DEVELOPMENT OF STATIONARY BANDS DURING ESCHERICHIA COLI SWIMMING

A Dissertation

by

KATRINA SIERRA HOFSTETTER

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Steve W. Lockless
Committee Members,	Deborah A. Siegele
	Joseph A. Sorg
	Sarah E. Bondos
Head of Department,	Thomas D. McKnight

May 2021

Major Subject: Biology

ABSTRACT

Bacteria have evolved to have different mechanisms of motility depending on the environment. On surfaces bacteria can locomote by swarming, a phenotype that is cell density dependent and requires cell coordination. As a result of the cell coordination several strains form stationary patterns. In a low agar environment, cells swim independently through the medium. Bacteria can respond to gradients of chemoeffectors in the environment and will localize to the steepest part of the gradient, creating traveling chemotaxis rings. The formation of stationary patterns, however, is rarely reported for swimming bacteria. Here we uncover a number of mutant strains in *Escherichia coli* K-12 that create a stationary bull's eye pattern in low percentage agar. This pattern consists of repeating regions of high and low cell density. We identified two mutations that can result in formation of this phenotype: one is a disruption of the transcription factor LrhA and the other is insertion of IS elements in the intergenic region between the divergently transcribed genes *flhD* and *uspC*. While mutations in these genes have been identified by others and shown to increase motility in TB swim plates at 30° C, we observed that these mutants form bull's eye patterns in swim plates made from LB and incubated at 37 °C. This pattern can form in a variety of *lrhA* disruptions, however, the location of the IS element in the intergenic *uspC-flhD* region correlates with the formation of either a bull's pattern or chemotaxis rings. IS elements that disrupt the *flhD* promoter correlate with chemotaxis patterns, and those outside of the promoter region correspond with banding pattern formation. We identify the

chemoreceptors necessary for pattern formation and show that the global regulator (p)ppGpp is essential for pattern formation.

DEDICATION

For my husband, without his support and encouragement this dissertation would not exist.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Steve Lockless, and my committee members Dr. Siegele, Dr. Sorg, and Dr. Bondos for their advice and feedback during the course of this research project. I would especially like to thank Dr. Lockless for his endless patience and support, especially when unexpected results and situations arose. I thank Dr. Siegele for inspiring and pushing me to become a better researcher, and Dr. Sorg for his willingness to always discuss experiments. I thank Dr. Bondos, for the many writing and presentation tips.

I am most appreciative of my lab mates: Shian Liu, Sarah Beagle, Jibran Khan, Ashley Hudson, who have become my second family.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Dr. Steve Lockless, advisor, and Drs. Deborah Siegele and Joseph Sorg of the Department of Biology and Dr. Sarah Bondos of the Department of Molecular and Cellular Medicine The concentric banding pattern was first observed by Sarah Beagle, who designed the conditions for the swim plate assay. The testing of amino acids was a collaborative effort between Sarah Beagle and the student. The pBAM1 transposon screen was designed, conducted by Jibran Khan. The screen was analyzed by Jibran Khan and Gregory Whitaker. Chemotaxis mutants were created with the help of Greg Baine, Cierra Weathers, and Rachel Porter. Bolo was designed by student, and, Hannah Gooden, Dr. Lockless, and Dr. Sze of the Department of Computer Science and Engineering. Cartoon figures were made using Biorender. All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by Texas A&M University startup funds to Steve Lockless. This work was also made possible in part by the Welch Foundation under Grant Number A1742, and by the NIH under Grant Number 1R01GM132436. Its contents are 7 solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

NOMENCLATURE

SMM-Na	Swim motility media Sodium Chloride
SMM-K	Swim motility media Potassium Chloride
WGS	Whole-genome sequencing
SCFA	Short chain fatty acids
GSO	Genome Sequencing Origin
AGSO	Adjusted Genome Sequencing Origin
VDB	Variant distance bias

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

Motile bacteria have the advantage to move to more favorable environments and can colonize new niches (1). There are six different categories of motility behaviors: swimming, swarming, gliding, twitching, sliding, and darting (2). For many bacteria, motility behavior is dependent on the composition of the media on which they are grown (3). Bacteria exist as single cells when grown on liquid or low agar conditions, but when incubated on surfaces, cells often coordinate to move together (2, 4).

Different patterns emerge from these collective movements (Figure 1). Pattern formation has been characterized in swarming bacteria (1, 5, 6). For swimming bacteria, movement is not dependent on cell coordination, as a consequence few patterning conditions for swimming bacteria have been identified (7–9). However, several synthetic bacteria can produce patterns (10–12).

Both swarming and swimming rely on the rotation of flagella (2). Here we review known factors that contribute to pattern formation for swarming and swimming bacteria, including the concentric zone pattern formation that occurs in *Proteus mirabilis*, the environmental conditions that are conducive for *Bacillus subtill*s pattern formation, the dendrite pattern formation in *Pseudomonas aeruginosa*, and described stationary swimming patterns in *Escherichia coli*.

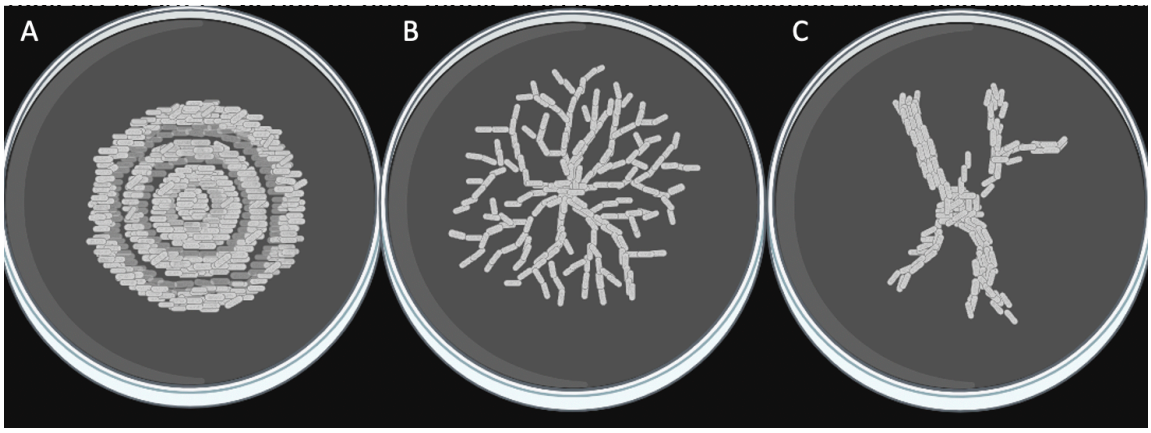

Figure 1. Cartoon of swarming patterns.
A) Bull's eye pattern B) Dense branching motility pattern C) Dendrite pattern

Swarming

Bacterial swarming is commonly observed in firmicutes, alpha-proteobacteria, and gamma-proteobacteria (3, 4). Swarming is triggered by responses to the surrounding environment, including extracellular signaling, cell-cell contact signaling or intercellular changes after cell contact with surfaces (13–15). This collective behavior is characterized by a side-by-side motion, powered by peritrichous or laterally expressed flagella. Due to the high energy demands, swarming occurs on energy rich media, and in a narrow range of agar concentrations. Media made with less than 0.3% agar promotes swimming and media with more than 1% agar will inhibit swarming in some species. Motility on the surface of agar is dependent on the reduction of surface tension of the agar. Bacteria can either use flagellar rotation to extract water from the media or create surfactants to reduce surface tension and move along the top of the agar (4).

During swarming initiation, the cells first undergo a lag phase where individual cells will differentiate into swarmer cells, generally characterized by elongated cells with increased lateral or peritrichous flagella and/or secrete surfactants to reduce the surface tension of the media to allow for swarming behavior (4). The lag phase can be shortened by overexpression of flagella or increasing surfactants depending on the species (4, 16–18). After the lag phase swarmer cells emerge to colonize the surface.

Swarming initiation cues in Vibrio parahaemolyticus

During lag phase cells differentiation into swarm cells. Swarmer cells are elongated, multi-nucleated and have an increase in flagella number. Cells must increase expression of peritrichous flagella or, if they contain a polar flagellum, such as in *Vibrio parahaemolyticus*, upregulate expression of lateral flagella (4). In *V. parahaemolyticus,* lateral flagella are synthesized by *laf* genes, which are upregulated when cells are grown on an agar surface compared to liquid culture (5, 18). Induction of lateral flagella in *V. parahaemolyticus* can be induced by inhibition of polar flagella rotation by increased viscosity of the media, or addition of a flagella antibody or sodium channel blocker, suggesting that flagella are responsible for sensing the surface and triggering the upregulation of lateral flagella (5, 19, 20). During swarming on LB agar, *V. parahaemolyticus* can form a target pattern formed by swarm fronts that stop and undergo consolidation phases that increase the cell density resulting in a visible band of cells (5). While work with *V. parahaemolyticus* has accumulated in a well understood

model of surface sensing, the mechanism behind pattern formation has not been well-studied.

Bull's eye pattern formation in Proteus

Both *Proteus mirabilis* and *P. vulgaris* form a target, or bull's-eye, pattern on swarm agar (21, 22). In the classic pattern, cells differentiate into swarmer cells that are elongated, hyper-flagellated and motile (21, 23). As the swarm front moves along the plate, the rate of spreading decreases before stopping. During the consolidation phase, continued growth occurs behind the swarm front and the swarmer cells dedifferentiate to short and non-motile cells, producing the thick band or terrace (22). From the terrace a new population of swarm cells move forward and the cycle repeats. This pattern of formation is most often found in *P. mirabilis but* has also been identified in *P. vulgaris* (21, 22).

An alternative bull's eye pattern not as readily studied is the ring pattern formation, more commonly found in *P. vulgaris* isolates. In this pattern, rings are thickened regions of growth that occur after the swarm front has passed. The formation of the rings was independent of the speed of the swarm front and could be formed after the swarm front had reached the edge of the plate (22, 24). Descriptions of *Proteus* throughout this dissertation reflect experiments to understand the classic pattern formation, in *P. mirabilis* and are referred to as *P. mirabilis* or *Proteus*.

Flagellar synthesis

Motility is a key part of pattern formation, which requires the synthesis of flagella. Flagellar synthesis is highly organized into three classes of operons Class I, II, III, each class of operons contain genes needed for expression of different flagella components (25). Expression of the flagella genes is regulated so that proteins synthesized in order to efficiently and properly assemble the flagella. The master operon, or Class I operon, in *E. coli* is composed of the *flhDC* genes, and is regulated by several repressors and activators (18, 25–27). FlhD and FlhC form a complex, $FlhD_4C_2$, which binds with sigma-70 upstream of Class II genes and activates transcription. Genes in the Class II operon form the basal body and also express the sigma factor 28 from *fliA* (25, 27) (Figure 2). Sigma-28 is bound in complex with another Class II protein, FlgM, an anti-sigma factor (25). After the basal body and hook are formed, FlgM is exported from the cells, freeing FliA which positively regulates the Class III genes (25, 27) (Figure 2). Class III genes include FliC, the structural unit of the flagellum, (FlaA in *Proteus*), and FliD (FlaD in *Proteus*) which caps FliC, along with chemotaxis genes and the motor proteins (25, 28).

For *Proteus*, the flagella, chemotaxis, and motor genes are clustered in one operon (28). This conformation is not common in bacteria typically, such as in *E. coli*, there are three clusters of operons that consist of different flagellar genes expressed in each class (26). The hierarchy of the Classes of gene expression still exist for *Proteus*, but the clustering of the genes may help to quickly regulate the flagella cells when they switch from swarmer to vegetative cells (28).

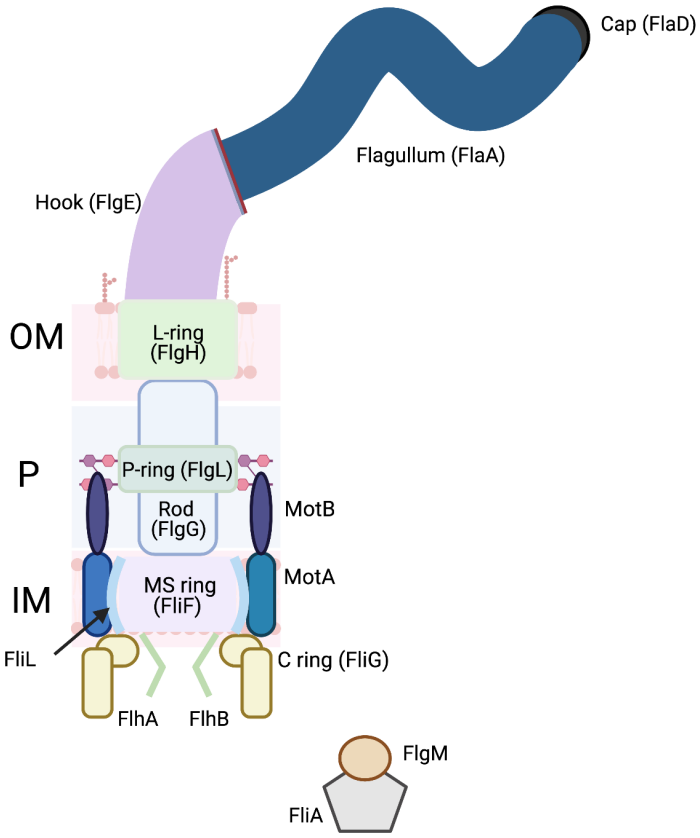

Figure 2. Flagellum structure.
Proteins that are under regulation of class II promoters are light colored, and proteins that are regulated by class III promoters are dark colored. After the basal body is assembled, FlgM is transported out of the cell, and the rest of the proteins under control of class II promoter are assembled. Membrane components are abbreviated outer membrane (OM), peptidoglycan layer (P), and inner membrane (IM). Proteins are named using the *Proteus* designation.

Regulation of flagellar genes

Umo. Four previously uncharacterized genes were identified that increase the expression of *flhDC* (Figure 3). These genes were named *umo* (upregulator of the master operon) *A-D* (29). *Umo* genes are upregulated during swarming, and while mutations in

any of the *umo* genes resulted in a decreased swarm diameter, cells with partial deletions in *umoB* and *umoD* fail to elongate and have reduced swarming diameters (23, 29).

umoB and *umoD* are predicted to work in the same pathway, with *umoD* upstream of *umoB* (29) (Figure 3). UmoD and UmoB work upstream of the Rcs (regulation of capsular polysaccharide synthesis) pathway, and lead to an increase in flagella (14). UmoB negatively regulates the Rcs pathway in *Proteus*, and is similar in sequence to *igaA*, a negative regulator of the Rcs pathway in *Salmonella typhimurium* (14, 29, 30).

Rcs. The Rcs phosphorelay system is composed of the response regulator RcsB, the sensor kinase RcsC, the phosphotransferase RcsD, and the outer membrane activator RcsF. RcsF interacts with IgaA (UmoB in *Proteus*) under outer membrane or peptidoglycan stress. This interaction allows RcsD to be phosphorylated by RcsC. RcsC auto-phosphorylates before RcsD transfers the phosphate group to RcsB (31).

Regulation of genes by the Rcs system can occur by phosphorylated RcsB alone, with a RcsB/RcsA complex, or with RcsB and other auxiliary proteins. The Rcs system can both positively regulate genes such as, *rprA*, a positive regulator of RpoS, and negatively regulating *flhDC* (31, 32). Negative regulation of *flhDC* transcription occurs when RcsB/RcsA directly bind to the +3 and +20 position in the promoter (31) (Figure 3).

Mutations in *rcsC*, *rscD*, or *rscB* result in increased swarming in *Proteus* (33, 34).. Belas et al.,1998 identified mutations in *rcsD* that, while increasing swarming, do

not form terraces or the bull's eye pattern, and predicted that this may occur from disruption to the consolidation phase (33). Clemmer and Rather isolated mutations in *rcsC* and *rcsD* that formed wider terraces (17). Potential differences in the strains could account for the discrepancies of the two phenotypes. The parental strains used in these studies were not identical, nor where the mutations were introduced. Previously strain differences in clinical isolates were observed to produce different patterns (35). However, both reports support that the Rcs system is important for pattern formation and swarming.

Disruption of the *waaL* gene results in strains that are unable to swarm, but mutations in *rscB* are able to rescue swarm defects (14, 34). WaaL is an O-antigen ligase that is needed in the synthesis of LPS (36). The swarm defect was also rescued by mutations in *umoD* or *umoB* (*igaA*), leading to discovery of an additional pathway that bypasses RcsF to phosphorylate RcsB under outer membrane stress. This suggests that the increase in expression of *flhDC* needed for swarming cells can arise from perturbations of the cell envelope (34).

Lrp. Lrp is a transcriptional regulatory protein that regulates amino acid synthesis, and peptide transport (37). Mutations in *lrp* disrupt swarming in *Proteus*, but this loss can be rescued by overexpression of *flhDC* (38). Lrp expression is increased in differentiating swarm cells before peak *flhDC* levels and may be important for integrating information about the nutrient availability during swarmer cell differentiation (18) (Figure 3).

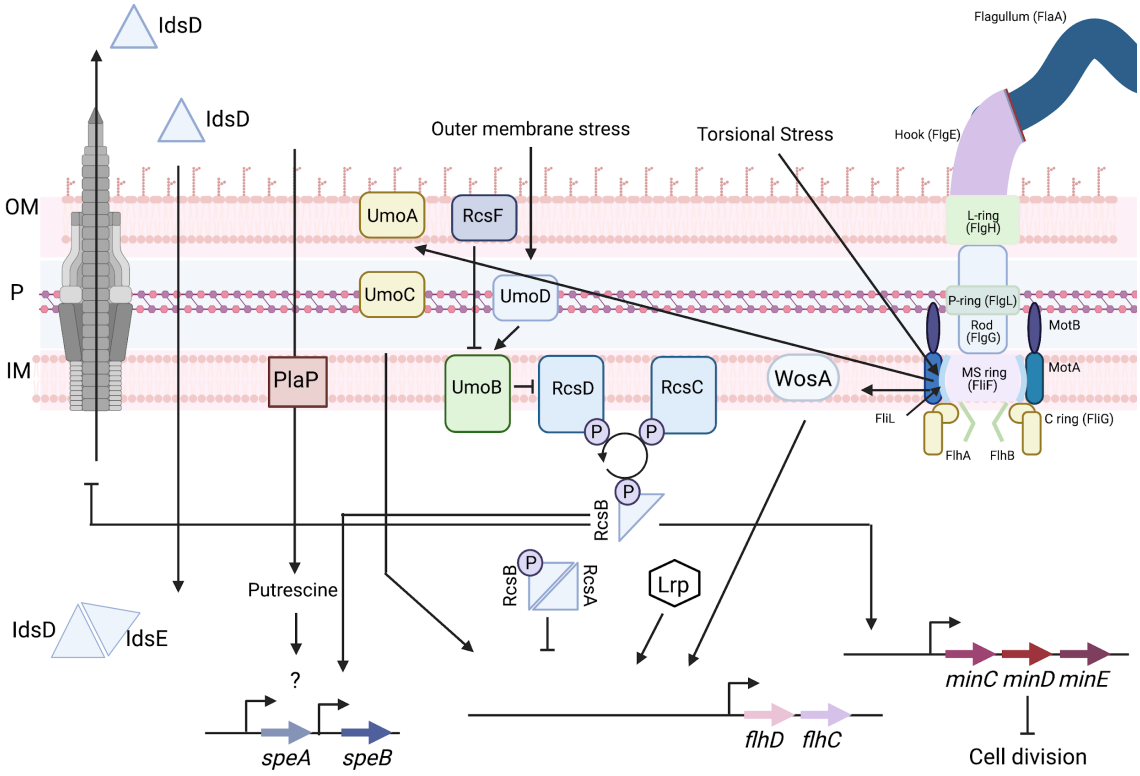

Figure 3. P. mirabilis swarming regulatory network.
IdsD is exported to adjacent *Proteus* by the Type VI secretion system (grey), compatible IdsD from neighboring cells needs to bind with IdsE for cells to be included in the swarm raft. The signaling pathway affected by putrescine is unknown. RcsB is an activator of *speB*. The Umo genes and Lrp are positive regulators of *flhDC*. Phosphorylated RcsB and RcsA negatively regulate *flhDC*. Under outer membrane stress UmoD is activated and leads to UmoB complexing with RcsD preventing phosphorylation from RcsC and decreasing phosphorylated RcsB. This relieves a negative regulator of *flhDC*. Phosphorylated RcsB can activate the *minCDE* operon which leads to inhibition of cell division. Torsional stress is sensed by flagella through FliL, this activates flagella synthesis possible through WosA or UmoA.

Differentiation into swarmer cells

Swarmer cell differentiation can occur on the surface of a semi-solid agar or other surfaces, such as a catheter (39). When the initial swarm front emerges from the inoculation site is dependent on the lag phase. From the inoculum, short cells

differentiate into swarmer cells, and then individual swarmer cells move back and forth until they form a raft with synchronized flagellar movements (40). The amount of time cells spend in lag phase is unaffected by growth rate but can be shortened by increasing the inoculum size or by overexpression of *flhDC* (16, 40).

In *V. parahaemolyticus,* inhibiting flagellar rotation can lead to swarm cell differentiation. It was hypothesized that flagella in *Proteus* would be able to sense the load that occurs when rotating in media conducive for swarming, and regulate differentiation to swarmer cells (5, 13, 41). This hypothesis was tested in *Proteus* where, similar to *V. parahaemolyticus*, inhibition of flagellar rotation with either an anti-FlaA antibody or increases the media viscosity using 15% polyvinylpyrrolidone resulted in differentiation into swarm cells (42).

To identify genes potentially involved in surface sensing, transposon-insertion mutants were screened for genes that resulted in elongated swarm cells under non-permissive conditions. Identified from the transposon screens were mutations in the motor-switch complex (*fliG*) or associated with the basal body (*fliL*) both of which caused swarmer cell differentiation to occur in liquid media (42, 43). FliL in *Salmonella typhimurium* is a cytoplasmic protein anchored in the inner membrane that is a component of the flagella basal body and interacts with the stator proteins MotA and MotB, and the proteins that comprise the M and C rings, FliF and FliG respectively (44, 45).

FliL is needed for swimming or swarming in *Caulobacter* and *Rhodobacter* (46, 47). The role of *fliL* in *Proteus* swarming is muddled by polar effects on downstream

genes caused by the TN5 insertions, or in the intron (TargeTron) insertion. Common to all *fliL* mutants, including the nonpolar deletion, was an elongation phenotype in liquid culture (39, 42, 48). The strains with insertions disrupting *fliL* were defective in swarming, and some were unable to synthesize flagella, likely caused by a large decrease in the amount of FlaA (42, 48). A nonpolar deletion of *fliL* still resulted in an elongated cell, but with upregulated flagella (49). Regulators of *flhDC* are upregulated in the background of both the Δ*fliL* (Y1006), the intron insertion made with the TargeTron knockout system (Y1003), and BB2204, one of the TN5-CM mutations (48–50). Expression of FlaA changes between Y1003 and Y1006. A large decrease of FlaA was observed in Y1003, while in the Y1006 strain, there was a small increase in amount of FlaA present (48, 49). Since both mutations would result in a nonfunctional protein, it is possible that the decrease in *flaA* expression results from an unknown regulatory function of the *fliL* RNA. Alternatively, as the Y1003 flagellar defect could be rescued with a plasmid containing *fliL* and *fliM,* it is possible that polar effects on *fliM* altered flagellar assembly (49). Class II flagellar genes in *Caulobacter* are not expressed in strains with *fliM* mutations (51). Upregulation of flagella in the Y1006 strain is likely to result from the increased regulators of *flhDC*, such as, *umoA* (49).

Another gene that affects flagella production under high viscosity is *wosA*. Based on the sequence of *wosA* it is expected to reside in the inner-membrane and cytoplasm, similar to *fliL*. *wosA* expression was increased in BB2204, and because of *wosA* predicted location in the cell, it is hypothesized that WosA interacts with FliL and is a part of the signaling pathway that triggers swarm cell differentiation (50) (Figure 3).

There is conflicting data regarding the role FliL plays under high torque in *E. coli*. It was hypothesized that FliL may help recruit or stabilize the stators under increased viscosity of the media and increasing torque. However, in contradiction to this finding, Partridge et al found that FilL mutants were not able to increase torque under high viscosity, and Chawla et al., found no defect in torque in Δ*fliL* strains (45, 52).

Loss of *fliL* in *Proteus* and *E. coli* results in a reduction in swimming speed. This reduction can be rescued by mutations made in the MotB plug region, which regulates proton flux (45, 49, 52). These *motB* mutations also restore the swarm defects present in a Δ*fliL Salmonella* strain (45). *motB* mutations also arise as suppressors in *Rhodobacter* that restore the swimming defect (46). These results suggest that FliL interacts with MotB to help regulate proton flux.

Cell elongation

Along with an increase in flagella, swarmer cells also inhibit cell division and filament into multinucleated elongated cells (21, 23, 24, 48, 49). Mutations that result in *Proteus* cells that are either always elongated or unable to elongate are unable to swarm, highlighting the importance of regulation of cell division during swarming and pattern formation (53). It was postulated that flagella genes were also responsible for the regulation of cell division based on the coordinated timing of cell elongation and the increase in flagella expression during swarming (16, 21, 23, 24).

Strains with disrupted flagella synthesis are unable to differentiate, but will elongate if *flhDC* is overexpressed, as shown in the *flhA* disruption strain (16). Increased

expression of *flhDC* also decreased the lag phase and resulted in cells which differentiated faster than wildtype (16, 17). FlhD was also predicted to regulate cell division in *E. coli* (54). However, the changes in cell division predicted to occur by FlhD were strain specific and not caused by the presence or absence of FlhD (55). ChipSeq experiments also failed to identify any binding sites upstream of known cell division genes in *E. coli* (27). Later it was tested if increasing *flhDC* was sufficient to differentiate swarmer cells. However, while the increase of *flhDC* did result in increased flagella but it did not produce differentiated swarm cells in liquid as *rcsD* and *fliL* mutants. This indicated that while upregulation of *flhDC* is necessary it is not sufficient to create swarmer cells (17, 49).

The *rcsD* mutations resulting in differentiated swarm cells in nonpermissive conditions suggest that the Rcs system may also regulate genes involved in cell division. The Rcs system regulates *ftsZ* and *ftsA* in *E. coli*, and was identified as a positive regulator of the Min operon in *Proteus* (56–58). The Min system is responsible for positioning the FtsZ ring in the center of the cell (59). Mutations that disrupt the *min* operon result in mini and larger cells due to uneven division, and increased *min* expression results in filamentous cells (60). RcsB binds in the promoter of *minCDE* to positively regulate expression and inhibit cell division (58) (Figure 3). Increased *min* expression was found to follow the same pattern as *flhDC* expression, high in swarm cells and low during early consolidation, before increasing along the edge of the consolidation terrace in preparation for the next swarm cycle (23, 58).

Consolidation phase

During the consolidation phase, cells dedifferentiate back to vegetative cells. The dedifferentiated cells are non-motile creating a stationary terrace (21, 23, 40). New swarm fronts emerge from the most recently formed terrace, creating a repeating cycle (40). For cells to transition from swarmer to vegetative cells, they must reduce expression of flagella and increase cell division. The signal that the cells sense to trigger dedifferentiation is currently unknown. Reduction in *flhDC* and flagellar gene expression has been correlated with the change to vegetative cells during consolation phases (13, 17, 61). Insertions that disrupt the *flhDC* promoter prevent downregulation of *flhDC* during consolation, likely due to disruption of negative regulatory binding sites (14).

During swarming, amino acids are depleted from the cells (24). In cells collected from the edge of a terrace prior to a swarm, genes involved in amino acid synthesis and uptake, and in cellular respiration are upregulated. It was proposed that the consolidation phase may be a rest period for cells to accumulate energy needed for the next swarm phase (23).

Proteus are able to identify other *Proteus* cells as self and, cells that are not recognized as self are unable to differentiate for the next swarm cycle (62). During swarming the Type VI secretion system (T6SS) is used to transport the IdsD protein into nearby cells (15, 62). The T6SS in *Proteus* may be negatively regulated by the Rcs system (63). In the cell that receives IdsD, the IdsE protein will bind to IdsD if the two proteins are compatible. If the two proteins, IdsD and IdsE, bind together both cells are able to differentiate and swarm. If the IdsE and IdsD do not bind together the cell that

receives IdsD experiences a growth arrest during the consolidation phase and is unable to swarm (62).

Media and Environmental Signals

Proteus can robustly swarm and form terraces on nutrient rich semi-solid media made with yeast extract, tryptone, and salt (40). Minimal media is non-permissible for swarming unless a swarming cue is added, normally CAS amino acids or defined set of amino acids including glutamic acid, aspartic acid, serine, proline, alanine, asparagine, glutamine (35, 64). Cells with insertions into the *rcsD* gene swarm on minimal media lacking swarm cues, but it is not understood how this mutation allows minimal media to be permissive for swarming (33). Swarm cues have different effects in different base media, and not all strains of *Proteus* respond to the same swarm cues (24).

Proteus swarms in a range of agar concentrations. Increasing the agar concentration shortened the length and velocity of the swarm front, but increased the length of the consolidation phase resulting in the same terrace cycle time (40). Increasing the concentration of agar does not change the number of terraces, but does increase biomass (24, 40). *Proteus mirabilis* has been reported to swarm on very high agar concentrations, and inhibition of swarming can require agar concentrations as high as 6% (1, 24). Under low concentrations of agar, *Proteus* swims as single cells through the media, however at an intermediate concentration strain BB2200 is unable to swim or swarm. The loss of motility for this strain occurs at agar concentrations of 0.5% and 0.6%. It is hypothesized that the agar concentration is too high for swimming but too low

to provide the increase the torque needed to trigger cells to undergo swarming differentiation (48).

Another nutrient required for *Proteus* to be able to swarm is putrescine. Two separate strains with mutations in putrescine synthesis have swarming defects (35, 65, 66). Putrescine is imported into *Proteus* by PlaP and synthesized from arginine by *speA* and *speB*, or from ornithine by *speF* (28, 65, 66). Mutations in *speA* or *speB* that disrupt gene function cause a delay in swarming, which can be rescued with the addition of putrescine (35, 65, 66). Activation of *speB* may be regulated by RcsB (63). Putrescine can also be produced from *speC*, and *speA speC* double mutants are unable to swarm. Double mutants in *speA* and *speC* have similar levels of flagellar expression, indicating that putrescine might be needed for cell elongation (14). It has also been suggested that putrescine functions as an extracellular signaling molecule, to allow cell-to-cell communication (65).

Patterns formed by Bacillus subtilis are dependent on environmental conditions

Swarming also occurs in the gram-positive *Bacillus subtilis*, where five unique patterns have been described each dependent on nutrient and agar conditions (6, 67). For agar conditions greater than 0.8%, *B. subtills* does not move by flagellar rotation. Instead, it uses production of extracellular matrix and surfactin to wet the surface of the agar. Cells are predicted to be pushed forward by cell growth (6, 68). In low nutrients environments, thin branches form a pattern called DLA-like. Increasing the nutrient concentration increases the number of branches and their widths compared to the DLA-

like. Inevitably, the large number of branches merge into a solid disk, called an Eden-like pattern (6, 67).

The other growth patterns require flagella (6). Flagella production for swarming in *B. subtilis* is regulated by SwrA and the response regulator DegU (18). SwrA and DegU-P from a complex that binds to regulate *fla/che* operon (69). Disruption of *swrA* inhibits swarming, but swarming can be rescued by upregulation of the *fla/che* operon that is responsible for flagellar basal body transcription. Expression of *swrA* is controlled by two promoters: a sigma-dependent promoter and a sigmaD-dependent promoter. Deletions of either promoter regions promote swarming. SigmaD expression is regulated by SwrA, but SigmaA is regulated by the response regulator DegU, which is phosphorylated by DegS (18). DegQ helps the phosphotransfer from DegS to DegU, and DegR helps stabilize phosphorylated DegU (DegU~P) (70, 71). DegU~P can lead to activation or inactivation of different pathways, depending on the amount present. Unphosphorylated DegU or low levels of DegU~P are needed for swarming motility, while medium levels are needed for biofilm formation. High levels of DegU~P inhibit both processes (70). Understandably, due to the fine tuning of the amount of DegU~P and the different effects on gene expression, the regulation of DegU~P is complex and is still being elucidated. However, glucose and nitrogen starvation are known to regulate *degU* expression (72, 73).

