice in length and then begin to migrate (Figure 2.2A). Next, for unknown reasons, cells stop moving and divide into 2-μm long cells in a period termed, "consolidation". Cells re-enter the swarm development cycle if still on a viscous or hard surface, resulting in a characteristic macroscale bullseye pattern (Figure 2.2B). The current model states that there is synchronization among *P. mirabilis* cells between each swarm development stage; it was developed largely based on photography, where cells were removed from the swarm conditions and imaged using wet mounts (91,92,111). Therefore, we first asked whether cells undergoing the swarm development cycle are synchronized by examining cell length.

We applied the Swarmetrics pipeline to these micrograph images, resulting in a dataset of nearly 80,000 cells (Supplement Figure S2.2). Box-and-whisker plots of cellular length and width, as well as the variance within the population, were generated using R and analyzed for statistical significance (Figure 2.2D - G). In the images from hours 2 and 3, at least 75% of the cells were shorter than 5 μm (Figure 2.2D), and the variance among them was small (Figure 2.4E). Hour 4 had a similar median size for the population's cell length; however, the variance increased, which might be explained by the long tail containing longer cells. Hours 5 and 6 had an increased median and variance. By contrast, at hour 7, the median cell length and variance decreased and then increased again at hour 8. There were statistically significant differences in cell length; all had significant differences ($p < 10^{-4}$) from the others in a pairwise Wilcoxon Rank Sum Tests, except for between hours 7 and 8 (Supplemental Figure S2.4). The population distribution for cell length and its variance gradually increased and decreased in a manner consistent with the micrographs visualizing each stage of the swarm development cycle. Unexpectedly, a population of short cells remained during all cycle stages.

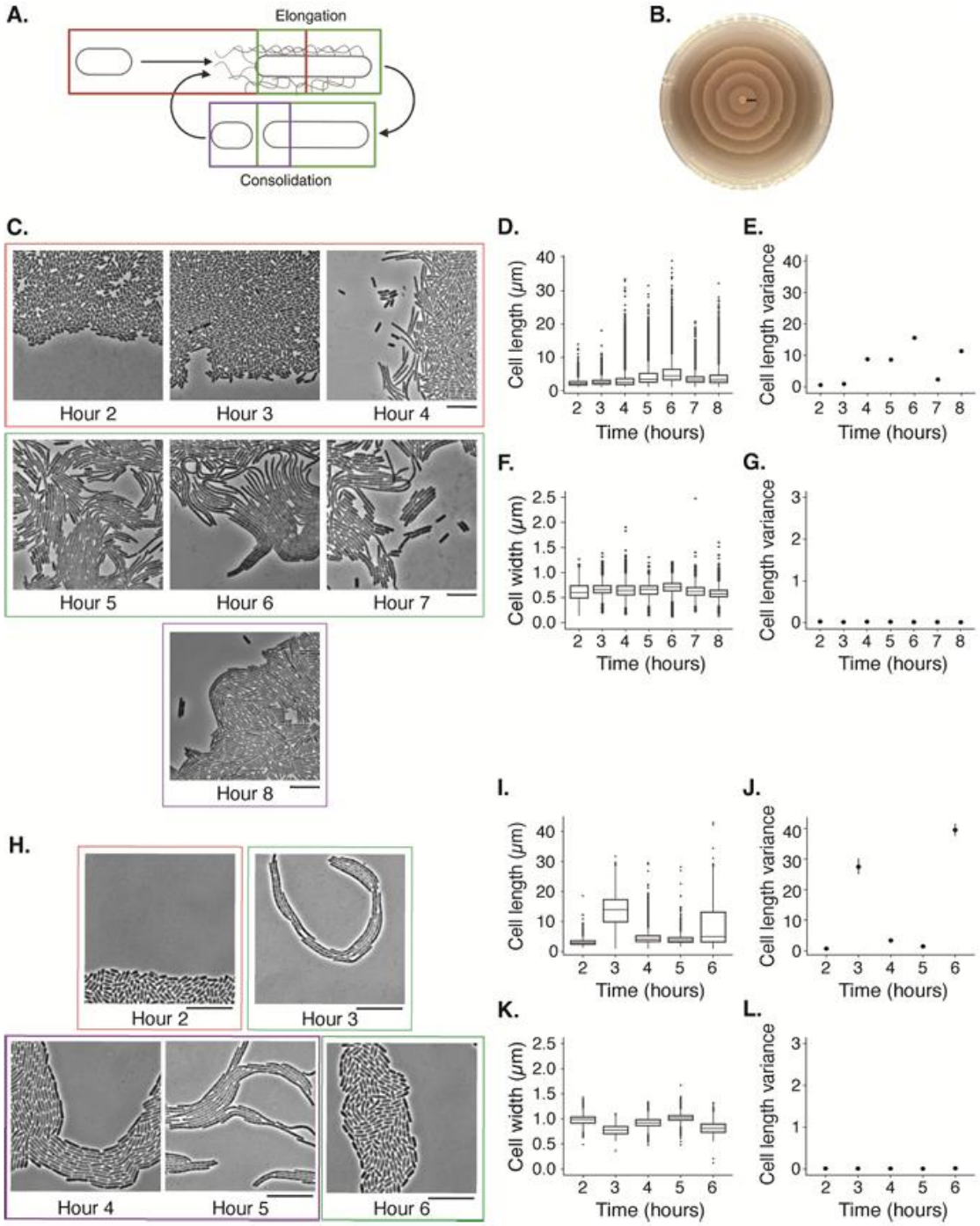


Figure 2.2 A P. mirabilis swarm contains two subpopulations of cells. (A) A diagram of the canonical swarm development cycle (Garling, E. (2025) https://BioRender.com/fdsg7i9). Cells elongate and become hyperflagellated during surface adaptation (red box), contributing to collective migration and visible population expansion during the swarming phase (green box). Cells likely divide during the consolidation phase (purple box). (B) For strain BB2000, after 18 h of growth on CM55 Agar at 37˚C, the population forms a bullseye pattern, covering a 10 cm diameter plate. The black line marks one full swarm development cycle and corresponds to the period of observation and analysis. (C and H) Representative phase-contrast images of strains BB2000 and ATCC 29906, respectively, taken hourly. Uncropped images are in Fig. S2.1. Scale bar, 10 µm. (D–G, I–L) We quantified the lengths of cells in the monolayer at the expanding edge of the swarm using the Swarmetrics pipeline. Each graph is a box-and-whisker

plot, each box represents the 1st–3rd quartile, and the middle line is the median. Whiskers indicate minimum and maximum; single dots are outliers. For strain BB2000 (D–G), sample sizes are as follows, sequentially from hours 2 to 8: 16,723; 20,726; 5,506; 9,367; 7,478; 10,865; and 9,156. For strain ATCC 29906 (I–L), sample sizes, sequentially from hours 2 to 6, are 7,740; 942; 10,142; 15,819; and 2,836. Shown are cell length (D and I), variance of length (E and J), width (F and K), and variance of width (G and L). The vertical bars on the variance plots show 95% confidence intervals. All time points were significantly different, except for length of BB2000 at hours 7 and 8. The population metrics and statistical tests for BB2000 and ATCC 29906 are in Fig. S2.2 and S2.3, respectively. Fig. S2.4 shows the length data from individual experiments.

The heterogeneity present in the swarm cycle and the effectiveness of Swarmetrics was also true for swarms of another *P. mirabilis* population. Strain ATCC 29906 swarms form macroscopic concentric rings that have wider swarm terraces than seen for strain BB2000. We performed equivalent swarm assays over a six-hour time course, followed by analysis using the Swarmetrics pipeline (Figure 2.2H, Supplemental Figure S2.2). As with strain BB2000, the distribution of cell lengths varied over the swarm cycle (Figure 2I) with the variance being largest during periods of active, collective migration and smallest during periods where cells were, on average, shortest (Figures 2.2J, Supplemental Figure S2.4). Cell width and its variance, by contrast, had minimal change over the six-hour time courses (Figures 2.2K, 2.2L), though the distributions of width were statistically significant between time points (Supplemental Figure S2.5). Based on this single-cell analysis, cells of strain ATCC 29906 appeared to undergo the swarm developmental cycle twice within the six-hour time course as opposed to the single cycle observed in strain BB2000. A population of short cells remained during the swarm development cycle for both strains ATCC 29906 and BB2000

2.3.3 Two subsets of populations persist across all the time points

We investigated whether at least two distinct populations were present throughout the swarm development cycle. Based on the literature, elongated cells are typically associated with active swarming literature 4,109). Therefore, we expected that short cells would be largely, if not fully, absent from periods of active swarming, while long cells would be absent from consolidation, i.e., periods of cell division. To query this prediction, we classified cells based on cell length: short (4 μm or smaller, which is the length of a cell doubling or less) and long (> 4 μm, which are cells elongated beyond cell doubling); see Figure 2.2A. The fraction of long cells increased at each time point until it became the majority population at 6 hours, followed by a decrease at hours 7 and 8 (Figure 2.3B). This timing matched that of the variance for cell length of the whole population (Figure 2.2E).

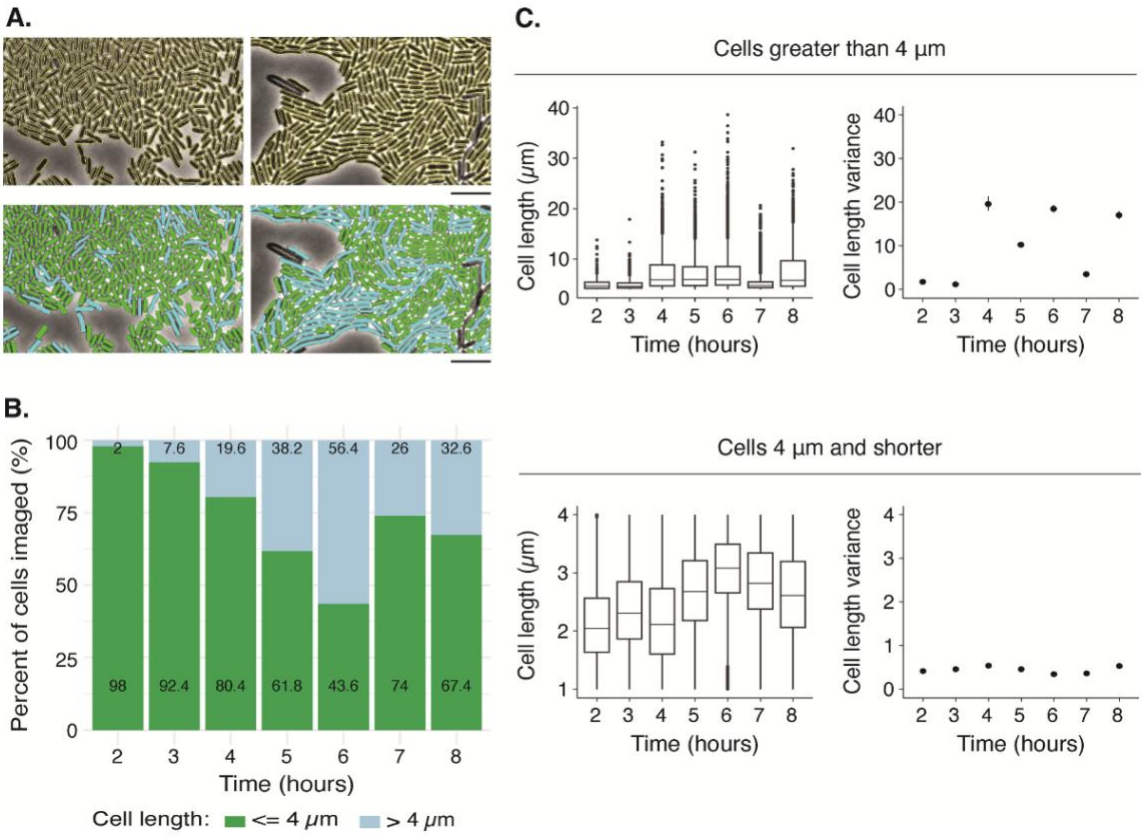


Figure 2.3 A subset of the population at the expanding edge is short throughout a complete swarm development cycle, including during active migration. (A) Top, representative phase contrast images from hours 4 and 5 are outlined yellow after being identified through overlay with their masked images. The bottom, 4 µm or less in length, was pseudo-colored green, and cells greater than 4 µm in length were pseudo-colored blue in the bottom panel of A. In each panel, a minor fraction of cells is not identified by the segmentation software. These cells usually lack the necessary pixel darkness that marks an intact cell or are visually indistinguishable from neighboring cells. Further, by manually investigating multiple images, we determined that all objects 1 µm or less in length were debris. scale bar, 10 µm. (B) The relative distribution of the total populations by hour for cells longer than 4 µm (blue) and 4 µm or shorter (green) in length. The number printed within each bar is the precise percentage value. (C) Top, the complete data set was subdivided in cells with lengths >4 µm; bottom, 4 µm or less and analyzed for cell lengths (left) and variance (right). Each graph is a box-and-whisker plot where the box shows the 1st–3rd quartile, and the middle line is the median. Whiskers indicate minimum and maximum; single dots are outliers. On the variance plots, the vertical bars show 95% confidence intervals. Fig. S2.1 contains the population metrics and statistical tests.

For the short cells, there was little change in length distributions and variance (Figure 2.3C). This result was likely due to the imposed boundaries (maximum of 4 µm and the minimum cell size of 1 µm). By contrast, the large cells had different cell length distributions at each time point over the eight-hour time course like the whole population (Figure 2.3C). The variance of the long cells, however, differed from the full population in hours 4, 5, 6, and 8 (Figure 2.32E and 2.3C). The variance was higher when considering only long cells than the full population, almost twice as large for hours 4, 6, and 8. This result indicated that long cells had a wide distribution at the same time when active swarming is visible by eye. Altogether, these findings underscore the heterogeneity in cell length during active collective migration and the potential for the variance of cell length to act as a marker for active swarming phases.

2.3.4 Cell fluorescence and morphologies can be analyzed simultaneously within dense populations during active behaviors

The observed variance in cell length, particularly amongst swarming cells, suggested that variance may be a marker for active swarming phases. However, swarming is a dynamic process that involves many biological factors. Therefore, we turned to gene expression of the flagellar regulon, which is known to play a role in swarming motility (4,109). We asked whether the image capture and analysis pipeline could be applied to epifluorescence microscopy and whether the gene expression pattern could provide further insights into *P. mirabilis* swarm dynamics.

We examined a *P. mirabilis* BB2000 strain engineered with a chromosomal *fliA* transcriptional reporter (112), because *fliA* is required for swarming and the flagellar class-1 transcriptional regulator, FlhDC, directly regulates *fliA* and other swarm-related genes (87–89,113). The Venus fluorophore folds rapidly in ~20 minutes in *E. coli*, and its measured half-life is ~100 minutes in eukaryotic cell lines (114,115). We captured phase contrast and paired epifluorescence images of this strain over eight hours (Figure 2.4A, Supplemental Figure S2.4). The fluorescence images in Figure 2.4A are not automatically scaled; the relative brightness differences between the time points reflects relative expression. By visual inspection, fluorescence is close to background levels in hours 2 and 6 and much brighter at hours 5 and 8, when cells are also visibly longer and moving across the surface. The hours 4 and 7 images had some of the most variability in visible brightness among individual cells. Altogether, the fluorescence associated with *fliA* visibly varied across the eight-hour time course, with the brightest fluorescence during time points containing elongated cells.

We used Swarmetrics to quantify the visible differences between time points. The phase contrast images were processed through the pipeline, and the binary-masks were then overlaid onto the fluorescent images within Microbe J. The intensity values for each cell were extracted from MicrobeJ, and the values were normalized to the background fluorescence per frame to account for autofluorescence. We calculated the mean fluorescence per cell (Figure 2.4) and the cell lengths (Figure S2.8B). The cell length distributions followed a similar pattern as the unlabeled wildtype (Figure S2.8B, S2.3A, S2.4A), with two similar subpopulations in the dataset: longer than 4 μm and 4 μm or shorter (Supplemental Figure S2.9).

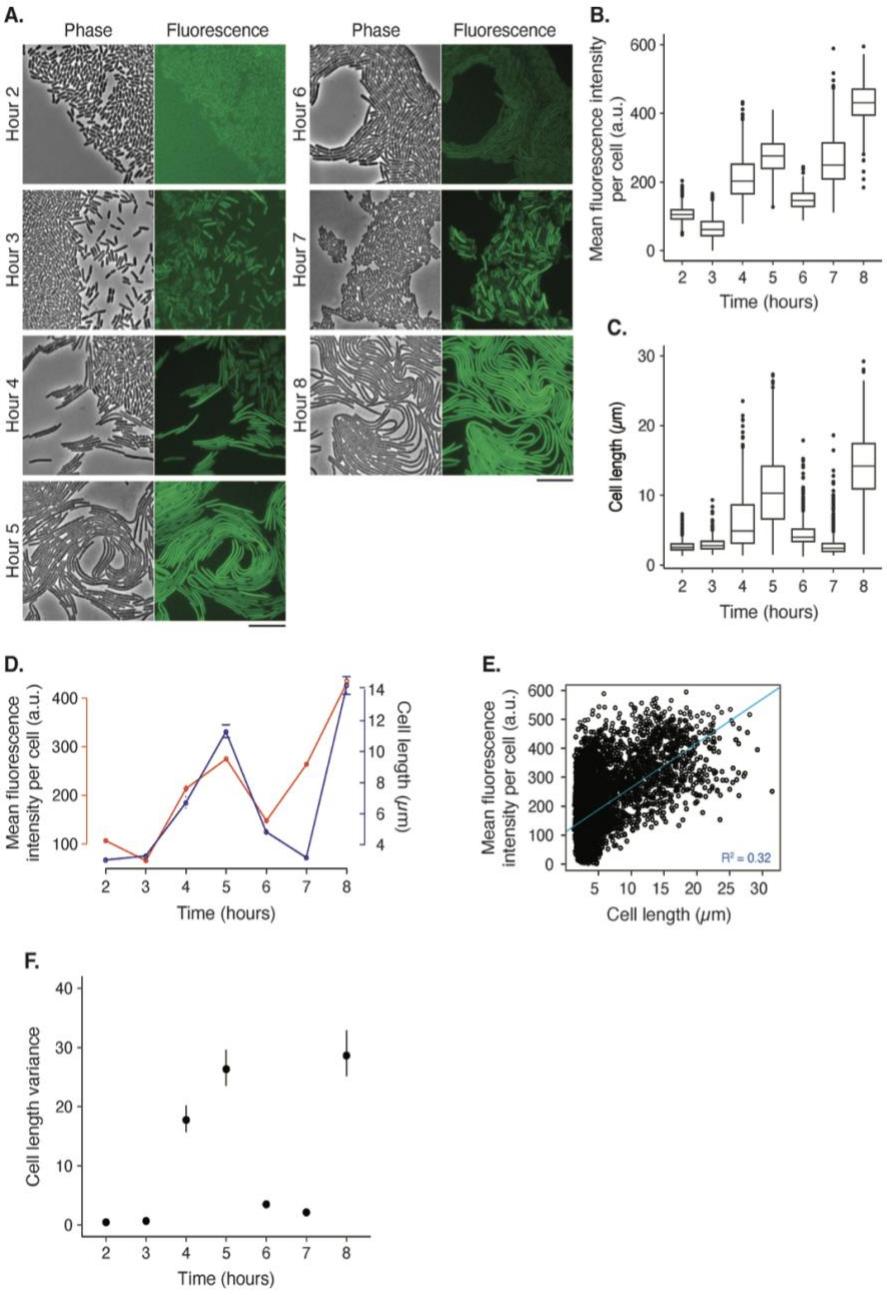


Figure 2.4 A swarming-associated genetic marker exhibits dynamic patterns that align with changes in cell length during a swarm cycle. (A) BB2000-derived cells with a fliA-venus reporter integrated at the fliA locus were examined during an 8-h time course and analyzed with Swarmetrics. Shown are representative images of phase contrast (Phase) and false-color (Fluorescence) images. No intensity rescaling was applied to the fluorescence images. Supplemental Figure 4.1 contains the uncropped images. Scale bar, 10 μm. (B and C) For the box-and-whisker plots, each box represents the 1st–3rd quartile, and the middle line is the median. Whiskers indicate minimum and maximum; single dots are outliers. See Fig. S4.1 and S4.2 for population metrics and statistical tests. Sample sizes are as follows, sequentially from hours 2 to 8: 1,603; 1,454; 506; 590; 864; 1,681; and 437. (B) The mean fluorescence intensity per cell was normalized to background fluorescence for all images; reported are arbitrary units (a.u.). (C) The lengths of cells at the expanding edge of the swarm were quantified for each time point. (D) The mean plot for fluorescence intensity per cell (red) was overlaid with the mean plot for cell length (blue) for the 8-h

assay. Error bars represent the 95% CI for each measurement at each time point. (E) The graph is of each cell in the eight-hour data set, plotted by length and its corresponding mean fluorescence intensity. The linear regression ($y = 0.0276x + 0.84336$) has an adjusted R^2 of 0.3177, suggesting a low correlation between average fluorescence intensity and cell length. (F) Plot of cell length variance in which The vertical bars indicate 95% CI.

Both the cell length and the fluorescence expression graphs showed a sinusoidal distribution of the means (Figure 2.2D). This dataset also showed two local minima and two local maxima in cell length and mean fluorescence intensity per cell, suggesting this reporter strain might have undergone more than one swarm development cycle during the observation period. The highest mean fluorescence corresponded to the distributions with the longest cell lengths, and the distributions of mean fluorescence intensity appeared to increase from hours 6 to 7 without a commensurate increase in the cell length at that transition point (similarly but less so from hours 3 to 4). These results were consistent with the observed weak correlation ($R^2 = 0.3177$, $p = 2.2^{-16}$) between cell length and average fluorescence intensity due to *fliA* expression (Figure 2.4E). Short cells had the same range of fluorescence intensities as long cells (Supplemental Figure S2.9). Altogether, these data agreed with earlier reports that *fliA* expression is a marker of swarm development stage (4). Notably, cell length variance and median cell length were aligned with the *fliA* fluorescence data (Figure 2.4F). The insights gleaned from this study advance an understanding of *P. mirabilis* swarm dynamics and underscore the broader importance of interrogating single-cell morphometrics in communities.

2.4 Conclusions and Perspectives

We have developed and applied the Swarmetrics pipeline to analyze flexible, elongated individual cells embedded in dense bacterial communities without use of fluorescent markers. Previous attempts to study such dense microbial communities were less successful. Software packages failed in multiple ways: (i) incorrectly identified elongated or S-shaped cells, (ii) inadvertently added division sites to elongated cells, (iii) inadequately separate adjacent cells, (iv) identified only a subset of homogeneously shaped cells, or (v) did not process multiple frames of an image from a single training set (94–97,116). However, by integrating the software tools Weka segmentation, MicrobeJ, and Fiji (101,102,104), we successfully extracted individual cell morphological data from multiple frames and statistically analyzed the resultant using R and associated packages (98–100). When examining *P. mirabilis* swarms, this image capture and analysis approach made it possible for population-wide analyses across a full swarm development cycle.

These results expand the established models of *P. mirabilis* swarm development (92,111,113) by demonstrating widespread heterogeneity throughout the population and decoupling individual cell length from life-stage in the swarm developmental cycle (Figure 2.5). Earlier studies of *P. mirabilis* swarming identified long, hyperflagellated cells as the source for collective migration and suggested synchronized transitions between long, swarming cells and short, dividing cells, e.g., when entering consolidation (86,117). However, this study's findings reveal significant heterogeneity in cell morphology throughout the swarm development cycle for strains BB2000 and ATCC 29906 (Figures 2.2 and 2.3). Differing cell morphologies were present even at the height of swarming, with short cells (4 μm or shorter) dominating the population at nearly every time point for strain BB2000 (Figure 2.3), which bring into question the notion of strict synchrony between active swarming and cell-dividing consolidation. The variance in cell length

at hours 4, 5, and 6 suggests that *P. mirabilis* cells may actively undergo cell division or elongation during swarming. These data also open the possibility that uniformity in cell shape and length is less critical for collective migration in *P. mirabilis* (Figure 2.5). How cells with a broad range of length morphology can align against one another for efficient coordination is unclear.

Questions about the metabolic and developmental state of short and long cells remain open but can start to be answered with this pipeline. In comparing single-cell fluorescence with cell morphology, we found that cell length and *fliA* expression are weakly correlated, which supports earlier hypotheses that flagellar regulation does not control cell shape and are instead co-regulated by the *rcs* regulon. Specifically, Rather and colleagues provided evidence that the RcsB protein regulates the genes *minCDE*, *flhDC*, and *fliA*, and that FlhDC may not regulate *minCDE* expression (118,119). The balance between cell length, DNA replication, and cell division is largely unexplored in *P. mirabilis* and other multinucleated elongated cells. Open questions include the location of FtsZ rings, the Min proteins in elongated cells, and the dynamics of these proteins during the swarm development cycle. Work in *V. parahaemolyticus* suggests that in elongated bacterial cells, the *min* system can function off-center, leading to asymmetric cell division and the emergence of longer and shorter cells within a migrating population (120,121). Asymmetric cell division in *V. parahaemolyticus* may preserve the elongated swarming phenotype, thereby allowing continued swarmer colony expansion while simultaneously permitting cell growth and division. The fate of the short cells remains unknown, and additional tracking technology is necessary to fully resolve this question for *V. parahaemolyticus*, *P. mirabilis*, and other migrating populations.

Fluorescent markers can help shed light on how alterations in swarm cycle dynamics influence the regulatory mechanisms underlying collective behaviors, especially when paired with phase contrast imaging. We found that the dynamics of the *fliA* expression are consistent with prior studies on gene expression during the swarm development cycle (4,28,109). However, we also saw an unexpected presence of short cells at every swarm development stage, some with high *fliA* expression, which led to the observation of a low correlation between absolute cell length and fluorescence associated with *fliA* (Figure 2.4). One interpretation is that the short cells might be active participants in the mass of cells migrating outward. Another potential explanation is that the expression level of *fliA* in the mother cell before cell division could influence the daughter cell fluorescence intensity but does not actually indicate if the cell is flagellated. Both possibilities need further extensive examination. The plots for the variance in cell length and mean *fliA* fluorescence intensity exhibited similar patterns with similar maxima and minima timing (Figures 2.4D, 2.4F), indicating that these two variables might indicate equivalent stages of the swarm development cycle. Given this observation, cell length variance might serve as a gene expression-independent marker for stages of the swarm cycle.
The need for non-fluorescence-based markers is critical, because the production of these proteins can change the metabolism, motility, and population dynamics of cells. This challenge was visible in this dataset with a single-copy chromosomally integrated fluorescence marker for gene expression (compare Figures 2.2 and 2.4). As the capabilities to integrate machine-learning grows, biologists also need to update experimental methods to better capture unmarked individuals in complex, multi-species communities. The Swarmetrics pipeline can serve as one such tool that reduces reliance on fluorescence markers.

The complexity of *P. mirabilis* swarm dynamics point to the need for advanced live-cell tracking in dense microbial communities to gain a more granular understanding of the changes in an individual cell's gene expression and morphology over time within a native context. Swarmetrics has limitations, including difficulties in segmenting cells that are not in a monolayer and the unbiased missed detection of some imaged particles during processing. However, the acquisition of sufficiently large datasets, such as the approximately 7,000 to 80,000 cells per strain analyzed in this manuscript, still provides a representative sample of the population and can produce robust results that overcome the day-to-day variability inherent to biological organisms (Supplemental Figure S2.6). Of note, data can enter the Swarmetrics pipeline at multiple stages, e.g., one can use an alternative cell detection algorithm to feed masks into MicrobeJ. We chose to use Weka Segmentation due to its accuracy in identifying elongated or irregularly shaped swarmer cells without introducing artificial division planes. The R code provides analysis tools for large datasets. Swarmetrics employs an easy-to-use GUI interface that can run locally on a laptop, thereby allowing for a lower barrier of entry into image analysis of large datasets. The use of a user-friendly GUI for nonexpert users cannot be understated, as segmentation and cell analysis are requisite for the successful analysis of microscopy images and is an issue that several people have raised within the biological fields (80,105,122). No single machine learning package can address every biological challenge. The Swarmetrics pipeline is well-suited for analyzing monolayers of densely packed cells, including those with irregular single-cell morphologies.

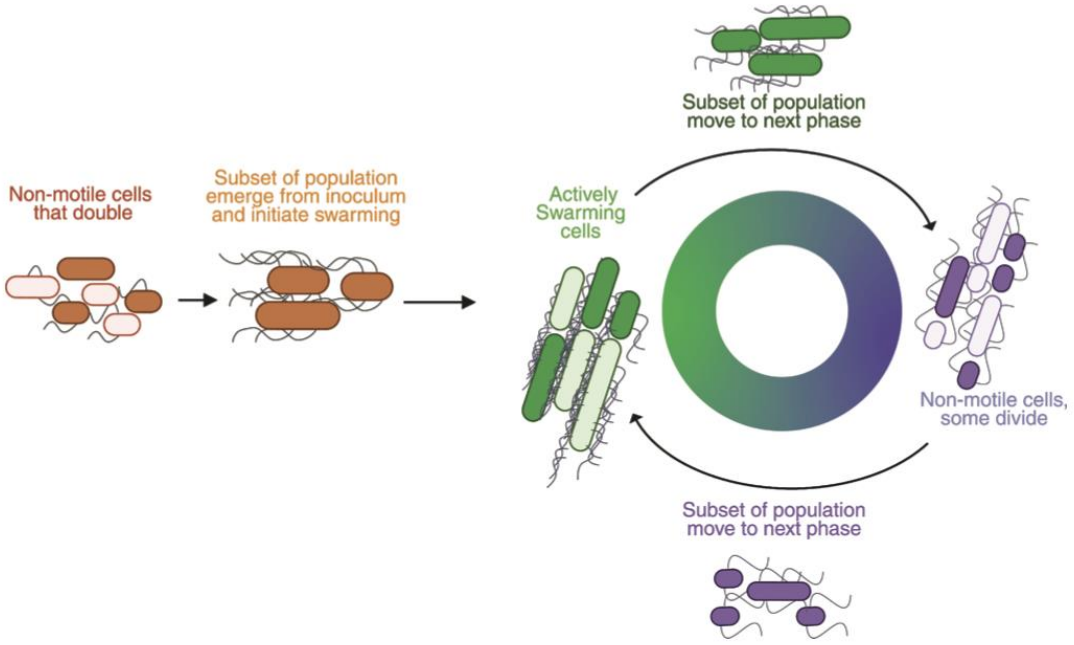


Figure 2.5 An updated model highlights intrinsic cell length heterogeneity throughout the *P. mirabilis* swarm development cycle. We propose that *P. mirabilis* cells display morphological heterogeneity even during active swarming. When swarming is initiated by a subset of the population, a gradient of single-cell morphologies emerges and actively participates in collective migration. Through each iterative stage of the swarm development cycle, a subset of the population initiates the forward momentum into the following life stage. The image was created in BioRender: Garling, E. (2025) https://BioRender.com/69b6ws0.

These discoveries also highlight that understanding how heterogeneity is generated and maintained in clonal bacterial populations is of critical importance. Heterogeneity, both at the morphological and genetic levels, has been described for many collective microbial populations, including biofilms where heterogeneity is shown to play an important role in collective assemblage, division of labor, and increased resistance to stress (81). However, due to the density and irregular morphologies of many cells in surface-constrained communities, morphometric analysis was stymied for these populations. This data shows that this pipeline overcomes some of the previously discussed limitations. Future iterations of Swarmetrics could serve as a solution for analyzing multi-layer microbial communities, such as biofilms, gut microbiomes, and soil ecosystems. Understanding how interactions between individual cells contribute to population-level behaviors is a fundamental question across biological fields, from microbiology to eukaryotic systems. The ability to analyze individual cells within dense microbial populations using Swarmetrics provides a powerful approach for integrating computational tools with microscopy techniques. As machine-learning approaches continue to evolve, their application in cellular biology and beyond promises to enhance the ability to capture, quantify, and interpret complex cellular behaviors at a granular resolution. In the future, bridging scales from single cells to entire populations could offer new opportunities to uncover mechanisms underlying collective behavior, adaptation, and structural organization in biological systems.

2.5 Materials and Methods

2.5.1 Strains and Media

P. mirabilis BB2000 (NCBI:txid1266738) (56), *P. mirabilis* BB2000 *fliA-venus* (10), and *P. mirabilis* ATCC 29906 (NCBI:txid525369) were used in this study. *P. mirabilis* was maintained on LSW^- agar (123). CM55 blood agar base (Oxoid, Basingstoke, UK) is the swarm-permissive agar. All liquid cultures were grown aerobically in LB broth (Lennox) at 37°C. *P. mirabilis* strains undergo equivalent swarm development cycles on CM55 media, LB agar, and other agar conditions, both on microscopy agar pads and on large diameter plates (11,86,91,92).

2.5.2 Swarming experiments & microscopy

Strains were grown aerobically overnight at 37°C in LB broth. Overnight cultures were diluted in LB broth (Lennox) to an optical density at 600 nm (OD_{600}) of 3.0, then inoculated onto a CM55 agar pad on a microscope slide. Microscopy was performed as previously described (25). Briefly, the expanding edge of the swarms and inoculum were imaged using a Leica DM5500B (Leica Microsystems, Buffalo Grove, IL) and a CoolSnap HQ2 cooled charge-coupled-device (CCD) camera (Photometrics, Tucson, AZ) for BB2000 and a Zeiss AxioObserver Z1 inverted fluorescence microscope with the Tucson sCMOS camera for ATCC 29906. A new slide was used for each time point; all were inoculated at the same time. MetaMorph version 7.8.0.0 (Molecular Devices, Sunnyvale, CA) and iVision (BioVision Technologies, Exton, PA) were used for image acquisition. Venus (maximum excitation, 515 nm; maximum emission, 528 nm) (114,124–126) was visualized in the green fluorescent protein (GFP) channel using a GFP ET filter cube (excitation, 470/40 nm; emission, 525/50 nm [Leica Microsystems, Buffalo Grove,

IL]). Fluorescence intensity and exposure time for each channel were equivalent across all microscopy experiments where applicable.

2.5.3 Image and statistical analyses

First, images were cropped if necessary to remove layered cells, or regions of the image that were oversaturated, out-of-focus, or contain excess debris to avoid obfuscating downstream analysis. Next, to segment images, stacks of images from each time point were loaded in the Trainable Weka Segmentation plugin v 3.3.4 (39–41) in Fiji 2.16.0/1.54g (National Institutes of Health, USA) (103,127). The algorithm was trained by manual annotations of cells, background, and the area between cells. Once the algorithm was trained, a classifier model was produced which was used for all analyzed images. Images were exported from Trainable Weka Segmentation as probability maps and converted to binary masks by adjusting the threshold and converting to binary masks in Fiji 2. To analyze phase contrast images, binary masks and phase contrast images were then loaded into the MicrobeJ v5.13h plugin in Fiji 2 (101). To false-color images within MicrobeJ based on cell length, the Type option field was selected and the argument "SHAPE.length> 4" was added as a parameter once the phase contrast images and their corresponding masks were loaded into MicrobeJ. Under this condition, when cells failed to meet the argument they would be colored green, and when cells met the argument, they would be colored blue. To analyze fluorescence images, phase contrast images and corresponding fluorescent images were loaded into the MicrobeJ plugin in Fiji.

Within MicrobeJ, several parameters were adjusted to filter out debris: area = 1-max, Circularity = 0.-0.9, length = 1-max. We manually investigated several images and found that all objects 1 μm or less in length were debris (Supplemental Figure S2.1). For this reason, we applied a filter cut-off of 1 μm. Additionally, the options of "exclude on edges", "shape descriptors", intensity, shape, and type were selected within MicrobeJ to allow for analysis of images and output of morphological information of *P. mirabilis*. Next, images were screened by hand in MicrobeJ to ensure that outliers were cells, not debris or improperly labeled cells. For fluorescent images, four 5x5 μm boxes were randomly placed on the cell-free areas of the image, and the fluorescence values for four boxes were averaged to achieve a mean background intensity value. This value was then subtracted from the fluorescence data before a mean fluorescence value was calculated. Data was then exported and analyzed with R Studio 2024.12.0+467 (98–100).

For the analysis length cutoff, the data was run through MicrobeJ without any length filters. We used the "Type" feature to visualize 1 μm and shorter particles or > 1 μm and then manually screened every image to make sure that no 1 μm or shorter particles were cells. The data was then exported and subjected to a secondary analysis in R to filter out the < 1 μm particles before analysis. For the box-and-whisker plots, each box represents the 1^{st} – 3^{rd} quartile, and the middle line is the median. Whiskers indicate minimum and maximum; single dots are outliers.
The cell detection algorithm identifies most cells. Failed particle detection is likely due to many factors including: (i) inaccuracies in cell detection, (ii) out-of-focus cell capture – often due to multi-layered cell populations, (iii) low-level cell death and debris, (iv) cells at the edges of the micrograph image, and (v) cells partially in the image. The ~89% detection rate of the visible cells and particles (as compared to counting by hand) is similar to published cell detection

algorithms (105–108). Particles are missed at all time points. The size of the final datasets (approximately 7,000 to 80,000 cells) supports that these data are robust, rigorous, and representative of the cell populations at the expanding edge of the swarm.

R for data management and statistical analysis. The pairwise Wilcoxon Rank Sum Test was selected, because this data is non-parametric, and the two populations are unmatched.

2.6 Acknowledgements

We would like to thank the members of Karine Gibbs' and Kimberley Seed's labs for thoughtful discussions and comments on the manuscript. STDHA and Alboukadel Kassambara's GGPlot2 Essentials: Great Data Visualization in R Edition 1, were especially helpful for the development of the R component of the Swarmetric pipeline. The eight-hour swarm time course was originally performed at Harvard University (K.H.) and then independently repeated at the University of California, Berkeley (E.E.G.). We thank the following organizations for funding this research: the Sloan Foundation (grant number G-2022-19509 to K.A.G.), the Packard Foundation, Harvard University, the National Institutes of Health (Training Grant number T32GM135143), the National Science Foundation (Graduate Research Fellowship to E.E.G.), and the University of California, Berkeley.

2.7 Author contributions

E.E.G. designed the project, performed the experiments and analysis, and co-wrote the manuscript. K.H. performed experiments (formerly at Harvard University), S.L. performed the ATCC 29906 experiments and analysis, and K.A.G designed the project, contributed to data analysis, and co-wrote the manuscript.

2.8 Data availability

Swarmetrics, which contains the imaging pipeline and R code used for data analysis, are available at *https://osf.io/mkeh4/overview.com*

2.9 Supplemental Figures and Tables

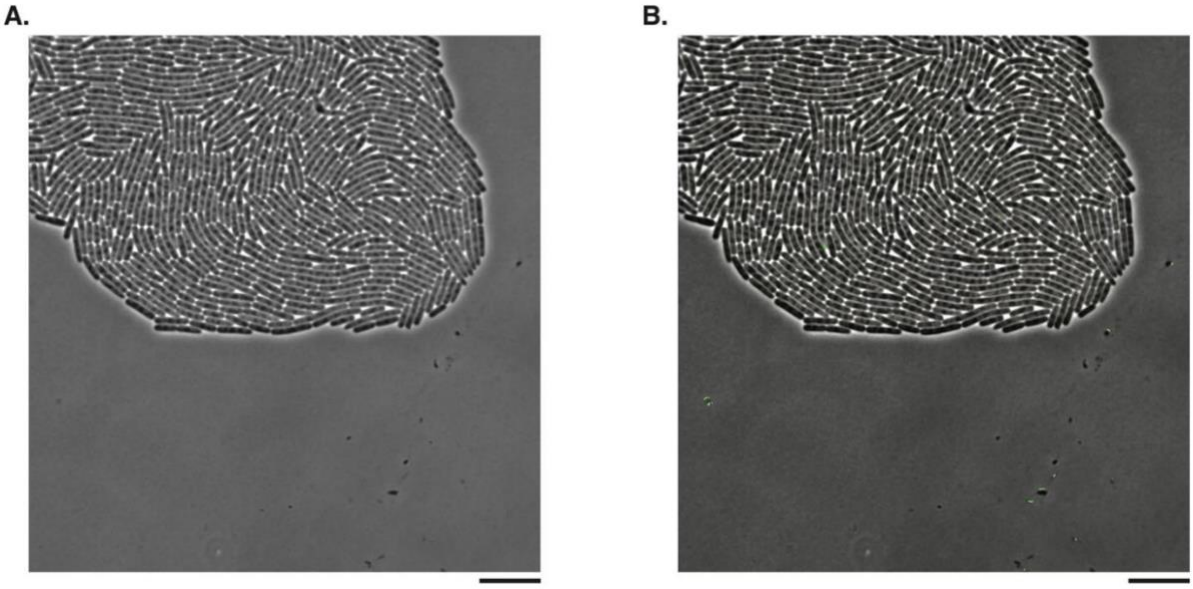


Figure S2.1 Particles 1 µm and shorter are not cells in these P. mirabilis swarms. (A) Micrograph of P. mirabilis strain ATCC 29906, replicate 1 at hour 4. This image is representative of all data. (B) Green marks particles of cell length 1 µm and shorter in the same micrograph as in A. We used the "Type" feature in MicrobeJ to highlight the particles. Scale bar, 10 µm.

Metric	Mean	Standard Deviation	Median	IQR	Max	Min
Length (µm)	6.42	4.25	5.10	3.38	33.9	1.85
Width (µm)	0.667	0.0827	0.669	0.108	0.958	0.326
Area (µm^2)	4.14	2.79	3.26	2.31	21.2	1.02
Perimeter (µm)	13.7	8.59	11.0	6.87	69.3	4.74
(1/sinuosity)	0.988	0.0226	0.994	0.00898	1.00	0.713
Curvature	0.0534	0.516	0.0389	0.0540	0.580	0.000109

Figure S2.2 Analysis of single-cell morphologies in the representative image in Figure 1. We generated this data using the Swarmetrics pipeline described in the Materials and Methods. A total of 665 cells from the image were analyzed; the characteristics for the cell population is recorded above for each morphological metric.

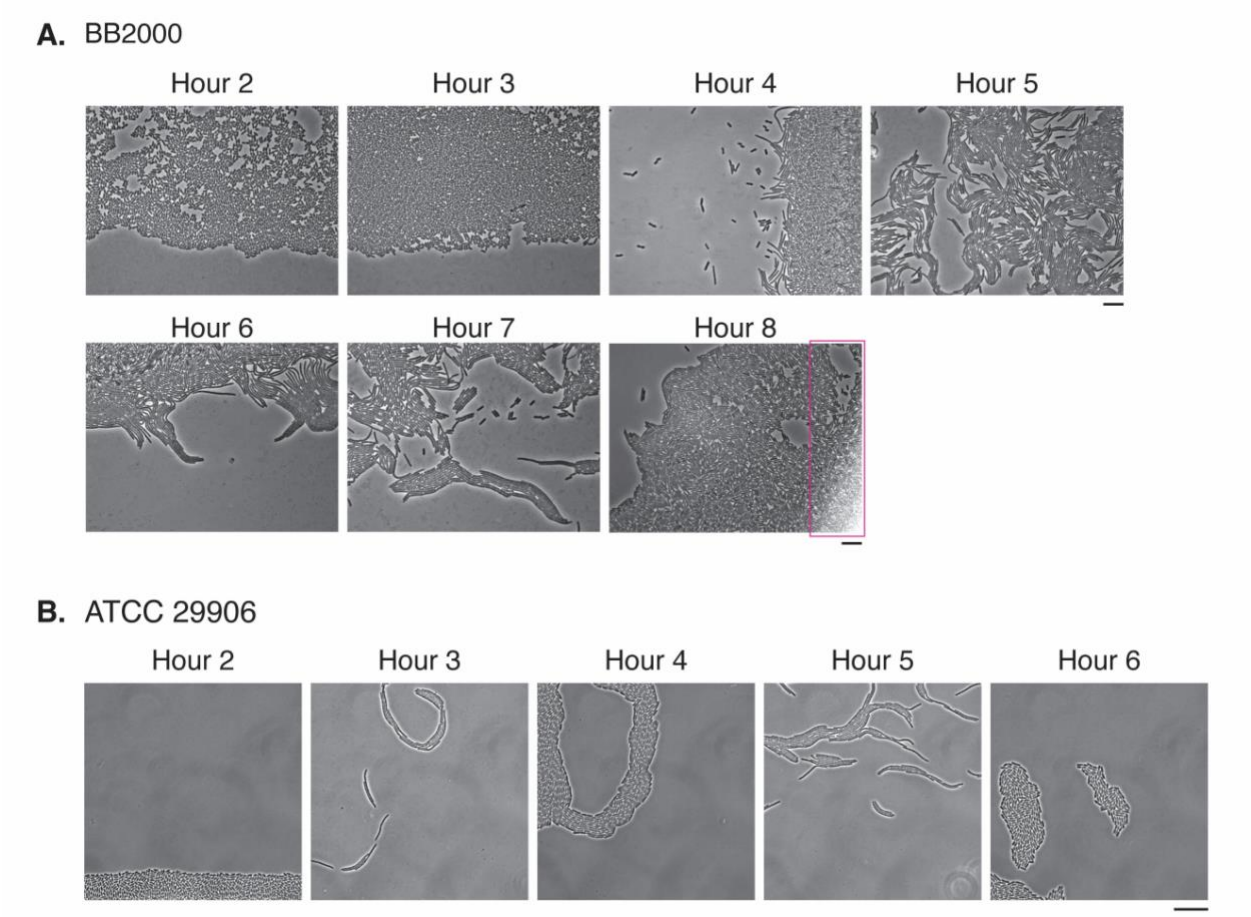


Figure S2.3 Full-frame images from P. mirabilis swarm development cycles. Displayed are the full-frame phase contrast images of the cropped pictures in Figure 2. Scale bar, 10 μm. (A) Strain BB2000. The magenta box at the hour 8 indicates where the image was cropped before proceeding with downstream analysis. We removed areas such as these that had overlapping cell layers, excess debris, or oversaturation within the image. (B) Strain ATCC 29906.

A.

Time (hours)	Mean length (µm)	Standard Deviation	Median length (µm)	IQR	Max length (µm)	Min length (µm)
2	2.21	0.783	2.07	0.964	13.8	1.07
3	2.58	0.982	2.39	1.14	17.9	1.07
4	3.30	2.96	2.44	1.77	33.2	1.07
5	4.37	2.94	3.37	2.54	31.3	1.07
6	5.59	3.95	4.32	3.10	38.7	1.07
7	3.50	1.53	3.18	1.50	20.6	1.07
8	4.26	3.37	3.19	2.20	31.9	1.07

B.

	2 hrs	3 hrs	4 hrs	5 hrs	6 hrs	7 hrs
3 hrs	2^{-16}	-	-	-	-	-
4 hrs	2^{-16}	1.9^{-5}	-	-	-	-
5 hrs	2^{-16}	2^{-16}	2^{-16}	-	-	-
6 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	-	-
7 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}	-
8 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}	0.13

C.

Time (hours)	Mean width (µm)	Standard Deviation	Median width (µm)	IQR	Max width (µm)	Min width (µm)
2	0.614	0.161	0.600	0.248	1.27	0.141
3	0.673	0.120	0.661	0.144	1.38	0.197
4	0.643	0.149	0.640	0.175	1.90	0.167
5	0.639	0.141	0.663	0.163	1.31	0.142
6	0.704	0.125	0.709	0.148	1.21	0.121
7	0.631	0.120	0.624	0.149	2.48	0.221
8	0.584	0.108	0.580	0.126	1.60	0.132

D.

	2 hrs	3 hrs	4 hrs	5 hrs	6 hrs	7 hrs
3 hrs	2^{-16}	-	-	-	-	-
4 hrs	2^{-16}	2^{-16}	-	-	-	-
5 hrs	2^{-16}	2^{-16}	0.00056	-	-	-
6 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	-	-
7 hrs	2^{-16}	2^{-16}	3.7^{-12}	2^{-16}	2^{-16}	-
8 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}

E.

Time (hours)	Mean length (µm)	Standard Deviation	Median length (µm)	IQR	Max length (µm)	Min length (µm)
2	2.18	0.817	2.07	0.995	13.8	0.08
3	2.47	1.04	2.23	1.18	17.9	0.07
4	2.24	2.53	1.73	2.06	33.2	0.04
5	2.69	2.65	2.36	3.10	31.3	0.06
6	4.00	4.07	3.31	4.62	38.7	0.06
7	2.84	1.73	2.70	1.81	20.6	0.07
8	2.71	3.03	2.06	2.62	31.9	0.4

F.

	2 hrs	3 hrs	4 hrs	5 hrs	6 hrs	7 hrs
3 hrs	2^{-16}	-	-	-	-	-
4 hrs	2^{-16}	2^{-16}	-	-	-	-
5 hrs	3^{-11}	2.7^{-15}	2^{-16}	-	-	-
6 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	-	-
7 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}	-
8 hrs	3^{-14}	2^{-16}	2^{-16}	5.1^{-8}	2^{-16}	2^{-16}

G.