Initiation of swarming in *B. subtilis* demonstrates a cell density-dependent lag phase, which may be required for cells to reach a density sufficient to secrete surfactant (4). Surfactant decreases the surface tension of the agar and allows for flagellar rotation,

and its production is dependent on quorum sensing (1, 4). Swarming in low agar concentrations and low nutrients produces many thin branches that only have actively growing cells at the tip of the branches. This pattern is called the Dense Branching Motile pattern or DBM (6, 67, 74, 75). These cells replicate DNA, synthesize peptidoglycan, and make rRNA, unlike the inactive cells in the stem (74). The inactive cells in the stem swirl, propelling the active cells in the tip forward (67, 76). Increasing the nutrient concentration allows cells to form a concentric pattern under a very narrow agar range (between 0.65% to 0.7%) (6, 74). The concentric pattern forms similar to *Proteus* and consists of a lag phase, and alternating swarming migration and consolidation phases. During the consolidation phase, cell division occurs in the terraces, and cell density increases at the edge of the recently formed terrace until the next swarm phase starts. At the higher end of the agar concentration spectrum, cells swarm for a shorter time, but undergo a longer consolidation phase (74). Decreasing the agar concentration to less than 0.65% forms a Disk-like pattern. Under these conditions, it is proposed that the swarm rate of the cells is faster than the nutrient concentration which results in a solid swarm front that never stops due to low nutrients (6).

B. subtilis pattern formation is based on varying nutrient and agar concentrations has been modeled based on the Fisher equation. Assumptions made in the model that there exist both active and inactive cell populations, where only the active cells can undergo cell-division and are motile, and inactive cells can never become active without the addition of nutrients. This is the basis for the reaction-diffusion model for all patterns. As the concentration of the nutrients available to the cells decreases, the

activity or number of active cells decreases (68). This is observed in the DBM pattern where only active cells are only located at the tip, where the most nutrients are located (75). This model predicts that in the concentric pattern, moving active cells stop when undergoing cell division. The resulting increase in cell numbers and it is this increase in cell numbers that pushes the cells forward and starts the next swarm phase (68). This wave of cells can be seen both under a microscope and when observing the plates of growing cells, and appears as a monolayer of active cells that forms an internal wave (74). The decrease in agar concentration increases the rate of swarming such that nutrients are never depleted which would cause the cells to switch to inactive cells. Furthermore, this model predicts that cell diffusion is dependent on the growth rate at a nutrient concentration (68).

Dendrite patterning in Pseudomonas aeruginosa

P. aeruginosa is a polar flagellated bacterium that is capable of swarming, swimming, and translocating via twitching motility. Twitching motility enables cells to spread over wet agar by using Type IV pili to pull themselves forward. Swimming for these strains occurs under low agar conditions, while swarming occurs on semi-solid surfaces, and twitching on solid surfaces (3).

Swarming *P. aeruginosa* cells produce tendrils that expand over the plate, creating a dendrite like pattern. The cells move from the inoculation center initially in a continuous swarm front, this creates the swarming center. Subsets of cells move from the swarming center in different directions, groups of each of these cells produces a tendril.

Cells along the swarm front are elongated, and most still possess a single flagellum (77–79). Unlike other swarm cells, *P. aeruginosa* does not become hyper flagellated or express lateral flagella (19, 23, 80). Expression of two polar flagella result in a hyper-swarmer strain that swarm faster than the ancestral strain and lacks tendrils (81). Tendrils do not intersect with tendrils from the same swarm center, and will change directions to avoid tendrils from another strain. The loss of interactions between the two populations is mediated by the production of rhamnolipids (79). Rhamnolipids are amphipathic molecules that have a hydrophobic lipid and a hydrophilic sugar, these allow for the surface tension to be reduced. Rhamnolipids are synthesized by first converting fatty acids are to HAA's (3-(3-hydroxyalkanoyloxy)alkanoic acids) by *rhlA*. Mono-rhamnolipids are made from rhamnose sugars and HAAs catalyzed by the rhamnosyltransferase RhlB. RhlC catalyzes the transfer of another rhamnose moiety onto the mono-rhamnolipids creating a di-rhamnolipid (82). The production of rhamnolipids is regulated by the quorum sensing systems LasI-LasR and RhlI-RhlR (83). LasR binds to its autoinducer and promotes expression of RhlR which along with its autoinducer promotes expression of *rhlA* and *rhlB* (82). Mutations in Las or Rhl are non-motile and wild type tendrils can interact (77, 79, 84). A *rhlC* mutant still produces mono-rhamnolipids, which allow for swarming but not tendril formation (79).

Swarming cells compared to cells grown in broth culture have upregulated virulence factors, type III secretion systems, and genes involved in rhamnolipid production, and fatty acid metabolism, and siderophores (pyochelin and pyoverdine) for iron uptake (84). Comparing the gene expression in cells in the tip to the swarm center

showed that cells in the tip had higher expression of ETC proteins, ATPases, and components of the ribosome and protein synthesis genes indicating that cells at the tip are more metabolically active (78). Rhamnolipid production genes were upregulated in the swarm center, but *rhlB* was downregulated in the swarm tip (78, 84). Rhamnolipid gene expression patterns correlated with observed rhamnolipids, which were found to be produced in much larger quantities in the swarm center (85). Iron limiting conditions can induce rhamnolipid production through activation of RhlL-RhlR (86). Pyochelin and pyoverdine, both siderophores that uptake iron, were downregulate in the tip while bacterioferritin was upregulated, suggesting that iron levels were high and had not been taken from the media yet. However, both siderophores were highly upregulated in the swarm center where iron may be limited (78). It is unknown if iron uptake is necessary for swarming since deletion of either siderophores did not have a swarm defect (84). Virulence factors are also under the control of quorum sensing systems and follow a similar pattern of expression with low expression in the tip cells, but high expression in the swarm center (78). Consistent with swarming cells only having one flagellum, flagellar genes were not upregulated compared to growth in liquid culture (84).

Type IV pili that mediate twitching behavior have been proposed to be necessary for swarming, however there are conflicting results on whether swarming is dependent on expression of type IV pili (77, 84, 87). Type IV genes were not upregulated in swarm cells compared to broth, and one operon needed for type IV pili formation was downregulated, suggesting that the requirement of type IV pili may be strain specific (84).

Swimming

E. coli can swarm, but requires low surface tension agar (4). *E. coli* can swim when grown in low agar conditions which allows the cells to swim through the media instead of on the top of the surface as when swarming (2). During swimming cells move a biased random walk, were cells alternate between forward movement (runs) and changes in direction (tumbles) (88). Overall forward movement of *E. coli* is mediated via their chemotaxis system, which comprises receptors located at the pole of the bacteria that are able to bind ligands in the environment (89). After the signaling domain binds a ligand, there is a conformational shift that prevents CheA from phosphorylated CheY. Phosphorylated CheY interacts with the flagellar motor cause flagella to switch to clockwise rotation (90). While moving in a gradient of chemoeffectors cells can alter the time spent in a run by decreasing the frequency of tumbling which causes movement in the desired direction (89).

Patterns formed by Swimming Escherichia coli

When grown in swim media on TCA cycle intermediates, succinate, fumarate, or malate, *E. coli* can form repeating cell aggregates that appear as spots along the plate. These spots are dependent on aspartate sensing, and the addition of aspartate or deletion of *tar,* the receptor that senses aspartate, inhibits this pattern formation. As cells in a swarm front utilize the TCA cycle intermediate, they excrete aspartate, which creates a localized gradient of aspartate (7, 8). As cells aggregate, they produce more aspartate that will attract local bacteria to the cell aggregate, depleting cells from the nearby

region and creating the spots (7). The spots remain stable, as cells in the periphery of the aggregate have an increased probability of tumbling compared to cells in the middle (9). Since cells in the edges of the aggregate tumble, they are unable to escape the aggregate. Only cells that enter the aggregate on a run have a chance of escaping the aggregate and catching up the swarm front. When the aggregates are initially formed, they are motile, but cells that escape rejoin the swarm front and the left behind spots are characterized as non-motile (9).

By using the quorum sensing systems found in *P. aeruginosa*, bacteria can be engineered to produce different patterns. Small clusters of bacteria can make Turing patterns, characterized by clusters of microscopic spots. By using the *P. aeruginosa* quorum sensing system, two strains are engineered that contain either dsRed or GFP under control of a different signaling molecule and transcriptional regulators. When the signaling molecule and the transcriptional regulator bind, they regulate the promoters that contain the dsRed and GFP genes. Since the signaling molecules have different diffusion rates spots of red and green form under the microscope due to cell aggregation (11).

A bull's eye pattern can be formed by creating a system in which cell motility is controlled by cell density. This system was created utilizing the quorum sensing system from *Vibrio fischeri.* The plasmids are designed so that a quorum sensing molecule accumulates when cell density is high. This molecule regulates a second plasmid and represses expression of *cheZ* (10). CheZ dephosphorylates CheY, creating a system where at high cell density a buildup of CheY~P occurs that causes the bacteria to tumble

(91). Concentric rings are formed by a buildup of quorum sensing signaling molecules behind the swarm front. After the amount of signaling molecules exceeds a threshold the increase in CheY~P causes the cells to be non-motile. These non-motile cells produce a stationary band of cells. In front of the band there are cells at a low cell density that have not accumulated CheY~P are continuing to move forward. The band continues to grow until nutrients are exhausted (10).

The engineered strain designed in (10) can be modified by introducing another regulating system. The *P. aeruginosa* quorum sensing system is coupled with the system from *Vibrio fischeri* to produce two sets of alternating stripes. The system is designed so each signaling molecule is produced in both strains under a constitutive promoter. However, each strain contains a different promoter controlling expression of *cheZ*, each expresses *cheZ* when the signaling molecule is bound to the corresponding response regulator. Co-inoculation of these strains produces concentric rings in a similar temporal space. To generate oscillating concentric rings, it was designed so that the strains would inhibit the motility of each other by repressing *cheZ* expression (12).

While fewer swimming patterns have been identified, engineered bacteria show that complex patterns can be formed by swimming bacteria.

Conclusions and Outstanding Questions

Previously, it was hypothesized that overexpression of *flhDC* was the primary regulator for swarm differentiation (13, 16). Though flagella expression is essential to swarm differentiation, the Rcs system integrates almost all genes known to be important

for swarming in *Proteus* (63). What is still unclear is how the conflicting systems needed for swarmer cell differentiation are integrated by the Rcs system. For example, RcsB/RcsA are negative inhibitors of *flhDC*, while RcsB is an activator of the *minCDE* operon. For swarmer cell differentiation, both a derepression of *flhDC* and activation of the *minCDE* system are needed. How the Rcs system integrates both of these systems coordinately under swarming is unknown.

Contrasting hypotheses for what genes are involved swarmer cell differentiation exist. While it is clear that FliL plays an important role in swarming cell differentiation for both *Proteus* and *Vibrio parahaemolyticus,* an alternative hypothesis is that the Rcs system is able to sense the surface by outer membrane stress (14, 34). It is important to determine whether FliL activation is under regulation from the Rcs system, or if these two systems of upregulating swarmer cell differentiation are independent.

In multiple strains of bacteria, we see the same patterns form, the bull's eye pattern and the dendritic pattern both form in two species (1, 4). While the bull's eye pattern forms in multiple strains including engineered *E. coli* (10). The mechanism behind this pattern formation is still unknown. A key component of terrace formation still not understood is what induces the switch between the swarming and consolidation phases.

Very few patterns have been identified on swimming agar, however complex patterns can emerge from engineered bacteria indicating that conditions conducive to pattern formation for swimming bacteria have not been reported. Further characterization of mechanisms important to these pattern formations can produce more

testable hypotheses to identify how bacteria are able to produce self-organized patterns. In this dissertation we describe that *E. coli* form a concentric pattern on nutrient rich swim media. To our knowledge this is the first characterization of this pattern in non-engineered swimming bacteria.

CHAPTER II

BULL'S EYE PATTERN FORMATION IN ESCHERICHIA COLI MUTANTS

Introduction

Bacterial motility is essential for colonization of environmental niches (1). While swimming *E. coli* K-12 move by flagellar rotation, and the directionality of movement, toward attractants and away from repellents, is mediated by the chemotaxis system (89). During the study of chemotaxis cells with increased motility were selected. This led to the identification of mutations in the promoter of the operon encoding the master regulator of flagella synthesis, *flhDC* (92). Mutations have also been identified in the intergenic region between *flhD* and *uspC*, and in known regulators of *flhDC* expression (92–94). Most previous work has been performed at non-physiological conditions, 30 °C typically using TB media.

Cells move in a biased random walk, characterized by short runs (forward movement) followed by tumbles which cause the cells to change direction (88). When in the presence of an attractant or repellent, cells can extend their run times by reducing the likelihood of a tumble. This results in a net gain of forward movement with respect to a chemoeffector gradient (89). The chemoreceptors Tar, Tsr, Tap, Trg, and Aer mediate chemotaxis toward specific chemoeffectors, including aspartate and maltose (Tar), serine (Tsr), dipeptides and pyrimidines (Tap), ribose and galactose (Trg), and oxygen (Aer) (91, 95–99). Ligands bind to chemoreceptors inducing a conformational change that inhibits the histidine kinase *cheA* from autophosphorylating, this prevents the

phosphorolation of *cheY* (90). The phosphorylation state of *cheY* ultimately determines if cells run by rotating flagella counterclockwise or tumble by rotating at least one flagellum clockwise, which occurs when P~CheY interacts with the flagellar motor (90, 100).

In a swim plate made with tryptone as the carbon and energy source, cells will form chemotaxis rings. As cells metabolize available attractants a spatial gradient is created. The cells will localize to the steepest part of the gradient, this can be easily visualized as a sharp ring of cells (101). The first of these rings is formed by the utilization of serine. In the region behind the serine front, cells will adapt and begin to use aspartate (102, 103). The third chemotaxis ring forms on the bottom of the plate due to a lack of oxygen, where cells anaerobically consume threonine (102). Repeated selection of *E. coli* K-12 cells from different locations behind a chemotactic front resulted in the isolation of strains that evolve to expand to different distances from a central inoculation point. These different strains were able to cohabitate, with each dominating in a different region on the plate. This may reflect a way into increase the number of niches available for colonization by the same species (104).

Here we report the isolation of mutants of *E. coli* K-12 (BW25113) with increased motility on LB swim medium at 37 °C. These mutants fell into two classes based on the pattern that formed on swim media. One pattern consisted of chemotaxis rings, formed by the utilization of different amino acids in the media (102). The other pattern consisted of wide bands of high cell density and narrow bands of lower cell density, creating a bull's eye pattern. We identified that for these mutants that the bull's

eye pattern does not form under typical conditions used to study chemotaxis, incubation at 30 °C on TB media.

While IS elements in the *flhD* promoter regions are correlated with chemotaxis patterns at 37 °C on LB swim media, IS elements in the intergentic region of *flhD-uspC* and mutations in *lrhA* resulted in band formation. Disruptions in *lrhA* and the IS elements in the intergenic *flhD-uspC* region each possess a different banding pattern. Here we described the expansion rates of both pattern and chemotaxis formation, along with describing the chemotaxis receptors needed for pattern formation.

Methods

Bacteria strains, and motility assays

The strains used in this study are in Table 1. Bacterial strains were streaked fresh for each experiment from 15% glycerol stocks. Stocks were stored at -80 °C. The growth medium used was LB which contained per Liter; 10 g tryptone (Bacto), 5 g yeast extract (Bacto) 10 g NaCl (Sigma). Kanamycin was used at the final concentration of 50 µg/mL. For swim plates: plates were made using 10 g tryptone, 5 g yeast extract, 10 g NaCl (170 mM) and 2.5 g agar per Liter, referred to here as Swim Motility Medium Sodium Chloride (SMM-Na). Swim Motility Medium Potassium Chloride (SMM-K) was made with 10 g tryptone, 5 g yeast extract, 12.4 g KCl (170 mM) and 2.5 g agar per Liter. TB plates were made using 10 g tryptone, 8 g NaCl and 2.5 g agar per Liter. EZ rich defined media was purchased (Teknova. Inc.) (105) and 10mM of either glucose or glycerol was added as a carbon source. Hepes pH 7 was added to a final concentration of 10mM and

was used to buffer the EZ rich defined media. When present amino acids were added to a final concentration of 10mM. Stocks of amino acids (1M) were adjusted to a pH of 7 using HCl, or KOH (for SMM-K) or NaOH (for SMM-Na). All additives to the media were added after autoclaving. Motility plates contained 25 mL of media, and with the exception of plates with added amino acids where dried unstacked for 90 minutes at room temperature, prior to inoculation. Plates were incubated at 37 °C for 12 hours unless otherwise noted. Plates with added amino acids were stacked and dried for an indeterminate amount of time.

Table 1. Strains used for experiments in Chapter II. Strains in bold have been sequenced.

Strain name	Genotype	Source
KH6	**BW25113 *dgcI*^{L323L} *wrbA*^{G96V} *nemA*^{Y359N}**	CGSC:7926
KH7	**MG1655**	CGSC:6300
KH1220	BW25113 IS***1*** between *flhD-uspC* (IS-320)	Isolate of KH6
KH8	BW25113 Δ*lrhA*::Kan	CGSC:11785
KH318	**BW25113 Δ*lrhA*::FRT**	CGSC:11785 (Kan removed)
KH707	**BW25113 IS*5* between *flhD-uspC* (IS-510)**	Isolate of KH6
KH1	**BW25113 Δ*trkA*::FRT *lrhA*670::IS*1* *yeeJ*^{T552I} *cyoD*104 Del**	CGSC:10452 (Kan removed)
KH10	**BW25113 Δ*kch*::Kan *ubiH*^{V223G} *lrhA*198::IS*5* A to G upstream *yciV***	CGSC:9121
KH1212	BW25113 Δ*lrhA*::FRT IS between *flhD-uspC*	Isolate of KH6
KH711	zfb-223::Tn10	CAG18484
KH761	BW25113 yfbO::tet Δ*trkA*::FRT *lrhA*670::IS*1*::*lrhA*^{wt} *yeeJ*^{T552I} *cyoD*104 Del	P1*vir*(KH711) x KH1
KH898	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	P1*vir*(KH8) x KH6
KH899	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	P1*vir*(KH8) x KH6
KH900	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	P1*vir*(KH8) x KH6
KH901	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	P1*vir*(KH8) x KH6
KH902	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	P1*vir*(KH8) x KH6
KH920	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	P1*vir*(KH8) x KH6
KH907	**MG1655 Δ*lrhA*::Kan**	P1*vir*(KH8) x KH7
KH924	**MG1655 Δ*lrhA*::Kan**	P1*vir*(KH8) x KH7
KH925	**MG1655 Δ*lrhA*::Kan**	P1*vir*(KH8) x KH7
KH818	BW25113 Δ*aer*::Kan	CGSC:12017
CW20	BW25113 Δ*tar*::*Kan*	CGSC:9561
KH946	BW25113 Δ*trg*::Kan	CGSC:9242
CW13	BW25113 Δ*tsr*::Kan	CGSC:11812
CW27	BW25113 Δ*tap*::Kan	CGSC:9560
CW24	BW25113 Δ*aer*::Kan Δ*lrhA*::FRT	P1*vir*(KH818) x KH318
RP079	BW25113 Δ*tar*::*Kan* Δ*lrhA*::FRT	P1*vir*(CW20) x KH318
CW26	BW25113 Δ*trg*::Kan Δ*lrhA*::FRT	P1*vir*(KH946) x KH318
CW16	BW25113 Δ*tsr*::Kan Δ*lrhA*::FRT	P1*vir*(CW16) x KH318
KH1200	BW25113 Δ*aer* IS*5* between *flhD-uspC* (IS-510)	P1*vir*(KH818) x KH707
KH1202	BW25113 Δ*tar* IS*5* between *flhD-uspC* (IS-510)	P1*vir*(CW20) x KH707
KH1198	BW25113 Δ*trg* IS*5* between *flhD-uspC* (IS-510)	P1*vir*(KH946) x KH707
KH1202	BW25113 Δ*tsr* IS*5* between *flhD-uspC* (IS-510)	P1*vir*(CW16) x KH707
CW17	**BW25113 Δ*lrhA*::FRT *nemA*^{Y359N}**	P1*vir*(KH8) x KH6
CW28	**BW25113 Δ*tap*::*Kan* Δ*lrhA*::*FRT rhsD*4066::IS*3***	P1*vir*(CW27) x CW17
RP087	**BW25113 Δtap::Kan ΔlrhA::FRT rhsD4066::IS3**	Isolate of CW28
RP089	**BW25113 Δtap::Kan ΔlrhA::FRT rhsD4066::IS3**	Isolate of CW28
RP091	**BW25113 Δ*tap*::*Kan* Δ*lrhA*::*FRT rhsD*4066::IS*3***	Isolate of CW28

Strain Generation

Deletions of genes were obtained from Keio strains (106) and moved into the desired strains using P1*vir* (107). Transductants were selected on LB agar containing kanamycin (50 µg/mL) and 10mM sodium citrate. Transductants were passaged an additional time on selective media to prevent phage contamination. Transductants were

single colony purified on selective media. Positive transductants were confirmed with PCR.

Isolating strains with increased motility

Cells were taken from an outgrowth and used to inoculate a fresh swim plate. After 12 hours of incubation at 37 °C, a sample was taken from approximately halfway between the center and edge of the plate and streaked for single colonies on LB agar. Cultures for freezer stocks were inoculated with single, isolated colonies. The motility phenotype of the stock cultures was confirmed.

Whole-genome sequencing

Genomic DNA was prepared for whole-genome sequencing (WGS) using a phenol-chloroform extraction protocol (108) with the following modifications: single colonies were used to inoculate 10ml of LB medium in 125mL flasks and grown for 4 h at 37 °C shaking at rpm 250. Cells were pelleted by centrifugation at 4,000g and washed with TE buffer. DNA was either extracted immediately or pellets were flash frozen in TE buffer with 20% sucrose and stored at -80 °C. After extraction genomic DNA was precipitated overnight at -20 °C in 300 mM Na acetate (pH 5.2) and 3 volumes of 90% EtOH.

DNA sequencing was performed at Texas A&M AgriLife Research Genomics and Bioinformatics Service, or the Microbial Genome Sequencing Center (MiGS). Adapters from reads were trimmed using TrimGalore (v0.6.3). Mapping of reads to the reference genome BW25113 (NZ_CP064677.1) was performed with Bowtie2 (v2.3.4.1)

(109)default settings, with the exception of distance between reads, which was set to a max of 3000 using flags –I 0 and –X 3000. Default settings were used for BWA MEM (v0.7.15) (110)alignments, and for SAMtools (1.7) (111) when sorting and indexing bam files. BCFtools (1.9) (112) (using consensus caller flags -cv) and Freebayes (v1.3.2) (113) variant callers were used with ploidy set to 1. IS elements were identified using ISMapper (v1) (114). Positive variants were defined as those with at least 80% of reads containing the alternative allele.

Images and Figures

Images were taken with a Canon M50 camera with settings: shutter speed 1/320, aperture F6.3, ISO 1600, and the EOS Utility application. Light was provided by a SYL 16CM ring light. Plates were imaged on black glass to prevent glare. Images were cropped and converted to greyscale using Microsoft PowerPoint, which was also used to adjust the brightness on images as needed. Genome structure cartoon was created in Abobe InDesign CS6.

Time-lapse images and movies

Time-lapse movies were created by imaging plates every minute for the duration of the pattern formation. Single colonies from fresh streaked plates were grown in 5 mL of LB media until an $O.D._{600}$ of 0.2-0.3. The cell culture was diluted to an $O.D._{600}$ and 0.2 ul was spotted on the center of an SMM-Na plate. The camera and SYL 16CM ring light were placed an incubator lined with black cardboard. The air flow was adjusted to keep the plate at a consistent 37 °C and prevent overheating due to the ring light.

Temperature was tracked using a digital thermometer. Images were taken every minute and to decrease the background noise, due to the LED ring light, the images from every 10-minute interval were averaged together using Matlab 2018b. Averaged images were used for time-lapse figures.

Expansion rates were measured using Matlab 2018b by first identifying the pixel intensity of the inoculation spot after 2.5 hours, before spreading had occurred. Pixel values greater than the pixel intensity of the inoculation spot at 2.5 hours were recorded for each image. Pixel values were normalized to the background pixel intensity from the first image. Likely due to the yellow color of the LB media, the green channel contained more background noise. Therefore, quantifications were made from pixel values recorded in the red channel.

Results

Mutants of E. coli K-12 form a periodic banding pattern on nutrient-rich soft agar

E. coli K-12 strain KH6, which is our lab isolate of BW25113 is non-motile at 37 °C on swim agar plates made with LB (referred to here as Swim Motility Media Sodium Chloride, SMM-Na) (Figure 4A). However, outgrowths of faster spreading cells were frequently seen arising from KH6 (Figure 4A), similar to those described for other poorly motile K-12 strains on motility agar made from TB at 30 °C (92–94). Passaging the outgrowths on fresh swim plates produced two distinct phenotypes: one consisted of chemotaxis rings (Figure 4B), and the other consisted of repeating regions of high cell

density, interspersed with low cell density regions resulting in a bull's eye pattern, referred to here as a banding pattern (Figure 4C). Similar bull's eye patterns have been observed for *Bacillus subtilis*, *Proteus mirabilis*, and *Vibrio parahaemolyticus*, when plated on nutrient-rich swarm agar where the bacteria are moving across a surface (1, 5, 74). However, to our knowledge this is the first reported description of this pattern in a non-engineered, swimming bacterial strain (10).

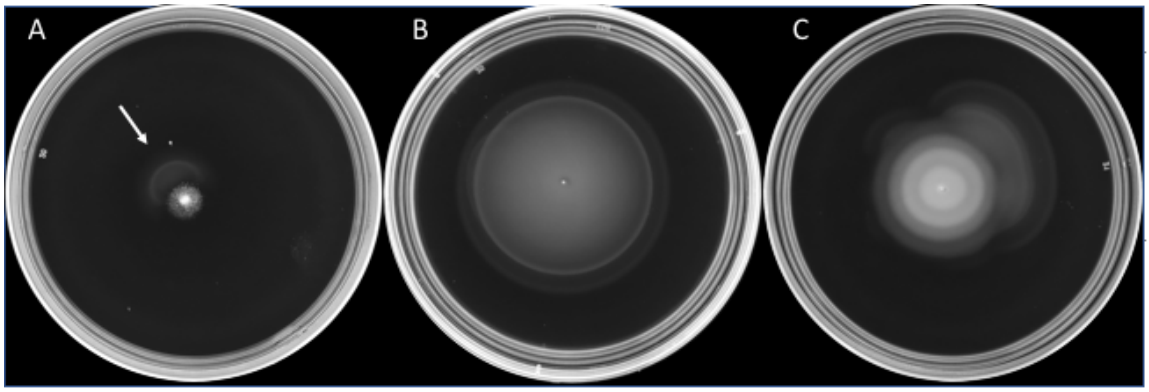

Figure 4. Motility phenotype of chemotaxis rings and banding pattern. Plates A and C were incubated for 12 hours whereas plate B was incubated for 8 hours in SMM-Na at 37 °C. A) Under the conditions tested the parental strain KH6 is non-motile. Outgrowths are indicated by the white arrow B) Canonical chemotaxis pattern observed in KH1220 C) Banding pattern observed in KH318

To identify genes responsible for formation of the banding pattern, a screen of mini-Tn5 transposon mutants in strain KH6 was performed. Screening of 32,004 transposon-positive colonies resulted in identification of 45 strains with increased motility, compared to KH6 which is non-motile at SMM-Na at 37 °C. From the 45 strains with increased motility, 20 formed a banding pattern (115). The location of the transposon insertions was determined using inverse PCR. Of these 20 strains, nine had a

transposon insertion in the gene encoding the transcription factor LrhA (Table 2). Mutations in *lrhA* have been previously identified to increase expression of *flhDC* and other downstream flagella genes, and result in increased motility (94, 116).

We hypothesized that the genes disrupted in the remaining 11 transposon-insertion mutants would be causal for formation of the banding pattern. However, the corresponding genes deletions strains obtained from the Keio collection (106) (Table 3) only five of the Keio strains showed a banding pattern. The deletion mutation from each of these strains was moved into strain KH6 by transduction with phage P1*vir*. The resulting transductants all showed a non-motile phenotype on SMM-Na, leading us to conclude that additional mutations were present and responsible for banding pattern formation in these strains. This result also led us to consider the possibility that some of the mini-Tn5-Kan insertion mutants that showed a banding pattern might also contain additional mutations. To test this possibility PCR screening was used to test for mutations in *lrhA* in the strains that did not contain a mini-TN5-Kan disruption in *lrhA*. This revealed that seven strains contained mutations in *lrhA* (Table 2). For some strains we PCR screened multiple isolates of the Keio stock. The four remaining mutants from the mini-transposon screen did not contain mutations in *lrhA*.

Due to the results just described, we tested the motility phenotype of all other Keio strains in our lab. This revealed 24 additional strains that formed the banding pattern (Table 4). We observed that moving the known gene deletion into KH6 resulted in transductants with a non-motile phenotype, with the exception of Δ*lrhA*771(del)::Kan the transductants of which resulted in a banding pattern. Indicating that these 23 strains

have additional mutations that are causal for the banding phenotype. We PCR screened *lrhA* from these 23 strains and identified that 13 had a mutation in *lrhA*. In summary of the 29 from the Keio banding strains, 18 contained a mutation that caused a banding phenotype *lrhA* (Table 3, Table 4). Of these, 18 strains contained a mutation in *lrhA*, strongly suggesting that mutation in *lrhA* are responsible for the banding phenotype.

Table 2. Genotypes of banding strains obtained from mini-Tn5 transposon screen. Grey boxes highlight the causal mutations for band formation. Strains in bold have been whole-genome sequenced. Strains are separated by banding type

Banding type (BT)	Strain	Strain Origin	*lrhA* Status	*flhD-uspC*
BT1				
	BW25113 mini-Tn5K between *syd-yqcC*	mini-Tn5Kan screen	IS element	
	BW25113 *hybD*432::mini-Tn5Kan	mini-Tn5Kan screen	IS element	
	BW25113 *ybcL*339::mini-Tn5Kan	mini-Tn5Kan screen	pBAM1 Kan insertion	
	BW25113 mini-Tn5Kan between *yegI-yegJ*	mini-Tn5Kan screen	IS element	
	BW25113 mini-Tn5Kan between *ydiP-ydiQ*	mini-Tn5Kan screen	9 BP INS	
	BW25113 *yihN*990::mini-Tn5Kan	mini-Tn5Kan screen	IS element	
	BW25113 *lrhA* 114::mini-Tn5Kan	mini-Tn5Kan screen	Kan insertion	
	BW25113 *lrhA* 112::mini-Tn5Kan	mini-Tn5Kan screen	Kan insertion	
	BW25113 *lrhA* 526::mini-Tn5Kan	mini-Tn5Kan screen	Kan insertion	
	BW25113 *lrhA* 826::mini-Tn5Kan	mini-Tn5Kan screen	Kan insertion	
	BW25113 *lrhA* 512::mini-Tn5Kan	mini-Tn5Kan screen	Kan insertion	
	BW25113 *lrhA* 512::mini-Tn5Kan	mini-Tn5Kan screen	Kan insertion	
	BW25113 *lrhA* 536::mini-Tn5Kan	mini-Tn5Kan screen	Kan insertion	
	BW25113 *lrhA* 112::mini-Tn5Kan	mini-Tn5Kan screen	Kan insertion	
	BW25113 *lrhA* 652::mini-Tn5Kan	mini-Tn5Kan screen	Kan insertion	
BT3				
	BW25113 *hscC*289::mini-Tn5Kan	mini-Tn5Kan screen	mini-Tn5Kan insertion	
	BW25113 mini-Tn5Kan between *yojI-mqo* *rpoD*^{A80S}	mini-Tn5Kan screen	WT Gene Seq	WT Gene Seq
	BW25113 *pqqL* 2613::mini-Tn5Kan	mini-Tn5Kan screen	WT Gene Seq	WT Gene Seq
	BW25113 *yjiM* 604::mini-Tn5Kan	mini-Tn5Kan screen	WT Gene Seq	WT Gene Seq
	BW25113 *xanQ* 605::mini-Tn5Kan	mini-Tn5Kan screen	WT Gene Seq	WT Gene Seq

Table 3. Genotypes of banding strains obtained from the Keio collection, corresponding to a gene disrupted by the mini-Tn5 transposon screen. Grey boxes highlight the causal mutations for band formation. Strains in bold have been whole-genome sequenced. Strains are separated by banding type. Multiple isolates of Banding type 1 strains were isolated from Keio strains.