Time (hours)	Mean width (µm)	Standard Deviation	Median width (µm)	IQR	Max width (µm)	Min width (µm)
2	0.641	0.172	0.649	0.247	1.28	0
3	0.657	0.146	0.656	0.147	1.84	0
4	0.509	0.257	0.583	0.341	1.54	0
5	0.522	0.289	0.637	0.508	2.53	0
6	0.526	0.296	0.646	0.548	1.98	0
7	0.565	0.172	0.581	0.191	2.48	0
8	0.466	0.216	0.529	0.248	1.60	0

H.

	2 hrs	3 hrs	4 hrs	5 hrs	6 hrs	7 hrs
3 hrs	2^{-16}	-	-	-	-	-
4 hrs	2^{-16}	2^{-16}	-	-	-	-
5 hrs	2^{-16}	2^{-16}	2^{-16}	-	-	-
6 hrs	2^{-16}	2^{-16}	2^{-16}	0.0042	-	-
7 hrs	2^{-16}	2^{-16}	2.2^{-13}	3.1^{-13}	2^{-16}	-
8 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}

Figure S2.4. P. mirabilis BB2000 cell length and width metrics. Images from the time course were analyzed using the Swarmetrics pipeline. The summary statistics are provided for the cell length (A, E) and width (C, G). Pairwise Wilcoxon Rank Sum Tests were performed, comparing each time point for length (B, F) and width (D, G). A – D are with the filter of 1 µm minimum in cell length. E – G are all identified objects (i.e., without the filter), which includes debris. All time points were statistically significantly different (< 0.05) from the others, except for length at hours 7 and 8. For filtered data (A – D), cell counts are as follows, sequentially from hours 2 to 8: 16,723; 20,726; 5,506; 9,367; 7,478; 10,865; 9,156. For all particles without a filter (E – H), sample sizes are as follows, sequentially from hours 2 to 6: 25,559; 22,245; 11,445; 15,145; 10,795; 14,878; 17,368.

A.

Time (hours)	Mean length (µm)	Standard Deviation	Median length (µm)	IQR	Max length (µm)	Min length (µm)
2	3.12	0.946	2.94	1.14	18.6	1.38
3	14.0	4.89	13.9	7.37	31.7	1.00
4	4.55	1.87	4.13	1.85	29.6	1.00
5	4.98	1.22	3.76	1.49	28.2	1.72
6	8.15	6.21	4.99	9.89	42.9	1.00

B.

	2 hrs	3 hrs	4 hrs	5 hrs
3 hrs	2^{-16}	-	-	-
4 hrs	2^{-16}	2^{-16}	-	-
5 hrs	2^{-16}	2^{-16}	2^{-16}	-
6 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}

C.

Time (hours)	Mean width (µm)	Standard Deviation	Median length (µm)	IQR	Max length (µm)	Min width (µm)
2	0.976	0.102	0.986	0.123	1.43	0.490
3	0.777	0.098	0.774	0.144	1.10	0.365
4	0.925	0.097	0.923	0.119	1.34	0.481
5	1.02	0.084	1.02	0.093	1.67	0.482
6	8.22	0.115	0.815	0.163	1.32	0.123

D.

	2 hrs	3 hrs	4 hrs	5 hrs
3 hrs	2^{-16}	-	-	-
4 hrs	2^{-16}	2^{-16}	-	-
5 hrs	2^{-16}	2^{-16}	2^{-16}	-
6 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}

E.

Time (hours)	Mean length (µm)	Standard Deviation	Median length (µm)	IQR	Max length (µm)	Min length (µm)
2	3.01	1.08	2.89	1.16	18.6	0.05
3	3.52	6.86	0.458	8.95	31.7	0.04
4	3.96	2.27	3.87	2.01	29.6	0.05
5	3.78	1.45	3.68	1.54	28.2	0.04
6	4.06	5.82	0.780	4.43	42.9	0.03

F.

	2 hrs	3 hrs	4 hrs	5 hrs
3 hrs	2^{-16}	-	-	-
4 hrs	2^{-16}	2^{-16}	-	-
5 hrs	2^{-16}	2^{-16}	2^{-16}	-
6 hrs	2^{-16}	5.5^{-8}	2^{-16}	2^{-16}

G.

Time (hours)	Mean width (µm)	Standard Deviation	Median width (µm)	IQR	Max width (µm)	Min width (µm)
2	0.941	0.197	0.981	0.136	1.43	0
3	0.322	0.323	0.163	0.611	1.10	0
4	0.813	0.297	0.905	0.148	1.34	0
5	0.973	0.224	1.02	0.010	1.67	0
6	0.449	0.373	0.301	0.717	1.32	0

H.

	2 hrs	3 hrs	4 hrs	5 hrs
3 hrs	2^{-16}	-	-	-
4 hrs	2^{-16}	2^{-16}	-	-
5 hrs	2^{-16}	2^{-16}	2^{-16}	-
6 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}

Figure S2.5 P. mirabilis ATCC 29906 cell length and width metrics show statistically significant differences. Images from the time course were analyzed using the Swarmetrics pipeline. These are the summary statistics for cell length (A, E) and width (C, G). Also shown are results of pairwise Wilcoxon Rank Sum Tests, comparing time points for length (B, F) and width (D, G). A – D are with the filter of 1 µm minimum in cell length. E – G are without the filter, which includes debris. All time points were statistically significantly different (< 0.05) from the others. For filtered data (A – D), sample sizes are as follows, sequentially from hours 2 to 6: 7,740; 942; 10,142; 15,819; 2,836. For all particles without a filter (E – H), sample sizes are as follows, sequentially from hours 2 to 6: 8,073; 3,083; 11,766; 16,852; 5,925.

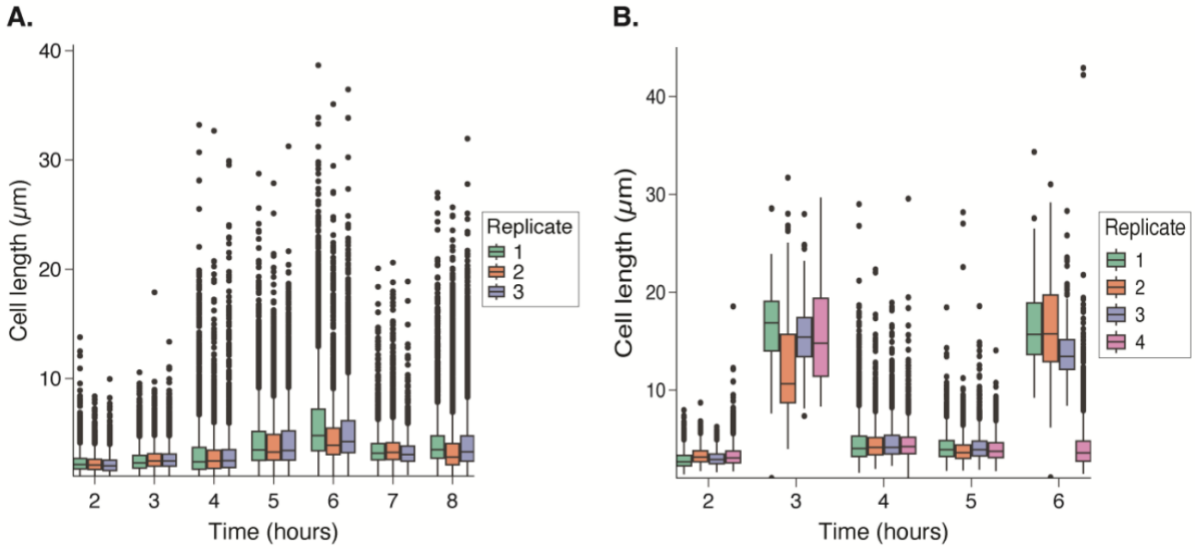


Figure S2.6. Data from independent biological replicates for the swarm development time course. Plotted are the resultant box-and-whisker plots using the Swarmetrics pipeline for each experiment. We quantified the lengths of cells at the expanding edge of the swarm. For the box-and-whisker plots, each box represents the 1 st – 3 rd quartile, and the middle line is the median. Whiskers indicate minimum and maximum; single dots are outliers. (A) BB2000 and (B) ATCC 29906. We used four replicates of ATCC 29906 to attain a minimal dataset of 30,000 cells and at least 750 cells per hour.

A.

Time (hours)	Count	Mean length (μm)	Standard Deviation	Median length (μm)	IQR	Max length (μm)	Min length (μm)
2	341	5.06	1.33	4.63	1.18	13.8	4
3	1568	4.93	1.09	4.55	1.02	17.9	4
4	1079	7.61	4.43	5.87	4.18	33.2	4
5	3576	7.04	3.20	5.91	3.74	31.3	4
6	4216	7.56	4.30	5.93	3.74	38.7	4
7	2826	5.32	1.87	4.72	1.15	20.6	4
8	2989	7.62	4.13	5.76	5.08	31.9	4

B.

	2 hrs	3 hrs	4 hrs	5 hrs	6 hrs	7 hrs
3 hrs	2^{-16}	-	-	-	-	-
4 hrs	1.9^{-6}	2^{-16}	-	-	-	-
5 hrs	5.8^{-10}	0.0801	3.1^{-16}	-	-	-
6 hrs	0.0015	7.4^{-7}	0.0801	5.6^{-6}	-	-
7 hrs	2^{-16}	2^{-16}	2^{-16}	4.2^{-13}	1.1^{-15}	-
8 hrs	0.527	7.4^{-12}	0.3742	2.1^{-7}	0.0801	2^{-16}

C.

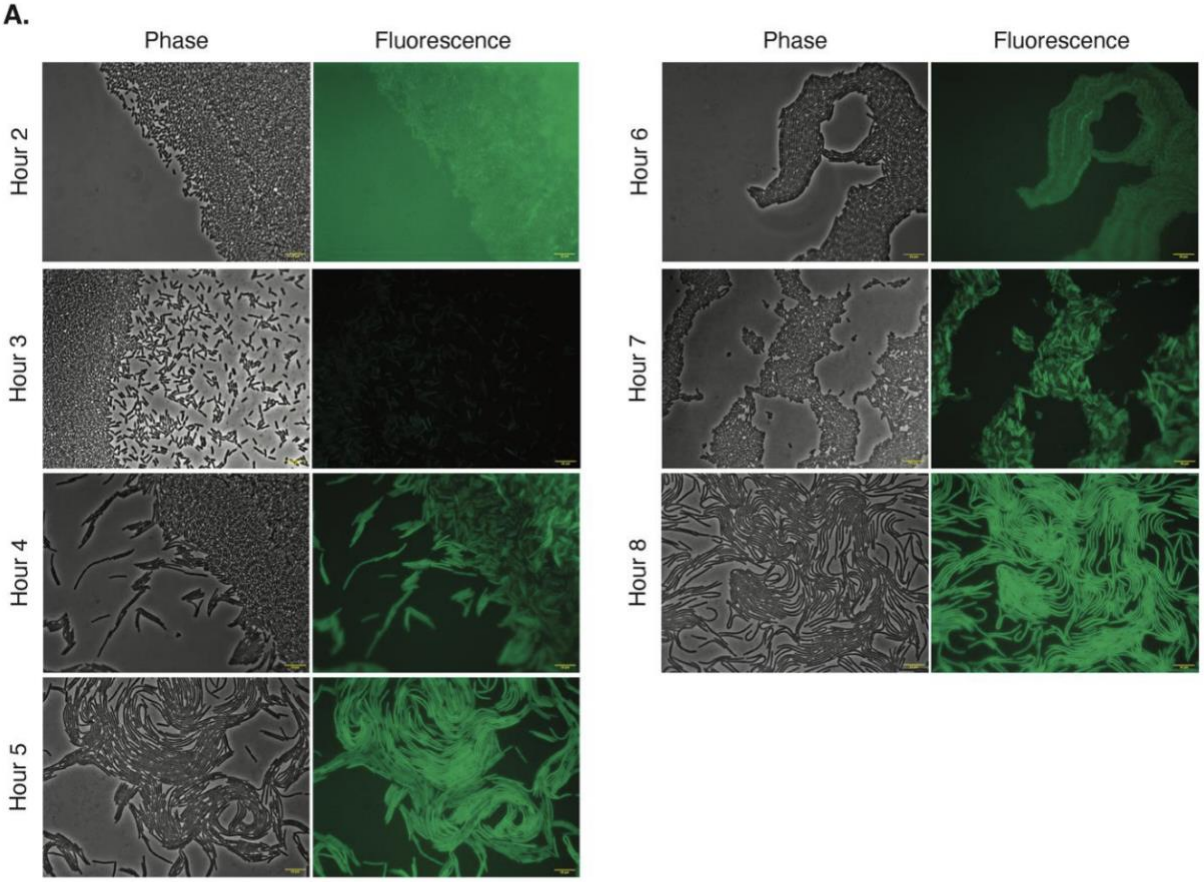
Time (hours)	Count	Mean length (μm)	Standard Deviation	Median length (μm)	IQR	Max length (μm)	Min length (μm)
2	16382	2.15	0.645	2.05	0.923	4	1.07
3	19158	2.39	0.677	2.31	0.981	4	1.07
4	4427	2.25	0.735	2.15	1.11	4	1.07
5	5791	2.71	0.677	2.70	1.01	4	1.07
6	3262	3.05	0.587	3.09	0.824	4	1.07
7	8039	2.86	0.603	2.83	0.954	4	1.07
8	6167	2.63	0.730	2.64	1.10	4	1.07

D.

	2 hrs	3 hrs	4 hrs	5 hrs	6 hrs	7 hrs
3 hrs	2^{-16}	-	-	-	-	-
4 hrs	7.6^{-12}	2^{-16}	-	-	-	-
5 hrs	2^{-16}	2^{-16}	2^{-16}	-	-	-
6 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	-	-
7 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}	-
8 hrs	2^{-16}	2^{-16}	2^{-16}	3.07^{-7}	2^{-16}	2^{-16}

Figure S2.7. Short and long P. mirabilis cells show differences over the duration of one swarm development cycle. The cell length dataset was subdivided into cells longer than 4 μm (A, B) or those 4 μm and shorter (C, D). Shown are summary statistics for cells > 4 μm (A) and cells 4 μm and shorter (C). Pairwise Wilcoxon Rank Sum Tests were performed that compared each time point for cells > 4 μm (B) and cells shorter (D).

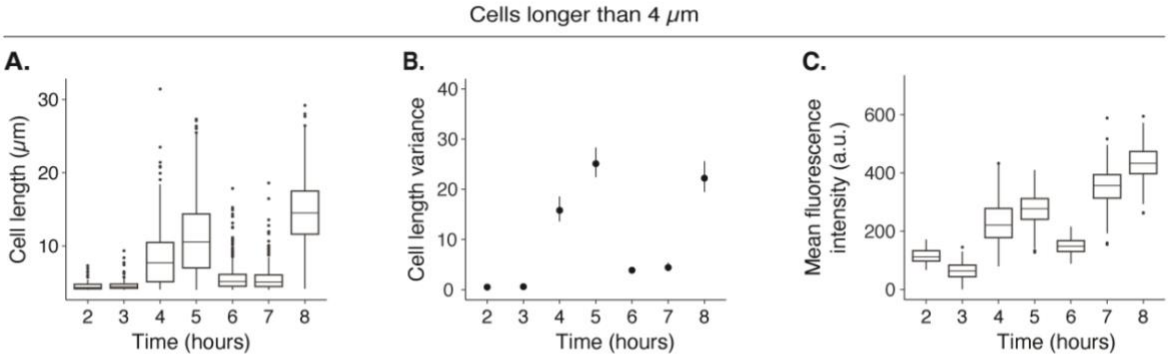
A.



B.

Time (hours)	Mean length (μm)	Standard Deviation	Median length (μm)	IQR	Max length (μm)	Min length (μm)
2	2.68	0.686	2.56	0.846	7.30	1.35
3	2.94	0.833	2.78	1.07	9.33	1.50
4	6.36	4.21	4.91	5.52	31.5	1.43
5	10.9	5.13	10.3	7.55	27.3	1.52
6	4.49	1.87	4.00	1.77	17.9	1.26
7	2.83	1.47	2.41	1.14	18.6	1.39
8	13.8	5.35	14.2	6.50	29.2	1.54

Figure S2.8. Heterogeneity persists during the swarm development cycle with *P. mirabilis* expressing a fliA-venus reporter. (A) These are the uncropped phase contrast (phase) and false-color (fluorescence) micrographs of those in Figure 4. Scale bar, 10 μm. (B) Summary statistics for cell length. Sample sizes are as follows, sequentially from hours 2 to 8: 1,603; 1,454; 506; 590; 864; 1,681; 437.

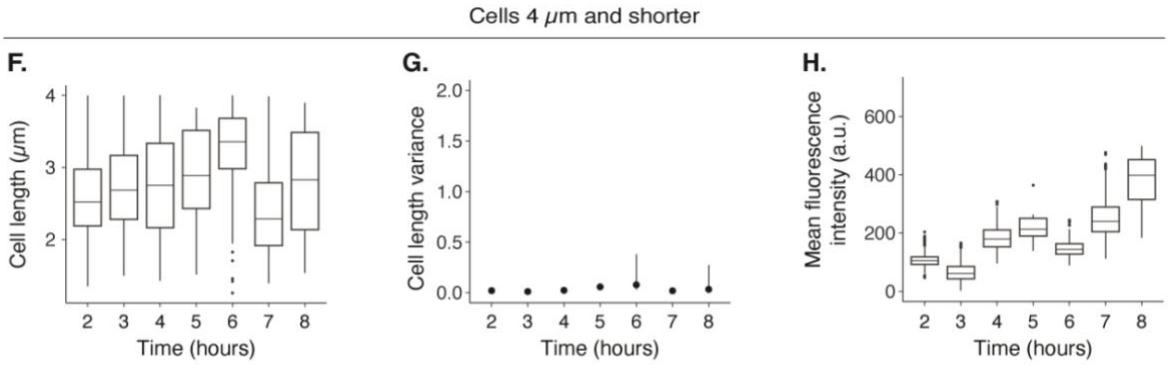


D.

Time (hours)	Count	Mean length (μm)	Standard Deviation	Median length (μm)	IQR	Max length (μm)	Min length (μm)
2	66	4.59	0.725	4.30	0.631	7.30	4.02
3	152	4.65	0.782	4.41	0.599	9.33	4.00
4	319	8.47	3.98	7.69	5.35	31.5	4.03
5	573	11.2	5.01	10.6	7.40	27.3	4.02
6	422	5.70	1.97	5.14	1.59	17.9	4.00
7	222	5.71	2.10	5.07	1.61	18.6	4.00
8	411	14.5	4.71	14.5	5.92	29.2	4.16

E.

	2 hrs	3 hrs	4 hrs	5 hrs	6 hrs	7 hrs
3 hrs	0.38	-	-	-	-	-
4 hrs	2^{-16}	2^{-16}	-	-	-	-
5 hrs	2^{-16}	2^{-16}	2.3^{-16}	-	-	-
6 hrs	1.0^{-10}	3.1^{-16}	2^{-16}	2^{-16}	-	-
7 hrs	1.8^{-7}	1.9^{-10}	2^{-16}	2^{-16}	0.53	-
8 hrs	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}	2^{-16}



I.

Time (hours)	Count	Mean length (μm)	Standard Deviation	Median length (μm)	IQR	Max length (μm)	Min length (μm)
2	1537	2.60	0.553	2.52	0.787	4.00	1.35
3	1302	2.74	0.569	2.68	0.889	3.99	1.50
4	187	2.76	0.712	2.75	1.17	4.00	1.43
5	17	2.88	0.728	2.89	1.08	3.83	1.52
6	424	3.29	0.486	3.36	0.699	4.00	1.26
7	1459	2.39	0.594	2.29	0.871	3.98	1.39
8	26	2.81	0.757	2.83	1.35	3.90	1.54

J.

	2 hrs	3 hrs	4 hrs	5 hrs	6 hrs	7 hrs
3 hrs	3.4^{-9}	-	-	-	-	-
4 hrs	0.036	1.00	-	-	-	-
5 hrs	0.474	1.00	1.00	-	-	-
6 hrs	2^{-16}	2^{-16}	2^{-16}	0.156	-	-
7 hrs	2^{-16}	2^{-16}	2.4^{-10}	0.049	2^{-16}	-
8 hrs	0.910	1.00	1.00	1.00	0.027	0.049

Figure S2.9 Analysis of short and long populations in the fliA-venus reporter strain. The dataset was divided into two groups based on cell length: cells > 4 μm (A – E) and 4 μm or shorter (F – J). Presented are the length distribution (A, F), variance with the vertical bars showing 95% confidence intervals (B, G), mean fluorescence intensity per cell per time point (C, H), summary statistics (D, I), and statistical analysis using Wilcoxon Rank Sum Tests (E, J). For the box-and-whisker plots, each box represents the 1st – 3 rd quartile, and the middle line is the median. Whiskers indicate minimum and maximum; single dots are outliers.

Chapter 3: Contact-Dependent Communication Shapes the Mesoscale Spatial Organization of Bacterial Swarms

This work is a manuscript in preparation

E. E. Garling[1,*], M. Byers[2,*], E. Bradley[2,3], J.D. Meiss[4], and K. A. Gibbs[1,3]

[1]Department of Plant and Microbial Biology, University of California, Berkeley
[2]Department of Computer Science, University of Colorado Boulder
[3]The Santa Fe Institute
[4]Department of Applied Mathematics, University of Colorado Boulder
* Co-first authors

II. Chapter 3 Preface

Given that much of the identity signaling pathway (*ids)* had been characterized painstakingly so by those before me and having discovered immense cell length heterogeneity present in WT BB2000 *P. mirabilis* swarms, I turned to examining mutant *P. mirabilis* swarming behaviors. Through 8-hour time courses of BB2000 cells lacking the ability to transmit identity signals while swarming (*T6S*$^{-}$), and strains lacking identity signaling altogether *(*Δ*ids*), one thing stood out in particular: the tight, coordinated rafts of cells I'd grown accustomed to seeing were not entirely absent, but certainly significantly reduced in the identity mutants. At one point, the best phrase I could use to describe what I was seeing by eye was: "think of pick-up sticks." Instead of clear, coordinated organization, I saw cells moving off on their own and even railroading, right on top of each other. While "pick-up sticks" renders a clear image in the minds of most, it was insufficient to capture the nuance necessary to describe how these mutants differed in their collective migration capabilities and organization compared to WT.

Dr. Gibbs and I were, and still are, very fortunate for the time, energy, and immense curiosity of Dr. Elizabeth Bradley, Dr. James Meiss, and Morgan Byers from CU Boulder. Together, we sought out to quantitatively describe the differences observed under the microscope between the two identity mutants versus wildtype. And, through a substantial amount of inspiration, collaboration, and teamwork, we describe a mathematical approach to understand the relatedness and interconnectedness of *P. mirabilis* swarming cells. Serendipity was certainly at play when we joined forces. And, as they say, "do you want to write a paper?" is the science equivalent of "we should start a band."

3.1 Abstract

Scientists use microscopy images to visualize and describe organisms and their social interactions. Microscale visual inspections of individual cells, paired with genetic approaches, have revealed key developmental steps for cells within microbial communities such as microbial swarms, biofilms, and fruiting bodies. However, there is a lack of formal, quantitative descriptions of these structures. This gap limits understanding of the relevance and behaviors of cell clusters, especially regarding cell development and cell-cell interactions. Here, we develop biologically informed mathematical methods to formally assess cell groupings that form in bacterial colonies under varying conditions. Our approaches to analyzing developmental stages of swarming, a type of collective migration, focus on cell length (as a proxy for developmental stage), cell-cell contacts, and cell groupings. These allow us to explore how identity signaling and cell-to-cell contact affect population structure at different scales. We found that contact-dependent, cell-cell communication and kin recognition govern population architecture during collective migration, such as swarming. By integrating microbiology, applied mathematics, and computer science, we open a new frontier for analyzing experimental data, leading to testable biological insights.

3.2 Introduction

Community structure relies on interactions between individuals whether animals, fungi, or bacteria. Kin recognition is one modifier of these interaction (105,128,129). In microbes, this structure is critical for colonization in diseases, fitness in mixed populations, and metabolism. Since Antonie van Leeuwenhoek, microbiologists have used increasingly higher magnifications to examine the activities of individual microorganisms, i.e., the microscale. However, a critical gap blocks forward progress in understanding individual cells in the context of microbial community structures: metric analysis of groups of cells, i.e., structure at the *mesoscale*. Closing this gap would allow one to answer longstanding questions—such as the role of cell-cell communication or kin identity in community structure and maintenance—which remain unanswered in multiple microbial systems.

This barrier has blocked progress on mesoscale analysis. Traditional techniques for image analysis often struggle to quantitatively assess cell connectedness and the spatial organization and architecture of bacterial communities (105,130,131). In addition, human visual bias can distort the interpretation of microscopy images (132). Recent studies have overcome some of these challenges in the context of populations with minimal to no motility (133–135), populations with cells that are uniform in shape (136), or subsets of a population (137,138). However, many microbial populations contain cells of varying lengths and shapes, that are motile, and that function collectively. To the best of our knowledge, universally applicable mathematical tools for mesoscale analysis of complex microbial communities have not yet been developed.

We move beyond these limitations using a combination of formal mathematical methods and targeted experiments. In Section (B), we discuss the opportunistic pathogenic bacterium,

Proteus mirabilis, studied in this paper, which utilizes contact-dependent kin identity as a regulator of collective migration (139). Using a genetics approach, combined with cell length analysis and novel mathematical analyses that focus on cell adjacency and the topology of cell groupings, we address the longstanding question: how does kin recognition shape the population's architecture? *P. mirabilis* cells move collectively as a group, also called a "swarm," on hard surfaces (140,141). This migration is dependent on a developmental cycle in which cells oscillate between being short (< 4 μm) and nearly uniformly rod-shaped to becoming elongated and snake-like (up to 40+ μm long) (7,92,142–144). For wild type populations, these distinct life stages can be reliably measured through quantitative changes in cell length distributions within a heterogeneous population (144). Kin recognition is active during this swarm development cycle: the current model is that kin identity acts either before or after collective migration to block access to colony benefits, such as shared molecules, nutrients or predator protection, by non-kin cells (28).

To the eye, microscopy images of swarms of non-kin cells appear to have structures that are distinct from the swarms of kin (19,25), but quantitative analysis to verify this distinction has been previously unavailable. Mathematical approaches from the field of computational geometry offer a promising path to formalize such qualitative observations, thereby delivering meaningful, quantifiable descriptions of the structures in complex images of polymorphic cell populations (see e.g., Figure 2.1A). In Section (C), we introduce mathematical tools that provide methods for quantifying the degree of cell-cell contact and the shapes of cell groupings in *P. mirabilis* colonies. Together, these methods provide an interaction network that quantifies the potential for communication among adjacent cells. We propose two methods to summarize the spatial inhomogeneity of contacts and to allow us to separate components of interacting cells. With these, we show that cell-cell communication and kin identity alter how neighbors interact and influence the structures that form in swarms. Our methods differ from existing order parameters used in the study of biological aggregation (e.g., packing fraction (145) or average nearest-neighbor distance (146)); indeed, the generality of the mathematical tools we present allows us to account for a collection of differently shaped individuals, and our results reveal patterns—and biological variability—across experiments. Our goal is to formally characterize *P. mirabilis'* population-level complexity, self-organization, immense morphological changes, collective dispersion, and coalescence.

3.3 Biological Methods

3.3.1 Strains and Media

The wild type *P. mirabilis* strain is BB2000 (123), from which two mutant strains are derived: Δ*ids* [*P. mirabilis BB2000 Δids::Tn-Cm(R)* (17)] and T6S^{-} [*P. mirabilis* BB2000 *tssA*$_{T95G}$, which was constructed and confirmed as described previously for CCS05 using the pKNG101-derived vector, pCS34 (25)]. *P. mirabilis* is cryogenically stored at -80°C and daily maintained on LSW^{-} agar at 4°C (147). For propagation, cells are grown overnight aerobically in LB broth (Lennox) at 37°C. CM55 blood agar base (Thermo Fisher Scientific, Waltham, MA) is the media for swarming.

3.3.2 Microscopy

Overnight cultures are diluted and normalized in LB broth (Lennox) to an optical density at 600 nm (OD_{600}) of 1.0 and then inoculated onto a CM55 agar pad on a microscope slide as previously described (17). Microscopy is performed as previously described (25). Briefly, the leading edge of the swarms and inoculum are imaged using a Leica DM5500B (Leica Microsystems, Buffalo Grove, IL) and a CoolSnap HQ2 cooled charge-coupled-device (CCD) camera (Photometrics, Tucson, AZ). For each biological replicate, a new slide is used for each time point and biological replicate; all are inoculated at the same time in parallel. MetaMorph version 7.8.0.0 (Molecular Devices, Sunnyvale, CA) is used for image acquisition.

3.3.3 Image Analysis

To avoid obfuscating downstream analysis, images are first cropped, if necessary, to remove layered cells, oversaturated regions, or areas of excess debris. The perimeter of the image is excluded, since it often includes cells that are partially out of the frame. Cell masks and length values are acquired using the Swarmetrics pipeline, which is described in Garling et al. (144). In the current Swarmetics pipeline, approximately 10% of cells in each microscopy image are missed (144), so the δ-adjacency graphs (see Section C) are artificially sparse—particularly at the edges of the image, where neighboring cells are cut off. This rate of detection is comparable to other bacterial detection algorithms (105). The resulting data is exported as .cvs files for the analysis described in Section C. Data availability reported below.

3.4 Results and Mathematical Models

Kin recognition through direct cell-to-cell communication appears to alter cell development—specifically, morphology or behaviors. For example, disruptions in cell-cell communication of self-identity reduce the width of *P. mirabilis* swarm colonies by as much as 80%, specifically through the transfer of Ids self-identity proteins via a contact-dependent T6SS (26,148). Comparative analyses of individual cells have not yet been pursued, though by eye, the communication-disrupted mutant strains appear to have different lengths and form atypical cell groupings (Figure 3.1A). In this study, we analyze three phases of the *P. mirabilis* swarm development cycle using phase contrast light microscopy. Using multiple independent isolates, we consider the potential role of cell-cell communication of kin identity in swarm structure. Figure 1A shows representative images of wild type strain BB2000 (WT) and its derived mutant strains: BB2000 Δ*ids (*Δ*ids*), which is completely disrupted in the communication of its non-lethal Ids self-identity but not in all communication (17), and BB2000 $vipA_{T95G}$ (T6S-), which completely lacks T6SS-dependent cell-cell communication (32, 33). These strains are shown at three sequential stages of swarm development: swarm initiation (entry in movement, hour 5), collective swarming (fastest movement, hour 6), and consolidation (cells stop moving and some divide, hour 7) (92). In the microscopy images, the structure of the groupings in the imaged populations exhibit interesting differences between strains and across time.

For comparative analysis across the three strains, we first need to assess whether each strain is at the same swarm development stage, i.e., progressing similarly to WT. This analysis also

examines how communication and kin identity impact cell morphology. We measure each cell's length using Swarmetrics, an image capture and analysis pipeline (144), resulting in a final dataset of 55,000 cells over all samples (Supplemental Table S3.1). For WT, both median cell lengths and their variances rise at hour 6 and then fall to their lowest values at hour 7 (Figures 3.1B, 3.1C), as previously reported (144). These results show that wild type cells begin to elongate during swarm initiation (hour 5), becoming (as a population) longest during collective swarming (hour 6), and shorten during consolidation (hour 7), likely due to cell division. Cell lengths for the Δ*ids* population follow the same pattern, except that the Δ*ids* cells are, on average, longer than wildtype during collective swarming (Figure 3.1B). By contrast, the T6S- population has longer cells than wildtype during both swarm initiation and collective swarming, but the distribution of short cells is similar to those for wildtype and the Δ*ids* strain during consolidation (Figure 3.1B). The variances for cell length within each mutant strain's population follow the same pattern as wildtype, with the greatest values during collective swarming and the lowest during consolidation; cells have the largest spread of cell lengths at hour 6 (Figure 3.1C). These data suggest that these three strains are all undergoing swarm development at roughly equivalent time periods, allowing for comparative analysis across strains.

A specific aspect of the cell-length data discussed above is the ratio of short to long cells, which is a marker of cell development. We previously found that long cells dominate populations of *P. mirabilis* during collective swarming, while short cells are dominant at the swarm initiation and consolidation stages (144). To further explore this issue, we partitioned the cells in the current experiments into two groups, using a 4 μm cutoff—double the length of a non-swarming cell (22). The percentages of the populations in these two groups (Figure 3.1D) show that wildtype and the Δ*ids* strain share a pattern in which the longer cells predominate during collective swarming. By contrast, the T6S- population has essentially the same fraction of long cells during swarm initiation and collective swarming (Figure 3.1D). Short cells dominate during consolidation for all strains, although less so for Δ*ids*. Altogether, this data shows that morphological changes in cells occur when communication is disrupted.

However, a challenge with this type of aggregate analysis of the cell lengths is that the data are spatially decontextualized. Figure 3.1D tells us about the overall size distributions but does not capture the geometry of the cell groupings that appear in the micrographs of Figure 3.1A. *In other words, we are missing some critical information about the mesoscale.* Formal, geometric assessment of this kind of structural information, we hypothesize, is key to analyzing and understanding the differences between the temporal evolution of the colony structure of these strains.

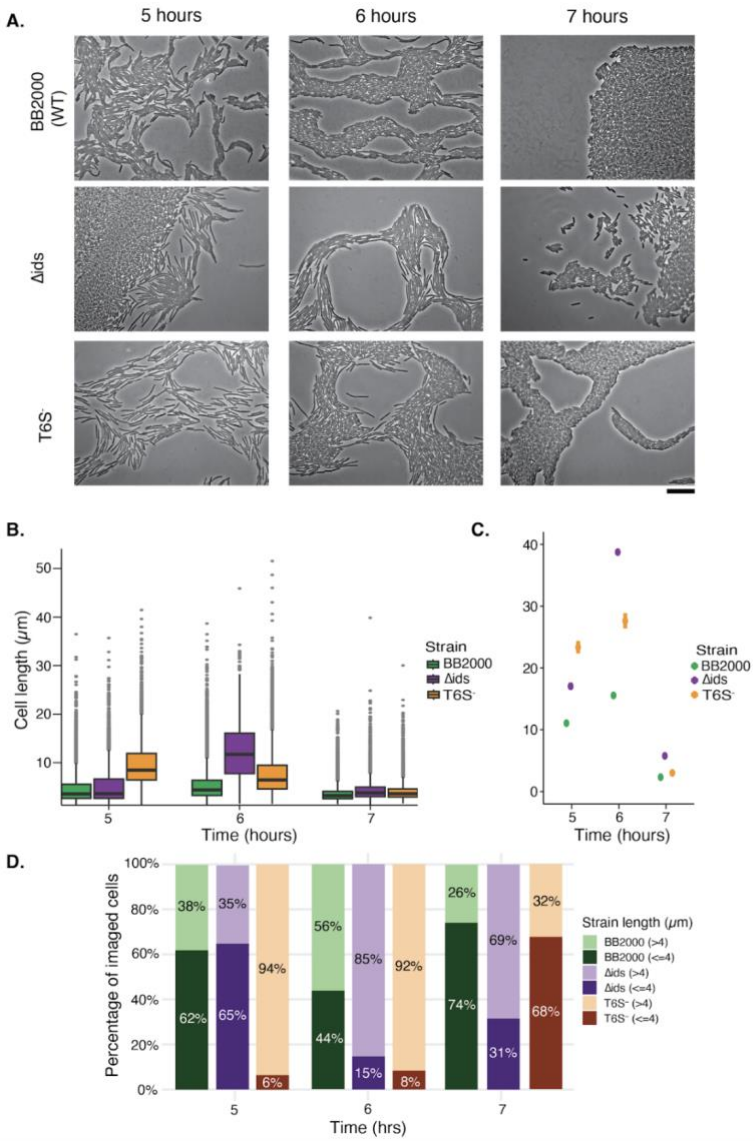
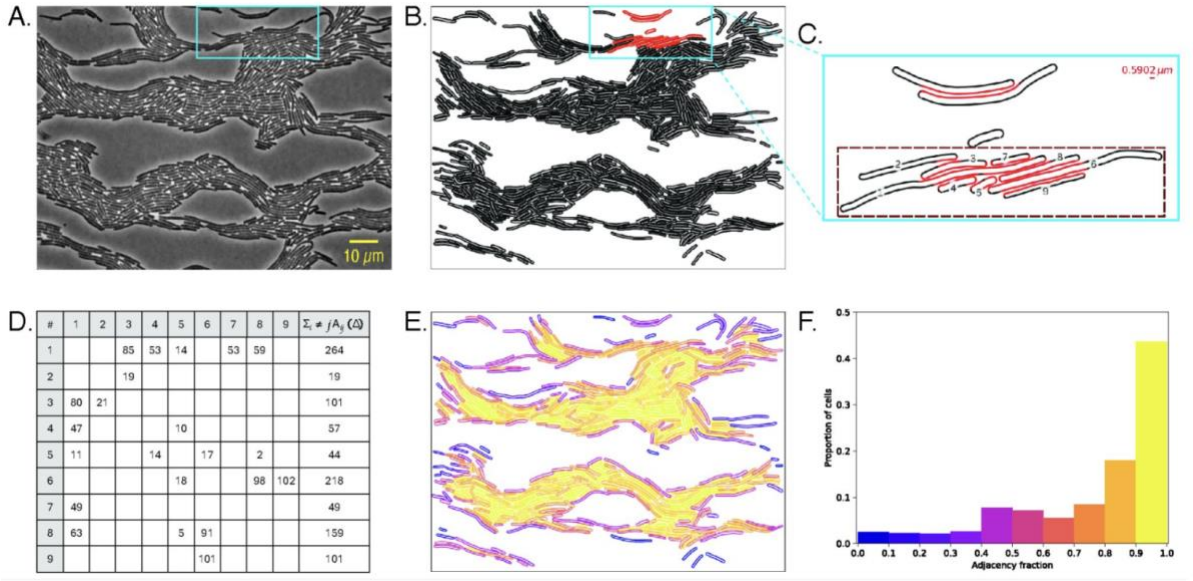


Figure 3.1 Disruptions in intercellular communication alter cell physiology and population structures. (A) Representative phase contrast images, at three timepoints, of swarms from *P. mirabilis* strain BB2000 (wild type) and two derived mutant strains with a disruptions in non-lethal identity signaling, BB2000 Δ*ids* (Δ*ids*), and in the cell-cell communication machinery, BB2000 $vipA_{T95G}$ (T6S-): swarm initiation at five hours, collective swarming at six hours, and consolidation at seven hours. Scale bar: 10 μm. (B) *P. mirabilis* cell lengths, computed using Swarmetrics, for the populations as labelled. Only cells present in a single layer were analyzed. Each graph is a boxplot where the box represents the 1st-3rd quartile and the middle line is the median. Single points above and below the box show the points outside these quartiles. A portion of the BB2000 dataset was previously published (144). (C) Variance of the cell lengths as a swarm development measure. The largest variance occurs during collective swarming; the smallest is during consolidation. (D) The relative distribution of the total populations by hour for cells with lengths > 4 μm (lighter shade) and < 4 μm (darker shade).

To quantify the mesoscale structure that is present in these micrographs (Figure 3.2A) we introduce a measure that we call **δ-adjacency**, which is designed to quantify cell-cell contact, a critical feature in communication. Direct physical contact is needed for the T6SS-dependent transfer of identity proteins like Ids (30, 31), for instance. In addition, efficient swarm migration requires cell-cell contact, as the bacterial flagella are thought to intertwine and allow for greater

force propulsion (7,85). To formally quantify this property, we define the δ-adjacency for each cell in a scatterplot (Figure 3.2B) to be the *fraction of its perimeter that is within a distance δ of some other cell*. The rationale for the choice of δ in this metric stems from image processing. The software in our pipeline creates cell masks that do not account for the bulkiness of intertwined flagella, thereby creating apparent gaps in the image between neighboring cells that are actually in physical contact. The parameter δ accounts for this shortcoming: cells are considered to be adjacent only if they are separated by a gap of at most δ. The value we choose, δ = 0.59 μm, is the average of the distance between pairs of visually observed neighboring cells across several images (Supplement Figure S3.1). We use this value in all the adjacency calculations reported in this paper.

To compute the δ-adjacency value for a cell, we consider each sample point along its perimeter to determine whether any are within δ of another cell. Since there are *N* cells in an image, the result is an *NxN* matrix, $A(\delta)$, the **δ-adjacency matrix**. Here each element, $A_{ij}(\delta)$, is the number of points on the perimeter of cell *i* that are within δ of cell *j*. In this matrix, there is one row and one column for each cell; the rows represent "outgoing" contact (who each cell is touching) and the columns represent "incoming" contact (who is touching each cell). Figure 3.2C demonstrates the results of this computation for the small patch of the scatterplot shown in Figure 2B; the δ-adjacent points are highlighted in red. For example, 47 of the perimeter points of cell 4 are δ-adjacent to cell 1 and 10 to cell 5, so $A_{4,1}(\delta) = 47$ and $A_{4,5}(\delta) = 10$. Figure 2D shows a portion of the δ-adjacency matrix corresponding to the nine cells in the highlighted region of Figure 2C. For the sake of readability, entries in the matrix that are zero (corresponding to pairs that are not δ-adjacent) are not shown. Note that this matrix will typically not be symmetric, since the different cells can have different sizes and different perimeter point spacing. For example, in the snapshot of Figure 4C, while $A_{4,1}(\delta) = 47$, $A_{1,4}(\delta) = 53$; similarly, while $A_{4,5}(\delta) = 10$, $A_{5,4}(\delta) = 14$.



#	1	2	3	4	5	6	7	8	9	$\Sigma_i \neq j A_{ij}(\Delta)$
1			85	53	14		53	59		264
2			19							19
3	80	21								101
4	47				10					57
5	11			14		17		2		44
6					18			98	102	218
7	49									49
8	63				5	91				159
9						101				101

Figure 3.2 Adjacency metric for quantifying cell-to-cell contact. (A) Phase contrast image of WT at 5 hours. (B) Scatter plot of the cells in Panel A. (C) Magnified image of the red-highlighted cells from Panel B. Points on each cell's perimeter that are within δ = 0.59 μm of any other cell are shown in red. The δ parameter allows us to identify contact in the face of imaging artifacts. (D) The δ-adjacency values are stored in a matrix whose i,j^{th} entry represents

the number of perimeter points of cell *i* that are δ-adjacent to cell *j*. The i^{th} row sum of this matrix represents the total number of perimeter points of cell *i* that are δ-adjacent to *any* other cell in the image. (E) Heatmap of the δ-adjacency fractions for the scatter plot in Panel B. (F) Histogram of adjacency fractions for panel E.

We can find each cell's *aggregate* δ-adjacency by computing the row sums of the δ-adjacency matrix: $\Sigma_{j=1\ldots n} A_{ij}(\delta)$ — that is, the total number of each cell's perimeter points that are within δ of *any* other cell. To allow for comparison between cells with different lengths, we normalize this quantity, dividing by the number of points along the perimeter of each cell, to give the **δ-adjacency fraction**, which runs from 0 to 1. The heatmap in Figure 3.2E demonstrates the results of this calculation on the scatterplot in Figure 2A, color-coding each cell according to its δ-adjacency fraction. Heatmaps like this can bring out important features of the colony structure that are not immediately apparent from a visual examination. For example, some of the regions in Figure 3.2E that appear equally dense to the eye in Figure 3.2B have a higher δ-adjacency fraction than others, indicating different degrees of cell-cell contact. Since this information is given for each cell, it can be used to reveal differences between regions in the colony—as well as patterns that could relate to important biological properties such as "rafting," where cells are "parallel" along their longest axis (7), which assumes a degree of contact. Such analysis goes beyond a qualitative description of scatter plots and provides an *explicit* measure of a cell's opportunity to communicate and share resources.

We can further aggregate the information in a δ-adjacency heatmap in the form of a histogram by binning the δ-adjacency scores, as shown in Figure 3.2F. In this example, 43.7% of the cells have an adjacency score between 0.9 and 1.0: *i.e.*, 90-100% of their perimeter is adjacent to some other cell (encoded as yellow in Figure 3.2F). While this lumped analysis does lose the geometric information that is present in Figure 3.2E, such high-level information about the overall proportion of cells that are in close contact can be useful in identifying large-scale differences between different phases and strains As an example, the more yellow we see on the histogram, the more opportunity there is for direct cell-to-cell communication across the population in the image, indicating possible benefits from group interactions. Histograms can then be aggregated across images so that consistent signals become visible above the natural variations of biological systems.

The results of applying the δ-adjacency metric to *the P. mirabilis* data sets are shown in Figure 3.3A. These bring out information about the micro- and mesoscale structure that is not apparent from the cell-length data or the microscopy images (see also Supplemental Figures S3.2 - 7). The top row of Figure 3.3A shows that δ-adjacency in the wild type strain increases as the population transitions from swarm initiation (hour 5) to the collective swarming phase (hour 6), where there is visible rafting, but then drops during consolidation (hour 7). This drop was unexpected, as consolidated cells visually appear to be densely packed, when viewed in comparison to hour 6: the δ-adjacency results indicate that visual perception of "dense packing" can be deceptive. The T6S- strain reaches its highest δ-adjacency during consolidation; the previous time points show a "pick-up sticks" organization with less cell-cell contact. For Δ*ids*, there is more variability between the three images, and the structure alternates between similarity to the wild type and T6S- strains. Notably, the δ-adjacency for this strain is much lower at hour 7 (consolidation) than the other strains. Similarities in the heatmaps are present across the dataset, yet the heat maps for the full series of images in the supplement show differences across replicates within a single

time-point, which is likely due to experimental and biological variability. Therefore, additional analyses are needed to deal with this noisy data.

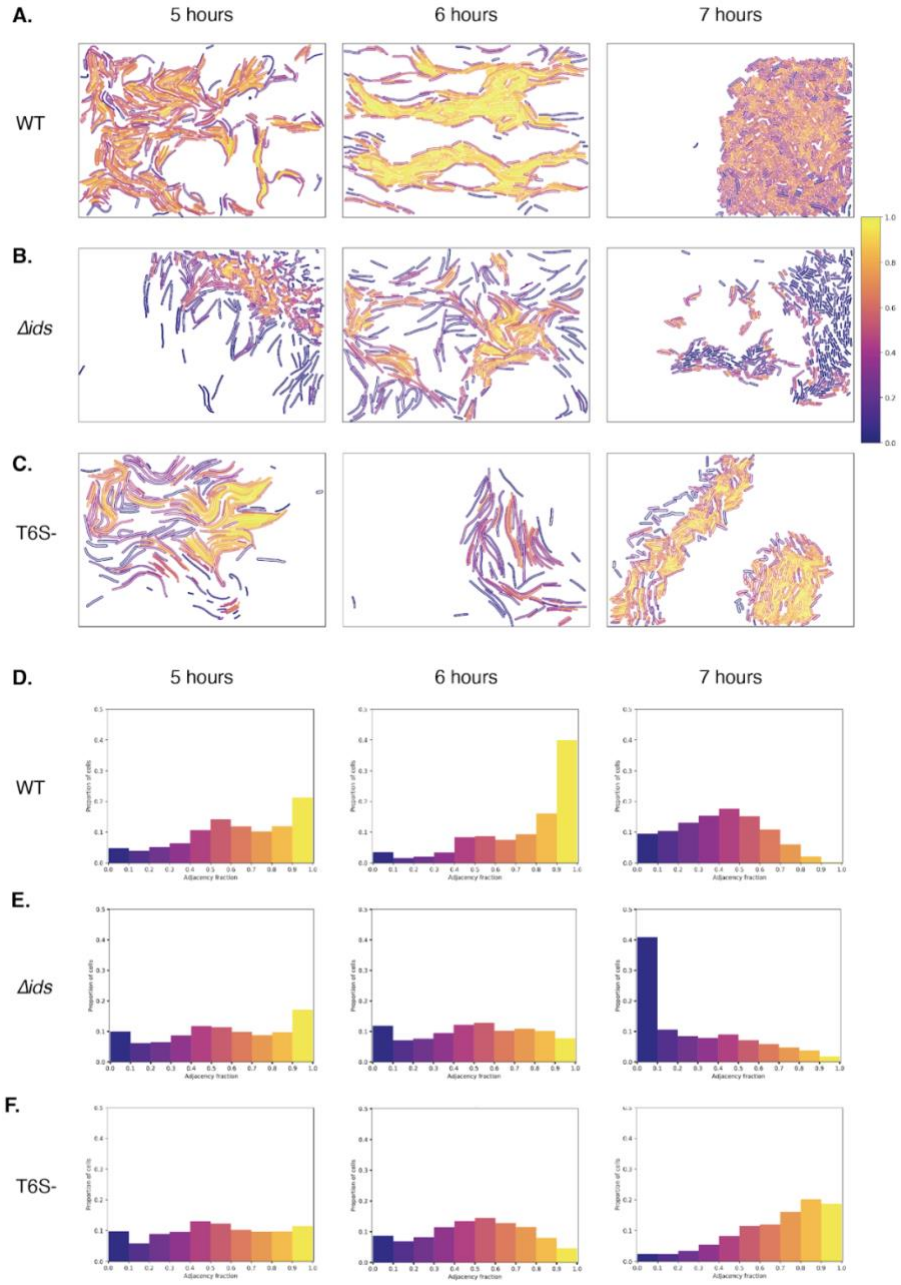


Figure 3.3 Adjacency reveals mesoscale structures. (A-C) Heatmaps of δ-adjacency for sample scatterplots for each strain at each timepoint. (D-F) Histograms of δ-adjacency fraction for all scatterplots for each strain.