Banding type (BT)	Strain	Strain Origin	*lrhA* Status	*flhD-uspC*
BT1				
	BW25113 Δ*ybcL*::Kan	JW0533-1	SNP T34S	
	BW25113 Δ*yegJ*::Kan (isolate B5)	JW2056-1	IS*1* element	
	BW25113 Δ*yegJ*::Kan (isolate P9) *lrhA*^{Q278X} A to G *trpE-yciV*	JW2056-1	SNP Q278X	WT Gene Seq
	BW25113 Δ*ydiQ*::Kan (Isolate #1)	JW5276-1	SNP Q278X	
	BW25113 Δ*ydiQ*::Kan (Isolate #10)	JW5276-1	SNP Q278X	
	BW25113 Δ*ydiQ*::Kan (isolate #2) *hemB*^{Q171L} *lrhA*^{Q278X}	JW5276-1	SNP Q278X	WT Gene Seq
	BW25113 Δ*ydiQ*::Kan (Isolate #5) hemB^{Q171L} lrhA^{Q278X}	JW5276-1	SNP Q278X	WT Gene Seq
BT2				
	BW25113 Δ*hscC*::Kan *wbbK*^{C237G} T-C between *rrlG-kgtP basRR*^{207L} IS*5* between *flhD-uspC*	JW0645-1	WT Gene Seq	IS5 -512
BT3				
	BW25113 Δ*xanQ*::Kan A to G *trpE-yciV* 4 base pair deletion between *flhD-uspC*	JW2850-3	WT Gene Seq	4 DEL -656-659
Non-Banding				
	BW25113 Δ*csrB*::FRT-Kan-FRT	ref	WT Gene Seq	WT Gene Seq
	BW25113 Δ*hybD*::FRT-Kan-FRT	JW2961-1	WT Gene Seq	WT Gene Seq
	BW25113 Δ*yhiN*::FRT-Kan-FRT	JW3459-1	WT Gene Seq	WT Gene Seq
	BW25113 Δ*mqo*::FRT-Kan-FRT	JW2198-1	WT Gene Seq	WT Gene Seq
	BW25113 Δ*pqqL*::FRT-Kan-FRT	JW1489-1	WT Gene Seq	WT Gene Seq
	BW25113 Δ*yjiM*::FRT-Kan-FRT	JW5786-1	WT Gene Seq	WT Gene Seq

Table 4. Genotypes of banding strains obtained from the Keio collection, not corresponding to mini-Tn5 transposon gene insertions. Strains are separated by banding type. Grey boxes highlight the causal mutations for band formation.

Strain obtained from Keio collection				
Banding type (BT)	**Strain**	**Strain Origin**	***lrhA* Status**	***flhD-uspC***
BT1				
	BW25113 Δ*lrhA*::FRT-Kan-FRT	JW2284-6	Deletion	
	BW25113 Δ*asr*::FRT-Kan-FRT	JW5826-1	SNP L74I	
	BW25113 Δ*cyoA*::FRT-Kan-FRT	JW0422-1	IS element	
	BW25113 Δ*proC*::FRT-Kan-FRT	JW0377-1	IS element	
	BW25113 Δ*purF*::FRT-Kan-FRT	JW2309-1	29-31 DEL	
	BW25113 Δ*qseB*::FRT-Kan-FRT	JW2993-1	IS*5* element	
	BW25113 Δ*sapB*::FRT-Kan-FRT	JW1286-2	IS element	
	BW25113 Δ*tar*::FRT-Kan-FRT	JW1875-5	IS*2* element	
	BW25113 Δ*yeeJ*::FRT-Kan-FRT	JW5833-1	SNP T18I	
	BW25113 Δ*asmA*::FRT-Kan-FRT *lrhA* 9bpDel A-T between *lysT-valT* T-C between *rrlG-kgtP*	JW2049-1	31-34 DEL	WT Gene Seq
	BW25113 Δ*atoS*::FRT-Kan-FRT *lrhA*^{W173R} *ubiI* 996::IS*1* *T-C between rrlG-kgtP*	JW2213-1	SNP W517R	WT Gene Seq
	BW25113 Δ*clsC*::FRT-Kan-FRT *yabP*^{T120T} *cyoD*^{M1I} *lrhA*^{R60S}	JW5150-1	SNP A178S	WT Gene Seq
	BW25113 Δ*kch*::FRT-Kan-FRT *ubiH*^{V223G} *lrhA198::IS5* A to G *trpE-yciV*	JW1242-1	IS*5* element	WT Gene Seq
	BW25113 Δ*trg*::FRT-Kan-FRTIS*1* -113 upstream *lrhA*	JW1417-1	IS*1* upstream lrhA	WT Gene Seq
	BW25113 Δ*fimB*::FRT-Kan-FRT *mazG*^{A188D} IS*30* between *flhD-uspC*	JW4275-1	WT Gene Seq	IS30 -533
	BW25113 Δ*luxS*::FRT-Kan-FRT T-C between *rrlG-kgtP* IS*1* between *flhD-uspC*	JW2662-1	WT Gene Seq	IS1 -625
	BW25113 Δ*pheA*::FRT-Kan-FRT *hemA*^{A321V} *mutS*^{A353S} *intD 1-48aa* DEL IS*1* between *flhD-uspC*	JW2580-1	WT Gene Seq	IS1 -476
BT2				
	BW25113 Δ*aer*::FRT-Kan-FRT T-C between *rrlG-kgtP* IS*5* between *flhD-uspC*	JW3043-1	WT Gene Seq	IS5 -512
	BW25113 Δ*cyoD*::FRT-Kan-FRT T-C between *rrlG-kgtP* IS*30* *mhp*T *IS*5 between *flhD-uspC*	JW0419-1	WT Gene Seq	IS5 - 512
	BW25113 Δ*metE*::FRT-Kan-FRT *phoA*::IS*5* *intQ*^{E212D} *fimE*::IS*2* IS*1* between *flhD-uspC*	JW3805-1	WT Gene Seq	IS1 -453
	BW25113 Δ*trkE*::FRT-Kan-FRT *fre*^{Q144X} IS*5* between *flhD-uspC*	JW1284-1	WT Gene Seq	IS5 -656
	BW25113 Δ*trkG*::FRT-Kan-FRT *mdtD*^{L441L} *hemG*^{L62Q} IS*1* between *flhD-uspC*	JW2062-1	WT Gene Seq	IS1 -576
	BW25113 Δ*tsx*::FRT-Kan-FRT IS*5* between *flhD-uspC*	JW0401-1	WT Gene Seq	IS1 -658

Banding types

We categorized the banding strains into three groups depending on the number of repeating bands of high cell density and the clarity of the pattern. Banding type 1 is characterized by a regular repeating pattern of bands with clearly demarcated regions of low cell density, i.e. the interbands (Figure 5C). Banding type 2 also has a repeating pattern of bands, but the boundaries between the bands and inter-bands are less distinct (Figure 5D). Banding type 3 is characterized by strains that formed one band (Figure 5E). Banding type 1 is correlated with mutations in *lrhA*. Banding type 2 is correlated with strains with mutations in the *flhDC-uspC* intergenic region. The banding type 3 strains did not contain mutations in *lrhA*. One banding type 3 strain contained a four base pair deletion in the intergenic region of *flhD-uspC*, and we have not identified the causal mutation in the remaining banding type 3 strains. The rest of the experiments to be reported focused on banding types 1 and 2.

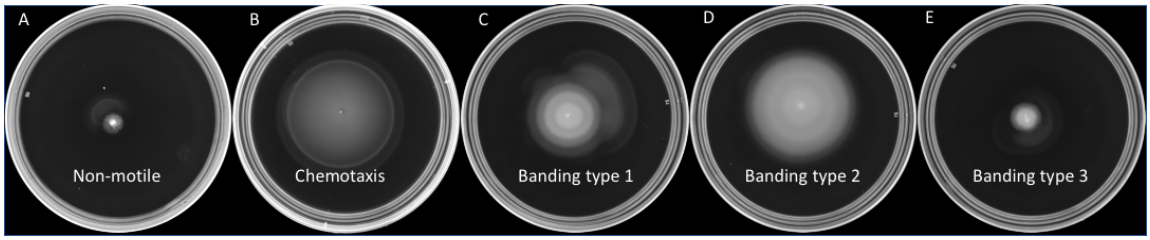

Figure 5. Motility phenotypes.
Plates A,C,D,E were incubated for 12 hours and plate B, was incubated for 8 hours in SMM-Na at 37 °C A) non-motile for BW25113 B) Canonical chemotaxis pattern observed in KH1220 C) Banding type 1 pattern representative strain KH318 D) Banding type 2 representative strain KH707 E) Banding type 3 representative strain KH938

Causal mutations for banding patterns

LrhA is a transcription factor that negatively regulates *flhDC*, the master regulator for flagellar synthesis (27, 116). Deletion of *lrhA* has been shown to increase the number of flagella by three-fold in MG1655 (117). Due to the frequency of *lrhA* mutations observed in our banding strains, and because the deletion of *lrhA* results in a banding phenotype, we hypothesized that disruption of *lrhA* was causal for band formation. Restoring the *lrhA*$^{+}$ allele in strains containing a *lrhA* mutation returned the wildtype non-motile phenotype (Figure 6). To confirm that *lrhA* mutations were sufficient to cause banding, the Δ*lrhA* (JW2284-6) was moved by transduction into wild-type strains. Wild-type strains MG1655 and BW25113 were used to test if the Δ*lrhA* resulted in a banding phenotype was strain specific. 10 independent transductions were preformed, three into MG1655 and seven into BW25113. Of 67 transductants were colony purified and the motility phenotype was tested, of the 67 transductants 66 of 67 formed a banding pattern (Figure 7). Whole genome sequenced (WGS) of seven independent transductants confirmed that the only mutation present in these strains was the Δ*lrhA* compared to our respective parental strains (Table 2). Confirming that the loss of *lrhA* alone is sufficient to induce the banding pattern.

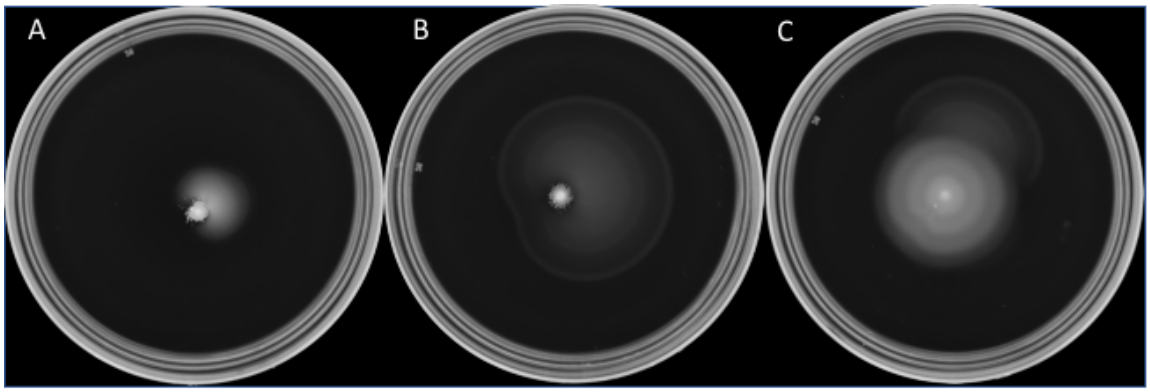

Figure 6. Restoring lrhA results in a parental phenotype.
A) KH6 B) KH1 with restored *lrhA* C) KH1 with *lrhA*670::IS*1*

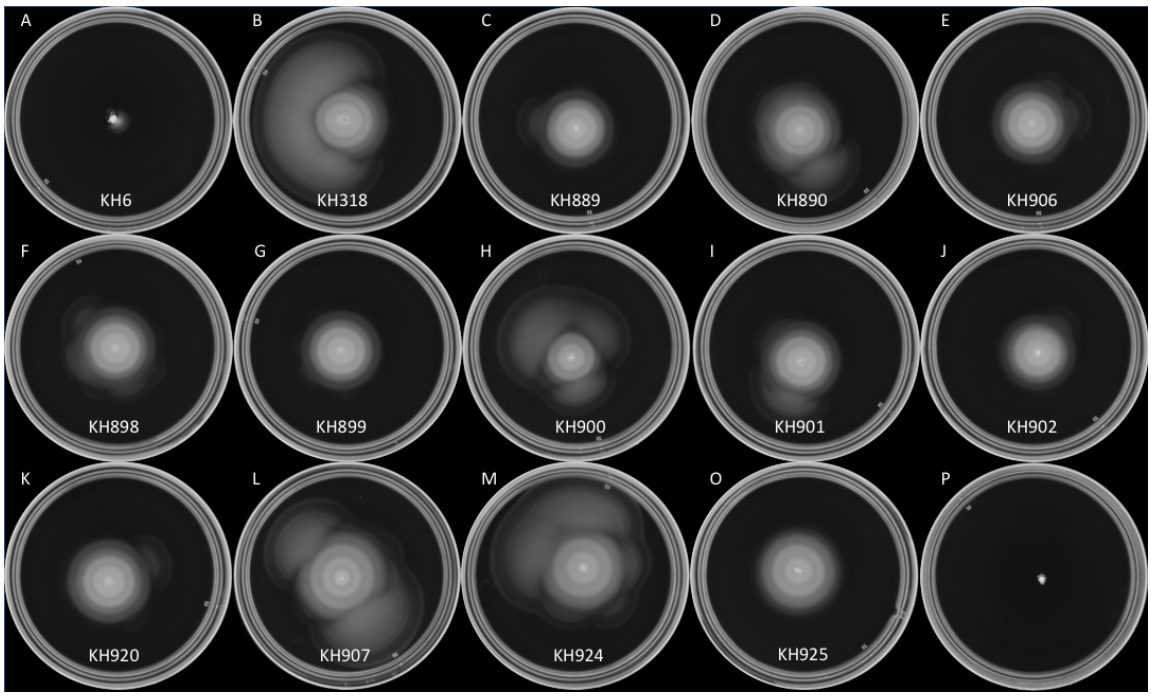

Figure 7. Deletions of lrhA exhibit a banding phenotype.
Plates were incubated 14 hours at 37 °C in SMM-Na. Strain names are listed below pictures. A) KH6 B) KH318 (BW25113 Δ*lrhA*::FRT) C-E) isolates from KH318 that were WGS F-K) BW25113 Δ*lrhA*::Kan isolates that were WGS L-O) MG1655 Δ*lrhA*::Kan isolates that were WGS. P) MG1655 (KH7)

Of the 51 banding strains that we had isolated 16 had a wild type *lrhA* allele. WGS of these 16 strains revealed that the majority (12/16) contained an IS element inserted into the intergenic region between the divergently transcribed genes *flhD* and *uspC* (Table 2,3,4; Figure 8). Nine strains with IS elements in the *flhD-uspC* region resulted in a banding type 2 phenotype. The IS insertion locations were clustered within 350 base pairs of the *uspC* start codon or more than 500 base pairs from *flhD* coding sequence (Figure 8, Table 5). It is unknown if these mutations affect expression of *uspC*, *flhD* or both. Inactivation in *uspC* results in the loss of flagella, however it is unknown if the effect is direct or indirect (118)

We selected for motile mutants from KH6 on SMM-Na and identified two banding and three chemotactic strains that contained IS elements between *flhD-uspC*. Strains that formed chemotaxis bands contained IS elements closer to *flhD*, while strains exhibiting banding contained IS elements near *uspC* (Figure 8; Figure 4 B and D).

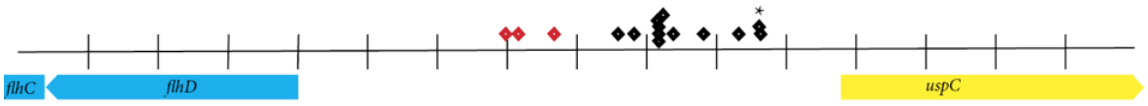

Figure 8. Schematic of mutations identified in *flhD-uspC* intergenic region. Distance between tick marks is equal to 100 base pairs. Diamonds represent location of IS element insertion. Red diamonds are IS element insertion locations for strains that exhibit chemotaxis rings, while black diamonds are locations for banding strains. Asterisk corresponds to nucleotide deletions sites

Table 5. Location of mutations identified in *flhD-uspC* intergenic region. Genomic location and distance mutations occurred from *flhD* and *uspC* are reported.

Strain name	Primary deletion	Location	Distance from *flhD*	Distance from *uspC*	Motility phenotype
KH707		1972964	-510	-270	Banding type 2
KH1213		1972969	-515	-265	Banding type 2
KH1217		1972746	-292	-488	Chemotatic
KH1218		1972817	-363	-417	Chemotatic
KH1220		1972774	-320	-460	Chemotatic
KH966	Δ*fimB*	1972987	-533	-247	Banding type 1
	Δ*luxS*	1973079	-625	-155	Banding type 1
KH408	Δ*pheA*	1972930	-476	-304	Banding type 1
KH970	Δ*trkG*	1973030	-576	-204	Banding type 2
KH896	Δ*metE*	1972907	-453	-327	Banding type 2
KH926	Δ*hscC*	1972966	-512	-268	Banding type 2
	Δ*aer*	1972966	-512	-268	Banding type 2
	Δ*tsx*	1973112	-658	-122	Banding type 2
KH921	Δ*trkE*	1973110	-656	-124	Banding type 2
KH869	Δ*cyoD*	1972966	-512	-268	Banding type 2
	Δ*xanQ*	1973110	-656-(-659)	-124-(-121)	Banding type 3

Pattern formation is dependent on yeast extract and incubation at 37 °C

The banding type 1 pattern formation only occurred on media made with yeast extract and incubated at 37 °C. Incubation at below 37 °C resulted in fainter bands with fainter demarcations in the bands, until the bands fully disappeared at 30 °C (Figure 9B).

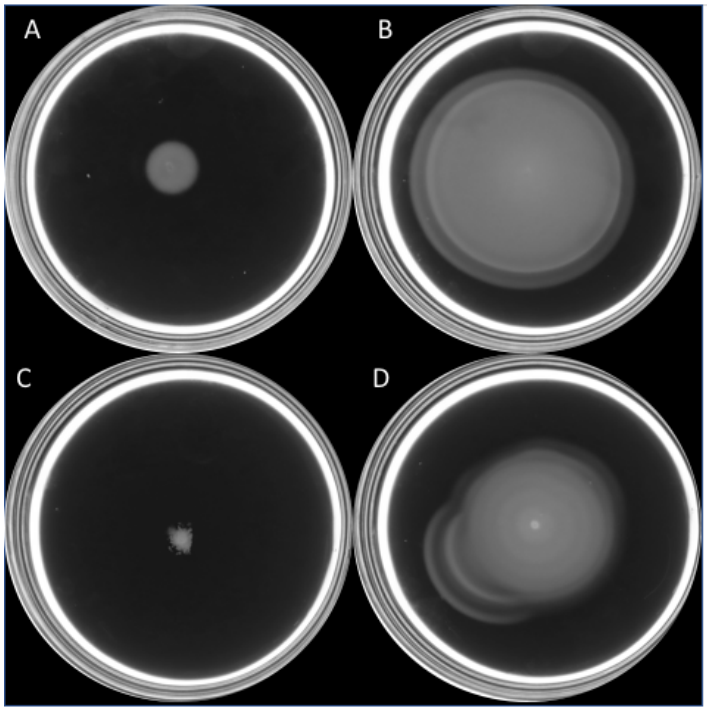

Figure 9. Banding type 1 pattern is dependent on incubation at 37 °C A-B) plates incubated at 30 °C C-D) plates incubated at 37 °C A and C) BW25113 (KH6) at 37 °C KH6 is non-motile but is motile at 30 °C B and D) KH318 is banding at 37 °C but is non-banding at 30 °C.

To test how individual components affected the banding pattern we tested for band formation on a rich defined media, however, it was not able to support banding (Figure 10). Based on this observation, we hypothesized that band formation is dependent on yeast extract. Media made with only yeast extract and salt (tryptone excluded from SMM-Na media) is sufficient for cells to form the banding pattern (Figure 11). In the absence of tryptone the bands became wider than those formed in SMM-Na (Figure 11). Increasing the amount of yeast extract in SMM-Na did alter the pattern but did not completely obscure the pattern (Figure 11). These experiments show

that the cells are utilizing a component of the yeast attract to form the banding pattern and that amino acids from tryptone can alter the pattern.

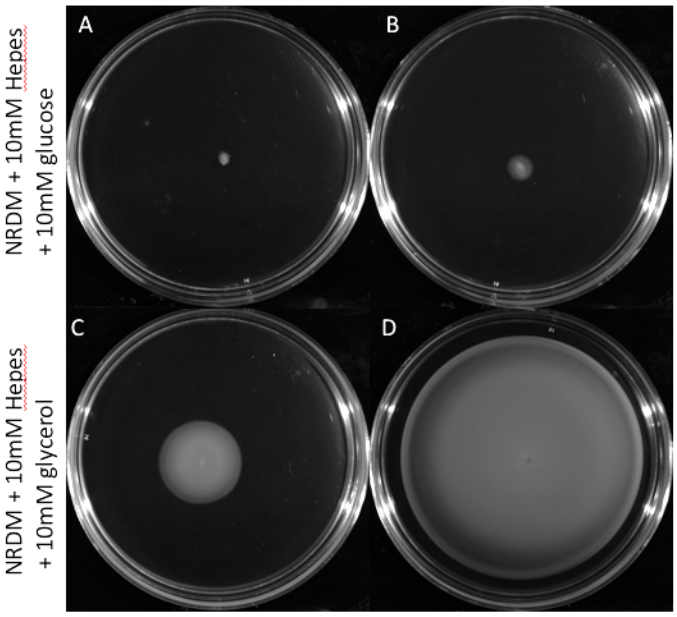

Figure 10. Banding type 1 strains are unable to band on EZ rich defined media. A-B) media made with glucose as a carbon source, both KH6 and KH318 were non-motile C-D) media made with glycerol as a carbon source A and C) KH6 media made glycerol support motility at 37 °C B and D) KH318 was non-banding at 37 °C.

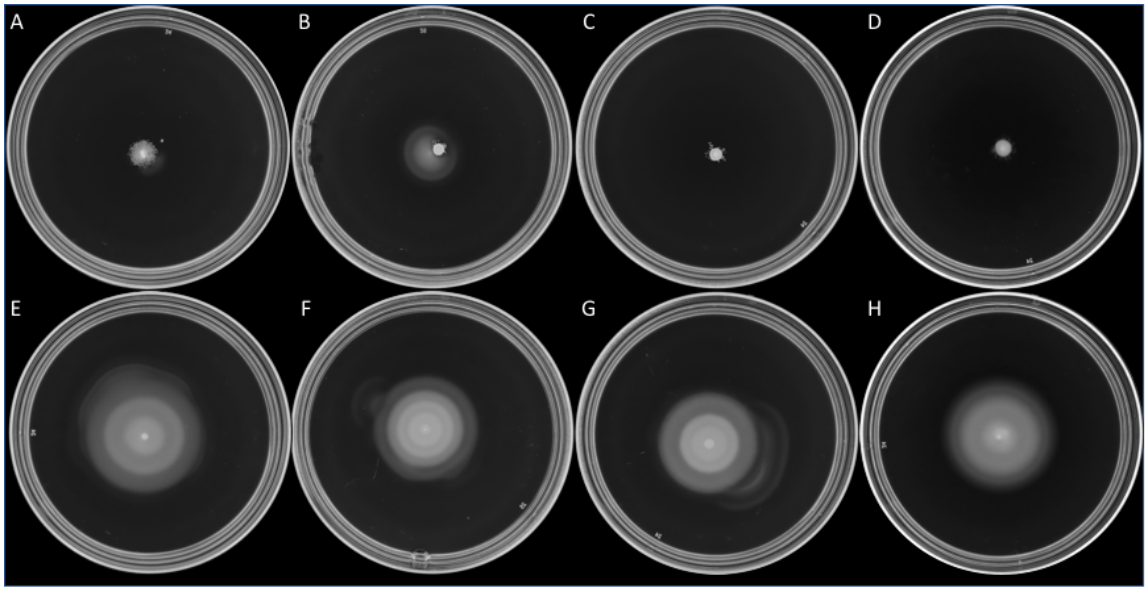

Figure 11. Increasing yeast extract concentration alters banding type 1 formation. A-C) KH6 D-G) KH318 (Δ*lrhA*) A and E) media made with yeast extract (5g/Liter) and NaCl (10g/Liter) supports banding B and F) SMM-Na C and G) SMM-Na with 2X yeast extract D and H) SMM-Na 10X yeast extract

To test for important components in yeast extract that could alter banding we added in excess nutrients present in yeast extract to SMM-Na. Due to the effect of banding with and without tryptone present in the media we tested which amino acids were able to alter the banding pattern. 19 amino acids were added individually to the SMM-K (Swim Motility Media Potassium Chloride) media. These experiments were done with KH10 (Δ*kch lrhA*198::IS*5*) and KH1 (Δ*trkA lrhA*670::IS*1*). The addition of 10mM serine inhibited the band formation, resulting in motile but non-banding phenotype. Serine was confirmed to affect both KH318 (Δ*lrhA*) and KH707 (*flhD* IS-510) (Figure 12). When KH1220 and KH707 was grown in the presence of added serine the serine chemotaxis ring was intensified. Perhaps, the addition of serine could change

the metabolism of the cell, as serine is the first amino acid utilized from the media, or it could change the chemotactic behavior of the cells (101, 103).

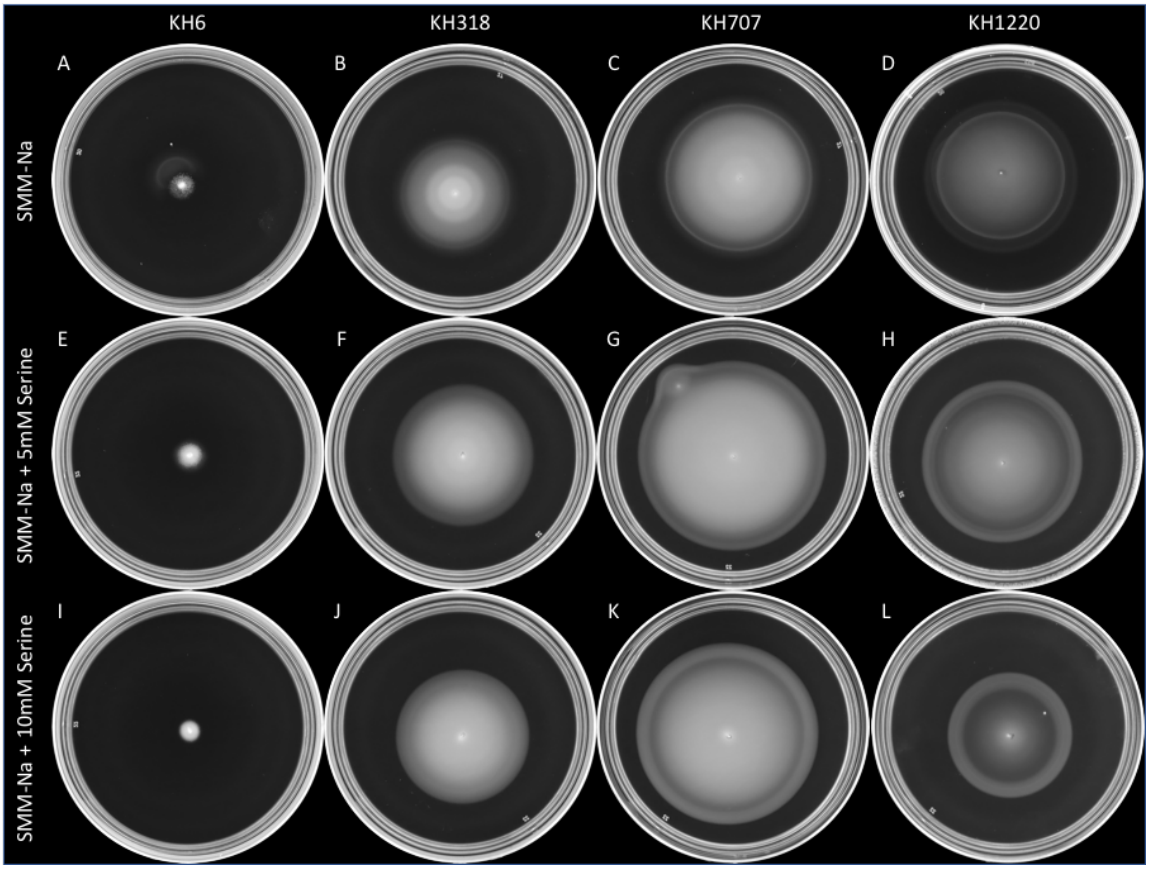

Figure 12. Supplementing media with serine obscures pattern formation. A-D) SMM-Na E-H) SMM-Na with 5mM serine I-L) SMM-Na with 10mM serine A,E and I) KH6 is non-motile under all conditions B,F and J) KH318 faint bands can be observed at 5mM serine (F) C,G and K) KH707 non-banding swimming can be seen at 5mM serine (G), presence of an expanded serine ring is seen at 10mM (K) D,H and L) KH1220 plates incubated for 8 hrs. Addition of 5mM serine (H) intensifies the serine chemotaxis ring. At 10mM this ring is expanded further.

Chemotaxis is necessary for band formation

The addition of serine to SMM-Na resulted in a motile by non-banding swimming pattern. To test if this was due to the sensing of serine or the utilization of serine we made a *tsr* deletion in the KH318 and KH707 background strains, banding types 1 and 2 respectively. The mutant strains resulted in a motile but non-banding swimming phenotype (Figure 13). It is possible that the additional serine saturated the Tsr receptor and prohibited the cells from using Tsr to sense gradients (89). To test if other chemoreceptors were essential to band formation, we individually deleted the four remaining receptors, *tap*, *tar*, *trg*, and *aer* from the two banding type strains and tested the motility phenotype of the double mutants. While we saw no change in the presence of banding in the absence of Aer or Trg, we identified that the deletion of *tap* resulted in non-motile phenotype for both strains (Figure 13). To ensure that the loss of *tap* merely did not slow the strains spreading, the motility plates were incubated until outgrowths appeared.

If the formation of the two banding phenotypes (bands 1 and bands 2) are the same, we predict that they would utilize the same chemoreceptors. Deleting the 5 chemoreceptors (*aer*, *tap*, *tar*, *tsr*, *trg*) in the KH707 background revealed that this pattern is dependent on the presence of *tar* in addition to, *tsr* and *tap* (Figure 13). In the KH318 background, *tar* does not cause a loss of banding, but does change the width of the bands. KH707 also colonizes a larger area of the plate than KH318, indicating that different chemoreceptors and ligands could be utilized by the strains to colonize different distances throughout the plates.