Overall, it is striking that the fraction of a cell's body that is in contact with others varies with the phase of the swarm development cycle. In view of the biological differences among these engineered strains, these results show that identity signaling and communication (via T6S$^-$) contribute to the degree of cell-cell contact. Disruptions in cell-cell communication appear to alter the proportional adjacency amounts and timing, as compared to wildtype. Note that the loss of communication (in the T6S$^-$ mutant strain) leads to *less* adjacency during collective swarming

and swarm initiation as compared to wildtype, raising the specific question as to whether cell-cell communication and identity signaling work to help "prime" cells for collective swarming. To begin answering this question, we need a formal quantification of the mesoscale structures in the colony: specifically, of the sizes and shapes of the *groups* of cells that are in contact with one another. Based on the biology, isolated cells have no immediate opportunity for direct, physical (T6S-dependent) cell-cell communication. In contrast, adjacent cells have the potential for T6S-dependent cell-cell communication (149,150). This may even reflect "network" potential for communication, similar to quorum sensing (151,152) but perhaps contact-dependent.

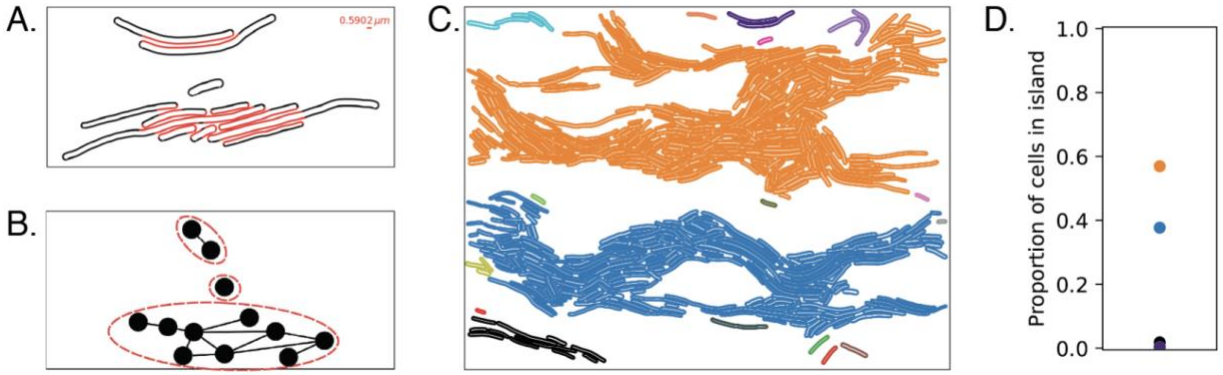


Figure 3.4 Building a graph from the δ-adjacency metric to identify cell groupings. (A) Duplicated from Figure 3.2C, this image shows δ-adjacent points highlighted in red. (B) The δ-adjacency graph for the 12-cells in panel A. Each vertex corresponds to a cell, and each edge connects a pair of cells that are δ-adjacent. A connected component of the graph—a group where all cells are δ-adjacent to at least one other cell in the group —is an *island*. There are two islands in this panel. The isolated cell between the islands is a *singleton*. (C) The island structure of the scatter plot of Figure 2. (D) Island size proportions for panel C, calculated as a fraction of the total number of cells. In addition, there are 11 singletons (not shown) that cumulatively make up 1.68% of the total cell count.

To formalize this, we use the information in the δ-adjacency matrix to identify groups of bacteria that are in contact with one another by building a graph that captures the clustering structure, see Figure 4. This graph contains a vertex for each cell, and a nonzero entry of $A_{ij}(\delta)$ indicates pairwise adjacency between cells *i* and *j* giving an edge between the corresponding vertices. We then analyze the clusters in this graph using the tools of topological data analysis (TDA) (153). In the TDA framework, points in a data set are defined to be *ε-connected* if they are within ε of each other and groups of ε-connected points are called `connected components.' In our context, ε = δ, and we call a connected component an *island*: a set of cells, each of which is δ-adjacent to at least one other cell in the set. For example, the zoomed-in region in Figures 3.4A and 3.4B contains two islands: one containing nine adjacent bacteria and one containing two. The final cell in this region is a *singleton*, an isolated cell that has no opportunity for communication or resource sharing. (In topological data analysis, this would be referred to as a δ-isolated cell: a connected component of size one.) Figure 3.4C shows the islands and singletons in the scatterplot from Figure 3.2B, distinguishing them by color. The largest islands (orange and blue) contain 56.95% and 37.71% of the 655 cells, respectively; the smaller islands, shown in various other colors, each contain between 0.31% and 1.83% of the total. In addition, there are 11 singletons that together comprise 1.68% of the total cell count.

To quantify this information, we make a plot that contains one point for each island, with the vertical position showing the proportion of cells that it contains (Figure 3.4D). We use

proportion—rather than the number of cells—so that it is possible to compare images with varying cell counts. Since these numbers are reported as proportions, they must sum to 1; thus, there can be at most one island with a proportion greater than 50%. We separately tabulate the number of singletons; in a series of experiments, it will be useful to make a box-and-whisker plot of the number of singletons in all of the corresponding frames, as we will show below.

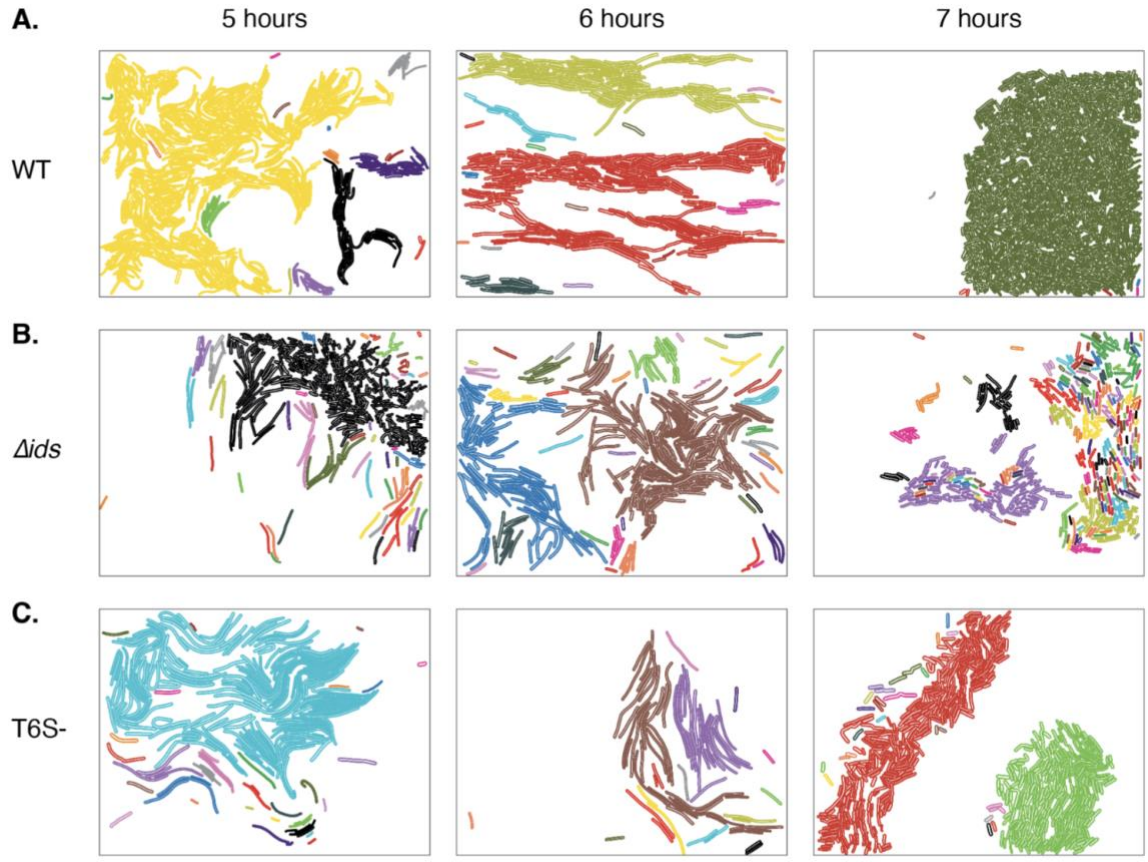


Figure 3.5. Identification of the islands and singletons for the scatterplots of Figure 3A, with each grouping shown in a different color. (A) Wildtype. (B) The Δids strain, which has partial communication. (C) The T6S$^-$ strain, which does not have T6S$^-$ dependent cell-cell communication.

This way of analyzing cell groupings can be useful in elucidating important biological information. Figure 3.5 shows the islands and singletons for the images of Figure 3.5, using color to distinguish groups. The number and size of the islands varies both with time and with strain. At hour 6, all three strains have two predominant islands, but the shapes of each of these larger islands, as well as the distribution of smaller islands and singletons are quite different. The wild type images also contain more singletons at hour 6 than the other timepoints. The compositions of the cell groupings in the Δ*ids* and T6S$^-$ mutant strains are visibly different from each other and from the wildtype strain.

While the images in Figure 3.5 are suggestive, they are not quantitative. To formalize the information in the island analysis, we use the size distributions introduced in Figure 3.4D to reveal and quantify patterns. These, as seen in Figures 3.6A-C, bring out both the patterns and

the biological variability across the image frames for each condition (strain and timepoint). These plots show island size proportions with a dot for each island, separating the image frames horizontally.

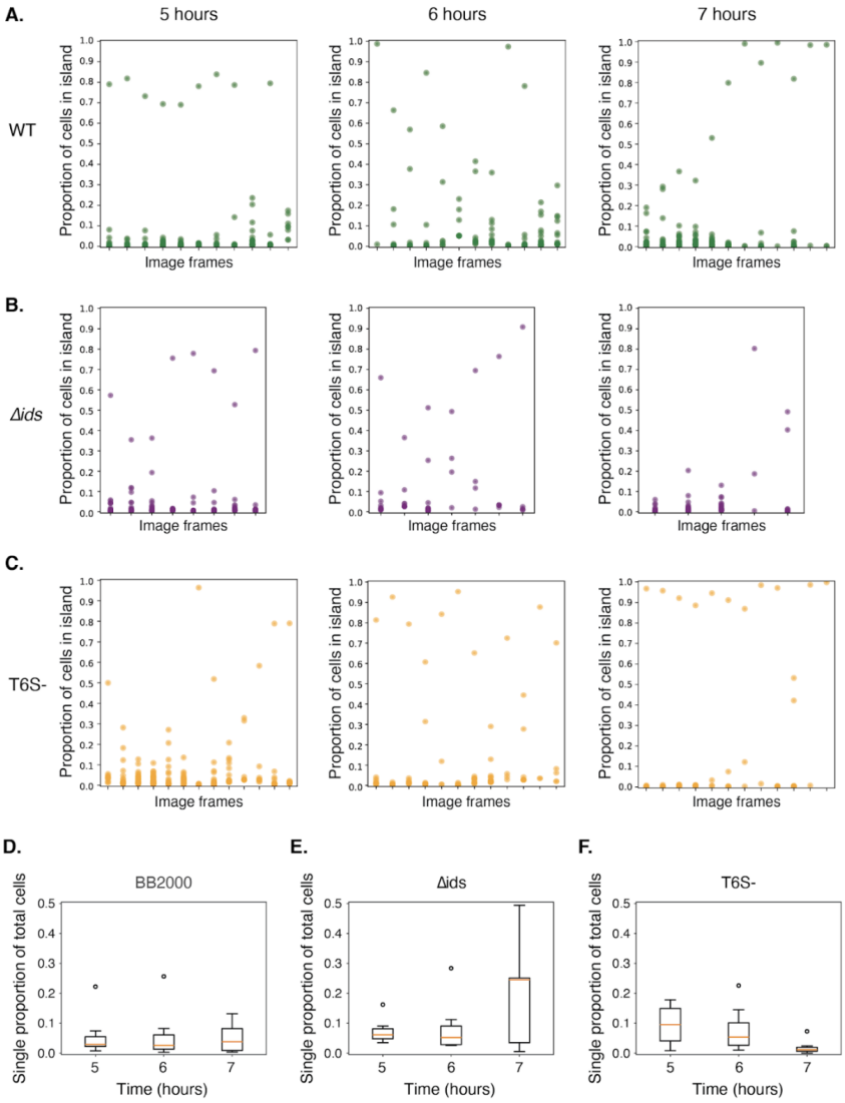


Figure 3.6. Analysis of the island sizes can formally quantify variations in the mesoscale structure between strains and frames at each timepoint. (A-C) The proportion of cells in each image that are contained in each island for each strain, trial, and timepoint. (D) Box-and-whisker plots of the singletons for each of these conditions.

This visualization reveals several interesting aspects of the mesoscale structure within a *P. mirabilis* strain. Island size distributions significantly vary with time for wildtype (Figure 6A) and the T6S- mutant strain (Figure 3.6C). By contrast, the Δ*ids* mutant strain has similar proportions of islands for each timepoint. Specifically, for the wild type population, the frames are dominated by a single large island containing ~70–80% of the cells. For hour 5, there are comparatively fewer small islands (under ~20%), and only two image frames contain only a collection of disconnected small islands (Figure 3.6A). At hour 6, there is no single, dominant island throughout the image frames, suggesting that collective swarming led to the "melting" or

dispersal of these large islands, producing a distribution of island sizes that spans the full range of possibilities. By hour 7, there are two dominant regimes for island structure: a single, very large island containing 80–90% of the cells, with fewer small islands (like hour 5 but with bigger islands) or—for about half the image frames—a spectrum of medium-size islands. This indicates that the mesoscale structure of the colony is not rigid, but that there is a continuum of possibilities.

These results support a hypothesis that contact-dependent communication plays a role in island formation: the island sizes are different when communication is disrupted. The Δ*ids* population shows little difference in island structure at the three timepoints: no single pattern in the distributions among image frames is visible (Figure 3.6B). Note however, that large islands (those containing over ~85% of cells) are mostly absent from the Δ*ids* populations; this suggests that this mutant strain forms smaller connected communities than the wild type strain.

The island structure of the T6S- mutant also shows patterns distinct from those of the wildtype (Figure 3.6C). At hour 5, its dominant distribution is a collection of small- and medium-sized islands, except for perhaps three exceptional image frames. As in wildtype and the Δ*ids* mutant strain, the island sizes vary considerably at hour 6, during collective swarming. However, at hour 7, except for one trial, all have a single, very large island containing more than 80% of the visible cells; the exception contains two islands of approximately 40% and 50% of the cells (almost 90% total). These data indicate that the T6S- mutant strain forms one large connected "mat" of cells during consolidation, to a larger degree than seen with wildtype. Taken together, these results suggest that the complete loss of communication *reduces* cell-cell connection during swarm initiation, is not impactful during collective swarming, and *increases* it during consolidation. By contrast, the loss of partial (non-lethal) communication in the Δ*ids* strain disrupts island structure, resulting in distributions like collective swarming throughout the swarm development cycle.

These analyses of the larger islands tell only a portion of the story: the singleton data are also revealing. Figure 3.6D shows box-and-whisker plots of proportion of singletons. These isolated cells comprise a minority of each population. For wildtype, singletons consistently represent a small percentage of the cells—often less than 10%. In other words, wild type cells tend to form groups. In contrast, both the mutant strains have a higher proportion of singletons at hours 5 and 6 than the wildtype. The T6S- strain has virtually no singletons at hour 7, during consolidation; in this phase, cells are generally found in comparatively large islands. The Δ*ids* strain has a higher proportion of singletons at every timepoint than the wildtype, and the greatest proportion during consolidation, which is consistent with the reduced δ-adjacency (Figure 3.3E). Taken together with the island size distributions, we see that disrupting cell-cell communication and kin identity impacts how cells maintain contact. This phenomenon would not happen without the involvement of the corresponding molecular mechanisms (and signaling outputs).

3.5 Discussion

Many existing studies of microbial collectives focus on the actions of individuals (the microscale) or the outcomes of the population (the macroscale), omitting that of intermediate sized groups: the mesoscale. The mesoscale is not only a critical bridge between the extremes,

but also a key player in the dynamics. To address this need, we focused on complementary approaches: cell length distributions and the geometry of cell groupings that form and dissolve during the developmental phases of a *P. mirabilis* colony. We developed and leveraged new methods to analyze the geometry of these structures, applying them to the results of targeted experiments to derive biological insights about cell-cell communication. We found that *P. mirabilis* colonies form cell groupings of different sizes and that patterns in these sizes vary over the stages of swarm development. Of particular note is our result that cell-cell communication—specifically the transfer of T6SS-dependent self-identity signals—regulates the morphology of cells and their degree of interaction with neighboring cells.

These findings not only agree with the well-known prediction that genetic relatedness promotes cooperation (28,154) but also support a new hypothesis about mesoscale structures that organize the network of connected cells within a swarm. Other biological studies have shown that bacterial populations can have increased fitness (growth, survival) in social groups: e.g., that larger group sizes have increased surfactant in biofilms, protecting them against threats like phage and antimicrobials, and that quorum sensing can contribute to one population's dominance (155–158). In *P. mirabilis* swarms, group size might increase access to nutrients through surface occupation or facilitate protection against stresses like phage. The data and analyses presented here, in context of the field, suggest a new biological model: kin recognition promotes populations to form highly adjacent, connected components (groups) that together form a *mesoscale* mesh network.

This application of computational mathematics to microbiology has really moved the needle. While previous work, which relied largely on qualitative descriptions derived from visual examination of microbes is certainly useful, it is essential to have a *formal* quantification of these structures if we are to truly understand the mechanisms that underlie organismal behavior and organization. There have been some recent developments along these lines in the context of uniformly shaped microbial cells with no motility (131)—as well as a significant body of work on other types of biological aggregations (145,158), where the individual agents can be approximated as points, circles, or ellipsoids. Our work is more general: our mathematics is designed to handle complex biological patterns and the variations that occur amongst experiments. Further, our experiments leverage those analyses to study biological mechanisms underlying group behaviors. Our methods enable automated analysis of large data sets. The results of this combined mathematical/biological approach are objective and reproducible, enabling validation of existing, as well as the formation of new, hypotheses about cell-cell communications—which can in turn inform the design of new experiments.

One of the most striking findings in the analysis of the wild type, Δ*ids* (disrupted non-lethal self-identity), and T6S- (no contact-dependent communication) strains over three phases of their life cycle is that visual density does not equate to physical contact. Perceiving groups of cells as functionally organized or coordinated when they are not was first termed by (159) as the "clustering illusion." Existing techniques for the analysis of biological aggregations, such as density and nearest-neighbor analysis (145,146), cannot resolve such an illusion. Our sharp focus on adjacency, however, allows us to reveal these visual density paradoxes quite clearly. For example, even though, to the human eye, there appears to be less "grouping" of T6S- cells at

all phases, when compared to wildtype, this is an illusion, as we saw using the δ-adjacency analysis (recall Figure 3.6).

A comparison of the wild type, Δ*ids,* and T6S- mutant strains suggests that communication dictates population architecture. The island analysis that we introduce in this paper—a new way to identify contact networks and thereby facilitate a specific understanding of the biologically-relevant structure of the swarm—is based on analysis of the connected components in a graph that captures the δ-adjacency relationships between cells in an image. The sizes of these groups bring out salient information about network structure in a novel way. For example, wild type populations are highly dynamic, forming large islands during swarm initiation, which then disperse during active swarming and then re-form during consolidation, perhaps suggesting a sort of dynamic coordination amongst the cells. In contrast, cells lacking all communication (T6S-) fail to aggregate into islands during swarm initiation and form oversized "mats" during consolidation. This suggests that communication may "prime" the cells for collective migration and prevent stagnation during consolidation. Finally, cell groupings in the Δ*ids* populations exhibit almost no change over time. Instead, the population appears to be stuck in a sort of "permanent migration state" consisting of disorganized small and medium islands. We posit that non-lethal identity signaling in this case may act as a signal that guides the transition into different life phases and/or organizes cells into rafts during collective migration.

Many opportunities for future work remain. A first goal would be to develop methods for characterizing other important attributes of the mesoscale structure, such as alignment. Although the term "alignment" is intuitive, it is surprisingly difficult to formalize, especially for cells with complicated shapes. Alignment and adjacency are related—two cells with high relative δ-adjacency should be designated as aligned, regardless of their shapes—but what constitutes a sufficiently "high" value is not obvious. Modifying the δ-adjacency graphs to incorporate directed or weighted edges could lead to a quantitative metric for this attribute. Another potentially interesting aspect of mesoscale structure relates to the holes that are visible in many of the *P. mirabilis* scatterplots in this paper. The ideas of *persistent homology,* another powerful tool from topological data analysis (160), could help quantify such structures. It could also be useful to vary the threshold at which an edge is added between two vertices of an adjacency graph to explore the effects on network structure.

Our analysis methods should extend smoothly to other collections of elongated bacteria that engage in collective migration, such as *Vibrio parahaemolyticus (153)*, many of which remain poorly understood due to restricted genetics, cultivation, and image analysis tools. They might also help understand the role that contact-mediated interactions between *Capnocytophaga gingivalis* cells play in multi-species communities (161). We are excited to apply our mesoscale-analysis techniques to systems that involve physical contact between heterogeneous agents, like those in biofilms, engineered soil crusts, or synthetic metabolic communities (81,162–164). Finally, we are looking forward to using these new methods to study how kin recognition and cell-cell communication regulate these structures, using genetic manipulations to further dissect the intracellular and intercellular pathways.

The methods presented here mark a significant shift from qualitative microscopy to a quantitative, spatial analysis of microbial spatial ecologies. Formal characterization of the

structures in a bacterial colony has many affordances: not just comparison and contrast, but real insights into biological mechanisms. Key to this is effective interleaving of mathematics and microbiology. We used data from laboratory experiments to develop, test, and refine the mathematical methods; these insights then led to new hypotheses and the design of laboratory experiments for their validation. This synergistic approach allowed us to go beyond the descriptive and led to insights about cell-cell communication and its role in the formation of multiscale structures that emerge during the developmental phases of the colony—insights that are relevant to network development, maintenance, and expansion.

3.6 Acknowledgements

We would like to thank Achala Chittor for construction of the T6SS-disrupted strain as well as members of the Gibbs lab for thoughtful discussions and comments on the manuscript. We thank the following funders of this research: the Sloan Foundation (grant number G-2022-19509 to K.A.G.), the National Science Foundation (Graduate Research Fellowship to E.E.G.), the University of California, Berkeley, the University of Colorado Boulder, and the Santa Fe Institute. We especially thank the David and Lucile Packard Foundation, who nucleated this collaboration by fortuitously having K.A.G. and E.B. present neighboring posters at a meeting.

3.7 Author Contributions

E.E.G. designed the project, performed the biological experiments, microscopy analysis, and co-wrote the manuscript. M.B. designed the project, created the code transforming biological insights into tractable mathematical analyses, and co-wrote the manuscript. E.B., J.D.M,, and K.A.G. designed the project, contributed to data analysis, and co-wrote the manuscript. Without the dedication and time of all collaborators, this project would not have been possible.

3.8 Supplemental Figures and Tables

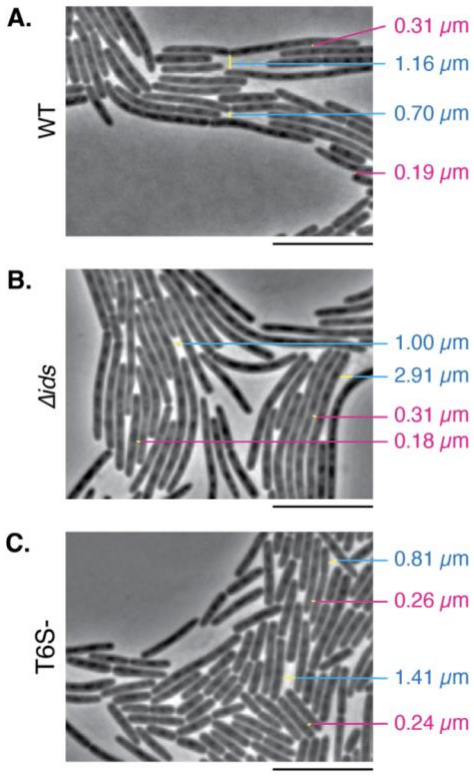


Figure S3.1 Measurement of δ-adjacency is based on calculations from phase contrast images. To determine the parameter δ, we measured several distances between adjacent (magenta labels) and non-adjacent (cyan labels) cells, as labeled in demonstration on this figure. The value we choose, δ = 0.59 μm, is the average of the distance between pairs of visually observed neighboring cells across several images. (A) WT. *(B)* Δ*ids*. (C) T6S-. Scale bar, 10 μm.

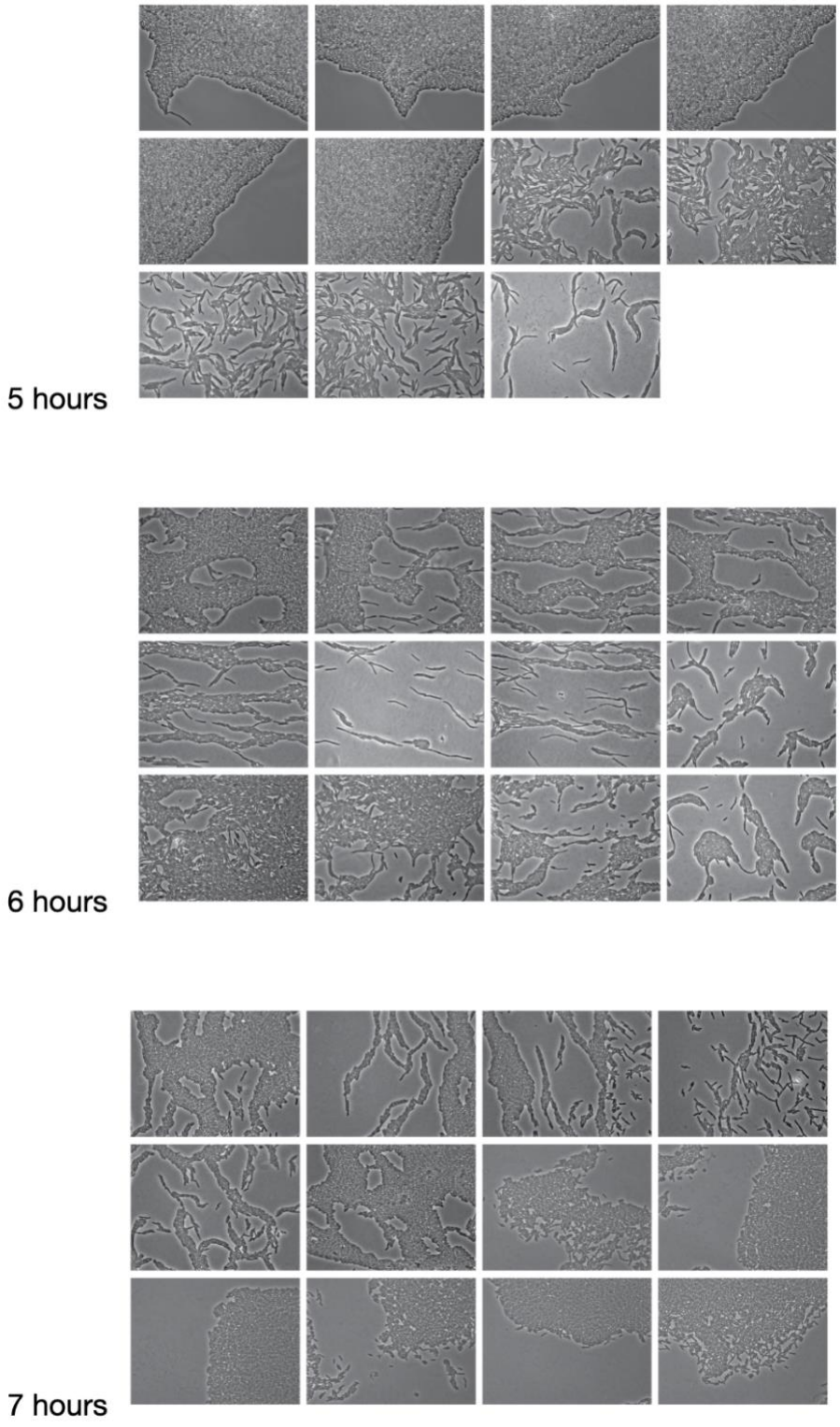
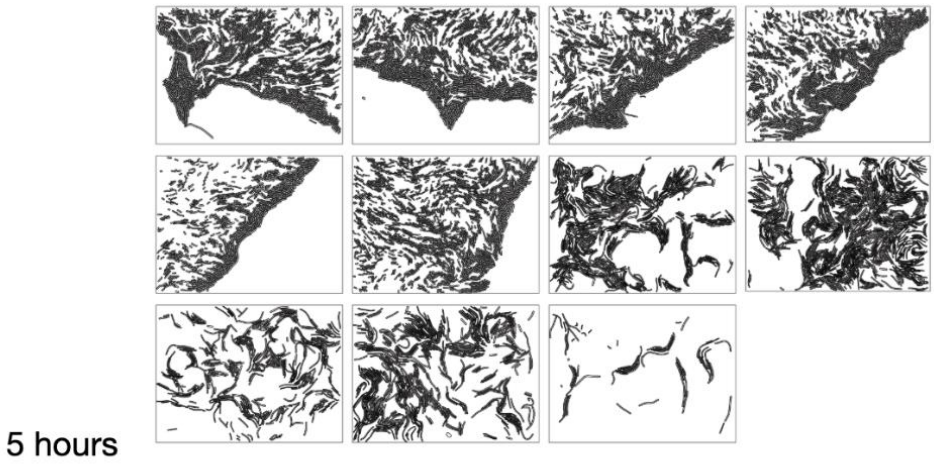
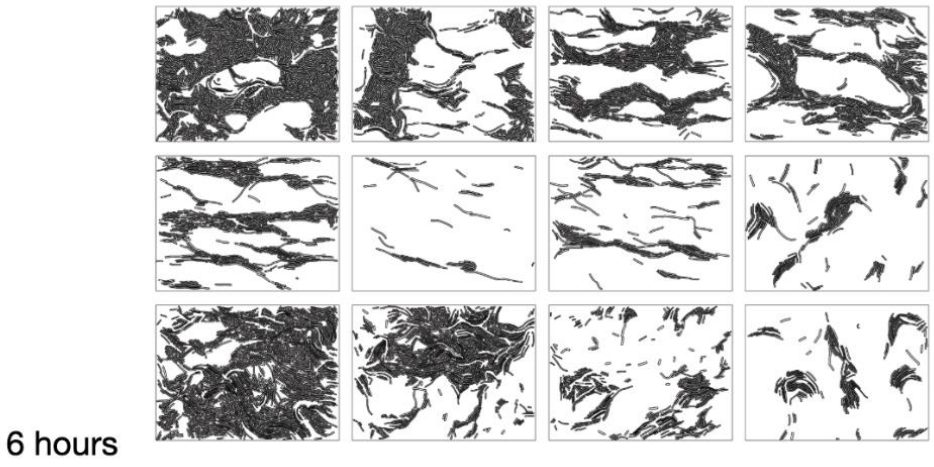


Figure S3.2 Phase contrast images of wild type strain BB2000 used for analysis. Each individual image is ~ 126.4 μm wide by ~ 94.4 μm tall.





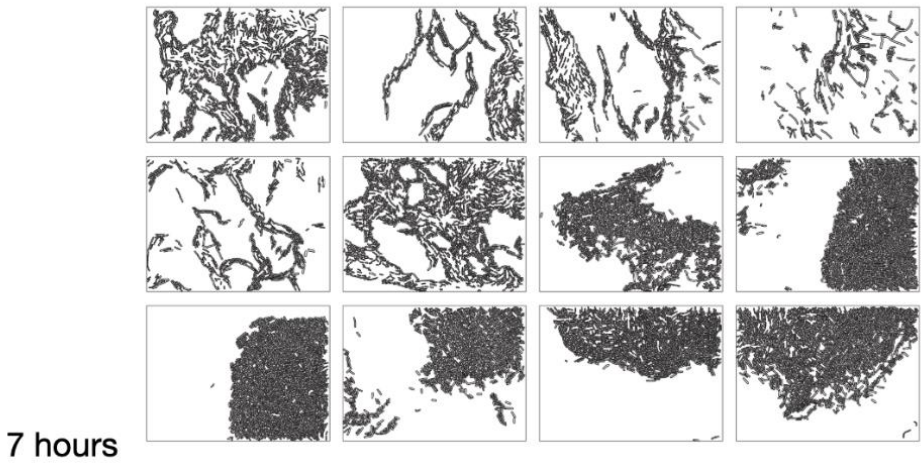


Figure S3.3 Scatterplots of wild type strain BB2000 used for analysis. Each individual image is ~ 126.4 μm wide by ~ 94.4 μm tall.

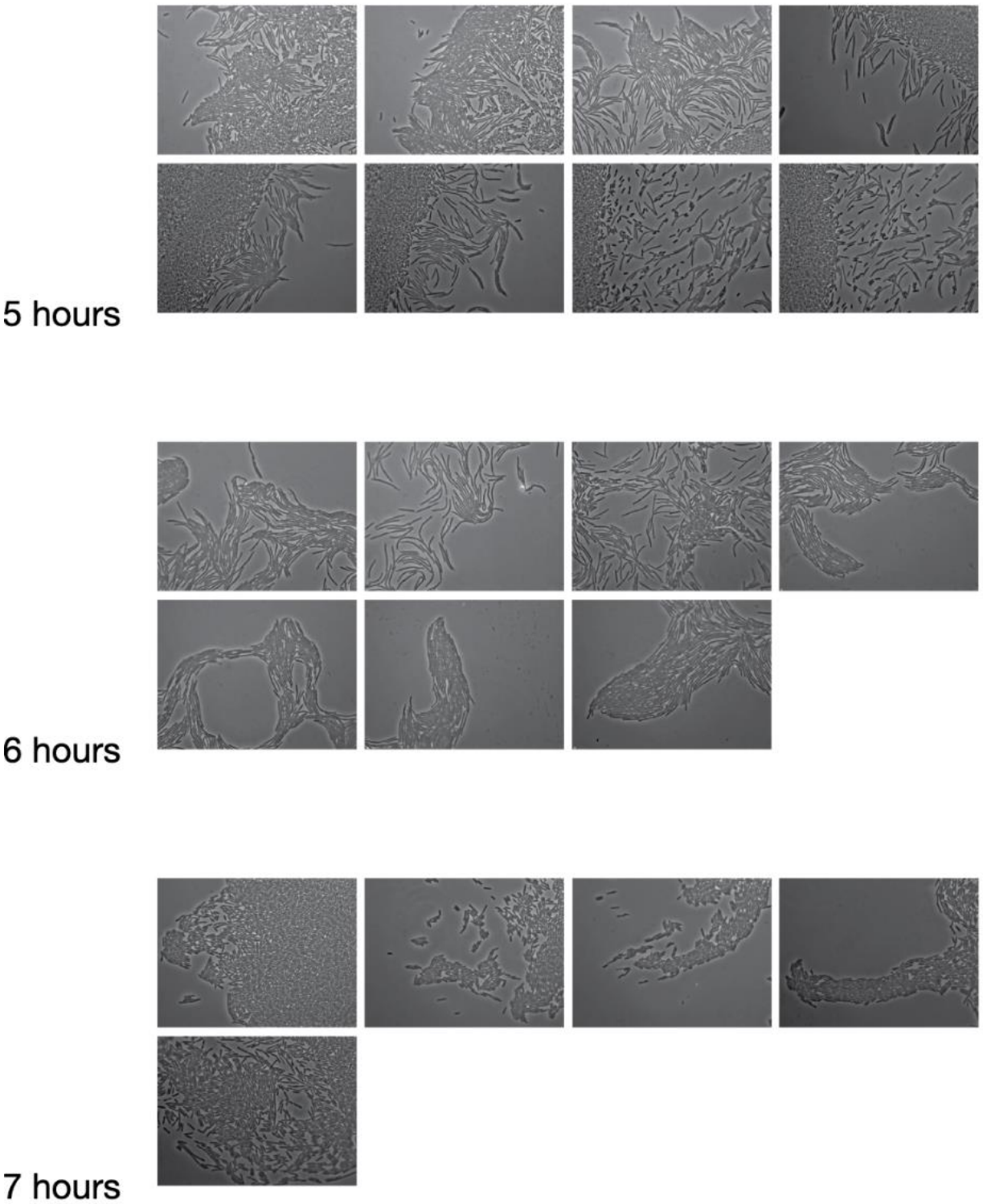


Figure S3.4 Phase contrast images of *P. mirabilis* strain Δ*ids* used for analysis. Each individual image is ~ 126.4 μm wide by ~ 94.4 μm tall.

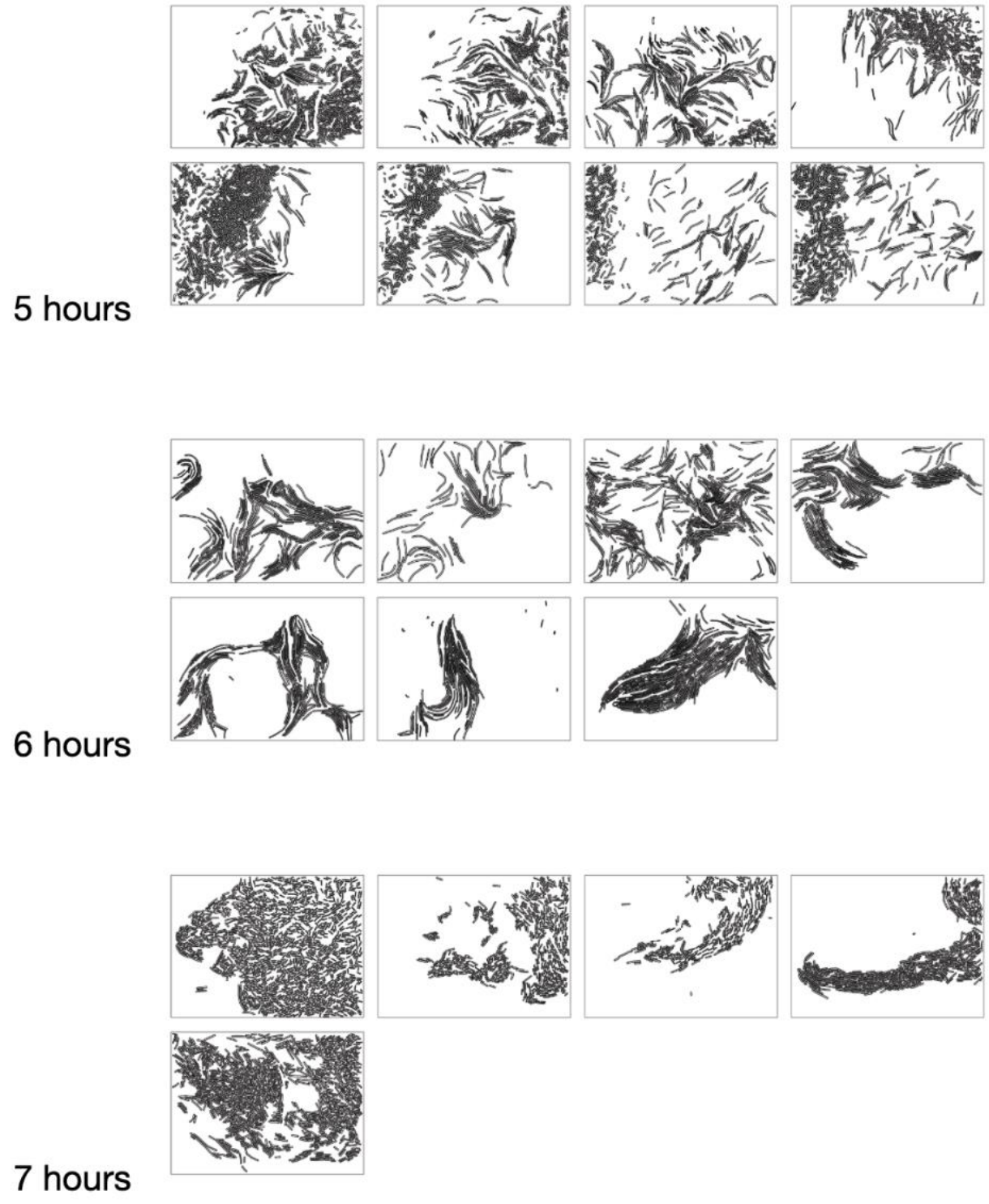


Figure S3.5 Scatterplots of *P. mirabilis* strain Δ*ids* used for analysis. Each individual image is ~ 126.4 μm wide by ~ 94.4 μm tall.

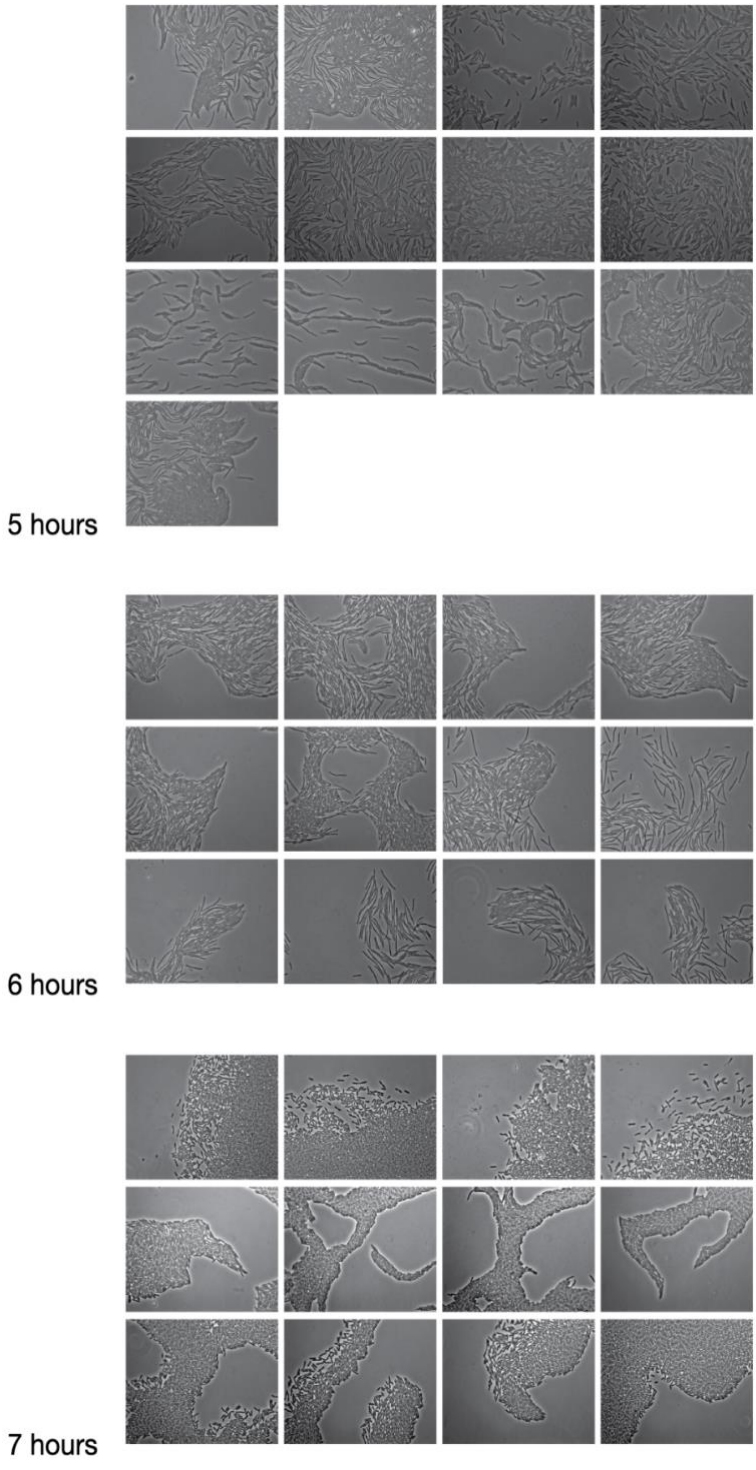


Figure S3.6 Phase contrast images of *P. mirabilis* strain T6S^{-} used for analysis. Each individual image is ~ 126.4 μm wide by ~ 94.4 μm tall.

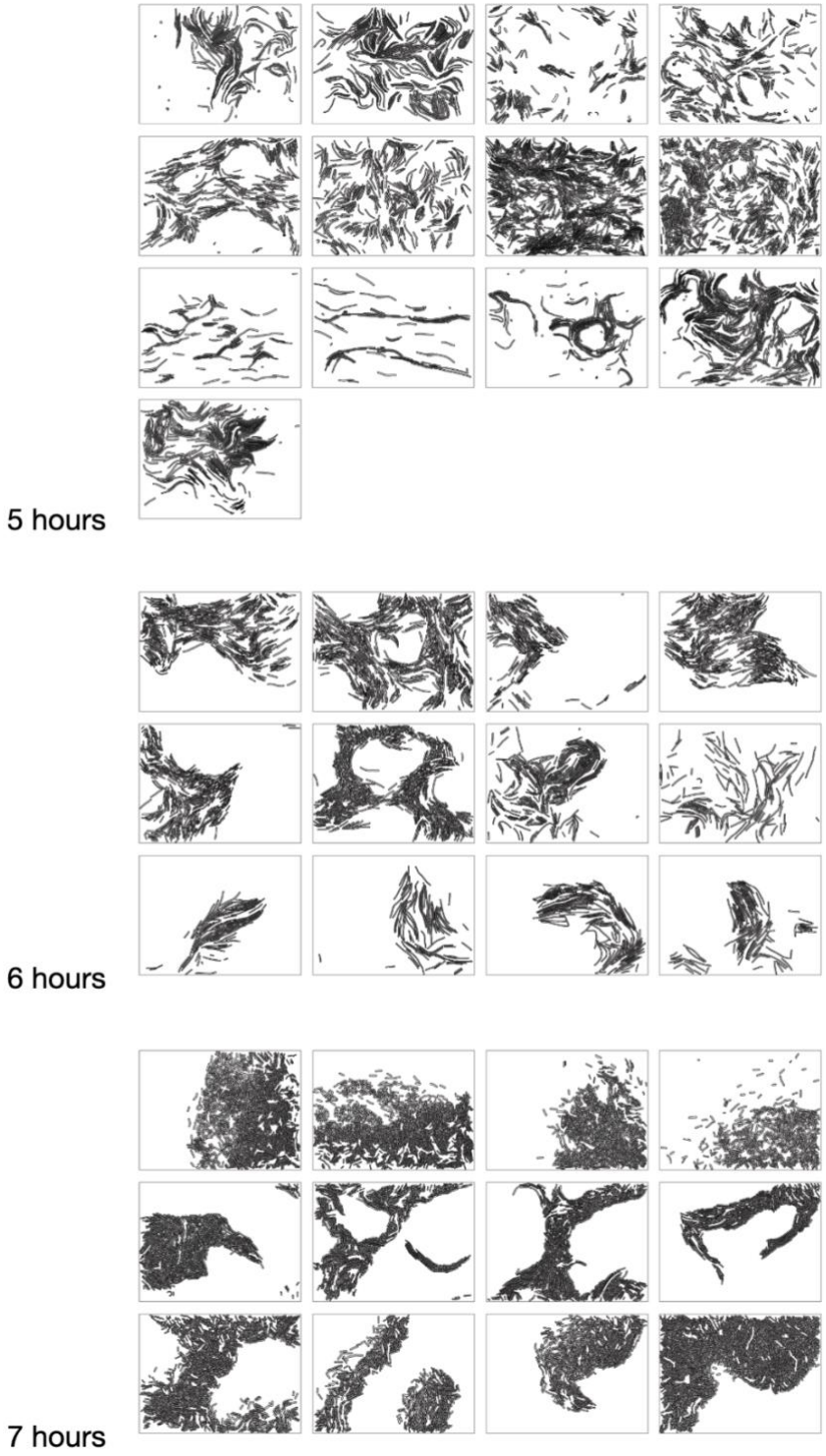


Figure S3.7 Scatterplots of *P. mirabilis* strain T6S^{-} used for analysis. Each individual image is ~ 126.4 μm wide by ~ 94.4 μm tall.

Table S3.1: Descriptive statistics for cell lengths (in µm) for all strains at all measured time points.

Strain	Hour	Count	Mean Cell Length (µm)	sd	median	IQR	max	min
BB2000	5	8258	4.94	3.34	3.72	3.01	36.5	2
BB2000	6	7320	5.68	3.95	4.38	3.13	38.7	2
BB2000	7	10275	3.60	1.52	2.36	1.47	20.6	2
Δids	*5*	*6082*	*5.56*	*4.16*	*3.75*	*4.27*	*35.7*	*2*
Δids	*6*	*1299*	*12.50*	*6.25*	*11.80*	*45.90*	*45.9*	*2*
Δids	*7*	*4553*	*4.43*	*2.40*	*3.77*	*39.90*	*39.9*	*2*
T6S-	5	4108	9.69	4.86	8.50	5.45	41.5	2
T6S-	6	3949	8.09	5.44	6.44	4.88	51.6	2
T6S-	7	9860	4.05	1.80	3.62	1.60	30.0	2

Chapter 4: Concluding Thoughts and Open Questions

4.1 WT *P. mirabilis* swarming collectives display a spectrum of single-cell morphologies

After successfully developing the Swarmetrics pipeline (144) presented in Chapter 2, a world opened, and enabled single-cell morphometrics (including cell length, width, sinuosity, curvature, to name a few) to be captured in dense, quickly migrating swarms. The dataset compiled for Chapter 2 consisted of metrics of nearly 80,000 cells through their life cycle.