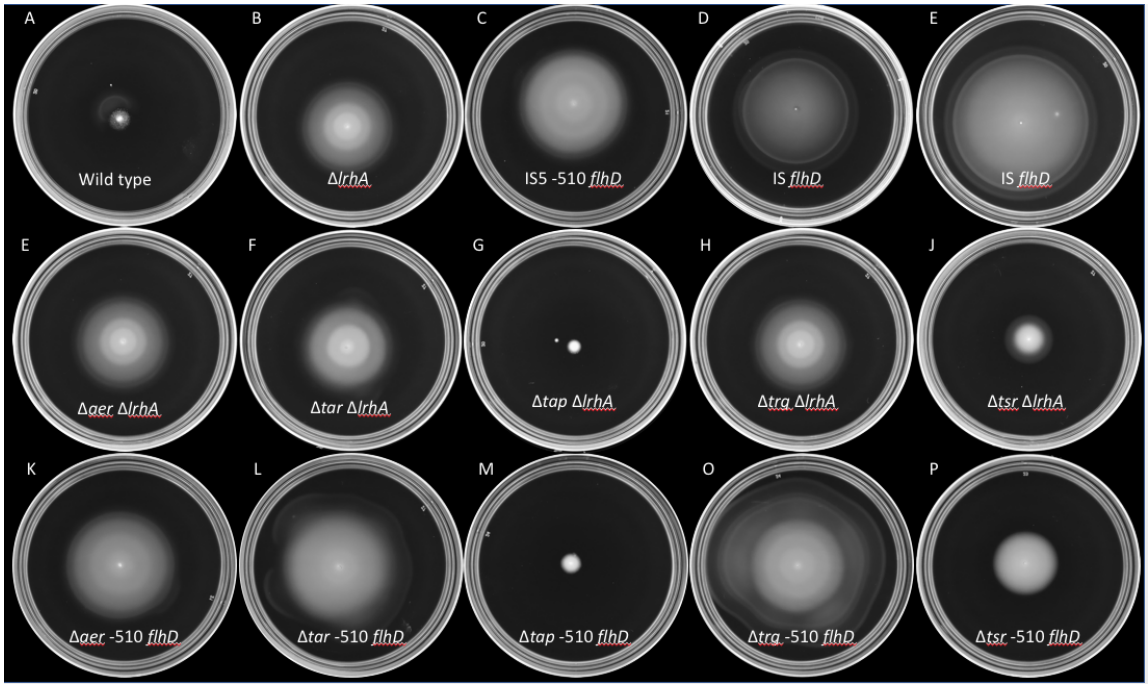

Figure 13. Tap and Tsr are needed for both banding type 1 and banding type 2 in SMM-Na.
Banding type 2 also requires Tar. Relevant genotypes listed in figure. A) KH6 B) KH318 (BW25113 Δ*lrhA*) C) KH707 IS5 (-510 *flhD*) D) KH1220 incubated 8 hours E) KH1212 (Δ*lrhA* +IS) incubated 8 hrs. F) CW24 (Δ*aer* Δ*lrhA*) G) RP079 (Δ*tar* Δ*lrhA*) H) CW28 (Δ*tap* Δ*lrhA*) I) CW26 (Δ*trg* Δ*lrhA*) I) CW16 (Δ*tsr* Δ*lrhA*) K) KH1200 (Δ*aer* Δ*lrhA*) L) KH1197 (Δ*tar* Δ*lrhA*) M) KH1202 (Δ*tap* Δ*lrhA*) O) KH1198 (Δ*trg* Δ*lrhA*) P) KH1202 (Δ*tsr* Δ*lrhA*)

To test if the phenotypic changes were due to the strains or to the media condition we tested all double mutations under traditional chemotaxis conditions, 30 °C in TB media (Figure 14). Here we see that KH318 and KH707 are non-motile in the absence of *tap* and non-banding by motile swimming in the absence of *tsr* similar to the SMM-Na phenotype. However, the loss of *tar* had the same effect in both strains and resulted in the loss of the second chemotaxis ring (Figure 14). The TB results indicate

that the loss of banding in Δ*tar* KH707 in SMM-Na is media specific and in the KH318 and KH707 strains the loss of Tap is detrimental to motility.

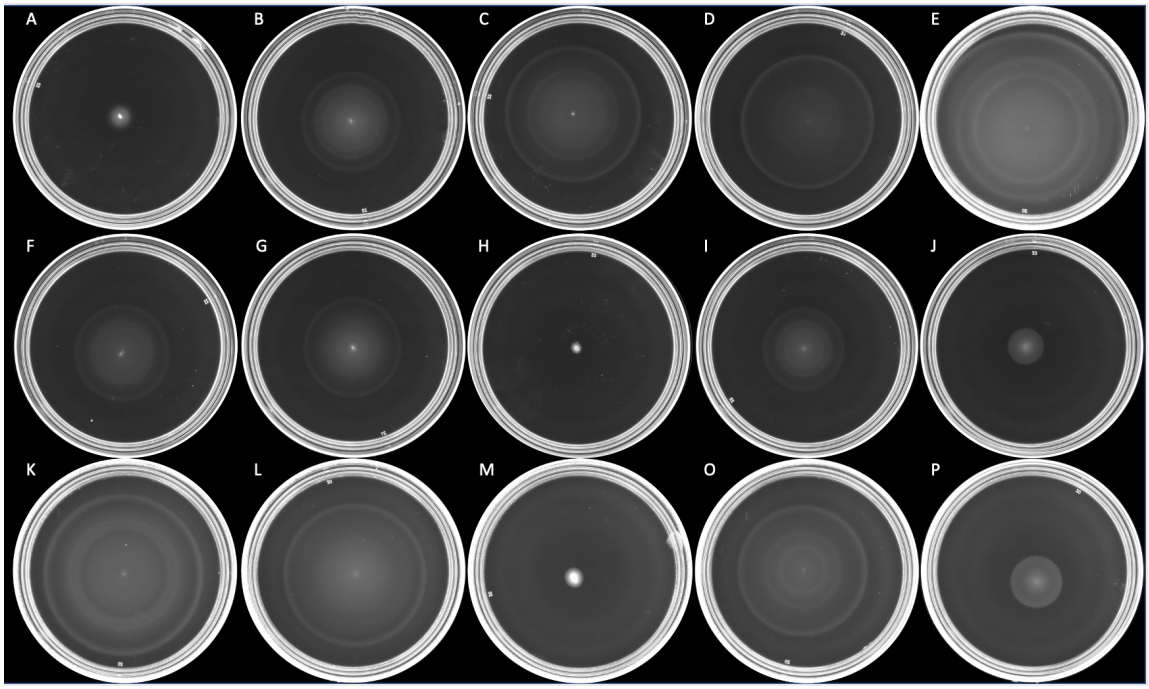

Figure 14. Tap and Tsr are needed for banding in TB 30 °C.
Plates incubated 8 hours at 30 °C in TB media. Relevant genotypes listed in figure. A) KH6 B) KH318 (BW25113 Δ*lrhA*) C) KH707 IS5 (-510) D) KH1220 E) KH1212 (Δ*lrhA* +IS) F) CW24 (Δ*aer* Δ*lrhA*) G) RP079 (Δ*tar* Δ*lrhA*) H) CW28 (Δ*tap* Δ*lrhA*) I) CW26 (Δ*trg* Δ*lrhA*) I) CW16 (Δ*tsr* Δ*lrhA*) K) KH1200 (Δ*aer* Δ*lrhA*) L) KH1197 (Δ*tar* Δ*lrhA*) M) KH1202 (Δ*tap* Δ*lrhA*) O) KH1198 (Δ*trg* Δ*lrhA*) P) KH1202 (Δ*tsr* Δ*lrhA*)

In addition to sensing pyrimidines, Tap senses dipeptides through interactions with the dipeptide transporter DppA (97, 99). To test what was affected by the Δ*tap* we passaged outgrowth from Δ*tap* Δ*lrhA* and selected banding reverents to be WGS. Sequencing of these strains identified an IS3 element in *rhsD*. The banding revertants, however, did not have any fixed mutations present. We are unable to conclude what caused the loss of banding in the Δ*tap* strains.

Tap is not functional as a sole chemoreceptor present in the cell, it requires the presence of either Tsr or Tar. A *tsr* mutant had the best response to dipeptides (97). Indicating that Tap may prefer to utilize Tar when sensing dipeptides, and that the loss of Tar in the two banding strains could affect Tap function. Tar senses both maltose and aspartate, and the addition of maltose to the SMM-Na media also results in wider bands in the KH318 strain, similar to RP079 (Δ*tar* Δ*lrhA*) (Figure 13, 15). Since a saturated chemotaxis receptor is not able to utilize a gradient of effector molecules (89). The changing in the bands in RP079 is likely due to a loss of Tar sensing likely aspartate. The loss of Tar in the KH707 caused a loss of bands, indicating that the chemoreceptors and ligands utilized for the banding pattern differ between bands type 1 and bands type 2.

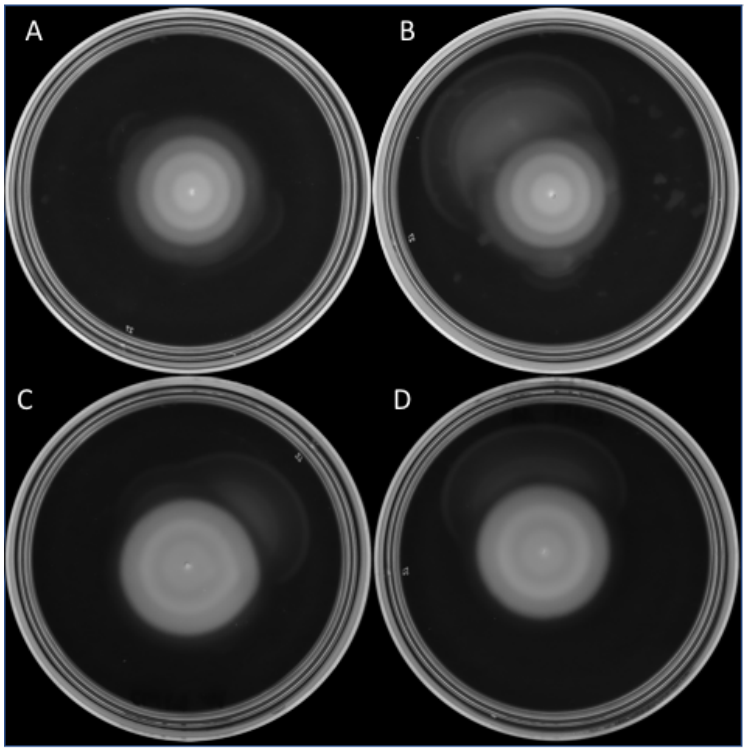

Figure 15. Addition of maltose alters banding pattern in banding type 1 formation. Strain CW17 (BW25113 Δ*lrhA::FRT*),. A-B) SMM-Na C-D) SMM-Na + 0.2% Maltose

Chemotaxis rings and banding types 1 and 2 have different expansion rates

Expansion rates of three types of motile mutants, banding type 1 (KH318), banding type 2 (KH707), and classic chemotaxis (KH1220). Each type of motility pattern resulted in a different expansion rate. Analysis of time-lapse movies (movies supplemental file) revealed that banding type 1 had the slowest expanse and had the longest lag time before moving from the inoculation site, while KH1220 had the fastest expansion rate, and had mostly colonized the plate prior to KH318 cells exited lag phase (Figure 16).

We repeatedly observed outgrowths from KH318 (Δ*lrhA*). When inoculated onto fresh SMM-Na these outgrowths would exhibit faster spreading than any banding strain, and chemotaxis rings were observed. Sequencing the intergenic region between *flhDC* and *uspC* revealed an IS element insertion. Previously, a double mutation of an IS element in the intergenic region between *flhDC* and *uspC* along with a *lrhA* mutation was reported to maintain faster spreading than either mutation alone (94). This result indicates that the chemotaxis pattern is dominant to the banding pattern (Figure 13E).

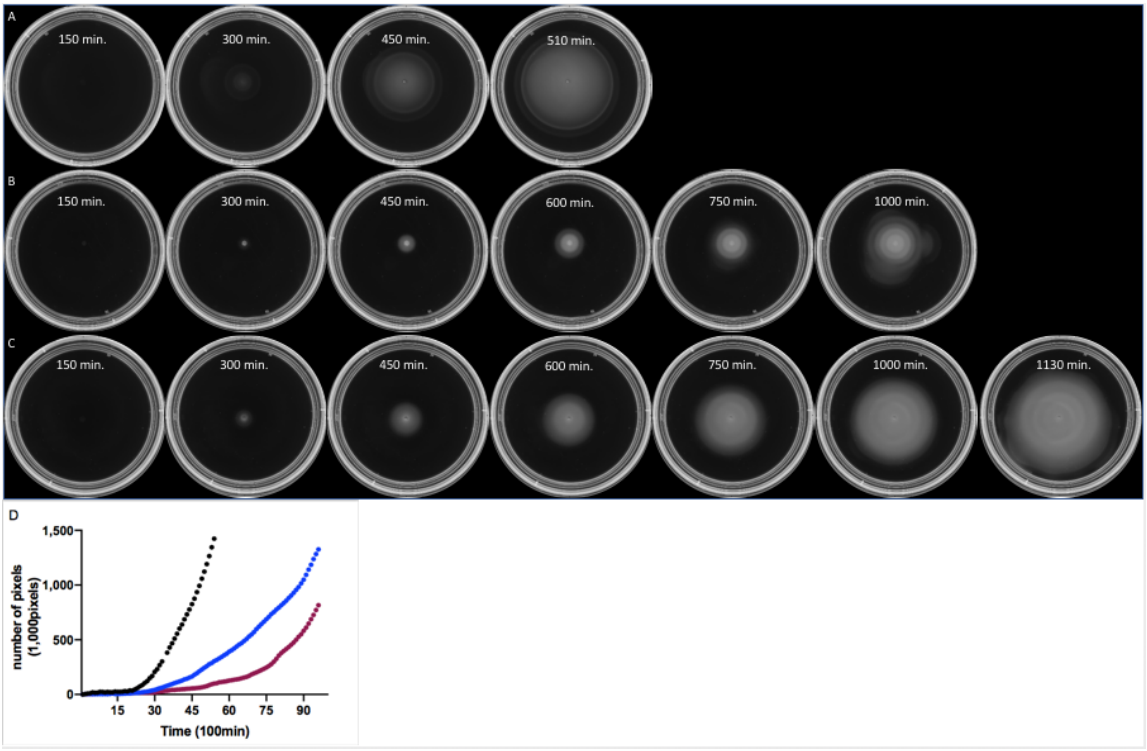

Figure 16. Time course of chemotactic ring and band formation.
A) Chemotaxis strain KH1220 B) Banding type 1 strain, KH318 C) Banding type 2 strain KH707. D) Quantification of cell expansion in time-lapse movies. Points correspond to the number of pixels in cell front. KH1220 (black), KH707 (blue), and KH318 (red).

Discussion

In addition to the well characterized chemotaxis pattern, here we described an alternative pattern that forms in rich swim media. This alternative pattern arises from mutations in strains previously characterized for their increased motility on media made with TB and incubated at 30 °C (92, 93, 119) but when grown at 37 °C on SMM-Na form a pattern consisting of repeating high and low cell density (Figure 3). Banding strains were isolated from motility media, transposon mutagenesis, or observed in Keio strains (Tables 1-4). They can be classified into four motility patterns: Banding type 1 containing repeating bands with a sharp demarcation, these strains mostly contained mutations or IS elements in *lrhA*, banding type 2 consisted of wider bands with fainter demarcation, and all contained IS elements within 350 base pairs of *uspC*, and more than 500 base pairs away from *flhD* (Figure 4 and 7 Table 5). Previously, strains with increased motility contain mutations between *flhD-uspC* and formed chemotaxis rings on TB media at 30 °C. However, these strains were reported to be tested on LB media at 37 °C (92, 94). We hypothesized that IS elements that were located near *uspC* would result in a banding phenotype, while mutations that clustered near *flhDC* would result in strains with chemotaxis rings. Banding type 3 were characterized as only forming one band, and canonical chemotaxis, consisting of three fronts of cells that move along amino acids spatial gradients (Figure 4).

Strains with chemotaxis rings were identified in strains with IS elements less than 400 base pairs away from *flhD* (Figure 7). LrhA is a reported transcription factor that directly binds in the *flhD* promoter to negatively regulate transcription (106), while

uspC also regulates flagella but the mechanism is unknown (108). By deleting chemoreceptors in a KH318 (Δ*lrhA*) and KH707 (*flhD* -510) strain we determined that these banding strains do not require the presence of the same chemoreceptors to form bands. These findings indicate that these patterns are formed by sensing different ligands in the media and responding to different spatial gradients.

Tsr senses serine, but can also sense other ligands (95, 120, 121). While the loss of the Tsr through a gene deletion or oversaturation of the receptor with 10mM serine resulted in a motile, but non-banding swimming phenotype (Figure 12), the loss of a functional *tsr* did not inhibit motility of the banding strains (Figure 13,14).

The loss of Tar, the chemotactic receptor that senses aspartate and maltose resulted in wider bands in KH318 (banding type 1) (Figure 13). The addition of maltose to KH318 also produces wider bands (Figure 15). It is possible that the addition of maltose saturates the receptor resulting in the same phenotype as the deletion of Tar. It has also appeared that the Δ*tar* Δ*lrhA* (RP079) strain forms bands at a slower rate than a Δ*lrhA* strain, but the rate of RP079 expansion has not been tested. Tar is more critical in the KH707 strain, as loss of *tar* resulted in a loss of bands. This suggests that KH318 and KH707 sense different ligands when migrating.

The loss of the Tap chemoreceptor is detrimental to motility in both banding strains, in SMM-Na at 37 °C and in standard TB media incubated at 30 °C. While the presence of Tap is essential to pattern formation and motility in the banding strains, it has previously been reported as having no obvious media defects in TB media (122). Tap senses pyrimidines and dipeptides for which sensing is dependent on the dipeptide

permease, DppA (97). Tap alone has poor chemoreceptor function due to weak phosphorylation of CheA, a mutant strain containing Tap as a lone chemoreceptor is unlikely to produce a strong enough response for chemotaxis (123). Tap needs the presence of either Tsr or Tar in the strain to function and may prefer Tar (97). Zhang et al., identified that after a cells group around a high concentration of aspartate, a subpopulation will move away from this attractant toward a provided TB source. The movement of the subpopulation of cells was dependent on the presence of Tap and Tar (124). We speculate that Tap and Tar may play a mechanistic role in band formation where we see recurring patterns of cells starting and stopping. The non-motile phenotype conferred by the Δ*tap* may not be due to the loss of the chemoreceptor. To identify processes affected in the Δ*tap* Δ*lrhA* strain we incubated these strains until outgrowths formed, several of these outgrowths banded when plated on fresh media. Sequencing these strains and the RP079 strain revealed an IS*3* element inserted into *rhsD* in all strains. Rhs proteins are shown in *Dickeya dadanti* to inhibit the growth of neighbor cells in a contact dependent manner (125). An evolution experiment that selected for different migration speeds on rich and minimal media identified several *rhs* genes that were mutated, but the effect of those mutations is not known (126). It is possible that the *rhsD* mutations are responsible for the non-motile phenotype in RP079.

LrhA directly regulates the master flagella regulator *flhDC* (116) which either directly or indirectly is responsible for regulation of the rest of the chemotactic genes and flagellar genes (27). UspC is also known to increase flagella production but it is possible due to the location of the IS elements in banding type 2 strains that the IS

elements affect *flhD* and not *uspC*. Further characterization is needed to determine if the IS elements affect production of *flhD*, *uspC* or both genes. Regardless, the upregulation of flagella in the KH707 is apparent due to the increased expansion of cells.

The different dependence on chemoreceptors for pattern formation indicates that the three strains, chemotactic and banding types 1 and 2 are utilizing different gradients of nutrients in the media to expand their populations. This would allow for the ability of the strains to expand into the same niche. Previously others have shown that two subpopulations of *E. coli* can cohabitate the same region by expanding their populations in two distinct distances from the center of the plate (104)This has also been seen in yeast where isogenic cells grown on rich complex media segregate over time into two distinct metabolic states used to occupy the same niche (127). The utilization of different nutrients could allow for a better colonization of a region by not competing for resources.

Future Directions

In the immediate future it will be imperative to confirm the role of Tap and RhsD on banding formation. Experiments have been started to test the motility phenotype of *rhsD* gene deletions in the background of KH318 and KH707. We have also started to test the role that DppA, that interacts with Tap to sense dipeptides, has on band formation. To confirm the role of Tsr, Tap, and Tar during band formation complementation experiments must be performed. To identify if growth rate or swimming speeds affect the chemotaxis mutants both growth curves and tethered cells assays would need to be performed. While no motility defects have been reported before

in Δt*ap* strains, the effect of Δtap and the other chemoreceptors on our KH1220 strain needs to be confirmed.

CHAPTER III

BULL'S EYE PATTERN FORMATION IS DEPENDENT ON (P)PPGPP SYNTHESIS

Introduction

Bacterial motility depends on environmental conditions including nutrient availability and a permissive surface for swarming or low agar conditions for swimming (2, 4). Various bacteria employ different tactics to swarm, resulting in characteristic patterns for different strains. Perhaps, the most well studied pattern, is the formation of concentric zones of cells that form while *Proteus mirabilis* swarms (22). During *Proteus* swarming differentiated swarmer cells will bundled together to form a raft (23). How these cells coordinate together has been an intriguing question for over a century. In contrast to swarming motility, swimming cells individually move through liquid or low agar (2). Previously, we described that deletion of *lrhA*, or IS elements inserted in the intergenic region between *flhD-uspC* resulted in a repeating pattern of high and low cell density under swimming conditions (chapter II); However, how these mutations lead to cell coordination and pattern formation is unknown.

Swimming, like swarming, is a flagellar-dependent phenotype (2). Expression of flagella in *E. coli* is under control of the class I operon *flhDC* (25). $FlhD_4C_2$ forms a transcription factor that activates class II genes, which are flagellar proteins that form the basal body and hook of the flagellum and other genes that regulate flagella synthesis. Those regulators activate transcription of the filament, motor and chemotaxis proteins (25, 27). Expression of *flhDC* is affected by many negative and positive regulators,

including LrhA (25). LrhA negatively regulates *flhDC* by directly binding to the *flhD* promoter (116). The LrhA binding site is frequently disrupted by an IS element and has been reported in several motile lab strains: RP437, W3110K, W3220, MC1000, W2637, YK410, MG1655-motile (93). Missense mutations in *lrhA* have also been found to increase the expansion rate of *E. coli* swarm cells, this is likely due to the increase in flagella (94, 117).

IS element insertions in the promoter region of *flhD* and between *flhD-uspC* have been identified previously, and lead to an increase in swarming diameter (92–94). The increase in motility likely is caused by disrupting the topology of the region and relieving the repression from negative regulators (26). We have identified that the location of the IS elements within the intergenic region between *flhD-uspC* affects the motility pattern on LB swim agar (Chapter II). The location of the IS elements suggests that disruptions close to *uspC* may be affecting the expression of *uspC* instead of *flhDC* as previously thought (92). Mutations in *uspC* are non-motile, but how *uspC* contributes to flagella synthesis is unknown (118). Expression of *uspC* is regulated by the global alarmone, (p)ppGpp, which activates the stringent response under stress conditions (128–131). During the stringent response, (p)ppGpp binds with DksA to RNA polymerase to modulate the expression of many genes to directly activate several amino acid biosynthesis genes and inhibit ribosomal promoters (129). Flagella synthesis and aggregation is also in part regulated by (p)ppGpp and DksA. Although differences in flagella synthesis were observed in a cell unable to synthesis (p)ppGpp (132–134).

Here we suggest a potential mechanism for pattern formation observed in *E. coli* mutants. We determine that (p)ppGpp production is essential for pattern formation and find that the addition of iron and short chain fatty acids affect motility and pattern formation.

Methods

Bacteria strains, motility plates

Bacterial strains were streaked fresh for each experiment from 15% glycerol stocks. For motility plates Swim Motility Media (SMM-Na) plates were made using 10 g tryptone, 5 g yeast extract, 5 g NaCl (170 mM) and 2.5 g agar per liter. All supplements added to the media were filter sterilized and added after autoclaving. Final concentrations: 2,3,5-Triphenyltetrazolium chloride 0.05g/L ; ferrous sulfate heptahydrate 10mM ; sodium acetate 20mM ; sodium butyrate 20mM ; sodium propionate 20mM ; sodium formate 20mM. Buffers were adjusted to pH 7 before adding to media: MOPS pH was raised to 7 with NaOH final concentration 50mM ; Phosphate buffer was made from combining Na_2HPO_4 and NaH_2PO_4, final concentration used 100mM. All motility plates contained 25 mL of media and were dried unstacked for 90 minutes prior to inoculation.

Imaging and Figure Generation

Images were taken with a Canon M50 camera with settings: shutter speed 1/320, aperture F6.3, ISO1600, and the EOS Utility application. Light was provided by a SYL 16CM ring light. Plates were imaged on black glass to prevent glare. Images were

cropped and converted to greyscale, using Microsoft PowerPoint, brightness on images was adjusted as needed. RNA polymerase structure image was created in PyMol (v2.3.2) using pdb file 5VSW (135).

Strain Generation

Deletions of genes were obtained from Keio strains (106) and moved into the desired backgrounds using P1*vir* (107). Transductants were selected on LB agar containing kanamycin (50 μg/mL) with 10mM sodium citrate. Transductants were passaged an additional time to prevent phage contamination. Strains were confirmed with PCR. Transductants containing both *relA* and *spoT* mutations were confirmed for auxotrophies on M9 0.2% glucose.

Banding suppressors were isolated by passaged once from outgrowths onto fresh SMM-Na plates. Cells were isolated from banding positive plates and single colonies were used to generate stocks.

Whole-genome sequencing

Cells were prepared for whole genome sequencing (WGS) following a phenol-chloroform extraction protocol (108). DNA was sequenced at Microbial Genome Sequencing Center (MiGS). Adapters from reads were trimmed using TrimGalore(v0.6.3). Mapping of reads was performed with Bowtie2 (v2.3.4.1) default settings, with the exception of distance between reads, which was set to a max of 3000 using flags –I 0 and –X 3000. Default settings were used for BWA MEM (v0.7.15)

alignments, and for SAMtools (1.7) when sorting and indexing bam files. BCFtools (1.9) (using consensus caller flags -cv) and Freebayes (v1.3.2) variant callers were used with ploidy set to 1. IS elements were identified using ISMapper (v1). Positive variants were defined as those with at least 80% of reads containing the alternative allele.

Table 6. Strains used for experiments in Chapter III

Strain name	Genotype	Source
KH6	**BW25113 *dgcI*^{L323L} *wrbA*^{G96V} *nemA*^{Y359N}**	CGSC:7926
KH7	**MG1655**	CGSC:6300
KH1220	BW25113 IS***1*** between *flhD-uspC* (IS-320)	This study
KH8	BW25113 Δ*lrhA*::Kan	CGSC:11785
KH318	**BW25113 Δ*lrhA*::FRT**	This study
KH707	**BW25113 IS*5* between *flhD-uspC* (IS-510)**	This study
KH1	**BW25113 Δ*trkA*::FRT *lrhA*670::IS*1* *yeeJ*^{T552I} *cyoD*104 Del**	CGSC:10452
KH10	**BW25113 Δ*kch*::Kan *ubiH*^{V223G} *lrhA*198::IS*5* A to G upstream *yciV***	CGSC:9121
KH1212	BW25113 Δ*lrhA*::FRT IS between *flhD-uspC*	This study
KH761	BW25113 yfbO::tet Δ*trkA*::FRT *lrhA*670::IS*1*::*lrhA*^{wt} *yeeJ*^{T552I} *cyoD*104 Del	This study
KH898	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	This study
KH899	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	This study
KH900	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	This study
KH901	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	This study
KH902	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	This study
KH920	**BW25113 Δ*lrhA*::Kan *nemA*^{Y359N}**	This study
KH907	**MG1655 Δ*lrhA*::Kan**	This study
KH924	**MG1655 Δ*lrhA*::Kan**	This study
KH925	**MG1655 Δ*lrhA*::Kan**	This study
CW24	BW25113 Δ*aer*::Kan Δ*lrhA*::FRT	This study
RP079	BW25113 Δ*tar*::*Kan* Δ*lrhA*::FRT	This study
CW26	BW25113 Δ*trg*::Kan Δ*lrhA*::FRT	This study
CW16	BW25113 Δ*tsr*::Kan Δ*lrhA*::FRT	This study
KH1200	BW25113 Δ*aer* IS*5* between *flhD-uspC* (IS-510)	This study
KH1202	BW25113 Δ*tar* IS*5* between *flhD-uspC* (IS-510)	This study
KH1198	BW25113 Δ*trg* IS*5* between *flhD-uspC* (IS-510)	This study
KH1202	BW25113 Δ*tsr* IS*5* between *flhD-uspC* (IS-510)	This study
CW17	**BW25113 Δ*lrhA*::FRT *nemA*^{Y359N}**	
CW28	**BW25113 Δ*tap::Kan* Δ*lrhA::FRT rhsD*4,066::IS*3***	This study
RP087	**BW25113 Δ*tap::Kan* Δ*lrhA::FRT rhsD*4,066::IS*3***	This study
RP089	**BW25113 Δ*tap::Kan* Δ*lrhA::FRT rhsD*4,066::IS*3***	This study
RP091	**BW25113 Δ*tap::Kan* Δ*lrhA::FRT rhsD*4,066::IS*3***	This study

Results

E. coli forms a periodic banding pattern while swimming

We have previously identified that disruptions to *lrhA* and mutations in the intergenic region between *flhD* and *uspC* result in a periodic banding pattern on swim

media made from LB (Swim Motility media or SMM-Na). These patterns vary slightly in appearance. While mutations in *lrhA* have sharp demarcations between bands of high cell density, and strains containing IS elements between *flhD-uspC* form wider bands with fainter demarcations (Chapter II). Similar banding patterns have been described in *Proteus mirabilis* and *Proteus vulgaris* during swarming on nutrient rich agar (21, 22). During swarming motility, only the top of the plate is colonized as cells join together to form rafts and secret surfactants to move along the top of the plate (4). We tested the ability of our cells to swim on a SMM-Na plate using a redox-sensitive dye 2,3,5-Triphenyltetrazolium chloride (TTC). TTC can be used as an electron acceptor during cellular respiration and turns from colorless to red (triphenyl formazan) when reduced (136). Using TTC we were able to confirm that cells (Figure 17) are swimming through the media and utilizing all dimensions of the plate. Since these cells are able to swim in three dimensions throughout the agar.

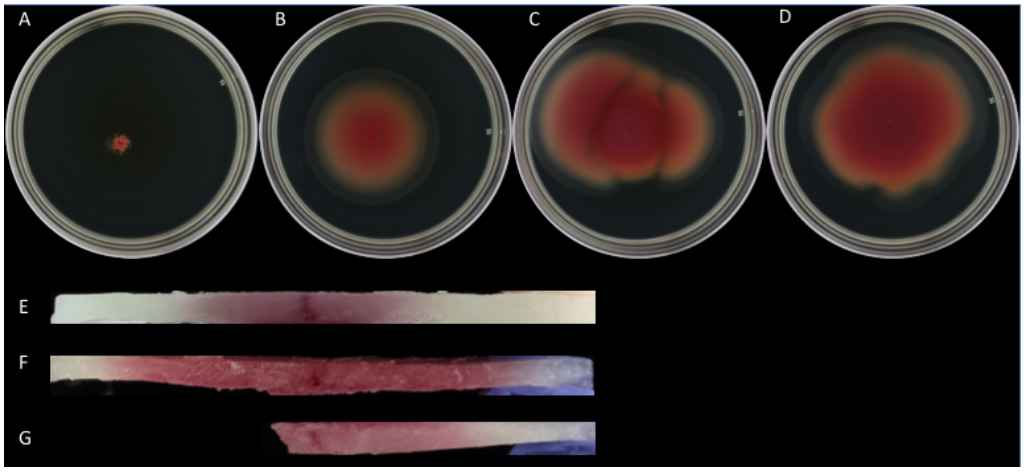

Figure 17. Banding pattern forms while swimming.
Plates were incubated at 37 °C in SMM-Na with TTC (0.05g/L). A) KH6 (BW25113) B) KH1220 (IS1 -320 *flhD*) C) KH318 (Δ*lrhA*) D) KH1193 (IS5 -510 *flhD*) E-G) Cross sections of plates through the inoculation center E) KH1220 (IS1 -320 *flhD*) F) KH318 (Δ*lrhA*) G) KH1193 (IS5 -510 *flhD*)

Increased flagella synthesis is necessary for band formation

LrhA directly negatively regulates flagellar expression and indirectly regulates expression of type 1 fimbriae, RpoS, and chemotaxis proteins (116, 117, 137, 138). RpoS, sigma-38, is induced under several stress conditions including carbon, phosphate and magnesium starvation and upon entry into stationary phase (139). Negative regulation of RpoS levels by *lrhA* occurs in two ways. LrhA increases the activity of RssB, an adapter protein which promotes RpoS degradation via the ClpXP protease, and it also represses transcription of the sRNA RprA, which activates RpoS translation. (137, 138, 140). In wild type cells, RpoS is degraded by ClpXP during exponential growth, but accumulates upon entry into stationary phase. In a *rssB* mutant strain RpoS is not degraded by ClpXP and will accumulate under all growth conditions (141). *LrhA*

mutants have a small increase in RpoS levels, and subsequent activity (137, 142). To determine whether an increase in RpoS could result in a banding pattern, we tested the motility of a Δ*rssB* strain, which has been shown to accumulate high levels of RpoS and increase activity of RpoS (141). We found Δ*rssb* to be non-motile, however since LrhA negatively regulates flagella synthesis (116) it is possible that banding required both an increase in flagella and RpoS activity. To confirm if band formation was dependent on *rpoS* we created a Δ*rpoS* Δ*lrhA* strain and found it did not alter the banding type 1 pattern (Figure 18).