With this data in hand, it was notable how much the approximations of *P. mirabilis* cell lengths during elongation and migration differed. The vast range of reported *P. mirabilis* cell lengths mirrors the mythology of Proteus itself, and ever-morphing figure whose form was difficult to define and likely arose through the propagation of decades of visually based approximations instead of quantitative measurement. This finding also emerged concurrently with a body of literature warning of the bias involved in visual interpretation alone (132). A close look at the literature yielded lengths of elongated *P.mirabilis* cells from 40 to over 100 μm (142,165). Of course, *P. mirabilis* strains BB2000 and ATCC were the only strains analyzed, and other strains may differ in their maximum or mean elongation length. However, the publication of this paper presented a possible solution for capturing individual cell morphometrics of cells that are typically difficult to segment. Indeed, Swarmetrics is capable of more accurately defining densely packed, oddly shaped, cell morphometrics, partially made possible with the addition of minor manual annotation in addition to machine learning algorithms.

Most notable from our findings in Chapter 1 was that cell length was not synchronized across the population, which is to say, not every cell elongated at the start of swarming. In fact, shorter cells ($\leq 4\mu m$) dominated the population at nearly every hour, except for hour 6 in the BB2000 timecourse, where 56.4% of imaged cells were >4 μm in length. The surprising amount of cell length heterogeneity observed across the *P. mirabilis* lifecycle enabled us to set forward a new model for the *P. mirabilis* life cycle. This model puts forth the idea that swarm development in *P. mirabilis* is governed, perhaps, by asynchronous and heterogenous cell differentiation, rather than by a uniformly distributed population-wide transition.

As machine-learning approaches improve, I am eager to see what other nuances are revealed regarding cooperative behaviors. Additionally, as technological advances continue to accelerate research in synthetic bacterial communities, biofilms, and microbiomes, it may become increasingly necessary for researchers to dynamically adjust the resolution at which population-wide behaviors are examined. Currently, technology remains a large barrier to smoothly transitioning from micro- to meso- to macroscale resolutions of group behaviors. However, advancements may allow researchers to probe subtle interactions between cells, and uncover the roles those interactions play in maintaining, participating in, and potentially imitating cooperative behaviors. Bridging these spatial and analytical scales could lead to a more comprehensive understanding of how collective behaviors emerge from interactions of individual cells can give rise to collective behaviors, cooperative organization, and, potentially, the evolutionary foundations of multicellularity.

4.2 Identity signaling impacts population architecture

In Chapter 3, I collaborated with mathematicians and computer scientists to turn loud, biological data – especially qualitative observations – into a mathematical and quantifiable measurement. Through this collaboration, I learned that cross-field collaboration requires a kind of bilingualism that I had not expected (given that I am bilingual). But speaking across disciplines, where a word like "network" holds such different significance, was exhilarating. Over many iterations, we were able to generate a measurement termed δ-adjacency. Defining this term was required for us to quantify cell-cell contact in a still-frame phase contrast image. And it brought the observations I saw under the microscope to life. The examination of adjacency mesoscale structures (Figure 3.3) highlights the overall proportion of cells in an image that are in close contact to other strains. Given that T6S, and the transmission of IdsD from one cell to another is contact-dependent, gaining this metric provided a clear picture of where opportunities for cell-cell communication were possible.

In addition, through a comparison of the wild type, Δ*ids* (disrupted non-lethal self-identity), and T6S$^-$ (no contact-dependent communication) strains at the mesoscale, it became clear that the population architecture was altered. The mathematical analysis was an example of the "clustering illusion," which states that groups of cells may be perceived as organized or coordinated, even when they're not (45). However, by focusing on adjacency, the paradox's veil was lifted. And, moved forward the hypothesis that wild type populations are dynamic through their life stages, perhaps coordinated by a process we do not yet understand. By comparison, T6S$^-$ cells do not form interlocked swarming structures, and collapse into large mono-layer cell "mats" during consolidation. It's possible that prior to collective migration cells may be "primed" or a subset may become "scouts," as a mechanism to induce collective migration, and, more importantly perhaps, keep the migration coordinated over time.

Due to the incredible speed of *P. mirabilis* cells during collective migration (2.45-3.77 mm/h) (7) I was not able to track individual cells as they progressed through their life cycle, nor into the next. However, this type of imaging paired with the mathematical analysis presented in Chapter , has immense potential in elucidating the role of (p)ppGpp in swarm winnowing (28), clarify how and/or what initiates swarming or other emergent collective behaviors, as well as further uncovering how direct cell-cell signaling can shape colony architecture spatially and temporally.

Chapter 5: References

1. Armitage JP, Smith DG, Rowbury RJ. Alterations in the cell envelope composition of Proteus mirabilis during the development of swarmer cells. Biochim Biophys Acta BBA - Gen Subj. 1979 May 16;584(3):389–97. doi:10.1016/0304-4165(79)90115-6

2. Manos J, Belas R. The Genera Proteus, Providencia, and Morganella. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E, editors. The Prokaryotes: A Handbook on the Biology of Bacteria Volume 6: Proteobacteria: Gamma Subclass [Internet]. New York, NY: Springer; 2006 [cited 2024 Oct 23]. p. 245–69. Available from: https://doi.org/10.1007/0-387-30746-X_12 doi:10.1007/0-387-30746-X_12

3. Wenner JJ, Rettger LF. A Systematic Study of the Proteus Group of Bacteria. J Bacteriol. 1919 Jul;4(4):331–53. doi:10.1128/jb.4.4.331-353.1919 PubMed PMID: 16558845; PubMed Central PMCID: PMC378813.

4. Pearson MM, Rasko DA, Smith SN, Mobley HL. Transcriptome of swarming Proteus mirabilis. Infect Immun. 2010;78(6):2834–45. doi:10.1128/iai.01222-09

5. Anderson JK, Smith TG, Hoover TR. Sense and sensibility: flagellum-mediated gene regulation. Trends Microbiol. 2010 Jan;18(1):30. doi:10.1016/j.tim.2009.11.001 PubMed PMID: 19942438; PubMed Central PMCID: PMC2818477.

6. HOENIGER JFMY 1965. Development of Flagella by Proteus mirabilis. Microbiology. 40(1):29–42. doi:10.1099/00221287-40-1-29

7. Jones BV, Young R, Mahenthiralingam E, Stickler DJ. Ultrastructure of Proteus mirabilis Swarmer Cell Rafts and Role of Swarming in Catheter-Associated Urinary Tract Infection. Infect Immun. 2004 Jul;72(7):3941–50. doi:10.1128/IAI.72.7.3941-3950.2004 PubMed PMID: 15213138; PubMed Central PMCID: PMC427392.

8. Armbruster CE, Mobley HLT, Pearson MM. Pathogenesis of Proteus mirabilis Infection. EcoSal Plus. 2018 Feb;8(1):10.1128/ecosalplus.ESP. doi:10.1128/ecosalplus.ESP-0009-2017 PubMed PMID: 29424333.

9. Armbruster CE, Hodges SA, Mobley HLT. Initiation of Swarming Motility by Proteus mirabilis Occurs in Response to Specific Cues Present in Urine and Requires Excess l-Glutamine. J Bacteriol. 2013 Mar;195(6):1305–19. doi:10.1128/JB.02136-12 PubMed PMID: 23316040; PubMed Central PMCID: PMC3591990.

10. Little K, Tipping MJ, Gibbs KA. Swarmer Cell Development of the Bacterium Proteus mirabilis Requires the Conserved Enterobacterial Common Antigen Biosynthesis Gene rffG. J Bacteriol. 200(18):e00230-18. doi:10.1128/JB.00230-18

11. Little K, Austerman J, Zheng J, Gibbs KA. Cell Shape and Population Migration Are Distinct Steps of Proteus mirabilis Swarming That Are Decoupled on High-Percentage Agar. J Bacteriol. 2019 May 8;201(11):10.1128/jb.00726-18. doi:10.1128/jb.00726-18

12. Dienes L. Reproductive Processes in Proteus Cultures. Proc Soc Exp Biol Med. 1946 Nov 1;63(2):265–70. doi:10.3181/00379727-63-15570

13. Budding AE, Ingham CJ, Bitter W, Vandenbroucke-Grauls CM, Schneeberger PM. The Dienes Phenomenon: Competition and Territoriality in Swarming Proteus mirabilis. J Bacteriol. 2009 Jun 15;191(12):3892–900. doi:10.1128/JB.00975-08

14. Gibbs KA, Wenren LM, Greenberg EP. Identity Gene Expression in Proteus mirabilis. J Bacteriol. 2011 Jul 1;193(13):3286–92. doi:10.1128/JB.01167-10 PubMed PMID: 21551301.

15. Senior BW. Investigation of the types and characteristics of the proteolytic enzymes formed by diverse strains of Proteus species. J Med Microbiol. 1999 Jul;48(7):623–8. doi:10.1099/00222615-48-7-623 PubMed PMID: 10403412.

16. Senior BW. Typing of Proteus strains by proticine production and sensitivity. J Med Microbiol. 1977 Feb;10(1):7–17. doi:10.1099/00222615-10-1-7 PubMed PMID: 320341.

17. Gibbs KA, Urbanowski ML, Greenberg EP. Genetic Determinants of Self Identity and Social Recognition in Bacteria. Science. 2008 Jul 11;321(5886):256–9. doi:10.1126/science.1160033 PubMed PMID: 18621670.

18. Cardarelli L, Saak C, Gibbs KA. Two Proteins Form a Heteromeric Bacterial Self-Recognition Complex in Which Variable Subdomains Determine Allele-Restricted Binding. mBio. 2015 Jul 1;6(3). doi:10.1128/mBio.00251-15 PubMed PMID: 26060269.

19. Zepeda-Rivera MA, Saak CC, Gibbs KA. A Proposed Chaperone of the Bacterial Type VI Secretion System Functions To Constrain a Self-Identity Protein. J Bacteriol. 2018 Jul 15;200(14):e00688-17. doi:10.1128/JB.00688-17 PubMed PMID: 29555703; PubMed Central PMCID: PMC6018363.

20. Wenren LM, Sullivan NL, Cardarelli L, Septer AN, Gibbs KA. Two Independent Pathways for Self-Recognition in Proteus mirabilis Are Linked by Type VI-Dependent Export. mBio. 2013 Aug 30;4(4). doi:10.1128/mBio.00374-13 PubMed PMID: 23882014.

21. Brunet YR, Hénin J, Celia H, Cascales E. Type VI secretion and bacteriophage tail tubes share a common assembly pathway. EMBO Rep. 2014 Mar;15(3):315–21. doi:10.1002/embr.201337936 PubMed PMID: 24488256; PubMed Central PMCID: PMC3989698.

22. Brunet YR, Zoued A, Boyer F, Douzi B, Cascales E. The Type VI Secretion TssEFGK-VgrG Phage-Like Baseplate Is Recruited to the TssJLM Membrane Complex via Multiple Contacts and Serves As Assembly Platform for Tail Tube/Sheath Polymerization. PLOS Genet. 2015 Oct 13;11(10):e1005545. doi:10.1371/journal.pgen.1005545

23. Salomon D, Kinch LN, Trudgian DC, Guo X, Klimko JA, Grishin NV, et al. Marker for type VI secretion system effectors. Proc Natl Acad Sci. 2014 Jun 24;111(25):9271–6. doi:10.1073/pnas.1406110111 PubMed PMID: 24927539.

24. Saak CC, Gibbs KA. The Self-Identity Protein IdsD Is Communicated between Cells in Swarming Proteus mirabilis Colonies. J Bacteriol. 2016 Dec 15;198(24):3278–86. doi:10.1128/JB.00402-16 PubMed PMID: 27672195.

25. Saak CC, Gibbs KA. The Self-Identity Protein IdsD Is Communicated between Cells in Swarming Proteus mirabilis Colonies. J Bacteriol. 2016 Dec 15;198(24):3278–86. doi:10.1128/JB.00402-16 PubMed PMID: 27672195.

26. Saak CC, Zepeda-Rivera MA, Gibbs KA. A single point mutation in a TssB/VipA homolog disrupts sheath formation in the type VI secretion system of Proteus mirabilis. PLOS ONE. 2017 Sep 26;12(9):e0184797. doi:10.1371/journal.pone.0184797

27. Tipping MJ, Gibbs KA. Peer pressure from a Proteus mirabilis self-recognition system controls participation in cooperative swarm motility. PLOS Pathog. 2019 Jul 19;15(7):e1007885. doi:10.1371/journal.ppat.1007885

28. Tipping MJ, Gibbs KA. Peer pressure from a Proteus mirabilis self-recognition system controls participation in cooperative swarm motility. PLOS Pathog. 2019 Jul 19;15(7):e1007885. doi:10.1371/journal.ppat.1007885

29. Reiness G, Yang HL, Zubay G, Cashel M. Effects of guanosine tetraphosphate on cell-free synthesis of Escherichia coli ribosomal RNA and other gene products. Proc Natl Acad Sci U S A. 1975 Aug;72(8):2881–5. doi:10.1073/pnas.72.8.2881 PubMed PMID: 1103124; PubMed Central PMCID: PMC432882.

30. Cashel M, Gallant J. Two compounds implicated in the function of the RC gene of Escherichia coli. Nature. 1969 Mar 1;221(5183):838–41. doi:10.1038/221838a0 PubMed PMID: 4885263.

31. Ronneau S, Caballero-Montes J, Coppine J, Mayard A, Garcia-Pino A, Hallez R. Regulation of (p)ppGpp hydrolysis by a conserved archetypal regulatory domain. Nucleic Acids Res. 2019 Jan 25;47(2):843–54. doi:10.1093/nar/gky1201 PubMed PMID: 30496454; PubMed Central PMCID: PMC6344854.

32. Traxler MF, Summers SM, Nguyen HT, Zacharia VM, Hightower GA, Smith JT, et al. The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. Mol Microbiol. 2008 Jun;68(5):1128–48. doi:10.1111/j.1365-2958.2008.06229.x PubMed PMID: 18430135; PubMed Central PMCID: PMC3719176.

33. Anderson BW, Fung DK, Wang JD. Regulatory Themes and Variations by the Stress-Signaling Nucleotide Alarmones (p)ppGpp in Bacteria. Annu Rev Genet. 2021 Nov 23;55:115–33. doi:10.1146/annurev-genet-021821-025827 PubMed PMID: 34416118.

34. Gordon DM. The Ecology of Collective Behavior. PLOS Biol. 2014 Mar 11;12(3):e1001805. doi:10.1371/journal.pbio.1001805

35. Gregor T, Fujimoto K, Masaki N, Sawai S. The onset of collective behavior in social amoebae. Science. 2010 May 21;328(5981):1021–5. doi:10.1126/science.1183415 PubMed PMID: 20413456; PubMed Central PMCID: PMC3120019.

36. Grosberg RK, Strathmann RR. The Evolution of Multicellularity: A Minor Major Transition? Annu Rev Ecol Evol Syst. 2007 Dec 1;38(Volume 38, 2007):621–54. doi:10.1146/annurev.ecolsys.36.102403.114735

37. Bertossa RC. Morphology and behaviour: functional links in development and evolution. Philos Trans R Soc B Biol Sci. 2011 Jul 27;366(1574):2056–68. doi:10.1098/rstb.2011.0035 PubMed PMID: 21690124; PubMed Central PMCID: PMC3130372.

38. Brogan DC, Hodgins JK. Group Behaviors for Systems with Significant Dynamics. Auton Robots. 1997 Mar 1;4(1):137–53. doi:10.1023/A:1008867321648

39. Thibaut JW, Kelley HH. The social psychology of groups. New Brunswick, U.S.A: Transaction Books; 1986. 313 p. (Social science classics series).

40. Kaveh K, Veller C, Nowak MA. Games of multicellularity. J Theor Biol. 2016 Aug 21;403:143–58. doi:10.1016/j.jtbi.2016.04.037 PubMed PMID: 27179461.

41. Lehmann L, Keller L, West S, Roze D. Group selection and kin selection: Two concepts but one process. Proc Natl Acad Sci. 2007 Apr 17;104(16):6736–9. doi:10.1073/pnas.0700662104 PubMed PMID: 17416674.

42. Ouellette NT, Gordon DM. Goals and Limitations of Modeling Collective Behavior in Biological Systems. Front Phys [Internet]. 2021 [cited 2022 Aug 18];9. Available from: https://www.frontiersin.org/articles/10.3389/fphy.2021.687823

43. Nakashima S, J. Kobayashi T. Population dynamics models for various forms of adaptation. Biophys Physicobiology. 2023 Sep 2;20(3):e200034. doi:10.2142/biophysico.bppb-v20.0034 PubMed PMID: 38124797; PubMed Central PMCID: PMC10728623.

44. Damas Sagamiko T. Optimal Control of a Threatened Wildebeest-lion Prey-predator System Incorporating a Constant Prey Refuge in the Serengeti Ecosystem. Appl Comput Math. 2015;4(4):296. doi:10.11648/j.acm.20150404.18

45. Packer C, Scheel D, Pusey AE. Why Lions Form Groups: Food is Not Enough. Am Nat. 1990 Jul;136(1):1–19. doi:10.1086/285079

46. Heinsohn R, Packer C. Complex cooperative strategies in group-territorial African lions. Science. 1995 Sep 1;269(5228):1260–2. doi:10.1126/science.7652573 PubMed PMID: 7652573.

47. Flemming HC, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. Biofilms: an emergent form of bacterial life. Nat Rev Microbiol. 2016 Sep;14(9):9. doi:10.1038/nrmicro.2016.94

48. Donlan RM. Biofilms: Microbial Life on Surfaces. Emerg Infect Dis. 2002 Sep;8(9):881–90. doi:10.3201/eid0809.020063 PubMed PMID: 12194761; PubMed Central PMCID: PMC2732559.

49. Verstraten P, Verweij G, Zwaneveld PJ. Complexities in the spatial scope of agglomeration economies. J Reg Sci. 2019;59(1):29–55. doi:10.1111/jors.12391

50. Lee CK, Vachier J, de Anda J, Zhao K, Baker AE, Bennett RR, et al. Social Cooperativity of Bacteria during Reversible Surface Attachment in Young Biofilms: a Quantitative Comparison of Pseudomonas aeruginosa PA14 and PAO1. mBio. 2020 Feb 25;11(1):e02644-19. doi:10.1128/mBio.02644-19 PubMed PMID: 32098815; PubMed Central PMCID: PMC7042694.

51. Chathoth K, Fostier L, Martin B, Baysse C, Mahé F. A Multi-Skilled Mathematical Model of Bacterial Attachment in Initiation of Biofilms. Microorganisms. 2022 Apr;10(4):4. doi:10.3390/microorganisms10040686

52. Yeor-Davidi E, Zverzhinetsky M, Krivitsky V, Patolsky F. Real-time monitoring of bacterial biofilms metabolic activity by a redox-reactive nanosensors array. J Nanobiotechnology. 2020 May 24;18(1):81. doi:10.1186/s12951-020-00637-y

53. Devaraj A, Buzzo JR, Mashburn-Warren L, Gloag ES, Novotny LA, Stoodley P, et al. The extracellular DNA lattice of bacterial biofilms is structurally related to Holliday junction recombination intermediates. Proc Natl Acad Sci. 2019 Dec 10;116(50):25068–77. doi:10.1073/pnas.1909017116

54. Jo J, Price-Whelan A, Dietrich LEP. Gradients and consequences of heterogeneity in biofilms. Nat Rev Microbiol. 2022 Oct;20(10):593–607. doi:10.1038/s41579-022-00692-2 PubMed PMID: 35149841; PubMed Central PMCID: PMC9590228.

55. Douglas AE. The microbial exometabolome: ecological resource and architect of microbial communities. Philos Trans R Soc B Biol Sci. 2020 Mar 23;375(1798):20190250. doi:10.1098/rstb.2019.0250

56. Pérez-Osorio AC, Williamson KS, Franklin MJ. Heterogeneous rpoS and rhlR mRNA levels and 16S rRNA/rDNA (rRNA gene) ratios within Pseudomonas aeruginosa biofilms, sampled by laser capture microdissection. J Bacteriol. 2010 Jun;192(12):2991–3000. doi:10.1128/JB.01598-09 PubMed PMID: 20348255; PubMed Central PMCID: PMC2901698.

57. Lyons NA, Kolter R. On The Evolution of Bacterial Multicellularity. Curr Opin Microbiol. 2015 Apr;24:21–8. doi:10.1016/j.mib.2014.12.007 PubMed PMID: 25597443; PubMed Central PMCID: PMC4380822.

58. Martin-Verstraete I, Stülke J, Klier A, Rapoport G. Two different mechanisms mediate catabolite repression of the Bacillus subtilis levanase operon. J Bacteriol. 1995 Dec;177(23):6919–27. doi:10.1128/jb.177.23.6919-6927.1995

59. Parisutham V, Guharajan S, Lian M, Ali MZ, Rogers H, Joyce S, et al. E. coli transcription factors regulate promoter activity by a universal, homeostatic mechanism. Science. 2025 Sep 11;389(6765):eadv2064. doi:10.1126/science.adv2064

60. Zhou Y, Vazquez A, Wise A, Warita T, Warita K, Bar-Joseph Z, et al. Carbon catabolite repression correlates with the maintenance of near invariant molecular crowding in proliferating E. coli cells. BMC Syst Biol. 2013 Dec 12;7:138. doi:10.1186/1752-0509-7-138 PubMed PMID: 24330501; PubMed Central PMCID: PMC3924228.

61. Hornef MW, Wick MJ, Rhen M, Normark S. Bacterial strategies for overcoming host innate and adaptive immune responses. Nat Immunol. 2002 Nov;3(11):1033–40. doi:10.1038/ni1102-1033

62. Ebrahimi A, Schwartzman J, Cordero OX. Cooperation and spatial self-organization determine rate and efficiency of particulate organic matter degradation in marine bacteria. Proc Natl Acad Sci. 2019 Nov 12;116(46):23309–16. doi:10.1073/pnas.1908512116

63. Tan L, Huang Y, Shang W, Yang Y, Peng H, Hu Z, et al. Accessory Gene Regulator (agr) Allelic Variants in Cognate Staphylococcus aureus Strain Display Similar Phenotypes. Front Microbiol. 2022 Feb 25;13. doi:10.3389/fmicb.2022.700894

64. Tal-Gan Y, Ivancic M, Cornilescu G, Blackwell HE. Characterization of Structural Elements in Native Autoinducing Peptides and Non-Native Analogues that Permit the Differential Modulation of AgrC-type Quorum Sensing Receptors in Staphylococcus aureus. Org Biomol Chem. 2016 Jan 7;14(1):113–21. doi:10.1039/c5ob01735a PubMed PMID: 26416476; PubMed Central PMCID: PMC4681674.

65. Li SI, Buttery NJ, Thompson CRL, Purugganan MD. Sociogenomics of self vs. non-self cooperation during development of Dictyostelium discoideum. BMC Genomics. 2014 Jul 21;15(1):616. doi:10.1186/1471-2164-15-616 PubMed PMID: 25048306; PubMed Central PMCID: PMC4118049.

66. Hirose S, Santhanam B, Katoh-Kurosawa M, Shaulsky G, Kuspa A. Allorecognition, via TgrB1 and TgrC1, mediates the transition from unicellularity to multicellularity in the social amoeba Dictyostelium discoideum. Dev Camb Engl. 2015 Oct 15;142(20):3561–70. doi:10.1242/dev.123281 PubMed PMID: 26395484; PubMed Central PMCID: PMC4631765.

67. Gruenheit N, Parkinson K, Stewart B, Howie JA, Wolf JB, Thompson CRL. A polychromatic ‘greenbeard’ locus determines patterns of cooperation in a social amoeba. Nat Commun. 2017 Jan 25;8(1):14171. doi:10.1038/ncomms14171

68. Evans CR, Smiley MK, Asahara Thio S, Wei M, Florek LC, Dayton H, et al. Spatial heterogeneity in biofilm metabolism elicited by local control of phenazine methylation. Proc Natl Acad Sci. 2023 Oct 24;120(43):e2313208120. doi:10.1073/pnas.2313208120

69. Gordon DM. The Ecology of Collective Behavior [Internet]. Princeton University Press; 2023 [cited 2025 Jan 10]. Available from: https://doi.org/10.1515/9780691232164 doi:10.1515/9780691232164

70. Giardina I. Collective behavior in animal groups: theoretical models and empirical studies. HFSP J. 2008 Aug;2(4):205–19. doi:10.2976/1.2961038 PubMed PMID: 19404431; PubMed Central PMCID: PMC2639936.