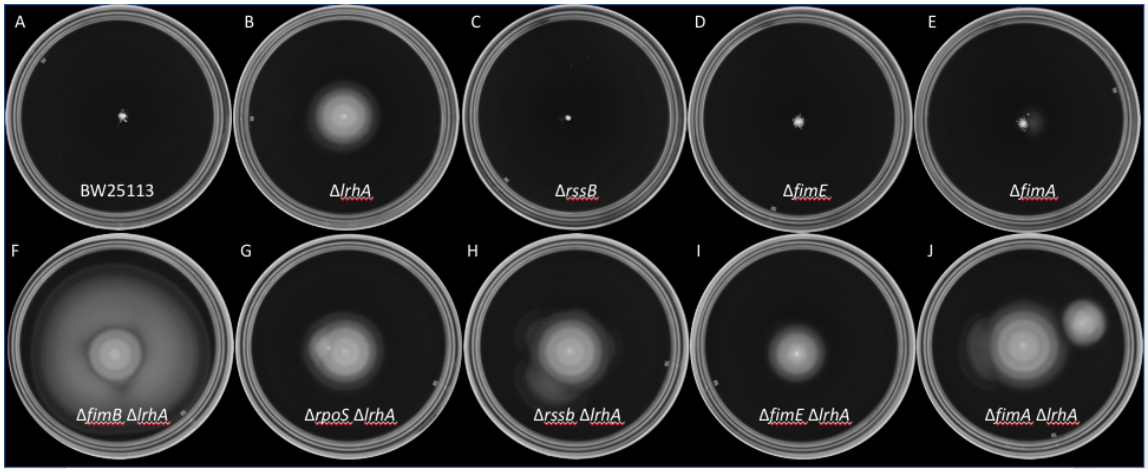

Figure 18. RpoS is not needed for band formation.
Plates were incubated at 37 °C in SMM-Na for 14 hours. Relevant genotypes are listed below the images. A) KH6 B) KH318 C) KH1001 D) KH1006 E) KH1007 F) KH986 G) KH116 H) KH984 I) KH1008 J) KH1009

LrhA is an inhibitor of expression of type 1 fimbriae and functions by the increasing the expression of negative regulator FimE (117). To test if type 1 fimbriae were necessary for band formation, we deleted *fimE* in our wildtype strain. We predicted the loss of *fimE* would increase type 1 fimbriae levels similar to those proposed to occur

in a Δ*lrhA* strain. We found our Δ*fimE* strain to be non-motile, due to the non-motile we could not rule out the involvement in a flagella dependent phenotype. To determine if type 1 fimbriae were involved in band formation, we deleted *fimA* in KH318. FimA is the major subunit for fimbriae, and we found that the loss of FimA did not affect the band formation (Figure 18).

LrhA negatively regulates *flhDC*, the master operon for flagella synthesis, by binding to the promoter (116). Mutations in *lrhA* lead to an increase in flagella synthesis in MG1655, and increase swarming distance (94, 117). We confirmed that the banding strain KH318 required both flagellar synthesis and a functional chemotaxis system (data not shown). Flagellar synthesis is regulated in part by temperature, where at high temperatures over 40 °C, very few flagella are synthesized (143). At 37 °C, wild-type BW25113 and MG1655 (CGSC #6300) strains are non-motile. The *E. coli* laboratory strains RP437, W3110K, W3220, MC1000, W2637, YK410 and MG1655-motile (CGSC #8237), all contain IS elements that disrupt the LrhA binding site in the *flhDC* promoter (93). We tested the motility of RP437 on SMM-Na and observed chemotaxis rings and an increase in speed of expansion over the plate (Figure 19). This led us to hypothesize that the bands may form due to a reaction-diffusion system, where speed at which the cells expand would determine the pattern. For slow swimming cells the migration front depletes the local nutrients below a threshold that causes the cells to become non-motile. For strains that expand faster, the migration front never depletes the local nutrients and consequently never stops migrating. This was supported by testing the motility phenotype in SMM-Na at 30 °C, conditions that allow for flagellar synthesis

in wild type strains. At 30 °C KH318 had only formed one very faint band and KH1193 had grown confluently without forming bands. While BW25113 and MG1655 had both produced two clear bands (Figure 20). To confirm that pattern at 30 °C was not specific to our parental strains, additional isolated of strains of BW25113 and MG1655 were tested. All wild type strains formed bands at 30 °C (Figure 20). These results support the hypothesis that the speed of expansion, potentially modulated by the number of flagella, creates the banding pattern due to nutrient limitation.

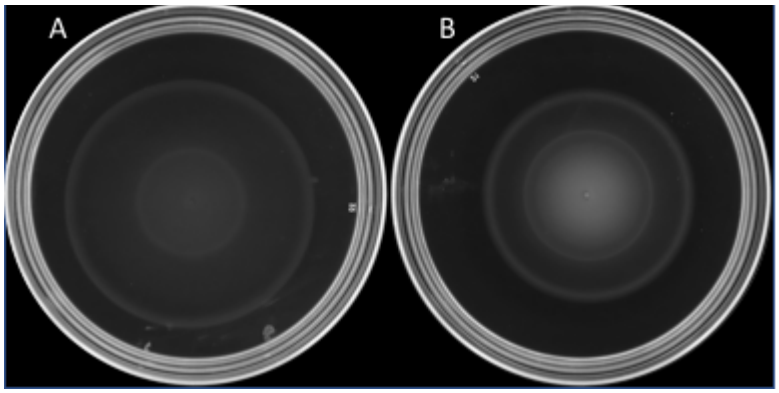

Figure 19. Motility phenotype of RP437
A) TB 30 °C B) SMM-Na 37 °C plates incubated 8 hours

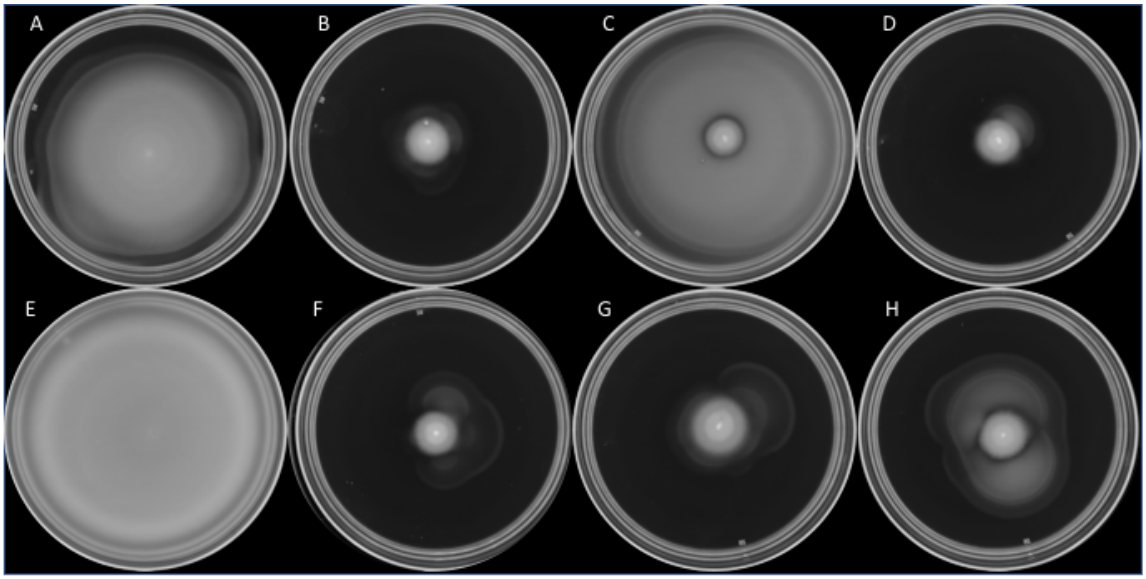

Figure 20. Wild type cells form a banding pattern in SMM-Na at 30 °C plates were incubated for 20 hrs. A) KH318 B) KH6 C) SDB262 D) SDB264 E) KH1197 F) SDB265 G) KH7 H) KH1024

(p)ppGpp is necessary for band formation

(p)ppGpp is a global regulator produced by two synthases, RelA and SpoT under nutrient and other stress conditions (130, 131, 144, 145). (p)ppGpp production induced by amino acid starvation is RelA-dependent while other stresses including fatty acid, iron, and carbon starvation produce (p)ppGpp in a SpoT dependent manner (131). Upon stress, (p)ppGpp binds to RNA polymerase (RNAP) to regulate the activity of genes (130). In addition to RNAP, (p)ppGpp can bind to other proteins and is involved in regulation of many cellular processes (146).

Strains containing IS elements upstream of *flhD* result in two phenotypes depending on the location of the IS element. IS elements located within 300 base pairs of

the *flhD* start codon led to a chemotaxis phenotype, which is characterized by the presence of expanding rings that correlate with the cells sensing and utilizing serine, aspartate and threonine (Figure 19). IS elements located more than 500 base pairs from *flhD*, that cluster closer to *uspC* result in a banding pattern (Chapter II). We hypothesized that IS elements closer to *uspC* increased *uspC* expression and affected flagella expression. UspC along with UspE contribute to flagellar production and adhesion (118) and both genes are positively regulated by (p)ppGpp (128).

We predicted that upregulation of *uspC* would lead to banding, and subsequently tested if loss of (p)ppGpp (designated ΔRS), a known positive regulator of *uspC*, would affect band formation (128, 147). We tested the effect of ΔRS on our KH318 (Δ*lrhA*) strain. This strain (KH1073) had reduced motility and was non-banding (Figure 21). While loss of *relA* in KH318 still formed bands, they were reduced in size and number (Figure 21). We were unable to test the motility phenotype of a Δ*lrhA spoT*- strain, as a *spoT* mutation alone is not viable, but is viable when made in a *relA*- background.

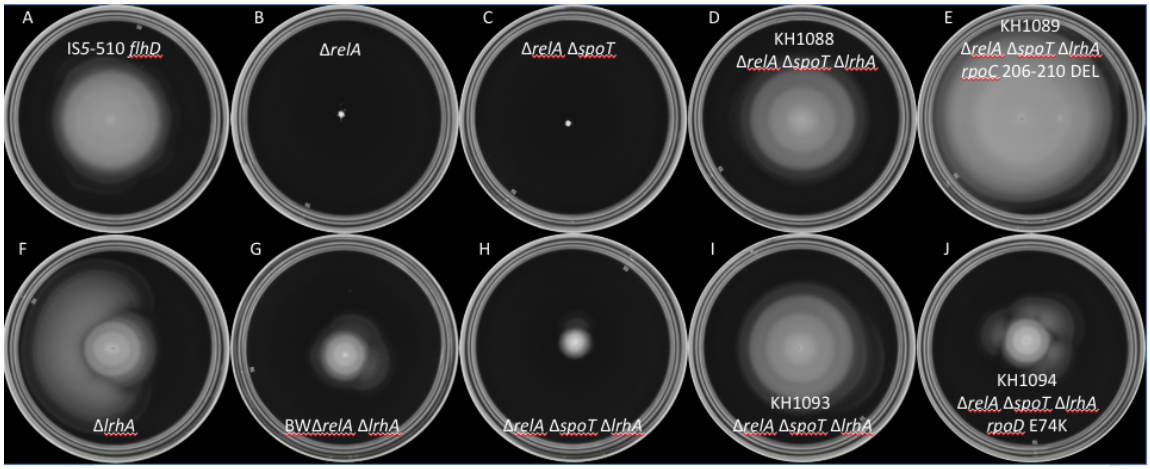

Figure 21. (p)ppGpp synthesis is required for banding type 1 formation. Plates were incubated at 37 °C in SMM-Na for 14 hours. Relevant genotypes are listed below the images. A) KH1193 B) KH1067 C) BW1071 D) KH1088 E) KH1089 F) KH318 G) KH1069 H) KH1073 I) KH1093 J) KH1094

We observed outgrowths on SMM-Na plate emanating from KH1073. We passaged these outgrowths onto fresh SMM-Na and identified four strains that suppressed the ΔRS phenotype and banded (Figure 21). From each of the four banding strains we whole-genome sequenced two isolates from each strain, and identified in one isolate a 5 amino acid deletion in *rpoC*, and in one isolate contained a E74K amino acid mutation in *rpoD* (Figure 22)

Other studies have identified suppressors of ΔRS strains in *rpoB*, *rpoC* and *rpoD* (M+ mutants) overcome the need for (p)ppGpp to regulate RNAP (129, 142, 148). We conclude from this data that (p)ppGpp is necessary for band formation.

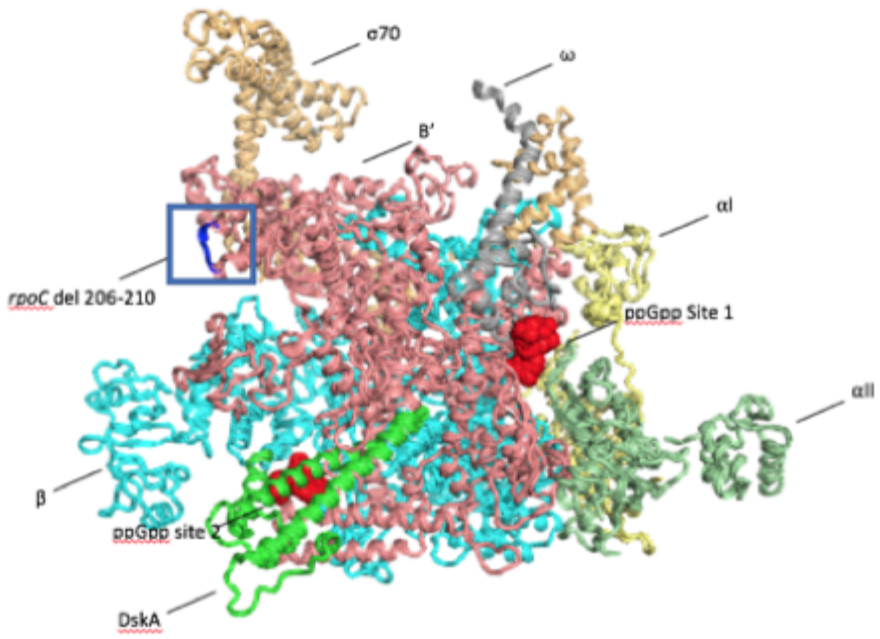

Figure 22. Mutations in RNA polymerase occur in banding ΔRS strains. Mutation isolated from KH1089 is highlighted in blue (blue box). Mutation in KH1094 is not crystallized in this structure. (adapted from pdb 5VSW)

Effect of carbon supplementation on pattern formation

(p)ppGpp is synthesized by RelA under amino acids starvation, and synthesized by SpoT under carbon, iron and fatty acid limitation (131, 149). We hypothesized that in our banding strains the slower migration compared to chemotaxis strains led to nutrient depletion which activated (p)ppGpp. Cells near the swim front should have access to nutrients, but cells behind the swim front could become nutrient limited as components are used from the media. Pattern formation is dependent on yeast extract being present in the media, and because of this we are not able to remove individual components of the media to test how banding formation responses. We instead added nutrients in excess to

SMM-Na to test for pattern modification. We predicted that providing nutrients in excess would relieve the stress and therefore the necessity of (p)ppGpp. We focused on stress conditions known to induce the stringent response and upregulated *uspC*, these included carbon, phosphate, iron, and acetate starvation (128, 131).

We first tested the effect of adding in various carbon sources on pattern formation. The addition of either glucose or maltose altered the pattern (Figure 23) while glycerol and fructose had no effect (Figure 23). Glucose has long been known to inhibit flagellar formation (143). We are unable to conclude that the reduction of motility with the addition of glucose is due to additional carbon and not inhibition of flagellar synthesis. Maltose increased the width of the bands, and we have previously shown that deletion of the chemoreceptor that senses maltose produces the same phenotype (chapter II). Therefore, we suggest that alteration of the banding pattern in the presence of maltose can alter the ability of the receptor to sense a gradient of chemoeffectors. The supplementation with fructose and glycerol did not affect pattern formation (Figure 23), suggesting banding formation is not due to carbon limitation.

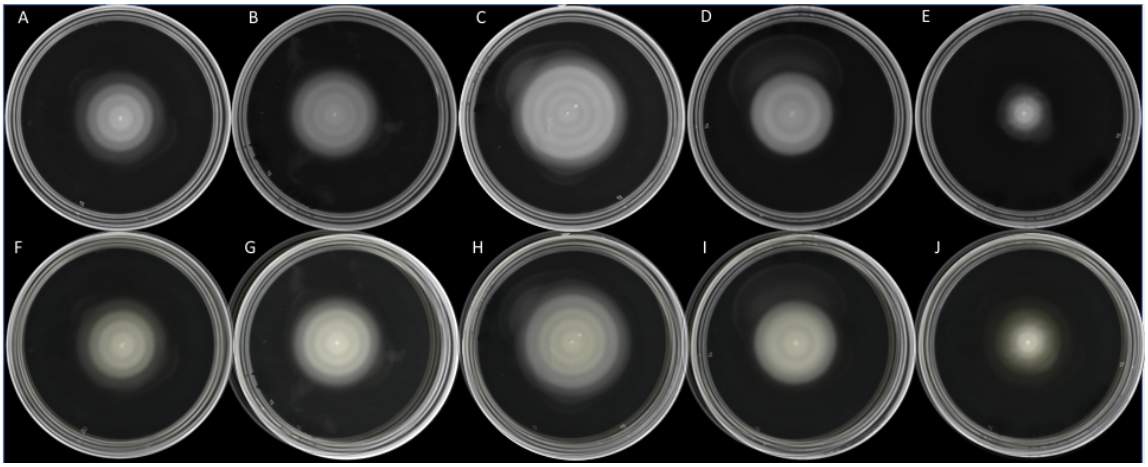

Figure 23. Addition of 0.2% Maltose and 0.2% glucose alter banding in Δ*lrhA*. CW17 (Δ*lrhA*::FRT) Plates incubated for 16 hours at 37 °C A) SMM-Na B) SMM-Na + 0.2% glycerol C) SMM-Na + 0.2% Fructose D) SMM-Na + 0.2% maltose E) SMM-Na + 0.2% glucose F-H) SMM-Na (transparency 65%) overlaid on each image from A-E

Phosphate supplementation does not affect pattern formation

Under phosphate starvation conditions, (p)ppGpp accumulation is SpoT dependent. The elevated level of (p)ppGpp induces RpoS through interactions with IraP and Rssb (129). We supplemented SMM-Na media with 100mM phosphate buffer and observed no effect on pattern formation (Figure 24).

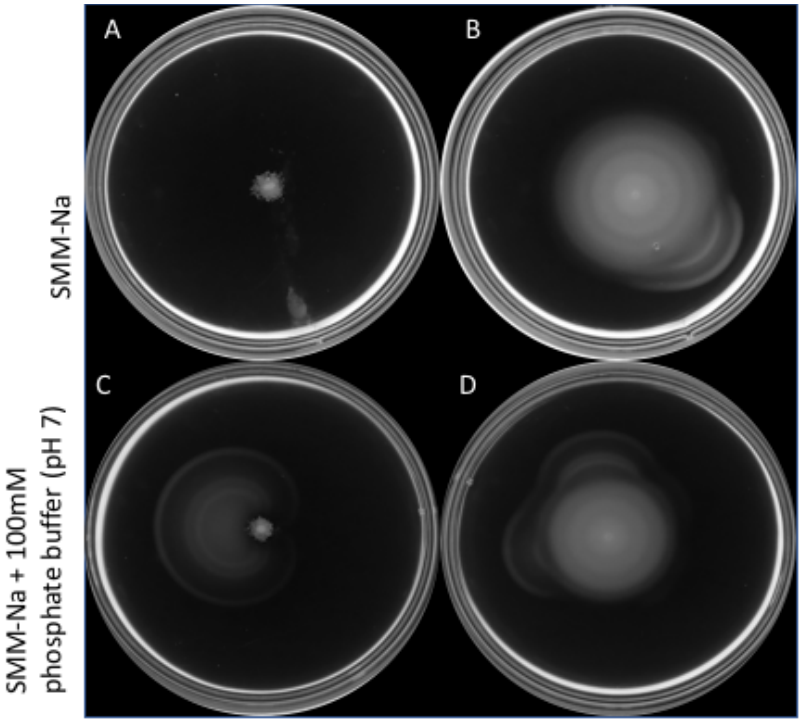

Figure 24. Addition of phosphate does not alter pattern formation.
Plates were incubated at 37 °C for 14 hours. A-B) SMM-Na C-D) SMM-Na + 100mM potassium phosphate buffer pH 7 A,C) KH6 BW25113 C,D) KH318 Δ*lrhA*

Effect of iron supplementation on pattern formation

When iron becomes limiting cells synthesis (p)ppGpp in a SpoT-dependent manner (150). (p)ppGpp upregulates production of the siderophore Ent that takes up iron (151). Regulation of Ent by (p)ppGpp is Fur independent (150). Under iron-sufficient conditions, Fe^{2+} binds Fur to negatively regulate iron acquisition genes (152). Mutations in *ent* and *fep* genes in *E. coli* W3110 have swarm deficiencies (153). The addition of less than 1mM ferrous sulfate did not alter the banding period of either KH318 or KH707, although the demarcations were sharper this could be due to the increased contrast of the media (Figure 25). The addition of 10mM ferrous sulfate returned a wild

type non-motile pattern, while KH1220 was still able to swim, although much slower. Due to the high concentration of iron needed to affect banding, we cannot exclude the possibility that toxic effects altered the pattern formation. While iron is not limiting in LB broth, iron uptake appears to be important for swarming (153). Further experiments are needed to determine if iron limitation triggers pattern formation during swimming.

Effect of short chain fatty acid supplementation on pattern formation

Under acetate starvation *uspC* is produced (128) and *E. coli* grown in 50mM acetate repress production of DksA (154). This indicates that acetate could influence the stringent response through DksA and also affect *uspC* expression. To test if fatty acids could alter banding, we added 20mM of different short chain fatty acids (SCFA) to SMM-Na buffered with 50mM MOPS. We found that the addition of acetate, propionate, and butyrate altered banding by producing a non-banding motility phenotype (Figure 26). Acetate, butyrate, and propionate have been shown to positively regulate flagella in Enterohemorrhagic *E. coli* (155) suggesting that the non-banding phenotype in SCFA-supplemented media could be due to upregulation of flagella. At 100mM concentrations, acetate, propionate, and butyrate all support band formation in BW25113 at 37 °C. We predicted this is due to an upregulation in flagella (data not shown). Growth on media supplemented with either acetate, propionate, or butyrate also affected the pattern formation of KH707. Under these conditions chemotaxis rings were observed and the plates were colonized at the same rate as the chemotaxis controls (Figure 26). The addition of formate did not affect the banding pattern in KH318, and KH707 was not tested the observed swimming behaviors of strains grown on either acetate,

propionate, or butyrate supports our hypothesis that cells with high migration rates move too fast to form bands.

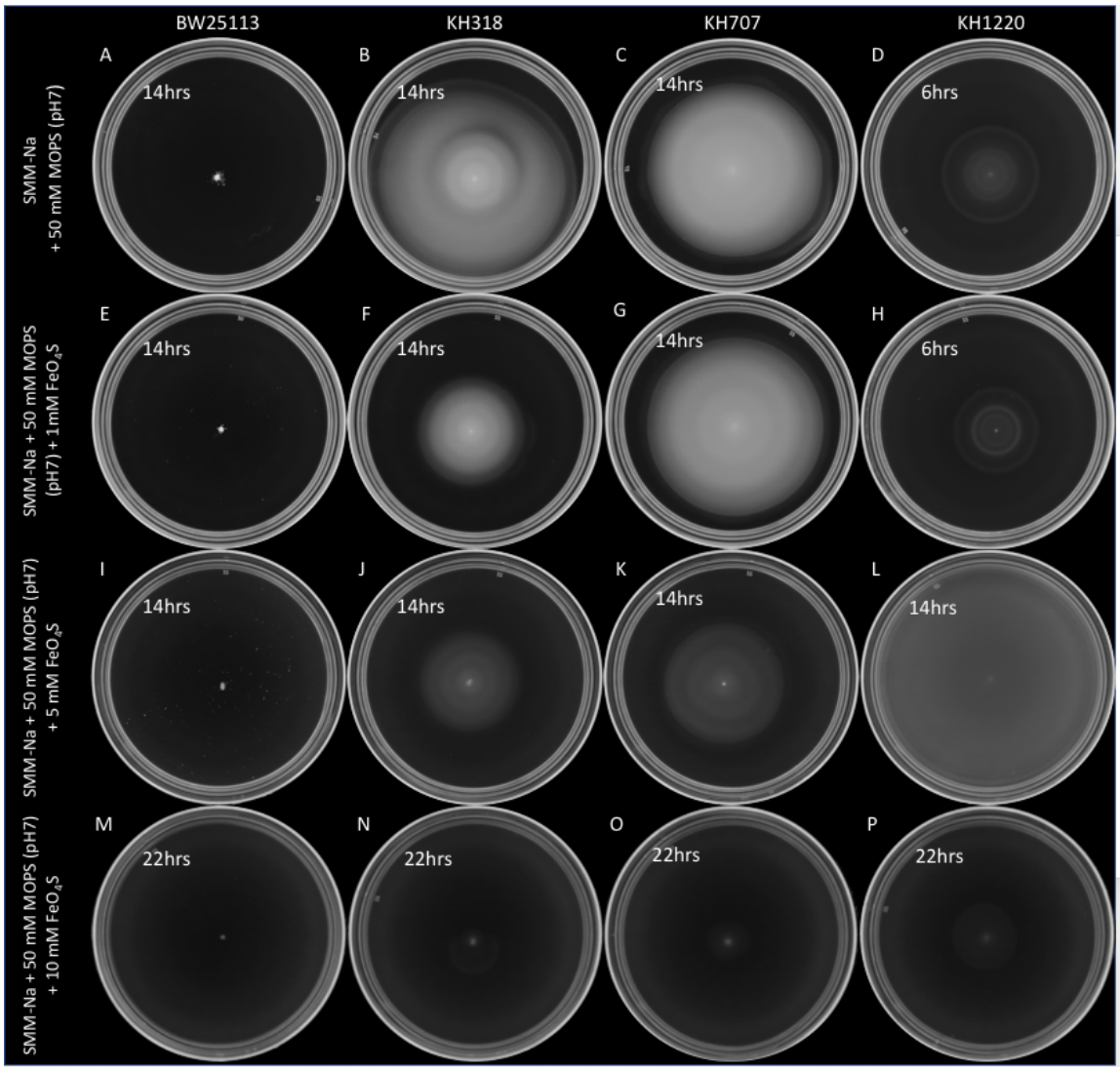

Figure 25. Addition of 10mM ferrous sulfate inhibits motility in KH318 and KH707.
Plates were incubated at 37 °C in SMM-Na. Time incubated is listed in figure. All plates contain 50mM MOPS (pH 7) with E-H) 1mM FeO_4S I-L) 5mM FeO_4S M-P) 10mM FeO_4S A,E,I,M) KH6 B,F,J,N) KH318 C,G,K,O) KH707 D,H,L,P) KH1220

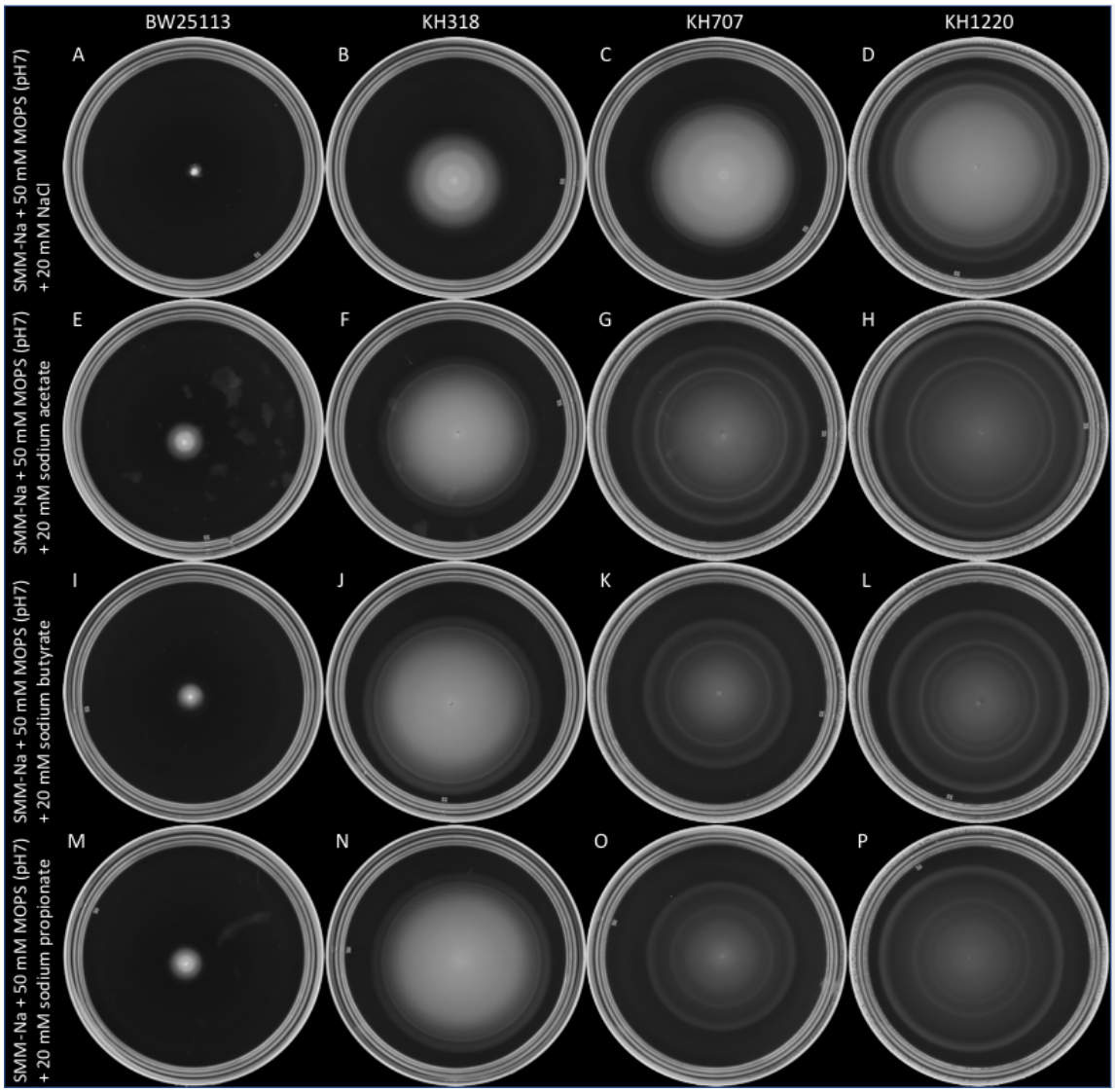

Figure 26. Short chain fatty acids obscure banding patterns.
Plates were incubated 12 hours at 37 °C unless otherwise noted A,E,I,M) BW25113 (KH6) B,F,J,N) Banding type 1 (KH318) C,G,K,O) Banding type 2 (KH707) G,K,O) incubated 6 hours D,H,L,P) Chemotaxis strain (KH1220) incubated 6 hours A-D) SMM-Na + 50mM MOPS pH7 + 20mM NaCl E-H) SMM-Na + 50mM MOPS pH7 + 20mM sodium acetate I-L) SMM-Na + 50mM MOPS pH7 + 20mM sodium butyrate M-P) SMM-Na + 50mM MOPS pH7 + 20mM sodium propionate

Acetate is converted acetyl-CoA, by two pathways one directly through Acs and the other produces acetyl phosphate as an intermediate product (AckA-Pta pathway.