71. Mizumoto N, Bourguignon T. Modern termites inherited the potential of collective construction from their common ancestor. Ecol Evol. 2020 Jun 2;10(13):6775–84. doi:10.1002/ece3.6381 PubMed PMID: 32724550; PubMed Central PMCID: PMC7381753.

72. Sumpter D. Collective animal behavior / David J.T. Sumpter. Princeton, N.J: Princeton University Press; 2010.

73. Seef S, Herrou J, de Boissier P, My L, Brasseur G, Robert D, et al. A Tad-like apparatus is required for contact-dependent prey killing in predatory social bacteria. Laub MT, Storz G, Sogaard-Andersen L, editors. eLife. 2021 Sep 10;10:e72409. doi:10.7554/eLife.72409

74. Murphy P, Comstock J, Khan T, Zhang J, Welch R, Igoshin OA. Cell behaviors underlying Myxococcus xanthus aggregate dispersal. mSystems. 2023 Sep 25;8(5):e00425-23. doi:10.1128/msystems.00425-23

75. Mendes-Soares H, Velicer GJ. Decomposing Predation: Testing for Parameters that Correlate with Predatory Performance by a Social Bacterium. Microb Ecol. 2013 Feb 1;65(2):415–23. doi:10.1007/s00248-012-0135-6

76. Thiery S, Turowski P, Berleman JE, Kaimer C. The predatory soil bacterium Myxococcus xanthus combines a Tad- and an atypical type 3-like protein secretion system to kill bacterial cells. Cell Rep. 2022 Sep 13;40(11):111340. doi:10.1016/j.celrep.2022.111340

77. Schaap P. Evolutionary crossroads in developmental biology: Dictyostelium discoideum. Dev Camb Engl. 2011 Feb 1;138(3):387–96. doi:10.1242/dev.048934 PubMed PMID: 21205784; PubMed Central PMCID: PMC3014629.

78. Devreotes P. Dictyostelium discoideum: a Model System for Cell-Cell Interactions in Development. Science. 1989 Sep 8;245(4922):1054–8. doi:10.1126/science.2672337

79. Kühn MJ, Talà L, Inclan YF, Patino R, Pierrat X, Vos I, et al. Mechanotaxis directs Pseudomonas aeruginosa twitching motility. Proc Natl Acad Sci. 2021 Jul 27;118(30):e2101759118. doi:10.1073/pnas.2101759118

80. Gómez-de-Mariscal E, García-López-de-Haro C, Ouyang W, Donati L, Lundberg E, Unser M, et al. DeepImageJ: A user-friendly environment to run deep learning models in ImageJ. Nat Methods. 2021 Oct;18(10):1192–5. doi:10.1038/s41592-021-01262-9

81. Jo J, Price-Whelan A, Dietrich LEP. Gradients and consequences of heterogeneity in biofilms. Nat Rev Microbiol. 2022 Oct;20(10):593–607. doi:10.1038/s41579-022-00692-2

82. Kobayashi K, Kanesaki Y, Yoshikawa H. Genetic Analysis of Collective Motility of Paenibacillus sp. NAIST15-1. PLoS Genet. 2016 Oct 20;12(10):e1006387. doi:10.1371/journal.pgen.1006387 PubMed PMID: 27764113; PubMed Central PMCID: PMC5072692.

83. Stickler D, Ganderton L, King J, Nettleton J, Winters C. Proteus mirabilis biofilms and the encrustation of urethral catheters. Urol Res. 1993 Dec 1;21(6):407–11. doi:10.1007/BF00300077

84. Hubbard TP, Chao MC, Abel S, Blondel CJ, Abel zur Wiesch P, Zhou X, et al. Genetic analysis of Vibrio parahaemolyticus intestinal colonization. Proc Natl Acad Sci. 2016 May 31;113(22):6283–8. doi:10.1073/pnas.1601718113

85. Kearns DB. A field guide to bacterial swarming motility. Nat Rev Microbiol. 2010;8(9):634–44. doi:10.1038/nrmicro2405

86. Belas R, Schneider R, Melch M. Characterization of Proteus mirabilis Precocious Swarming Mutants: Identification of rsbA, Encoding a Regulator of Swarming Behavior. J Bacteriol. 1998 Dec;180(23):6126–39. PubMed PMID: 9829920; PubMed Central PMCID: PMC107696.

87. Howery KE, Şimşek E, Kim M, Rather PN. Positive autoregulation of the flhDC operon in Proteus mirabilis. Res Microbiol. 2018;169(4–5):199–204. doi:10.1016/j.resmic.2018.02.005 PubMed PMID: 29581007.

88. Furness RB, Fraser GM, Hay NA, Hughes C. Negative feedback from a Proteus class II flagellum export defect to the flhDC master operon controlling cell division and flagellum assembly. J Bacteriol. 1997 Sep;179(17):5585–8. doi:10.1128/jb.179.17.5585-5588.1997 PubMed PMID: 9287017; PubMed Central PMCID: PMC179433.

89. Matsuyama T, Takagi Y, Nakagawa Y, Itoh H, Wakita J, Matsushita M. Dynamic aspects of the structured cell population in a swarming colony of Proteus mirabilis. J Bacteriol. 2000;182(2):385–93.

90. Gygi D, Rahman MM, Lai HC, Carlson R, Guard-Petter J, Hughes C. A cell-surface polysaccharide that facilitates rapid population migration by differentiated swarm cells of Proteus mirabilis. Mol Microbiol. 1995 Sep;17(6):1167–75. doi:10.1111/j.1365-2958.1995.mmi_17061167.x PubMed PMID: 8594335.

91. Armitage JP, Rowbury RJ, Smith DG. The effects of chloramphenicol, nalidixic acid and penicillin on the growth and division of swarming cells of Proteus mirabilis. J Med Microbiol. 1974 Nov;7(4):459–64. doi:10.1099/00222615-7-4-459 PubMed PMID: 4615151.

92. Rauprich O, Matsushita M, Weijer CJ, Siegert F, Esipov SE, Shapiro JA. Periodic phenomena in Proteus mirabilis swarm colony development. J Bacteriol. 1996 Nov 1;178(22):6525–38. doi:10.1128/jb.178.22.6525-6538.1996

93. Morrison RB, Scott A. Swarming of Proteus—A Solution to an Old Problem? Nature. 1966 Jul 1;211(5046):255–7. doi:10.1038/211255a0

94. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. 2006 Oct 31;7(10):R100. doi:10.1186/gb-2006-7-10-r100

95. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods. 2021 Jan;18(1):1. doi:10.1038/s41592-020-01018-x

96. Panigrahi S, Murat D, Le Gall A, Martineau E, Goldlust K, Fiche JB, et al. Misic, a general deep learning-based method for the high-throughput cell segmentation of complex bacterial communities. Xiao J, Storz G, Hensel Z, editors. eLife. 2021 Sep 9;10:e65151. doi:10.7554/eLife.65151

97. Greenwald NF, Miller G, Moen E, Kong A, Kagel A, Dougherty T, et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. Nat Biotechnol. 2022 Apr;40(4):4. doi:10.1038/s41587-021-01094-0

98. Wickham H. ggplot2: Elegant Graphics for Data Analysis [Internet]. Springer-Verlag New York; 2016. Available from: https://ggplot2.tidyverse.org

99. RStudio Team. RStudio: Integrated Development Environment for R [Internet]. Boston, MA: RStudio, Inc.; 2019. Available from: http://www.rstudio.com/

100. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: https://www.R-project.org/

101. Ducret A, Quardokus EM, Brun YV. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. Nat Microbiol. 2016 Jun 20;1(7):16077. doi:10.1038/nmicrobiol.2016.77 PubMed PMID: 27572972; PubMed Central PMCID: PMC5010025.

102. 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 Jun 28;9(7):10.1038/nmeth.2019. doi:10.1038/nmeth.2019 PubMed PMID: 22743772; PubMed Central PMCID: PMC3855844.

103. Schindelin J, Rueden CT, Hiner MC, Eliceiri KW. The ImageJ ecosystem: An open platform for biomedical image analysis. Mol Reprod Dev. 2015;82(7–8):518–29. doi:10.1002/mrd.22489

104. Arganda-Carreras I, Kaynig V, Rueden C, Eliceiri KW, Schindelin J, Cardona A, et al. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. Bioinformatics. 2017 Aug 1;33(15):2424–6. doi:10.1093/bioinformatics/btx180

105. Ma J, Xie R, Ayyadhury S, Ge C, Gupta A, Gupta R, et al. The multimodality cell segmentation challenge: toward universal solutions. Nat Methods. 2024 Jun;21(6):1103–13. doi:10.1038/s41592-024-02233-6

106. Riendeau JM, Gillette AA, Guzman EC, Cruz MC, Kralovec A, Udgata S, et al. Cellpose as a reliable method for single-cell segmentation of autofluorescence microscopy images. Sci Rep. 2025 Feb 14;15(1):5548. doi:10.1038/s41598-024-82639-6 PubMed PMID: 39952935; PubMed Central PMCID: PMC11828867.

107. Pang M, Roy TK, Wu X, Tan K. CelloType: a unified model for segmentation and classification of tissue images. Nat Methods. 2025 Feb 1;22(2):348–57. doi:10.1038/s41592-024-02513-1

108. Torro R, Díaz-Bello B, El Arawi D, Dervanova K, Ammer L, Dupuy F, et al. Celldetective: an AI-enhanced image analysis tool for unraveling dynamic cell interactions [Internet]. 2025 Mar 11. doi:10.7554/elife.105302.1

109. Allison C, Emody L, Coleman N, Hughes C. The role of swarm cell differentiation and multicellular migration in the uropathogenicity of Proteus mirabilis. J Infect Dis. 1994;169(5):1155–8. doi:10.1093/infdis/169.5.1155

110. Gmiter D, Kaca W. Into the understanding the multicellular lifestyle of Proteus mirabilis on solid surfaces. Front Cell Infect Microbiol. 2022 Sep 2;12. doi:10.3389/fcimb.2022.864305

111. Fraser GM, Hughes C. Swarming motility. Curr Opin Microbiol. 1999;2(6):630–5. doi:10.1016/S1369-5274(99)00033-8

112. Little K, Austerman JL, Zheng J, Gibbs KA. Swarming bacteria respond to increasing barriers to motility by increasing cell length and modifying colony structure [Internet]. bioRxiv; 2018 [cited 2022 Apr 29]. p. 398321. Available from: https://www.biorxiv.org/content/10.1101/398321v1 doi:10.1101/398321

113. Belas R, Suvanasuthi R. The Ability of Proteus mirabilis To Sense Surfaces and Regulate Virulence Gene Expression Involves FliL, a Flagellar Basal Body Protein. J Bacteriol. 2005 Oct;187(19):6789–803. doi:10.1128/jb.187.19.6789-6803.2005

114. Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nat Biotechnol. 2002 Jan;20(1):87–90. doi:10.1038/nbt0102-87 PubMed PMID: 11753368.

115. Daniel K, Icha J, Horenburg C, Müller D, Norden C, Mansfeld J. Conditional control of fluorescent protein degradation by an auxin-dependent nanobody. Nat Commun. 2018 Aug 17;9(1):3297. doi:10.1038/s41467-018-05855-5

116. Lee MY, Bedia JS, Bhate SS, Barlow GL, Phillips D, Fantl WJ, et al. CellSeg: a robust, pre-trained nucleus segmentation and pixel quantification software for highly multiplexed fluorescence images. BMC Bioinformatics. 2022 Jan 18;23(1):46. doi:10.1186/s12859-022-04570-9

117. Henrichsen J. Bacterial surface translocation: a survey and a classification. Bacteriol Rev. 1972 Dec;36(4):478–503. PubMed PMID: 4631369; PubMed Central PMCID: PMC408329.

118. Howery KE, Şimşek E, Kim M, Rather PN. Positive autoregulation of the flhDC operon in Proteus mirabilis. Res Microbiol. 2018 May 1;169(4):199–204. doi:10.1016/j.resmic.2018.02.005

119. Howery KE, Clemmer KM, Şimşek E, Kim M, Rather PN. Regulation of the Min Cell Division Inhibition Complex by the Rcs Phosphorelay in Proteus mirabilis. J Bacteriol. 2015 Aug 1;197(15):2499–507. doi:10.1128/JB.00094-15 PubMed PMID: 25986901; PubMed Central PMCID: PMC4518839.

120. MacCready JS, Vecchiarelli AG. In long bacterial cells, the Min system can act off-center. Mol Microbiol. 2018;109(3):268–72. doi:10.1111/mmi.13995

121. Muraleedharan S, Freitas C, Mann P, Glatter T, Ringgaard S. A cell length-dependent transition in MinD-dynamics promotes a switch in division-site placement and preservation of proliferating elongated Vibrio parahaemolyticus swarmer cells. Mol Microbiol. 2018;109(3):365–84. doi:10.1111/mmi.13996

122. Belevich I, Jokitalo E. DeepMIB: User-friendly and open-source software for training of deep learning network for biological image segmentation. PLoS Comput Biol. 2021 Mar;17(3):e1008374. doi:10.1371/journal.pcbi.1008374 PubMed PMID: 33651804; PubMed Central PMCID: PMC7954287.

123. Belas R, Erskine D, Flaherty D. Transposon mutagenesis in Proteus mirabilis. J Bacteriol. 1991 Oct;173(19):6289–93. doi:10.1128/jb.173.19.6289-6293.1991 PubMed PMID: 1655704; PubMed Central PMCID: PMC208382.

124. Rekas A, Alattia JR, Nagai T, Miyawaki A, Ikura M. Crystal structure of venus, a yellow fluorescent protein with improved maturation and reduced environmental sensitivity. J Biol Chem. 2002 Dec 27;277(52):50573–8. doi:10.1074/jbc.M209524200 PubMed PMID: 12370172.

125. Kremers GJ, Goedhart J, van Munster EB, Gadella TWJ. Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Förster radius. Biochemistry. 2006 May 30;45(21):6570–80. doi:10.1021/bi0516273 PubMed PMID: 16716067.

126. Balleza E, Kim JM, Cluzel P. Systematic characterization of maturation time of fluorescent proteins in living cells. Nat Methods. 2018 Jan;15(1):47–51. doi:10.1038/nmeth.4509

127. 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. doi:10.1038/nmeth.2019

128. Hamilton WD. The genetical evolution of social behaviour. I. J Theor Biol. 1964 Jul 1;7(1):1–16. doi:10.1016/0022-5193(64)90038-4

129. Waldman B. The ecology of kin recognition. Annu Rev Ecol Syst. 1988;19:543–71. doi:10.1146/annurev.es.19.110188.002551

130. Natan G, Worlitzer VM, Ariel G, Be'er A. Mixed-species bacterial swarms show an interplay of mixing and segregation across scales. Sci Rep. 2022 Oct 3;12(1):16500. doi:10.1038/s41598-022-20644-3

131. Wang H, Zhu Q, Ding L, Shen Y, Yang CY, Xu F, et al. Scalable volumetric imaging for ultrahigh-speed brain mapping at synaptic resolution. Natl Sci Rev. 2019 Oct;6(5):982–92. doi:10.1093/nsr/nwz053 PubMed PMID: 34691959; PubMed Central PMCID: PMC8291554.

132. Lee RM, Eisenman LR, Khuon S, Aaron JS, Chew TL. Believing is seeing – the deceptive influence of bias in quantitative microscopy. J Cell Sci. 2024 Jan 10;137(1):jcs261567. doi:10.1242/jcs.261567

133. Copeland R, Zhang C, Hammer BK, Yunker PJ. Spatial constraints and stochastic seeding subvert microbial arms race. PLoS Comput Biol. 2024 Jan;20(1):e1011807. doi:10.1371/journal.pcbi.1011807 PubMed PMID: 38277405; PubMed Central PMCID: PMC10849242.

134. Nadell CD, Drescher K, Foster KR. Spatial structure, cooperation and competition in biofilms. Nat Rev Microbiol. 2016 Sep;14(9):589–600. doi:10.1038/nrmicro.2016.84

135. van den Berg N, Thijssen K, Nguyen TT, Cordero M, García Vázquez A, Mitarai N, et al. Emergent collective alignment gives competitive advantage to longer cells during range expansion. Nat Commun. 2025 Dec 19;17(1):1039. doi:10.1038/s41467-025-67791-5 PubMed PMID: 41419731; PubMed Central PMCID: PMC12847927.

136. Rombouts S, Mas A, Le Gall A, Fiche JB, Mignot T, Nollmann M. Multi-scale dynamic imaging reveals that cooperative motility behaviors promote efficient predation in bacteria. Nat Commun. 2023 Sep 11;14(1):5588. doi:10.1038/s41467-023-41193-x

137. Wilbert SA, Mark Welch JL, Borisy GG. Spatial ecology of the human tongue dorsum microbiome. Cell Rep. 2020 Mar 24;30(12):12. doi:10.1016/j.celrep.2020.02.097

138. Mark Welch JL, Dewhirst FE, Borisy GG. Biogeography of the oral microbiome: The site-specialist hypothesis. Annu Rev Microbiol. 2019 Sep 8;73:335–58. doi:10.1146/annurev-micro-090817-062503 PubMed PMID: 31180804; PubMed Central PMCID: PMC7153577.

139. Yaniv K, Gibbs KA. Kin Recognition Systems and Their Role in Multicellular Behaviors. Annu Rev Microbiol. 2025 Oct;79(1):475–95. doi:10.1146/annurev-micro-051724-092527 PubMed PMID: 40834433.

140. Norsworthy AN, Pearson MM. From catheter to kidney stone: the uropathogenic lifestyle of Proteus mirabilis. Trends Microbiol. 2017;25(4):304–15. doi:10.1016/j.tim.2016.11.015

141. Schaffer JN, Norsworthy AN, Sun TT, Pearson MM. Proteus mirabilis fimbriae- and urease-dependent clusters assemble in an extracellular niche to initiate bladder stone formation. Proc Natl Acad Sci U S A. 2016 Apr 19;113(16):4494–9. doi:10.1073/pnas.1601720113 PubMed PMID: 27044107; PubMed Central PMCID: PMC4843424.

142. Jansen AM, Lockatell CV, Johnson DE, Mobley HLT. Visualization of Proteus mirabilis Morphotypes in the Urinary Tract: the Elongated Swarmer Cell Is Rarely Observed in Ascending Urinary Tract Infection. Infect Immun. 2003 Jun;71(6):3607–13. doi:10.1128/IAI.71.6.3607-3613.2003 PubMed PMID: 12761147; PubMed Central PMCID: PMC155743.

143. Rather PN. Swarmer cell differentiation in Proteus mirabilis. Environ Microbiol. 2005 Aug;7(8):1065–73. doi:10.1111/j.1462-2920.2005.00806.x PubMed PMID: 16011745.

144. Garling EE, Li S, Ho K, Gibbs KA. Population heterogeneity revealed in morphometric analysis of densely populated microbial swarm collectives. mBio. 2026 Jan 14;17(1):e0334225. doi:10.1128/mbio.03342-25 PubMed PMID: 41363448.

145. Cavagna A, Cimarelli A, Giardina I, Orlandi A, Parisi G, Procaccini A, et al. New statistical tools for analyzing the structure of animal groups. Math Biosci. 2008 Jul 1;214(1–2):32–7. doi:10.1016/j.mbs.2008.05.006

146. Ballerini M, Cabibbo N, Candelier R, Cavagna A, Cisbani E, Giardina I, et al. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. Proc Natl Acad Sci. 2008 Jan 29;105(4):1232–7. doi:10.1073/pnas.0711437105

147. Gokhale S, Conwill A, Ranjan T, Gore J. Migration alters oscillatory dynamics and promotes survival in connected bacterial populations. Nat Commun. 2018 Dec 10;9(1):5273. doi:10.1038/s41467-018-07703-y

148. Alteri CJ, Mobley HLT. The Versatile Type VI Secretion System. Microbiol Spectr. 2016 Apr;4(2):10.1128/microbiolspec.VMBF-0026–2015. doi:10.1128/microbiolspec.VMBF-0026-2015 PubMed PMID: 27227310; PubMed Central PMCID: PMC4887148.

149. Basler M, Pilhofer M, Henderson GP, Jensen GJ, Mekalanos JJ. Type VI secretion requires a dynamic contractile phage tail-like structure. Nature. 2012;483(7388):182–6. doi:10.1038/nature10846

150. Basler M, Mekalanos JJ. Type 6 secretion dynamics within and between bacterial cells. Science. 2012;337(6096):815. doi:10.1126/science.1222901

151. Bridges AA, Prentice JA, Wingreen NS, Bassler BL. Signal Transduction Network Principles Underlying Bacterial Collective Behaviors. Annu Rev Microbiol. 2022 Sep 8;76:235–57. doi:10.1146/annurev-micro-042922-122020 PubMed PMID: 35609948; PubMed Central PMCID: PMC9463083.

152. Parsek, Greenberg. Sociomicrobiology: the connections between quorum sensing and biofilms. Trends Microbiol. 2005;13(1):27–33.

153. Edelsbrunner H, Harer J. Persistent homology: a survey. In: Goodman JE, Pach J, Pollack R, editors. Surveys on discrete and computational geometry: Twenty years later. Providence: American Mathematical Society; 2008. p. 257–82. (Contemporary mathematics).

154. Foster KR, Parkinson K, Thompson CRL. What can microbial genetics teach sociobiology? Trends Genet TIG. 2007 Feb;23(2):74–80. doi:10.1016/j.tig.2006.12.003 PubMed PMID: 17207887; PubMed Central PMCID: PMC3942651.

155. Popat R, Cornforth DM, McNally L, Brown SP. Collective sensing and collective responses in quorum-sensing bacteria. J R Soc Interface. 2015 Feb 6;12(103):20140882. doi:10.1098/rsif.2014.0882 PubMed PMID: 25505130; PubMed Central PMCID: PMC4305403.

156. Kraemer SA, Velicer GJ. Social complementation and growth advantages promote socially defective bacterial isolates. Proc R Soc B Biol Sci. 2014 Apr 22;281(1781):20140036. doi:10.1098/rspb.2014.0036

157. Daniels M, van Vliet S, Ackermann M. Changes in interactions over ecological time scales influence single-cell growth dynamics in a metabolically coupled marine microbial community. ISME J. 2023 Mar;17(3):406–16. doi:10.1038/s41396-022-01312-w

158. Mathijssen AJTM, Culver J, Bhamla MS, Prakash M. Collective intercellular communication through ultra-fast hydrodynamic trigger waves. Nature. 2019 Jul;571(7766):560–4. doi:10.1038/s41586-019-1387-9

159. Costea PI, Hildebrand F, Arumugam M, Bäckhed F, Blaser MJ, Bushman FD, et al. Enterotypes in the landscape of gut microbial community composition. Nat Microbiol. 2018 Jan;3(1):8–16. doi:10.1038/s41564-017-0072-8 PubMed PMID: 29255284; PubMed Central PMCID: PMC5832044.

160. Gilovich T, Vallone R, Tversky A. The hot hand in basketball: On the misperception of random sequences. Cognit Psychol. 1985 Jul 1;17(3):295–314. doi:10.1016/0010-0285(85)90010-6

161. McCarter L, Silverman M. Surface-induced swarmer cell differentiation of Vibrio parahaemolyticus. Mol Microbiol. 1990 Jul;4(7):1057–62. doi:10.1111/j.1365-2958.1990.tb00678.x PubMed PMID: 2233248.

162. Velasco Ayuso S, Giraldo Silva A, Nelson C, Barger NN, Garcia-Pichel F. Microbial Nursery Production of High-Quality Biological Soil Crust Biomass for Restoration of Degraded Dryland Soils. Appl Environ Microbiol. 2017 Jan 17;83(3):e02179-16. doi:10.1128/AEM.02179-16

163. Shrivastava A, Patel VK, Tang Y, Yost SC, Dewhirst FE, Berg HC. Cargo transport shapes the spatial organization of a microbial community. Proc Natl Acad Sci. 2018 Aug 21;115(34):8633–8. doi:10.1073/pnas.1808966115

164. Du H, Li M, Liu Y. Towards applications of genome-scale metabolic model-based approaches in designing synthetic microbial communities. Quant Biol. 2023 Mar 1;11(1):15–30. doi:10.15302/J-QB-022-0313 PubMed PMID: 41674535; PubMed Central PMCID: PMC12807142.

165. Schaffer JN, Pearson MM. Proteus mirabilis and Urinary Tract Infections. Microbiol Spectr. 2015 Oct;3(5):10.1128/microbiolspec.UTI. doi:10.1128/microbiolspec.UTI-0017-2013 PubMed PMID: 26542036.