Using the AckA-Pta pathway, acetate is converted to acetyl phosphate with AckA and then to acetyl-CoA using the EutD and Pta acetyltransferases (156). Pyruvate can be converted to acetyl-CoA in both anaerobic and aerobic conditions (154). We tested if the effect of acetate on the band formation could be due to an increase in acetyl-CoA. When 10mM sodium pyruvate was added to SMM-Na we observed a change in the banding pattern, including fainter demarcations of the inter-bands and an alteration of the width of the bands (Figure 27). This led us to test if acetate being converted to acetyl-CoA was necessary for banding. We deleted the different enzymes needed for this conversion. While the single deletions of *acs*, *eutD*, and *pta* had no effect on banding in a Δ*lrhA* strain, Δ*ackA* Δ*lrhA* caused a reduction in the banding phenotype (Figure 28). Acetyl phosphate has been shown to regulate over 100 genes, including *flhDC*, it is possible that the change in flagellar regulation is what is contributing to the loss of the pattern (156). We were unable to make a Δ*eutD* Δ*pta* Δ*lrhA* mutant to answer if conversion from acetyl phosphate to acetyl-CoA affects band formation.

Discussion

Here we propose that cell coordination, and, therefore, pattern formation occurring in swimming bacteria, is dependent on nutrient availability. When grown at 37 °C wild type *E. coli* strains BW25113 and MG1655 are non-motile. When grown in SMM-Na at 30 °C which is a permissive temperature for flagella synthesis (143)these strains will form a banding pattern (Figure 20). We conclude that the banding formation is a normally occurring phenotype in swimming bacteria under permissive conditions.

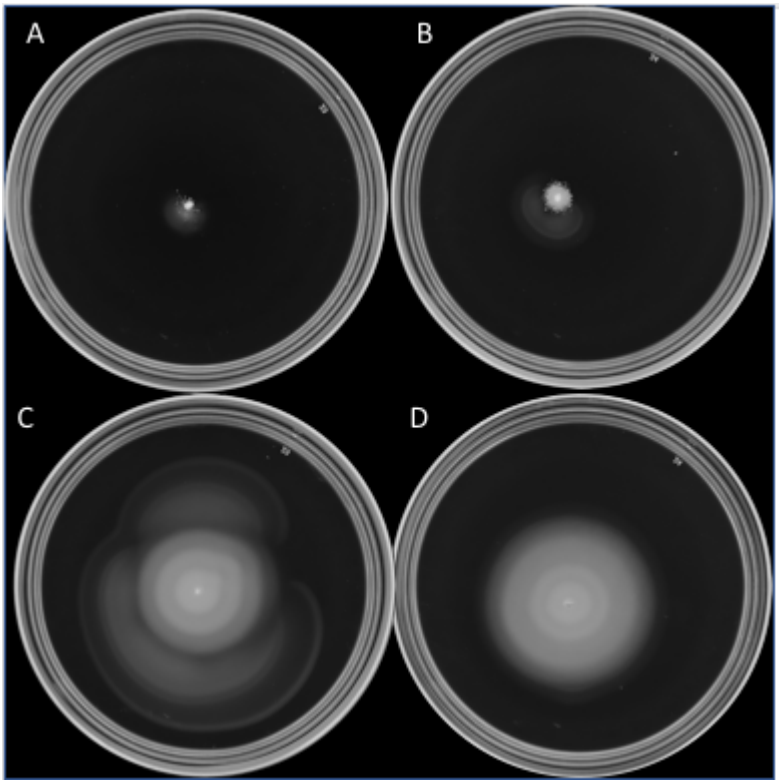

Figure 27. Addition of 10mM sodium pyruvate modifies the banding pattern. Plates were incubated at 37 °C in SMM-Na for 14 hours. A-B) KH6 (BW25113) C-D) KH318 (Δ*lrhA*)

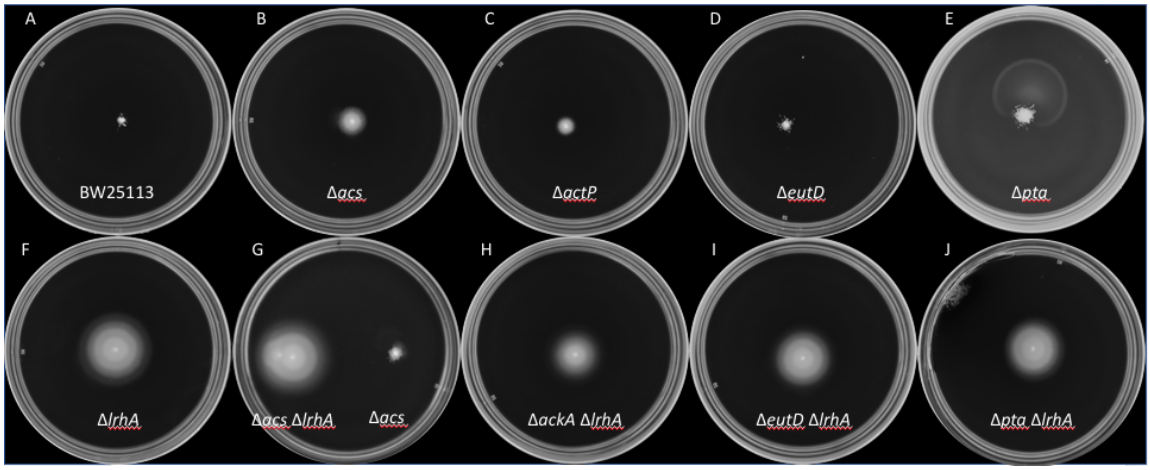

Figure 28. Deletion of ackA alters band formation.
Plates were incubated at 37 °C in SMM-Na for 14 hours. A) KH6 B) KH1129 C) KH1130 D) KH1134 E) KH1126 F) KH318 G) KH1104 (left) KH1129 (right) H) KH1136 I) KH1135 J) KH1126

Two types of mutations have been identified by our lab to form these patterns under growth at 37 °C on swim plates. Both mutations in *lrhA* and IS elements inserted upstream of *flhD-uspC* are involved in flagellar production (92, 94). We predict that an increase in flagellar synthesis is what allows the banding formation to be visible at 37 °C when wild type cells are non-motile. These strains do not form patterns when incubated at 30 °C, while pattern formation by these strains could be temperature dependent it could also be a result of an increase in flagella. Strains with increased migration rate such as those that display a chemotaxis pattern, never pause and in this model would move too fast to fully deplete nutrients. Bacteria that have faster migration fronts in semi-solid agar tumble more often and decrease the time spent running (91, 126).

While swarming *Proteus* alternates between swarming and consolidation phases (40). During swarming, amino acids are depleted, and during the consolidation phase,

cells increase amino acid biosynthesis (23, 24). It is hypothesized that the consolidation phase is a period of time for the cells to accumulate energy needed for the next swarm phase (23). Disruption of *lrhA* in *Proteus* had a minimal effect on swarming but, *hexA* a homolog of *lrhA*, is upregulated during *Proteus* consolidation, and is needed for infection along the urinary tract in mice (23).

We focused on the mechanism for band formation using KH318 (Δ*lrhA*). In *E. coli* LrhA is not only a negative regulator of *flhDC*, but also negatively regulates type 1 fimbriae and RpoS (116, 117, 137, 138). We are able to conclude that RpoS is not involved in banding formation, a double mutant of *lrhA* and *rpoS* did not affect band formation (Figure 18). Although, the loss of the major subunit of type 1 fimbriae, FimA did not affect the banding pattern formation we cannot conclude that type 1 fimbriae are not important for band formation. In the absence of FimA, small type 1 fimbriae can form composed of FimG and FimH and maintain some functional binding activity (157). It is possible, that this residual binding activity is contributing to the formation of the banding pattern. LrhA is regulated by RcsB and Lrp both of which have been found to affect swarming in *Proteus* (34, 38, 138). Lrp is a global regulator that regulates metabolism and motility in response to nutrient conditions (158)Lrp during swarming in *Proteus* positively regulates flagella synthesis, and *lrp* expression is highest corelates with upregulation of *flhDC* (38). In *E. coli* regulation of *flhDC* by Lrp is predicted to occur through LrhA (37). Expression of Lrp is dependent on (p)ppGpp synthesis (158). (p)ppGpp synthesis is necessary for our pattern to form, we speculate that a possible mechanism is regulated through Lrp in response to nutrient stress.

In *Proteus* (p)ppGpp accumulates in cells that are unable to be recognized as self and inhibits *flhDC* excluding cells from joining a swarm raft (62). In *E. coli* (p)ppGpp has been shown to negatively regulate *flhDC* in liquid culture (133, 134) but on swim plates made from LB positively regulates flagella synthesis (132). A loss of (p)ppGpp in our KH318 strain results in a non-banding phenotype with reduced motility (Figure 21). We conclude from this that (p)ppGpp is necessary for the banding pattern to form, and while we speculate that this is response to nutrient stress occurring by the slow migration of the banding cells, we have not been able to show that nutrient stress occurs.

We attempted to relieve any potential nutrient stress by supplementing SMM-Na with nutrients that when limiting can trigger the stringent response (131). The addition of iron and SCFA affected the banding pattern. However, both iron and SCFA regulate flagella synthesis, making it unable for us to conclude if the alteration in the pattern resulted from increased or decreased flagellar synthesis or a change to the nutrient level of the plates. In iron limiting conditions YdiV and SlyD bind to Fur and alter the DNA binding ability, depressing iron acquisition genes (159). YdiV is upregulated under iron limiting conditions and can regulate $FlhD_4C_2$ by binding and preventing transcription of class II flagella operons and by targeting $FlhD_4C_2$ for degradation via ClpX (160). Acetate, butyrate, and propionate all upregulate flagella synthesis in Enterohemorrhagic *E. coli* (155).

We provide preliminary data that pattern formation can be affected by addition of iron and acetate, butyrate, and propionate. Both iron and short chain fatty acids can

affect flagellar regulation, and further testing is needed to identify the mechanism that contributes to loss of band formation.

Future directions

We are currently testing the effect of the known regulators of LrhA, Lrp and RscB on pattern formation. RscB inhibits *lrhA* (138) in a Δ*rscB* strain we would anticipate higher levels of *flhDC* due to an increase in LrhA and possible regulation on *flhDC*. Therefore, we predict that a double Δ*rscB* Δ*lrhA* strain would have a reduction in motility. Lrp positively regulates LrhA (37), therefore we predict that a Δlrp strain may band due to decrease in LrhA negative regulation on *flhDC*. We are testing the effect of loss of (p)ppGpp on the banding type 2 pattern formation using KH1193. First we need to test if the IS elements more that 500 base pairs away from *flhD* positively effect *uspC* expression, which can be tested with qPCR. (p)PpGpp is known to positively regulate *uspC* (128). If the IS element increases *uspC* expression then we would predict no effect on pattern formation in a ΔRS strain, and that these mutations form bands independently of (p)ppGpp.

CHAPTER IV

BOLO:RESTORATION OF DATA LOSS DURING ALIGNMENT OF CIRCULAR GENOMES

Introduction

Illumina-based sequencing is the dominant Next-Generation Sequencing (NGS) platform used for the analysis of whole genomes, in part because of the precipitous decrease in cost (161). The development of Next-Generation Sequencing (NGS) techniques, facilitated several discoveries in bacterial research, advancing the understanding of evolution (162) pathogenicity (163) and epidemiology (164). NGS has revolutionized the way biological research is conducted. Due to large increase in sequencing (161) the need for fully automated analysis is needed to handle the large amounts of data (165).

Alignment algorithms have been designed to work with linear chromosomes and, as a consequence, the analysis of circular genomes poses a challenge for existing software. To use standard analysis programs the circular genome sequence needs to be transformed into an artificially linear sequence. This is done by introducing a cut site in the reference genome, we refer to this as the Genome Sequencing Origin (GSO). There is a systematic loss of data near the GSO. The data loss comes from the inability of mappers to correctly align sequencing reads that cross the GSO. Instead, these sequences are trimmed or altogether ignored, leading to lower quality signals in those regions. This can make finding variants in this region challenging. While effort was made to place the

GSO in innocuous regions, some were placed in locations of importance. For example, the GSO in mitochondrial genomes is placed in the hypervariable region (166). Although this region was seemingly inconsequential at the time of sequencing, it contains information on variants important to discovering humans' geographically heritage (167, 168).

While short-read sequencing has a higher accuracy and lower costs compared to long-read approaches. Benefits exist for sequencing using long-read platforms, such as identifying structural variants (169). Due to the long length of the reads this sequencing platform can correctly map over the GSO, and software programs have been developed to correctly circularize these reads (170–172). However, due to the higher cost and lower accuracy of reads, Illumina remains the dominate platform for sequencing (161). While solutions exist for mapping across the GSO using short reads (172) what is lacking is an open-source program where read recovery is fully automated.

Here we describe our software program Bolo, which is designed to recover read loss that occurs near the GSO and allow callers to identify variants in this region. Bolo is indifferent towards mapping and variant calling programs, making it compatible future software programs. Here we show that Bolo can recover read depth near the GSO, and correctly identify variants near the GSO in organisms with one or multiple circular chromosomes.

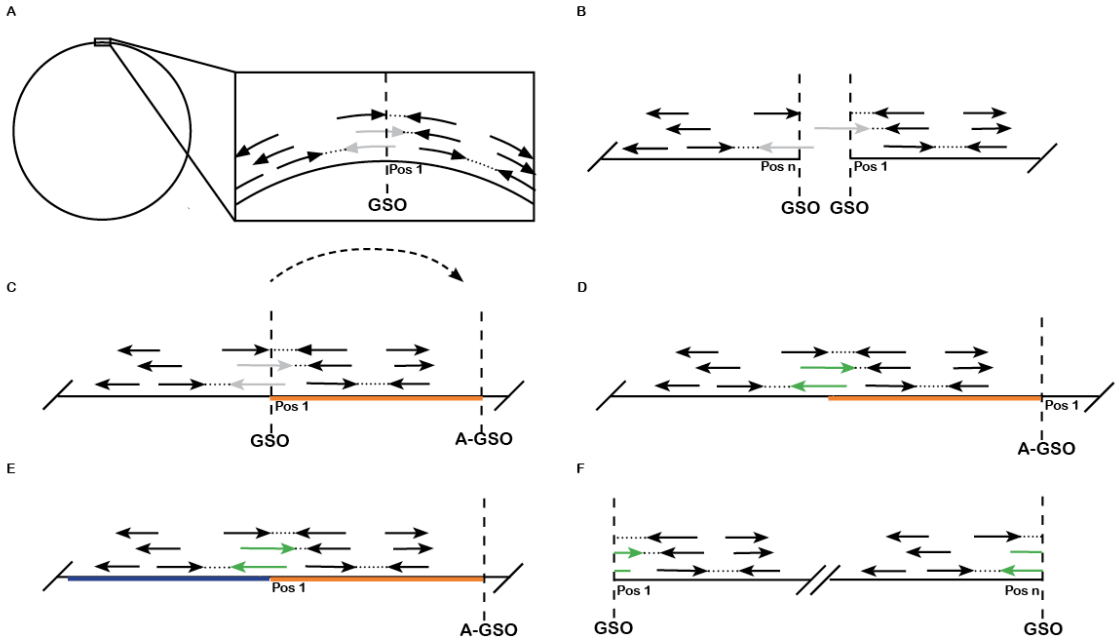

Figure 29. Workflow for Bolo.
A) Reads that align to the GSO in a circular genome. Reads that are aligned are shown in black, reads that are unable to align due to proximity to the GSO are shown in grey. B) To linearize a circular genome the reference sequence is cut at the GSO. (C) Reads aligned to the original genome. For the rotated alignment the GSO is shifted length d (orange line) to the end of the genome, creating the A-GSO. D) Reads are aligned to the rotated genome and original genomes are shown in black. Reads that were previously unable to map are now mapped (green). E) Reads that map between d and w (blue line) are compared between the original alignment and the pair with the higher quality score is selected to move to the output file. Reads moved from the adjusted to the original file have the position changed to match the original file. F) Reads that are determined to cross the GSO are split and moved to the output file, correctly aligned to the beginning and end of the genome (green).

Implementation

To process these circular genomes, an artificial origin must be made in the reference sequence, effectively linearizing the circular genome. Reads that align near the GSO are lost, as not enough base pairs matched to the reference genome occur to allow

for a successful mapping (Figure 29). Bolo recovers the resulting data loss near the GSO by rotating the reference genome, the user realigns the reads to the adjusted reference genome and Bolo then identifies and splits reads that cross the GSO.

Bolo contains a two-part algorithm, one part rotates the reference genome, and the second splits reads that cross the GSO. The rotated genome is produced by advancing the GSO forward a set number of base pairs (***d***) to produce a reference genome with an adjusted genome sequencing origin (A-GSO). The length that the A-GSO is advanced is recommended to be greater than the length of the reads to prevent the same reads from being lost when aligning to the rotated genome. The default for ***d*** is set to 1000. If the reference file contains multiple chromosomes the user may select the chromosomes to rotate using the "-p" option to either list chromosomes to rotate or the "-p all" option to rotate all chromosomes. Illumina reads aligned to both the rotated genome and the original genome serve as inputs for the second part of the Bolo algorithm (Figure 29 C and D). Bolo compares reads mapped to the rotated genome to the original genome within the total length of the genome minus ***d*** + ***w***, where ***w*** is at minimum the length of the reads and ending at total length of the genome minus d (Figure 29 D). The value of w needs to be at a minimum the length of the read, however, if indels occur near the GSO, ***w*** must be long enough to allow for these additional base pairs, therefore the default is set to 1500. Reads that map between these two locations are assessed to select which read pair contains the higher MAPQ score. Whether a read crosses the GSO is determined using the mapping location. If the read from the rotated alignment does cross the GSO it is split into two read pairs. In the case of a singleton

read, the first N base pairs are paired with a zero-length read to complete pair A, and the last L-N base pairs of the singleton read are paired with a zero-length read to complete pair B (Figure 30, Table 7, case 1). When the GSO occurs in the first read of a pair the first N base pairs of the first read are paired with a zero-length read to complete pair A. The last L-N base pairs of the first read are paired with the second read to complete pair B (Figure 30, Table 7, case 2). When the GSO occurs N base pairs into the second read. The first read is paired with the first N base pairs of the second read to complete pair A. The last L-N base pairs of the second read are paired with a zero-length read to complete pair B (Figure 30, Table 7, case 3). If the GSO occurs in both reads, N1 base pairs into the first read and N2 base pairs into the second read. The first N1 base pairs of the first read and the first N2 base pairs of the second read are paired to complete pair A. The last L-N1 base pairs of the first read and the last L-N2 base pairs of the second read are paired to complete pair B (Figure 30, Table 7, case 4). If multiple chromosomes are being split the -p option can be used to list chromosomes or -p all will split reads at the GSO for all chromosomes. These split reads are combined with the remaining reads from the original alignment into the output bam file, which can then be processed further and used for variant calling.

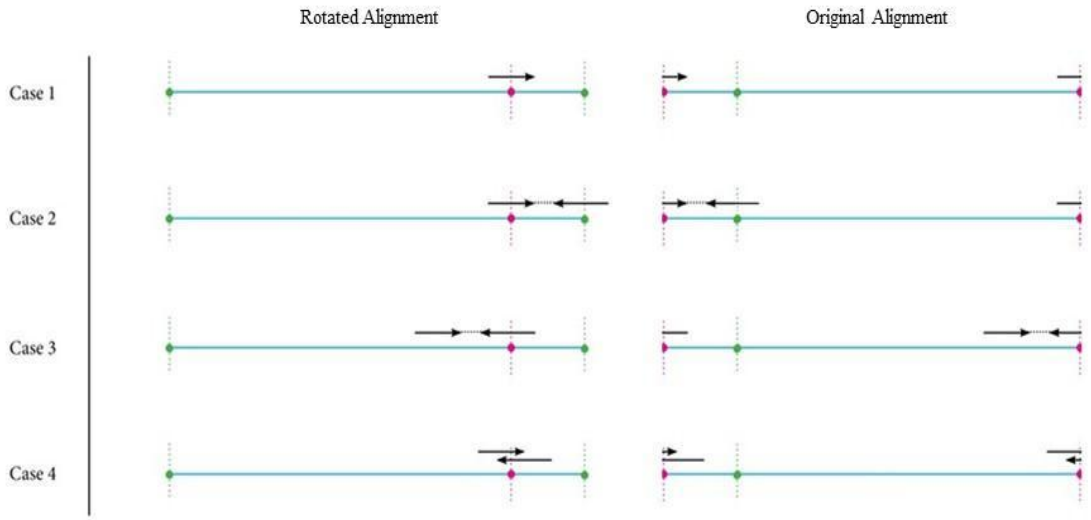

Figure 30. Bolo splitting strategy dependent on read location to A-GSO.

Table 7. Read splitting strategy dependent on read location to the A-GSO.

FWD	First	Add A				First N	0	First N	First N
REV	First	Add A	77	0	0	*	0	*	*

Pair B	**Read Copied**	**QNAME**	**FLAG**	**POS**	**MAPQ**	**CIGAR**	**TLEN**	**SEQ**	**QUAL**
FWD	First	Add B		Add N		Remaining L-N	0	Remaining L-N	Remaining L-N
REV	Second	Add B	77				0		

Manipulations for a Case 2 pair.

Pair A	**Read Copied**	**QNAME**	**FLAG**	**POS**	**MAPQ**	**CIGAR**	**TLEN**	**SEQ**	**QUAL**
FWD	First	Add A				First N	0	First N	First N
REV	First	Add A	77	0	0	*	0	*	*

Pair B	**Read Copied**	**QNAME**	**FLAG**	**POS**	**MAPQ**	**CIGAR**	**TLEN**	**SEQ**	**QUAL**
FWD	First	Add B		Add N		Remaining L-N	Subtract N	Remaining L-N	Remaining L-N
REV	Second	Add B					-(Subtract N)		

Manipulations for a Case 3 pair.

Pair A	**Read Copied**	**QNAME**	**FLAG**	**POS**	**MAPQ**	**CIGAR**	**TLEN**	**SEQ**	**QUAL**
FWD	First	Add A					Subtract N		
REV	Second	Add A				First N	-(Subtract N)	First N	First N

Pair B	**Read Copied**	**QNAME**	**FLAG**	**POS**	**MAPQ**	**CIGAR**	**TLEN**	**SEQ**	**QUAL**
FWD	Second	Add B		Add N		Remaining L-N	0	Remaining L-N	Remaining L-N
REV	Second	Add B	77	0	0	*	0	*	*

Manipulations for a Case 4 pair.

Pair A	**Read Copied**	**QNAME**	**FLAG**	**POS**	**MAPQ**	**CIGAR**	**TLEN**	**SEQ**	**QUAL**
FWD	First	Add A				First N1	N1	First N1	First N1
REV	Second	Add A				First N2	-(N1)	First N2	First N2

Pair B	**Read Copied**	**QNAME**	**FLAG**	**POS**	**MAPQ**	**CIGAR**	**TLEN**	**SEQ**	**QUAL**
FWD	First	Add B		Add N1		Remaining L-N1	Subtract N1	Remaining L-N1	Remaining L-N1
REV	Second	Add B		Add N2		Remaining L-N2	-(Subtract N1)	Remaining L-N2	Remaining L-N2

Methods

Read depth

Reads from BW25113 sequenced (Chapter II) to an average coverage of approximately 400 were down sampled using BBMap reformat.sh and varying sample rates were used to create the depths tested. Reads were aligned with Bowtie2 (v2.3.4.1) globally, and local alignments were done using the –local flag. For all Bowtie2 alignments default settings were used with the exception of distance between reads, which was set to a max of 3000 using flags –I 0 and –X 3000. Default settings were used for BWA MEM (v0.7.15) alignments, and for SAMtools (1.7) when sorting and indexing bam files. BCFtools (1.9) (using consensus caller flags -cv) and freebayes (v1.3.2) variant callers were used with ploidy set to 1. A bedfile or --region was used to call variants present in the first and last 500 base pairs. Positive variants were defined as those with at least 80% of reads containing the alternative allele.

Table 8. Genomes used in Chapter IV.

Strain name	genotype	obtained/created
KH6	**BW25113 *dgcI*^{L323L} *wrbA*^{G96V} *nemA*^{Y359}**	CGSC:7926
SDB262	**BW25113**	Gift from Siegele Lab
SDB264	**BW25113 *rrsA* A196T**	Gift from Siegele Lab
SDB265	**BW25113 *rrsA* A196T**	Gift from Siegele Lab

Variants

To assess Bolo's ability to identify different mutations, five different variants were made in the BW25113 reference file, downloaded January 2020 from NCBI. A single base pair

change was mimicked from changing the first base from a T to an A. To test a single base pair deletion a T was inserted at position 1 in the reference file, a 10 base pair deletion was evaluated by inserting the random sequence TGGGGACTGA before the first base pair of the reference genome. Single and 10 base pair deletions were emulated by deleting 1 and 10 base pairs, respectively, from the beginning of the reference file. To keep the sequence upstream and downstream of the variants constant and minimize changes in read depth, we introduced variants starting at position 1 in the reference genome. We then rotated the variant forward 1 base pair for each iteration to determine if callers were able to identify variants with and without Bolo.

Downloaded Data

Salmonella typhimurium (DRR079316), *Arabidopsis thaliana* (DRR001194) and *Homo sapiens* (SRR3560942) genomes were downloaded from the European Nucleotide Archive. To rotate both chromosomes of *S. typhimurium,* bolo was used with the "-p all" option. To rotate the chloroplast genomes from *A. thaliana*, the option -p NC_037304.1,NC_000932.1 was used.

Software

Bolo utilizes the SeqAn (2.0+) library (173).

Results

Bolo recovers read depth

Read depth reduction occurs most drastically near the GSO if aligners are unable to determine that reads crossing the GSO can map to both the beginning and end of the alignment. In a global alignment, portions of the reads that map to the other side of the GSO will be treated as gaps and the ability of the read to map is dependent on the severity of the gap. For example, a read that maps 60% to beginning of the genome (right of the GSO) will not map because 40% of the read cannot be mapped. While reads that map almost entirely to one side of the GSO are able to be mapped. The region that is unable to be mapped is treated like a gap, and the number of base pairs this can be and still result in a proper mapping is dependent on software programs and settings.

In a local alignment 20 base pairs are first aligned and then extended to give the highest possible quality score. Local aligners allow trimming of the reads to produce the best quality scores, while global aligners have to map the entire read, this is predicted to result in global aligners having less reads map near the GSO. Aligning with BWA MEM (110)is predicted to result in the least amount of data loss because it permits read splitting, enabling both parts of the read to map to different locations in the genomes, including either side of the GSO. For the split reads the length of matches must be longer than the seed region to ensure the ability to map, as such BWA MEM resulted in the highest read depth of the aligners tested near the GSO (Figure 31). Read depth decrease near the GSO was quantified by down-sampling BW25113 wildtype reads to an average coverage of 100, and were aligned to the reference sequence using Bowtie2 (109)both

local and global aligners, as well as BWA MEM. Read depth decreased near the GSO for all three alignments (Figure 31). As expected, the most severe data loss occurred in the global alignment, and the least data loss occurred with BWA MEM. To validate that Bolo was able to recover reads mapped near the GSO, coverage near the GSO was compared pre- and post-bolo alignments (Figure 31). The range of recovery surrounding the GSO demonstrated dependence on read length. For reads 250 nucleotides in length aligned with a global aligner, the read depth increased 250 base pairs on either side of the GSO, while for reads with a nucleotide length of 150 read depth increased an average of 150 base pairs from the GSO (Figure 31).

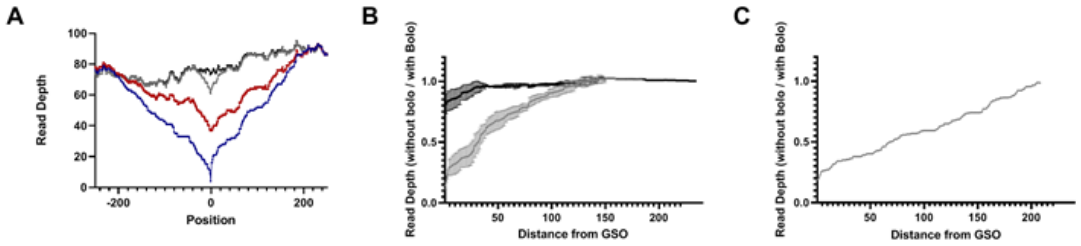

Figure 31. Bolo recovers read depth near GSO for multiple mappers.
A) Read depth loss near the GSO for BWA MEM (grey), Bowtie2 local (red), Bowtie2 end to end (blue) and read depth recovered by Bolo (black). B) Read depth recovered for read length 150, aligned with BWA (black), and Bowtie2 global (grey). Average of three genomes normalized to the Bolo read depth. Error bars standard deviation C) Read depth recovered for read length 250 aligned with Bowtie2 global (grey), n=1.

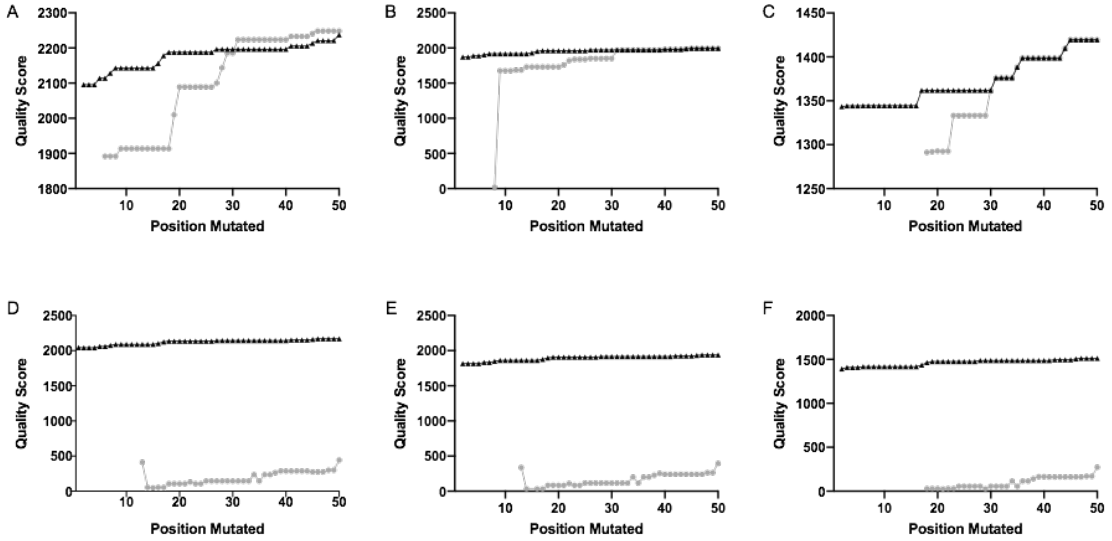

Figure 32. Variants are correctly identified near GSO using Freebayes with Bolo Bolo (black) and without Bolo (grey). (A-C) BWA MEM aligner (D-F) Bowtie2 global aligner. (A and D) Single base pair change from T to A. (B and E) Single base pair insertion, insertion of a T. (C and F) 10 base pair insertion, AGCTTTTCAT.

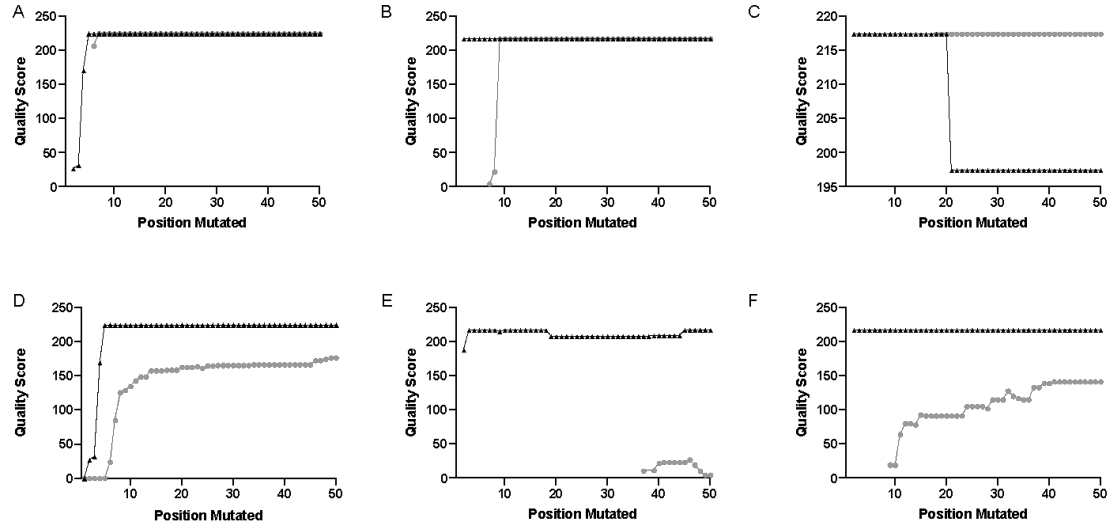

Figure 33. Variants are correctly identified near GSO using BWA MEM with Bolo. Bolo (black) and without Bolo (grey). (A-C) BWA MEM aligner (D-F) Bowtie2 global aligner. (A and D) Single base pair change from T to A. (B and E) Single base pair insertion, insertion of a T. (C and F) 10 base pair insertion, AGCTTTTCAT.

Using BCFtools (112) the base pair change can be identified at position 2 and 3 with a low-quality score (Figure 33). At these positions, the variant distance base (VDB) is low. The callers can weigh the position of the variant, for example, mismatches at the end of a read are score more stringently because the end of the reads usually have a higher error rate. The VDB accounts for the location of the variant within the read. This suggests that proximity of the variant to the end of the read causes the low-quality score. Based on how the programs are constructed, Freebayes (113)is more accurate immediately on either side of the GSO.

We tested the callers' ability to identify single base pair insertions and deletions with and without Bolo. These variants were correctly identified with high quality scores by both callers at all positions tested (Figure 32, 32 B and E). Since the variants occur at position 1 the callers both call an insertion or deletion at position 1 when the insertion or deletion is in position 2. Freebayes contains a realigner designed to find small insertions and deletions, enabling this caller to find the variants sooner than BCFtools. Larger deletions and insertions had a similar pattern where use with Bolo was able to find the variants earlier than without Bolo, and using Bolo increased the read depth and quality scores of those variants when the variants were closer to the GSO (Figure 32,33, C,F)

Overall, these findings demonstrate Bolo enables identification of variants including insertions, deletions, and single base pair mutations which would otherwise be lost using current genome analysis software.

Read depth needed to identify a variant

Bolo enables callers to find variants closer to the GSO as a result of the increase in read depth. To determine the maximum read depth for which Bolo is beneficial to the user, we used a set of reads sequenced originally to a depth of 400, and down sampled that set to depths of 200, 100, 50, and 25. Using again both aligners and callers, we tested the effect of read depths on identification of the single base pair change variant. At an average depth of 25, Bolo was able to identify variants which were only called at a depth of 400 without Bolo. Increasing the average read depth resulted in an increase in BCFtools quality scores up to a depth of 50, past which Bolo did not alter the quality scores with BCFtools (Figure 34).

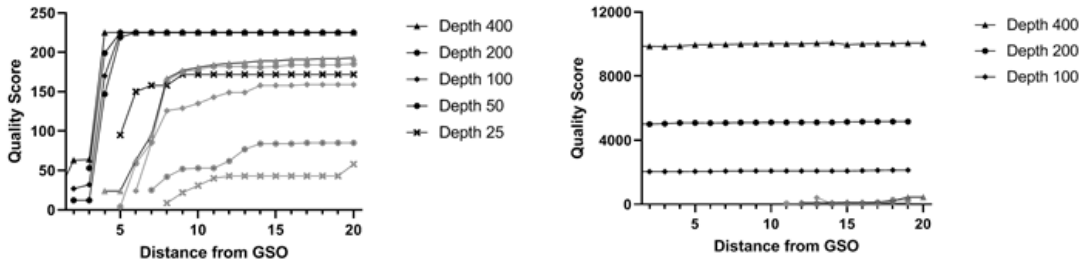

Figure 34. Coverage needed to identify a variant.
A) Down sampled reads aligned with Bowtie2 global and called with BCFtools. B) Down sampled reads aligned with Bowtie2 and called with freebayes. With Bolo (black), without Bolo (grey).

Bolo recovers read depth in multiple genomes

While nearly all organisms contain a circular chromosome, many contain multiple circular chromosomes, in the case of bacterial plasmids, and plastids and mitochondria in plants. To confirm that Bolo is able to recover read depth from multiple chromosomes, we compared coverage pre- and post-Bolo for *Salmonella typhimurium* LT2 which contains an approximate 100 kb plasmid. Read depth is recovered for both circular chromosomes with the use of Bolo (Figure 35). To test that Bolo splits reads of the circular chromosomes but not reads that cross the edges of linear chromosomes, we tested *Arabidopsis thaliana* which contains linear chromosomes and a mitochondria and chloroplast genomes. Reads were recovered for both circular chromosomes near the GSO, while the edges of the linear chromosomes were unaltered. The use of Bolo aligned many reads to the adjusted reference sequence near the GSO that in the original alignment had mapped to the linear chromosomes. Since Bolo selected the read pair with the higher MAPQ score this shows an additional benefit of Bolo to increase the accuracy of the alignment even in the linear chromosomes.

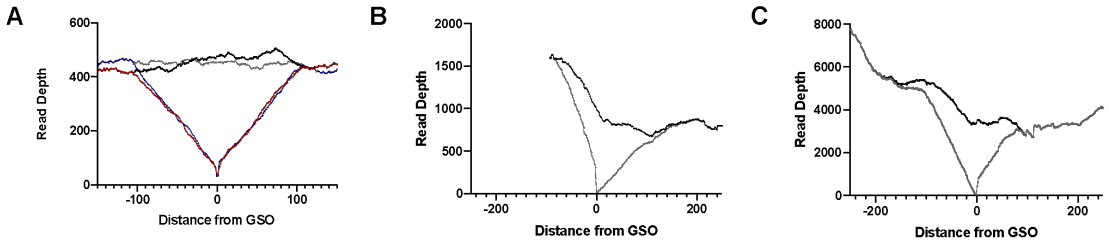

Figure 35. Bolo recovers read depth for multiple chromosomes.
A) *Salmonella typhimurium* LT2 aligned with Bowtie2 global. Read depth with Bolo shown in black (chromosome) and grey (plasmid) and without Bolo red (chromosome) and blue (plasmid). B) Reads mapped to the *Arabidopsis thaliana* mitochondria and C) chloroplast with Bolo (black) without Bolo (grey).

Discussion

The use of Bolo can increase read depth for local and global aligners and enhances the split function of BWA MEM. Due to a decrease in read depth that occurs near the GSO callers are unable to correctly identify variants unless the organism has been sequenced to a high read depth (Figure 34). The ability to sequence to a lower depth and retain the ability to identify variants is provided by utilizing Bolo within a pipeline.

We confirm that Bolo recovers read depth for both global and local aligners and enhances two popular callers BCFtools and Freebayes to correctly identify variants near the GSO. For variants near the GSO we have found BCFtools to be at a disadvantage due to accounting for the VDB. Although it is known that Freebayes has a high false positive rate (174). However, both callers fail to identify single base pair insertions and deletions at the first position in the reference file due to the callers needing a base pair prior to the altered base. Manual examination of the alignment after Bolo can be done to quickly identify a mutation at position 1, but future work is needed to create or adapt callers to identify these variants.

Bolo is designed to work with standard bam files, enabling it to be incorporated into current and future pipelines. Allowing for automation of short read genome sequencing, and increased data coverage for circular genomes.

CHAPTER V

CONCLUSIONS

The goal of this dissertation was to identify the causal genes responsible for band formation during growth on swim media (SMM-Na) at 37 °C. Here we find that mutations in *lrhA* or mutations distal to the start codon of *flhD* result in a banding pattern on SMM-Na. Several groups have reported an increased motility phenotype when IS elements disrupt the intergenic region between *flhD* and *uspC* (92–94). Here we show that the location of the IS element affects the motility phenotype. Insertions clustered near *uspC* occur in band forming strains, while mutations that are close to *flhD* form chemotaxis rings, as previously reported. The location of the IS element present in the *flhD-uspC* intergenic region should be accounted for when used for chemotaxis and motility studies.

Pattern formation is dependent on the presence of flagella. At 30 °C, a permissive temperature for flagellar synthesis, wild type *E. coli* will band on SMM-Na, but is non-motile at 37 °C, this is likely due to a lack of flagellar synthesis. We identified two mutations are predicted to increase flagellar synthesis and form bands at 37 °C. Strains that migrate faster than the banding strains form chemotaxis rings instead of bands, which we hypothesize is due to increased flagellar synthesis. Alternatively, faster migration in swim agar can be accomplished by increasing tumbling frequency which free cells from agar. Cells that have a faster migration rate have a longer doubling time (126).

We have shown that the global regulator (p)ppGpp, which is activated under nutrient stress, is essential for banding formation in KH318 (Δ*lrhA*). The ΔRS Δ*lrhA* strain has a reduction in swim diameter and fails to form bands. Mutations in RNA polymerase can return a banding phenotype to KH318, and similar suppressor mutations have been previously identified to suppress a ΔRS phenotype (129). We hypothesis that (p)ppGpp responding to a sensing stress is required for the increased swim diameter in our KH318 strain. We have not yet tested whether (p)ppGpp is also necessary for band formation in KH1193. Although (p)ppGpp can induce expression of *uspC*, the location of the IS elements clustered near *uspC* could potentially increase expression of *uspC* in a (p)ppGpp independent manner.

In both *E. coli* and *Proteus*, (p)ppGpp accumulation has been shown to inhibit *flhDC* (62, 133, 134). However, there is conflicting evidence in *E. coli* that shows that (p)ppGpp is necessary for flagellar production on LB swim plates (132). Likewise, in *P. aeruginosa*, (p)ppGpp synthesis is required for swarming (175). Synthesis of (p)ppGpp is also needed for accumulation of Lrp (149), which can directly regulate *lrhA* expression (37). Lrp contributes to repression of flagellar genes in liquid culture with rich media, while causing a slight upregulation in minimal media. The proposed mechanism is that under nutrient rich conditions, cells will repress motility to stay in an environment conducive for rapid growth, and when challenged with nutrient limitation, will move to find a better niche (37). This is in contrast to another hypothesis that due to the high energy cost of synthesizing flagella, cells under nutrient limitation will conserve resources by not making flagella. In the banding pattern, we can speculate that both

hypotheses may be correct. The requirement of (p)ppGpp suggests that cells are moving under stress conditions to find a better environment, but cells behind the swim fronts are non-motile. The switch from motile to non-motile in the progenitors located behind migrations front have been proposed to be in response to nutrient level depletion (68, 104).

We propose a model where cells inoculated in the center of the plate grow and divide rapidly depleting the local nutrients. In conditions permissive for flagella synthesis motile cells will migrate from the inoculation center. The migration speed is regulated by cells growth and chemotaxis. As cells migrate they deplete the nutrient level below a threshold that causes a switch from motile to non-motile cells. The non-motile cells form a stable band. Motile cells emerge from the band and repeat the cycle

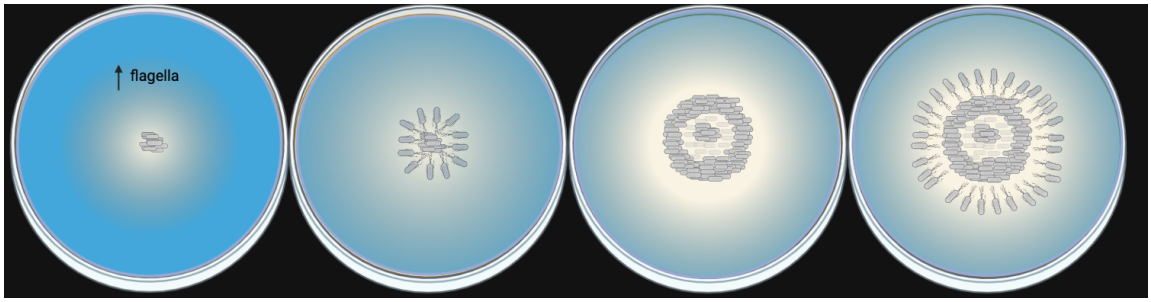

Figure 36. Proposed model of band formation.
A) Cells inoculated in the center of the plate grow rapidly and begin to consume nutrient (represented by a blue gradient) B) After a period of time motile cells leave the inoculation center and migrate in accordance to chemoeffectors present in the plate. C) As cells migrate they deplete the nutrient level below a threshold and switch from motile to non-motile cells. D) This cycle is repeated and forms a bull's eye pattern.

The stress that (p)ppGpp is sensing in SMM-Na is still unknown. While we provide preliminary data supplementation with 10mM ferrous sulfate did slow the expansion of the chemotaxis strain, while returning the banding strains to a non-motile wild type state. During swarming in *P. aeruginosa,* cells behind the swarm front increase expression of siderophores that accumulate iron in the cell (84). Mutations in the *E. coli* siderophore enterobactin (Ent) or in the transport system that moves Ent bound with iron back into the cytoplasm cause swarm defects in *E. coli* (153). It is possible that iron becomes limiting behind the swim front as the cells on the edge of the swim front uptake iron from the media. However, this requires further confirmation.

The addition of 20 mM short chain fatty acids (SCFA) causes a non-banding but motile phenotype. The addition of SCFA can affect flagellar synthesis in addition to, feeding into central metabolism (156, 176). It is unknown which process is contributing to the loss of banding with the addition of SCFA. The addition of sodium pyruvate to SMM-Na alters the banding pattern, suggesting that changes to central metabolism can affect banding. Supplementation with 100mM of each sodium acetate, butyrate, and propionate all induce banding in wild type strains at 37 °C, supporting that SCFA alter flagellar expression.

We propose an advantage that the strains that can band migrate at different rates. We propose this as a way for multiple strains to cohabitate in the same region. The different banding types have varying migration rates, where KH707 (banding type 2) moves faster that KH318 (banding type 1). Both strains migrate slower than KH1220 which does not form bands but chemotaxis rings on SMM-Na. Faster migration in swim

agar can be due to cells that increase their tumble frequency and decrease their run time (91, 126). This is opposite of how faster migration occurs in liquid conditions where cells increase their runtime and suppress tumbling (95). This can be explained by the presence of the agar matrix that can trap cells, cells that tumble more frequently can free themselves from the agar and spend more time moving. We have yet to perform single cell tracking on our banding strains to test if they have an altered tumbling frequency or run time.

The different banding strains seem to utilize different chemoreceptors. In KH318 deletions of *tsr* and *tap*, formed a non-banding and a non-motile phenotype respectively. While the same phenotype existed for deletion of *tsr*, *tap* in KH707 the deletion of *tar* resulted in a non-banding phenotype. The deletion of *tar* in KH318 produces wider bands. We see that the deletion of *tar*, or growth on maltose gives us the same pattern for Kh318. Tar is able to sense maltose and also aspartate. The addition of maltose could affect Tar activity, although it is unknown if this occurs by saturation of the Tar receptor making it unable to sense a gradient or changes to the run and tumble frequencies.

Pattern formation in both strains is dependent on a functional Tsr receptor, which senses serine, although the addition of serine has slightly different effects on KH318 and KH707. In the KH707 strain, we see an intense band at the edge of the swim front that is also seen in the KH1220 strain. This allows us to propose that the addition of serine alters the migration speed of the cells by changing the serine gradient.

In both strains, the loss of *tap* resulted in a non-motile phenotype, suggesting that strains move forward by sensing dipeptides. However, these strains also contained IS elements disrupting *rhsD*. Experiments are needed to determine if the non-functional *rhsD* or dipeptide sensing is important for band formation.

Evolution experiments in Liu et al., found that repeatedly selecting for cells on swim plates that different distances from the inoculation center resulted in strains that had evolved to swim to that distance. These strains were able to cohabitate together when competing against each other. Cells evolved on swim plates that did not contain chemoattractants always increased their expansion rate and did not evolve to migrate to a set distance (103). We propose that chemotaxis plays a commanding role in this process to direct cells to evolve to use resources not already consumed by faster migrating cells.

During the course of this dissertation, we routinely whole-genome sequenced strains to identify mutations. In making the analysis of the genome sequences more complete we created a program and automates read recovery at the Genome Sequencing Origin (GSO). This allows us to include the reads on either side of the GSO in our analysis. While others have created programs similar to ours, ours provides a fully automated and open source program that can be easily added to current analysis pipelines.

Future directions

We still lack a complete gene regulatory network, linking (p)ppGpp, *lrhA* and *uspC*. Although we propose a link between (p)ppGpp and *lrhA* involving Lrp. It is not clear if Δ*lrhA* and IS5 -510 *flhD* (KH707/KH1193) are in the same pathway or represent

two distinct pathways for pattern formation. We see several differences in the motility phenotypes between the two strains including response to loss of chemoreceptors, pattern formation in the presence of serine and maltose, and the motility response in the presence of short chain fatty acids. We hypothesize that deletion of *lrhA* in the KH1193 strain will have an additive effect on the migration rate. We propose this based on the observation that non-banding outgrowths always occur in our KH318 strain after incubation on SMM-Na, which appear to have an increased expansion rate and occur when IS elements insert in the *flhD-uspC* intergenic region. (p)PpGpp has been shown to upregulate uspC therefore, we are testing the effect of a ΔRS strain now in the KH1193 background. We predict that the location of the IS element may be a way to cause banding in a (p)ppGpp-independent fashion.

We lack confirmation of what makes the banding strains motile in SMM-Na at 37 °C. We propose to use electron microscopy to quantify flagellar numbers in banding strains compared to BW25113. We propose to measure tumbling rate and swimming speed, using single cell tracking and cell tethering assays. We propose to quantify flagellar number, swimming speed, and tumbling frequency in the banding strains KH318 and KH1193 but compared to the chemotaxis strain KH1220 and parental BW25113 in the presence of aspartate, maltose, serine, and at 37 °C and 30 °C.

While Δ*uspC* strains have been reported as unable to produce flagella, how *uspC* modulates flagellar synthesis has not been yet discovered. We have identified that motile, but non-banding outgrowths form in the Δ*uspC* strain. Whole-genome sequencing these strains could identify the role of *uspC* in flagellar regulation. To

determine if IS elements in the *flhD-uspC* intergenic region affect expression of *flhD*, *uspC*, or both due to topographically changes to the intergenic region. We propose to express the intergenic region containing IS elements from KH1220 and KH1193 along with the protein coding region of *flhD* or *uspC.* These would be expressed in either a Δ*flhD* and Δ*uspC* respectively. We anticipate the intergenic region from the *uspC* start codon through *flhD* containing an IS -320 (KH1220) will form chemotaxis rings when expressed in a Δ*flhD* strain, while the construct containing the IS -510 will be non-motile. While the construct from *flhD* start codon through the protein coding region of *uspC* containing the IS -510 expressed in a ΔuspC strain will result in a banding pattern, while the IS -320 will be non-motile. It is plausible that due to disruption of the topology of intergenic region that both the IS elements affect expression of both genes. Expression changes can be measured using qPCR with the above constructs and strains.

In addition to the strains utilized for experiments in this dissertation, we have uncovered a number of banding strains in Keio mutant strains. Several of these strains have the ability to modify the pattern. Such as Δ*cyoD*, (KH869) where the bands are only apparent if held in the right lighting. The gene that contributes to this phenotype is not known; banding is likely caused by the presence of an IS*5* element -512 base pairs from *flhD*. This strain also contains a T-C base pair change 6 nucleotides from the *kgtP* start codon (*rrlG-kgtP*), and an IS30 element that disrupts *mhpT*. It is unknown what background mutation modifies the banding phenotype. *cyoD* is a subunit of the cytochrome BO oxidase, which is the oxidase used during growth under aerobic conditions (177). While *kgtP* is an alpha-ketoglutarate permease (178). These mutations

suggest that bands may require aerobic growth conditions to form bands. We propose looking at band formation during microaerobic and aerobic conditions to test if band formation is altered.

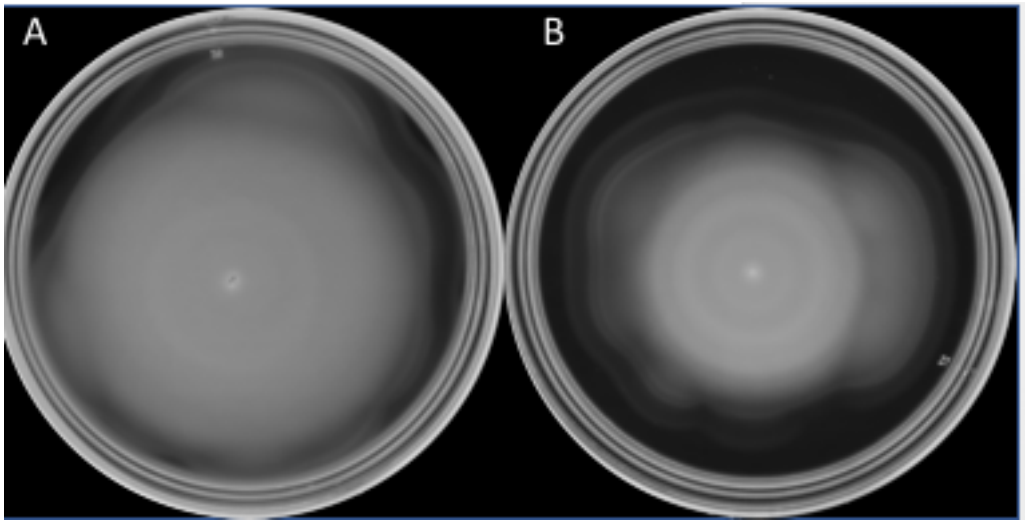

Figure 37. Motility phenotype is altered in a Δ*cyoD* mutant.
SMM-Na plates were incubated at 37 °C for 14 hours A) KH869 B) KH1193

REFERENCES

1. Verstraeten N, Braeken K, Debkumari B, Fauvart M, Fransaer J, Vermant J, Michiels J. 2008. Living on a surface: swarming and biofilm formation. Trends Microbiol 16:496–506.
2. Harshey RM. 2003. Bacterial motiliyt on a surface: Many Ways to a Common Goal.
3. Mattingly AE, Weaver AA, Dimkovikj A, Shrout JD. 2018. Assessing travel conditions: Environmental and host influences on bacterial surface motility. J Bacteriol 200:1–17.
4. Kearns DB. 2011. A field guide to bacterial swarming motility. Nat Rev Micro 8:634–644.
5. McCarter L, Hilmen M, Silverman M. 1988. Flagellar dynamometer controls swarmer cell differentiation of V. parahaemolyticus. Cell 54:345–351.
6. Tasaki S, Nakayama M, Shoji W. 2017. Morphologies of Bacillus subtilis communities responding to environmental variation. Dev Growth Differ 59:369–378.
7. Budrene EO, Berg HC. 1991. Complex patterns formed by motile cells of Escherichia coli. Group 349:630–633.
8. Mittal N, Budrene EO, Brenner MP, Van Oudenaarden A. 2003. Motility of Escherichia coli cells in clusters formed by chemotactic aggregation. Proc Natl Acad Sci U S A 100:13259–13263.
9. Budrene EO, Berg, Berg HC. 1995. Dynamics of formation of symmetrical

patterns by chemotactic bacteria. Nature.

10. Liu C, Fu X, Liu L, Ren X, Chau CKL, Li S, Xiang L, Zeng H, Chen G, Tang LH, Lenz P, Cui X, Huang W, Hwa T, Huang JD. 2011. Sequential establishment of stripe patterns in an expanding cell population 334:238–241.
11. Karig D, Martini KM, Lu T, DeLateur NA, Goldenfeld N, Weiss R. 2018. Stochastic Turing patterns in a synthetic bacterial population. Proc Natl Acad Sci 115:201720770.
12. Curatolo AI, Zhou N, Zhao Y, Liu C, Daerr A, Tailleur J, Huang J. Cooperative pattern formation in multi-component bacterial systems through reciprocal motility regulation. Nat Phys.
13. Fraser GM, Hughes C. 1999. Swarming motility. Curr Opin Microbiol 2:630–635.
14. Morgenstein RM, Szostek B, Rather PN. 2010. Regulation of gene expression during swarmer cell differentiation in Proteus mirabilis. FEMS Microbiol Rev 34:753–763.
15. Schaffer JN, Pearson MM. 2016. Proteus mirabilis and Urinary Tract Infections . Urin Tract Infect 3:383–433.
16. Furness RB, Fraser GM, Hay NA, Hughes C. 1997. Negative feedback from a Proteus class II flagellum export defect to the flhDC master operon controlling cell division and flagellum assembly. J Bacteriol 179:5585–5588.
17. Clemmer KM, Rather PN. 2007. Regulation of flhDC expression in Proteus mirabilis. Res Microbiol 158:295–302.
18. Patrick JE, Kearns DB. 2012. Swarming motility and the control of master

regulators of flagellar biosynthesis Joyce. Bone 23:1–7.

19. McCarter L. 1999. The multiple identities of Vibrio parahaemolyticus. J Mol Microbiol Biotechnol 1:51–57.

20. Belas R, Simon M, Silverman M. 1986. Regulation of lateral flagella gene transcription in Vibrio parahaemolyticus. J Bacteriol 167:210–218.

21. Bisset KA, Douglas c. w. I. 1975. A continous study of morphological phase in the swarm of Proteus. 9:229–231.

22. Douglas CWI, Bisset KA. 1976. Development of concentric zones in the Proteus sarm colony. 9:497–500.

23. Pearson MM, Rasko DA, Smith SN, Mobley HLT. 2010. Transcriptome of swarming Proteus mirabilis. Infect Immun 78:2834–2845.

24. Williams FD, Schwarzhoff RH. 1978. Nature of the swarming phenomenon in Proteus. Annu Rev Microbiol 32:101–122.

25. Osterman IA, Dikhtyar YY, Bogdanov AA, Dontsova OA, Sergiev P V. 2015. Regulation of flagellar gene expression in Bacteria. Biochem 80:1447–1456.

26. Soutourina OA, Bertin PN. 2003. Regulation cascade of flagellar expression in Gram-negative bacteria. FEMS Microbiol Rev 27:505–523.

27. Fitzgerald DM, Bonocora RP, Wade JT. 2014. Comprehensive Mapping of the Escherichia coli Flagellar Regulatory Network. PLoS Genet 10.

28. Pearson MM, Sebaihia M, Churcher C, Quail MA, Seshasayee AS, Luscombe NM, Abdellah Z, Arrosmith C, Atkin B, Chillingworth T, Hauser H, Jagels K, Moule S, Mungall K, Norbertczak H, Rabbinowitsch E, Walker D, Whithead S,

Thomson NR, Rather PN, Parkhill J, Mobley HLT. 2008. Complete genome sequence of uropathogenic Proteus mirabilis, a master of both adherence and motility. J Bacteriol 190:4027–4037.

29. Dufour A, Furness RB, Hughes C. 1998. Novel genes that upregulate the Proteus mirabilis flhDC master operon controlling flagellar biogenesis and swarming. Mol Microbiol 29:741–751.

30. Cano DA, Domínguez-Bernal G, Tierrez A, Garcia-del Portillo F, Casadesús J. 2002. Regulation of capsule synthesis and cell motility in Salmonella enterica by the essential gene igaA. Genetics 162:1513–1523.

31. Wall E, Majdalani N, Gottesman S. 2018. The Complex Rcs Regulatory Cascade. Annu Rev Microbiol 72:111–139.

32. McCullen CA, Benhammou JN, Majdalani N, Gottesman S. 2010. Mechanism of positive regulation by DsrA and RprA small noncoding RNAs: Pairing increases translation and protects rpoS mRNA from degradation. J Bacteriol 192:5559–5571.

33. Belas R, Schneider R, Melch M. 1998. Characterization of Proteus mirabilis precocious swarming mutants: Identification of rsbA, encoding a regulator of swarming behavior. J Bacteriol 180:6126–6139.

34. Morgenstein RM, Rather PN. 2012. Role of the Umo proteins and the rcs phosphorelay in the swarming motility of the wild type and an O-antigen (waaL) mutant of Proteus mirabilis. J Bacteriol 194:669–676.

35. Armbruster CE, Hodges SA, Mobley HLTT. 2013. Initiation of swarming motility

by Proteus mirabilis occurs in response to specific cues present in urine and requires excess L-glutamine. J Bacteriol 195:1305–1319.

36. Han W, Wu B, Li L, Zhao G, Woodward R, Pettit N, Cai L, Thon V, Wang PG. 2012. Defining function of lipopolysaccharide O-antigen ligase waal using chemoenzymatically synthesized substrates. J Biol Chem 287:5357–5365.

37. Kroner GM, Wolfe MB, Freddolino PL. 2019. Escherichia coli Lrp regulates one-third of the genome via direct, cooperative, and indirect routes. J Bacteriol.

38. Hay NA, Tipper DJ, Gygi D, Hughes C. 1997. A nonswarming mutant of Proteus mirabilis lacks the Lrp global transcriptional regulator. J Bacteriol 179:4741–4746.

39. Mobley HLT, Belas R. 1995. Swarming and pathogenicity of Proteus mirabilis in the urinary tract. Trends Microbiol 3:280–284.

40. Rauprich O, Matsushita M, Weijer CJ, Siegert F, Esipov SE, Shapiro JA. 1996. Periodic phenomena in Proteus mirabilis swarm colony development. J Bacteriol 178:6525–6538.

41. Alavi M, Belas R. 2001. Surface sensing, swarmer cell differentiation, and biofilm development. Methods Enzymol 336:29–40.

42. Belas R, Suvanasuthi R. 2005. The ability of Proteus mirabilis to sense surfaces and regulate virulence gene expression involves fliL, a flagellar basal body protein. J Bacteriol 187:6789–6803.

43. Belas R, Goldman M, Ashliman K. 1995. Genetic analysis of Proteus mirabilis mutants defective in swarmer cell elongation. J Bacteriol 177:823–828.

44. Schoenhals GJ, Macnab RM. 1999. FliL is a membrane-associated component of the flagellar basal body of Salmonella. Microbiology 145:1769–1775.

45. Partridge JD, Nieto V, Harshey RM. 2015. A new player at the flagellar motor: FliL controls both motor output and bias. MBio 6:1–11.

46. Suaste-Olmos F, Domenzain C, Mireles-Rodríguez JC, Poggio S, Osorio A, Dreyfus G, Camarena L. 2010. The flagellar protein FliL is essential for swimming in Rhodobacter sphaeroides. J Bacteriol 192:6230–6239.

47. Jenal U, White J, Shapiro L. 1994. Caulobacter flagellar function, but not assembly, requires FliL, a non-polarly localized membrane protein present in all cell types. J Mol Biol.

48. Lee YY, Patellis J, Belas R. 2013. Activity of Proteus mirabilis FliL is viscosity dependent and requires extragenic DNA. J Bacteriol 195:823–832.

49. Lee YY, Belas R. 2015. Loss of FliL alters proteus mirabilis surface sensing and temperature-dependent swarming. J Bacteriol 197:159–173.

50. Hatt JK, Rather PN. 2008. Characterization of a novel gene, wosA, regulating FlhDC expression in Proteus mirabilis. J Bacteriol 190:1946–1955.

51. Yu J, Shapiro L. 1992. Early Caulobacter crescentus genes fliL and fliM are required for flagellar gene expression and normal cell division. J Bacteriol 174:3327–3338.

52. Chawla R, Ford KM, Lele PP. 2017. Torque, but not FliL, regulates mechanosensitive flagellar motor-function. Sci Rep 7:1–9.

53. Belas R, Erskine D, Flaherty D. 1991. Proteus mirabilis mutants defective in

swarmer cell differentiation and multicellular behavior. J Bacteriol 173:6279–6288.

54. Prub BM, Nelms JM, Park C, Wolfe AJ. 1994. Mutations in NADH:Ubiquinone Oxidoreductase.

55. Siegele DA, Bain S, Mao W. 2010. Mutations in the flhD gene of Escherichia coli K-12 do not cause the reported effect on cell division. FEMS Microbiol Lett 309:94–99.

56. Gervais FG, Phoenix P, Drapeau GR. 1992. The rcsB gene, a positive regulator of colanic acid biosynthesis in Escherichia coli, is also an activator of ftsZ expression. J Bacteriol 174:3964–3971.

57. Carballès F, Bertrand C, Bouché JP, Cam K. 1999. Regulation of Escherichia coli cell division genes ftsA and ftsZ by the two-component system rcsC-rcsB. Mol Microbiol 34:442–450.

58. Howery KE, Clemmer KM, Şimşek E, Kim M, Rather PN. 2015. Regulation of the min cell division inhibition complex by the Rcs phosphorelay in Proteus mirabilis. J Bacteriol 197:2499–2507.

59. Pichoff S, Lutkenhaus J. 2001. Escherichia coli division inhibitor mincd blocks septation by preventing Z-ring formation. J Bacteriol 183:6630–6635.

60. de Boer PAJ, Crossley RE, Rothfield LI. 1989. A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli. Cell 56:641–649.

61. Allison C, Lai H, Hughes C. 1992. Co-ordinate expression of virulence genes

during swarm-cell differentiation and population migration of Proteus mirabilis 6:1583–1591.

62. Tipping MJ, Gibbs KA. 2019. Peer pressure from a Proteus mirabilis self-recognition system controls participation in cooperative swarm motility. PLoS Pathog 15:1–24.

63. Howery KE, Clemmer KM, Rather PN. 2016. The Rcs regulon in Proteus mirabilis: implications for motility, biofilm formation, and virulence. Curr Genet 62:775–789.

64. Jones HE, Park RW. 1967. The influence of medium composition on the growth and swarming of Proteus. J Gen Microbiol 47:369–78.

65. Sturgill G, Rather PN. 2004. Evidence that putrescine acts as an extracellular signal required for swarming in Proteus mirabilis. Mol Microbiol 51:437–446.

66. Kurihara S, Kato K, Asada K, Kumagai H, Suzuki H. 2010. A putrescine-inducible pathway comprising PuuE-YneI in which γ-aminobutyrate is degraded into succinate in Escherichia coli K-12. J Bacteriol 192:4582–4591.

67. MATSUYAMA T, MATSUYAMA M. 2002. Population Morphogenesis by Cooperative Bacteria. Forma 16:307–326.

68. Mimura M, Sakaguchi H, Matsushita M. 2000. Reaction-diffusion modelling of bacterial colony patterns. Phys A Stat Mech its Appl 282:283–303.

69. Serena M, Osera C, Marini S, Scavone F, Bellazzi R, Galizzi A, Calvio C. 2013. The Role of SwrA , DegU and P D3 in fla / che Expression in B . subtilis 8:1–11.

70. Kobayashi K. 2007. Gradual activation of the response regulator DegU controls

serial expression of genes for flagellum formation and biofilm formation in Bacillus subtilis. Mol Microbiol 66:395–409.

71. Mukai K, Kawata-Mukai M, Tanaka T. 1992. Stabilization of phosphorylated Bacillus subtilis DegU by DegR. J Bacteriol 174:7954–7962.

72. Yasumura A, Abe S, Tanaka T. 2008. Involvement of nitrogen regulation in Bacillus subtilis degU expression. J Bacteriol 190:5162–5171.

73. Ishii H, Tanaka T, Ogura M. 2013. The bacillus subtilis response regulator gene DegU is positively regulated by CcpA and by Catabolite-Repressed Synthesis of ClpC. J Bacteriol 195:193–201.

74. Wakita JI, Shimada H, Itoh H, Matsuyama T, Matsushita M. 2001. Periodic Colony Formation by Bacterial Species Bacillus subtilis. J Phys Soc Japan 70:911–919.

75. Hamouche L, Laalami S, Daerr A, Song S, Holland IB, Séror SJ, Kassem Hamze, Putzera H. 2017. Bacillus subtilis Swarmer Cells Lead the Swarm, Multiply, and Generate a Trail of Quiescent Descendants 8:1–14.

76. Ariel G, Rabani A, Benisty S, Partridge JD, Harshey RM, Be'Er A. 2015. Swarming bacteria migrate by Lévy Walk. Nat Commun 6.

77. Kohler T, Curty LK, Barja F, Van Delden C, Pechere JC. 2000. Swarming of Pseudomonas aeruginosa is dependent on cell-to-cell signaling and requires flagella and pili. J Bacteriol 182:5990–5996.

78. Tremblay J, Déziel E. 2010. Gene expression in Pseudomonas aeruginosa swarming motility. BMC Genomics 11:587.

79. Caiazza NC, Shanks RMQ, O'Toole GA. 2005. Rhamnolipids modulate swarming motility patterns of Pseudomonas aeruginosa. J Bacteriol 187:7351–7361.

80. Mukherjee S, Bree AC, Liu J, Patrick JE, Chien P, Kearns DB. 2015. Adaptor-mediated Lon proteolysis restricts Bacillus subtilis hyperflagellation. Proc Natl Acad Sci U S A 112:250–255.

81. van Ditmarsch D, Boyle KE, Sakhtah H, Oyler JE, Nadell CD, Déziel É, Dietrich LEP, Xavier JB. 2013. Convergent evolution of hyperswarming leads to impaired biofilm formation in pathogenic bacteria. Cell Rep 4:697–708.

82. Soberón-Chávez G, Lépine F, Déziel E. 2005. Production of rhamnolipids by Pseudomonas aeruginosa. Appl Microbiol Biotechnol 68:718–725.

83. Smith RS, Iglewski BH. 2003. P. aeruginosa quorum-sensing systems and virulence. Curr Opin Microbiol 6:56–60.

84. Overhage J, Bains M, Brazas MD, Hancock REW. 2008. Swarming of Pseudomonas aeruginosa is a complex adaptation leading to increased production of virulence factors and antibiotic resistance. J Bacteriol 190:2671–2679.

85. Morris JD, Hewitt JL, Wolfe LG, Kamatkar NG, Chapman SM, Diener JM, Courtney AJ, Leevy WM, Shrout JD. 2011. Imaging and analysis of Pseudomonas aeruginosa swarming and rhamnolipid production. Appl Environ Microbiol 77:8310–8317.

86. Glick R, Gilmour C, Tremblay J, Satanower S, Avidan O, Déziel E, Greenberg EP, Poole K, Banin E. 2010. Increase in rhamnolipid synthesis under iron-limiting

conditions influences surface motility and biofilm formation in Pseudomonas aeruginosa. J Bacteriol 192:2973–2980.

87. Rashid MH, Kornberg A. 2000. Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of Pseudomonas aeruginosa. Proc Natl Acad Sci U S A 97:4885–4890.

88. Berg HC, Brown DA. 1972. Chemotaxis in E coli analysed by 3D tracking. Nature 239:500–504.

89. Sourjik V, Wingreen NS. 2012. Responding to chemical gradients: Bacterial chemotaxis. Curr Opin Cell Biol 24:262–268.

90. Baker MD, Wolanin PM, Stock JB. 2006. Signal transduction in bacterial chemotaxis. BioEssays 28:9–22.

91. Wolfe AJ, Berg HC. 1989. Migration of bacteria in semisolid agar. Proc Natl Acad Sci U S A 86:6973–6977.

92. Fahrner KA, Berg HC. 2015. Mutations that stimulate flhDC expression in Escherichia coli K-12. J Bacteriol 197:3087–3096.

93. Barker CS, Prüß BM, Matsumura P. 2004. Increased motility of Escherichia coli by insertion sequence element integration into the regulatory region of the flhD operon. J Bacteriol 186:7529–7537.

94. Lee C, Park C. 2013. Mutations upregulating the flhDC operon of Escherichia coli K-12. J Microbiol 51:140–144.

95. Hedblom ML, Adler J. 1980. Genetic and biochemical properties of Escherichia coli mutants with defects in serine chemotaxis. J Bacteriol 144:1048–1060.

96. Manson MD, Blank V, Brade G, Higgins CF. 1986. Peptide chemotaxis in E. coli involves the Tap signal transducer and the dipeptide permease.

97. REBBAPRAGADA* A, JOHNSON MS, HARDING GP, ZUCCARELLI AJ, FLETCHER HM, ZHULIN IB, TAYLOR BL. 1997. The Aer protein and the serine chemoreceptor Tsr independently sense intracellular energy levels and transduce oxygen , redox , and energy signals for Escherichia coli behavior 94:10541–10546.

98. Liu X, Parales RE. 2008. Chemotaxis of Escherichia coli to pyrimidines: A new role for the signal transducer tap. J Bacteriol 190:972–979.

99. Berg HC. 2003. The rotary motor of bacterial flagella. Annu Rev Biochem 72:19–54.

100. Mears PJ, Koirala S, Rao C V., Golding I, Chemla YR. 2014. Escherichia coli swimming is robust against variations in flagellar number. Elife 2014.

101. Stramer B, Wood W. 2009. Chemotaxis Methods and ProtocolsMethods in molecular biology (Clifton, N.J.).

102. Berg HC. 2004. E. coli in MotionE. coli in Motion.

103. Adler J. 1966. Chemotaxis in Bacteria 153.

104. Liu W, Cremer J, Li D, Hwa T, Liu C. 2019. An evolutionarily stable strategy to colonize spatially extended habitats. Nature.

105. FREDERICK C. NEIDHARDT, PHILIP L. BLOCH ADFS. 1974. Culture Medium for Enterobacteria. Microbiology 119:736–747.

106. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA,

Tomita M, Wanner BL, Mori H. 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2.

107. Miller JH. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory.

108. Laurent Bouillaut, Shonna M. McBride and JAS. 2011. Genetic Manipulation of C Diff. Curr Protoc Microbiol 2:1–20.

109. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.

110. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM 00:1–3.

111. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. Gigascience 10:1–4.

112. Narasimhan V, Danecek P, Scally A, Xue Y, Tyler-Smith C, Durbin R. 2016. BCFtools/RoH: A hidden Markov model approach for detecting autozygosity from next-generation sequencing data. Bioinformatics 32:1749–1751.

113. Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing 1–9.

114. Hawkey J, Hamidian M, Wick RR, Edwards DJ, Billman-Jacobe H, Hall RM, Holt KE. 2015. ISMapper: identifying transposase insertion sites in bacterial genomes from short read sequence data. BMC Genomics 16:667.

115. Khan JA. 2021. ENHANCING TRANSCRIPTION IN ESCHERICHIA COLI

BW25113 AND PSEUDOMONAS PUTIDA KT2440 USING
BACTERIOPHAGE LAMBDA ANTI- TERMINATOR PROTEIN Q.

116. Lehnen D, Blumer C, Polen T, Wackwitz B, Wendisch VF, Unden G. 2002. LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in Escherichia coli. Mol Microbiol 45:521–532.

117. Blumer C, Kleefeld A, Lehnen D, Heintz M, Dobrindt U, Nagy G, Michaelis K, Emödy L, Polen T, Rachel R, Wendisch VF, Unden G. 2005. Regulation of type 1 fimbriae synthesis and biofilm formation by the transcriptional regulator LrhA of Escherichia coli. Microbiology 151:3287–3298.

118. Nachin L, Nannmark U, Nystrom T. 2005. Differential Roles of the Universal Stress Proteins of Escherichia coli in Oxidative Stress Resistance, Adhesion, and Motility Laurence. Microbiology 187:6265–6272.

119. Lee J, Park YH, Kim YR, Seok YJ, Lee CR. 2015. Dephosphorylated NPr is involved in an envelope stress response of Escherichia coli. Microbiology 161:1113–1123.

120. Hegde M, Englert DL, Schrock S, Cohn WB, Vogt C, Wood TK, Manson MD, Jayaraman A. 2011. Chemotaxis to the quorum-sensing signal AI-2 requires the Tsr chemoreceptor and the periplasmic LsrB AI-2-binding protein. J Bacteriol 193:768–773.

121. Pasupuleti S, Sule N, Cohn WB, MacKenzie DS, Jayaraman A, Manson MD. 2014. Chemotaxis of Escherichia coli to norepinephrine (NE) requires conversion of NE to 3,4-dihydroxymandelic acid. J Bacteriol 196:3992–4000.

122. Slocum MK, Parkinson JS. 1985. Genetics of methyl-accepting chemotaxis proteins in Escherichia coli: Organization of the tar Region. J Bacteriol 163:586–594.

123. Weerasuriya S, Schneider BM, Manson MD, Manson MD. 1998. Chimeric Chemoreceptors in Escherichia coli : Signaling Properties of Tar-Tap and Tap-Tar Hybrids 180:914–920.

124. Zhang X, Si G, Dong Y, Chen K, Ouyang Q, Luo C, Tu Y. 2018. Escape band in Escherichia coli chemotaxis in opposing attractant and nutrient gradients.

125. Koskiniemi S, Lamoureux JG, Nikolakakis KC, De Roodenbeke CTK, Kaplan MD, Low DA, Hayes CS. 2013. Rhs proteins from diverse bacteria mediate intercellular competition. Proc Natl Acad Sci U S A 110:7032–7037.

126. Fraebel DT, Mickalide H, Schnitkey D, Merritt J, Kuhlman TE, Kuehn S. 2017. Environment determines evolutionary trajectory in a constrained phenotypic space. Elife 6:1–32.

127. Varahan S, Walvekar A, Sinha V, Krishna S, Laxman S. 2019. Metabolic constraints drive self-organization of specialized cell groups. Elife 8.

128. Gustavsson N, Diez AA, Nyström T. 2002. The universal stress protein paralogues of Escherichia coli are co-ordinately regulated and co-operate in the defence against DNA damage. Mol Microbiol 43:107–117.

129. Potrykus K, Cashel M. 2008. (p)ppGpp: Still magical? Annu Rev Microbiol 62:35–51.

130. Gourse RL, Chen AY, Gopalkrishnan S, Sanchez-Vazquez P, Myers A, Ross W.

2018. Transcriptional Responses to ppGpp and DksA. Annu Rev Microbiol 72:163–184.

131. Ronneau S, Hallez R. 2019. Make and break the alarmone: regulation of (p)ppGpp synthetase/hydrolase enzymes in bacteria. FEMS Microbiol Rev 43:389–400.

132. Magnusson LU, Gummesson B, Joksimović P, Farewell A, Nyström T. 2007. Identical, independent, and opposing roles of ppGpp and DksA in Escherichia coli. J Bacteriol 189:5193–5202.

133. Durfee T, Hansen AM, Zhi H, Blattner FR, Ding JJ. 2008. Transcription profiling of the stringent response in Escherichia coli. J Bacteriol 190:1084–1096.

134. Lemke JJ, Durfee T, Gourse RL. 2010. DksA and ppGpp Directly Regulate Transcription of the Escherichia coli Flagellar Cascade 74:1368–1379.

135. Molodtsov V, Sineva E, Zhang L, Hauang X, Cashel M, Ades SE, Murakami KS, Huang X, Cashel M, Ades SE, Murakami KS. 2018. Allosteric effector ppGpp potentiates the inhibition of transcript initiation by DksA. Mol Cell 69:828–839.

136. Praveen-Kumar, Tarafdar JC. 2003. 2,3,5-Triphenyltetrazolium chloride (TTC) as electron acceptor of culturable soil bacteria, fungi and actinomycetes. Biol Fertil Soils 38:186–189.

137. Gibson KE, Silhavy TJ. 1999. The LysR homolog LrhA promotes RpoS degradation by modulating activity of the response regulator SprE. J Bacteriol 181:563–571.

138. Peterson CN, Carabetta VJ, Chowdhury T, Silhavy TJ. 2006. LrhA Regulates

rpoS Translation in Response to the Rcs Phosphorelay System in Escherichia coli 188:3175–3181.

139. Battesti A, Majdalani N, Gottesman S. 2010. The RpoS-Mediated General Stress Response in Escherichia coli . Annu Rev Microbiol 65:189–213.

140. Zhou Y, Gottesman S, Hoskins JR, Maurizi MR, Wickner S. 2001. The RssB response regulator directly targets σS for degradation by ClpXP. Genes Dev 15:627–637.

141. Zhou Y, Gottesman S. 1998. Regulation of proteolysis of the stationary-phase sigma factor RpoS. J Bacteriol 180:1154–1158.

142. Zhou YN, Jin DJ. 1998. The rpoB mutants destabilizing initiation complexes at stringently controlled promoters behave like “stringent” RNA polymerases in Escherichia coli. Proc Natl Acad Sci U S A 95:2908–2913.

143. Adler J, Templeton B. 1967. The effect of environmental conditions on the motility of Escherichia coli. J Gen Microbiol 46:175–184.

144. Metzger S, Schreiber G, Aizenman E, Cashel M, Glaser G. 1989. Characterization of the relA1 mutation and a comparison of relA1 with new relA null alleles in Escherichia coli. J Biol Chem 264:21146–21152.

145. Xiao H, Kalman M, Ikehara K, Zemel S, Glaser G, Cashel M. 1991. Residual guanosine 3’,5’-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations. J Biol Chem 266:5980–5990.

146. Kanjee U, Ogata K, Houry WA. 2012. Direct binding targets of the stringent response alarmone (p)ppGpp. Mol Microbiol 85:1029–1043.

147. Kvint K, Nachin L, Diez A, Nystrom T. 2003. The bacterial universal stress protein : function and regulation 140–145.

148. Murphy H, Cashel M. 2003. Isolation of RNA Polymerase Suppressors of a (p)ppGpp Deficiency. Methods Enzymol 371:596–601.

149. Magnusson LU, Farewell A, Nyström T. 2005. ppGpp: A global regulator in Escherichia coli. Trends Microbiol 13:236–242.

150. Vinella D, Albrecht C, Cashel M, D'Ari R. 2005. Iron limitation induces SpoT-dependent accumulation of ppGpp in Escherichia coli. Mol Microbiol 56:958–970.

151. Garénaux A, Caza M, Dozois CM. 2011. The Ins and Outs of siderophore mediated iron uptake by extra-intestinal pathogenic Escherichia coli. Vet Microbiol 153:89–98.

152. Lee JW, Helmann JD. 2007. Functional specialization within the fur family of metalloregulators. BioMetals 20:485–499.

153. Inoue T, Shingaki R, Hirose S, Waki K, Mori H, Fukui K. 2007. Genome-wide screening of genes required for swarming motility in Escherichia coli K-12. J Bacteriol 189:950–957.

154. Kirkpatrick C, Maurer LM, Oyelakin NE, Yoncheva YN, Maurer R, Slonczewski JL. 2001. Acetate and formate stress: Opposite responses in the proteome of Escherichia coli. J Bacteriol 183:6466–6477.

155. Tobe T, Nakanishi N, Sugimoto N. 2011. Activation of motility by sensing short-chain fatty acids via two steps in a flagellar gene regulatory cascade in

enterohemorrhagic Escherichia coli. Infect Immun 79:1016–1024.

156. Wolfe AJ. 2005. The Acetate Switch 69:12–50.

157. Jones CH, Pinkner JS, Roth R, Heuser J, Nicholes A V., Abraham SN, Hultgren SJ. 1995. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. Proc Natl Acad Sci U S A 92:2081–2085.

158. Landgraf JR, Jingcai WU, Calvo JM. 1996. Effects of nutrition and growth rate on Lrp levels in Escherichia coli. J Bacteriol 178:6930–6936.

159. Zhang F, Li B, Dong H, Chen M, Yao S, Li J, Zhang H, Liu X, Wang H, Song N, Zhang K, Du N, Xu S, Gu L. 2020. YdiV regulates Escherichia coli ferric uptake by manipulating the DNA-binding ability of Fur in a SlyD-dependent manner. Nucleic Acids Res 48:9571–9588.

160. Takaya A, Erhardt M, Karata K, Winterberg K, Yamamoto T, Hughes KT. 2012. YdiV: A dual function protein that targets FlhDC for ClpXP-dependent degradation by promoting release of DNA-bound FlhDC complex. Mol Microbiol 83:1268–1284.

161. Land M, Hauser L, Jun SR, Nookaew I, Leuze MR, Ahn TH, Karpinets T, Lund O, Kora G, Wassenaar T, Poudel S, Ussery DW. 2015. Insights from 20 years of bacterial genome sequencing. Funct Integr Genomics 15:141–161.

162. Lenski RE. 2017. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. ISME J 11:2181–2194.

163. Laabei M, Recker M, Rudkin JK, Aldeljawi M, Gulay Z, Sloan TJ, Williams P, Endres JL, Bayles KW, Fey PD, Yajjala VK, Widhelm T, Hawkins E, Lewis K,

Parfett S, Scowen L, Peacock SJ, Holden M, Wilson D, Read TD, Van Den Elsen J, Priest NK, Feil EJ, Hurst LD, Josefsson E, Massey RC. 2014. Predicting the virulence of MRSA from its genome sequence. Genome Res 24:839–849.

164. Sundermann AJ, Chen J, Miller JK, Saul MI, Shutt KA, Griffith MP, Mustapha MM, Ezeonwuka C, Waggle K, Srinivasa V, Kumar P, Pasculle AW, Ayres AM, Snyder GM, Cooper VS, Van Tyne D, Marsh JW, Dubrawski AW, Harrison LH. 2020. Outbreak of Pseudomonas aeruginosa Infections from a Contaminated Gastroscope Detected by Whole Genome Sequencing Surveillance . Clin Infect Dis 15261:1–5.

165. Petit RA, Read TD. 2020. Bactopia: A flexible pipeline for complete analysis of bacterial genomes. bioRxiv 5:1–18.

166. Anderson S, Bankier AT, Barrell BG, Bruijn MHL de, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG. 1981. Sequence and organization of the human mitochondrial genome 290:1–18.

167. Rishishwar L, Jordan IK. 2017. Implications of human evolution and admixture for mitochondrial replacement therapy. BMC Genomics 18:1–11.

168. Stoneking M. 2000. Hypervariable sites in the mtDNA control region are mutational hotspots. Am J Hum Genet 67:1029–1032.

169. Huddleston J, Chaisson MJP, Steinberg KM, Warren W, Hoekzema K, Gordon D, Graves-Lindsay TA, Munson KM, Kronenberg ZN, Vives L, Peluso P, Boitano M, Chin CS, Korlach J, Wilson RK, Eichler EE. 2017. Discovery and genotyping

of structural variation from long-read haploid genome sequence data. Genome Res 27:677–685.

170. Hunt M, Silva N De, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: Automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:1–10.

171. Liao YC, Cheng HW, Wu HC, Kuo SC, Lauderdale TLY, Chen FJ. 2019. Completing Circular Bacterial Genomes With Assembly Complexity by Using a Sampling Strategy From a Single MinION Run With Barcoding. Front Microbiol 10:1–10.

172. Iliopoulos CS, Okanlawon TA, Pissis SP, Tischler G. 2010. Mapping short reads to a genomic sequence with circular structure. Proc IEEE/EMBS Reg 8 Int Conf Inf Technol Appl Biomed ITAB.

173. Reinert K, Dadi TH, Ehrhardt M, Hauswedell H, Mehringer S, Rahn R, Kim J, Pockrandt C, Winkler J, Siragusa E, Urgese G, Weese D. 2017. The SeqAn C++ template library for efficient sequence analysis: A resource for programmers. J Biotechnol 261:157–168.

174. Sandmann S, De Graaf AO, Karimi M, Van Der Reijden BA, Hellström-Lindberg E, Jansen JH, Dugas M. 2017. Evaluating Variant Calling Tools for Non-Matched Next-Generation Sequencing Data. Sci Rep 7:1–12.

175. Xu X, Yu H, Zhang D, Xiong J, Qiu J, Xin R, He X, Sheng H, Cai W, Jiang L, Zhang K, Hu X. 2016. Role of ppGpp in Pseudomonas aeruginosa acute pulmonary infection and virulence regulation. Microbiol Res 192:84–95.

176. Cummings JH, Pomare EW, Branch WJ, Naylor CPE, Macfarlane GT. 1987. Short chain fatty acids in human large intestine, portal, hepatic, venous blood. Gut 28:1221–1227.

177. Nakamura H, Saiki K, Mogi T, Anraku Y. 1997. Assignment and functional roles of the cyoABCDE gene products required for the Escherichia coli bo-type quinol oxidase. J Biochem 122:415–421.

178. Seol W, Shatkin AJ. 1992. Escherichia coli α-ketoglutarate permease is a constitutively expressed proton symporter. J Biol Chem 267:6409–6413.

179. Beagle SD, Lockless SW. 2020. Unappreciated Roles for K+ Channels in Bacterial Physiology. Trends Microbiol 1–9.

180. McLaggan D, Jones MA, Gouesbet G, Levina N, Lindey S, Epstein W, Booth IR. 2002. Analysis of the kefA2 mutation suggests that KefA is a cation-specific channel involved in osmotic adaptation in Escherichia coli. Mol Microbiol 43:521–536.

181. Sukharev S. 1999. Mechanosensitive channels in bacteria as membrane tension reporters. FASEB J 13 Suppl:S55–S61.

182. Liu J, Prindle A, Humphries J, Gabalda-Sagarra M, Asally M, Lee DYD, Ly S, Garcia-Ojalvo J, Süel GM. 2015. Metabolic co-dependence gives rise to collective oscillations within biofilms. Nature 523:550–554.

183. Epstein W, Kim BS. 1971. Potassium Transport Loci in Escherichia coli K-12. J Bacteriol 108:639–644.

184. Cao Y, Jin X, Huang H, Derebe MG, Levin EJ, Kabaleeswaran V, Pan Y, Punta

M, Love J, Weng J, Quick M, Ye S, Kloss B, Bruni R, Martinez-Hackert E, Hendrickson WA, Rost B, Javitch JA, Rajashankar KR, Jiang Y, Zhou M. 2011. Crystal structure of a potassium ion transporter, TrkH. Nature 471:336–340.

185. Zhang H, Pan Y, Hu L, Hudson MA, Hofstetter KS, Xu Z, Rong M, Wang Z, Prasad BVV, Lockless SW, Chiu W, Zhou M. 2020. TrkA undergoes a tetramer-to-dimer conversion to open TrkH which enables changes in membrane potential. Nat Commun 11.

186. Manson MD, Tedesco PM, Berg HC. 1980. Energetics of flagellar rotation in bacteria. J Mol Biol 138:541–561.

187. Bossemeyer D, Borchard A, Dosch DC, Helmer GC, Epstein W, Booth IR, Bakker EP. 1989. K+-transport protein TrkA of Escherichia coli is a peripheral membrane protein that requires other trk gene products for attachment to the cytoplasmic membrane. J Biol Chem 264:16403–16410.

188. Cao Y, Pan Y, Huang H, Jin X, Levin EJ, Kloss B, Zhou M. 2013. Gating of the TrkH ion channel by its associated RCK protein TrkA. Nature 496:317–322.

189. Dinnbier U, Limpinsel E, Schmid R, Bakker EP. 1988. Transient accumulation of potassium glutamate and its replacement by trehalose during adaption of growing cells of Escherichia coli K-12 to elevated sodium chloride concentrations. Arch Microbiol 150:348–357.

190. Epstein W. 1986. Osmoregulation by potassium transport in Escherichia coli. FEMS Microbiol Lett 39:73–78.

191. Rhoads DB. 1978. Cation transport in Escherichia coli. IX. Regulation of K

transport. J Gen Physiol 72:283–295.

192. Liu Y, Ho KK, Su J, Gong H, Chang AC, Lu S. 2013. Potassium transport of Salmonella is important for type III secretion and pathogenesis. Microbiol (United Kingdom) 159:1705–1719.

193. Su J, Gong H, Lai J, Main A, Lu S. 2009. The potassium transporter trk and external potassium modulate Salmonella enterica protein secretion and virulence. Infect Immun 77:667–675.

194. Bakker EP, Mangerich WE. 1981. Interconversion of components of the bacterial proton motive force by electrogenic potassium transport. J Bacteriol 147:820–826.

APPENDIX A

UNDERSTANDING THE ROLE OF THE TRK SYSTEM DURING PATTERN FORMATION

Introduction

Ion channels have been studied extensively in eukaryotic cells. However, the role for ion channels in bacterial physiology is an emerging field (179). In *E. coli*, several mechanosensitive non-selective channels have been characterized: *mscL*, *mscS*, *mscM*, and *mscK* (180, 181). MscL, MscS, MscM open in response to osmotic shock, and MscK opens under high extracellular potassium (180, 181) The only *E. coli* potassium-selective channel, *kch*, has been proposed to regulate membrane potential, but conclusive evidence has not been yet identified (179). In *B. subtilis* however, YugO a K^+ selective channel, regulates long range electrical signaling mediated by nutrient stress (179, 182). The Trk system is previously identified as a transporter of K^+,(183)but has since been discovered as an non-selective channel TrkH (184). The opening of the Trk channels causes an alteration in membrane potential (185). Membrane potential is important for several processes in bacteria, including flagellar rotation, where protons are pumped through MotA and MotB to rotate the flagella (186). Here we investigated the role that the non-selective channels of the Trk system play in motility of *E. coli*.

The Trk system made of two ion channels TrkG and TrkH, these are regulated by TrkA, which is assisted in attaching to the channels with TrkE (183, 187, 188) TrkA is an RCK (regulator of potassium conductance) that is attached to TrkH or TrkG to

regulate the flow of ions (183, 185, 187, 188). In *Vibrio parahaemolyticus* only components of the Trk system are constrained to TrkH and TrkA, the loss of TrkA results in TrkH opening much more frequently (188). The Trk system can help the cells maintain turgor, by accumulating K^+ under osmotic stress (189–191). Many other systems in bacteria are affected by the Trk system, inducing Type III secretion in *Salmonella* (192, 193) proton motive force (192, 194) and membrane potential (185).

Results and Discussion

At the beginning of this dissertation project we observed a unique pattern for strains containing ΔtrkA, in that strains were unable to band on SMM-Na but would readily form band on SMM-K (Figure 38). We created a ΔTrk Δ*lrhA* mutant strain that lacked all components of the Trk system. This failed to return a banding pattern on SMM-Na, while at the time we concluded that TrkA could have additional regulatory roles besides the Trk system. We later serendipitously discovered that the TrkA phenotype was dependent on the bottle of yeast extract. With a bottle of the same Lot number of yeast extract bands could easily be observed in the Δ*trkA* strains. Great care was taken to ensure that experiments presented in this dissertation are not due to a containment of

present in the earlier studies.

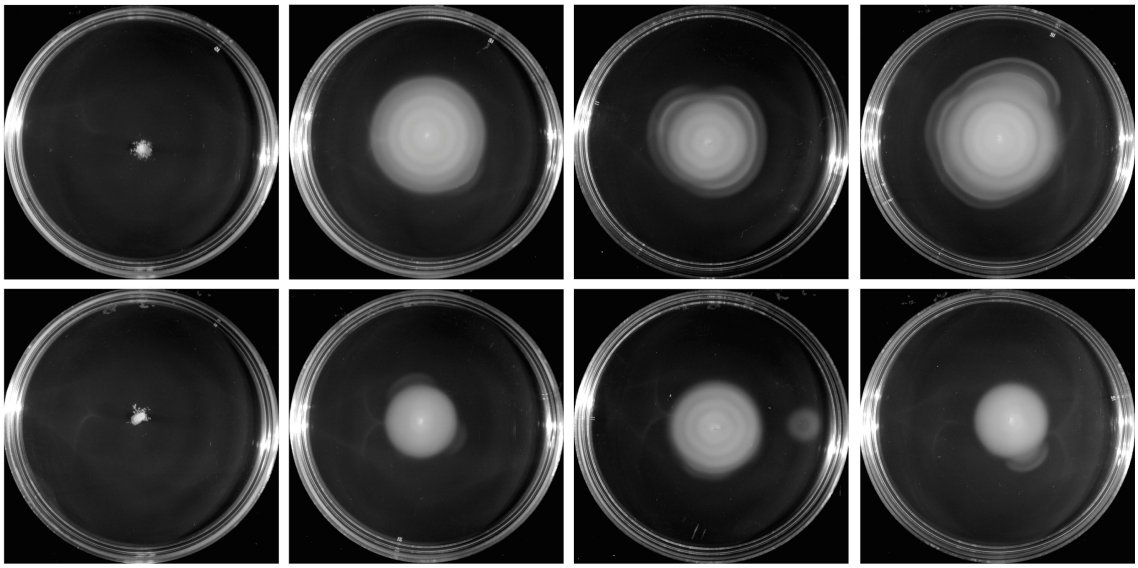

Figure 38. Deletion of trkA has a unique motility phenotype and exhibits banding only in SMM-K.
Experiment preformed 11/2017 Motility phenotypes of ion channel mutants in SMM-K and SMM-Na media. A-E) SMM-K media, I-J) SMM-Na media. A,I) KH6 (BW25113); B,G); KH318 (BW25113 Δ*lrhA*::FRT); C,H) KH10 (BW25113 Δkch::Kan ubiH^{V223G} lrhA198::IS5 A to G upstream yciV); D,I) KH1 (BW25113 ΔtrkA::FRT lrhA670::IS1 yeeJT552I cyoD104 Del); E-J) KH49 (BW25113 Δ*trkA*::Kan)

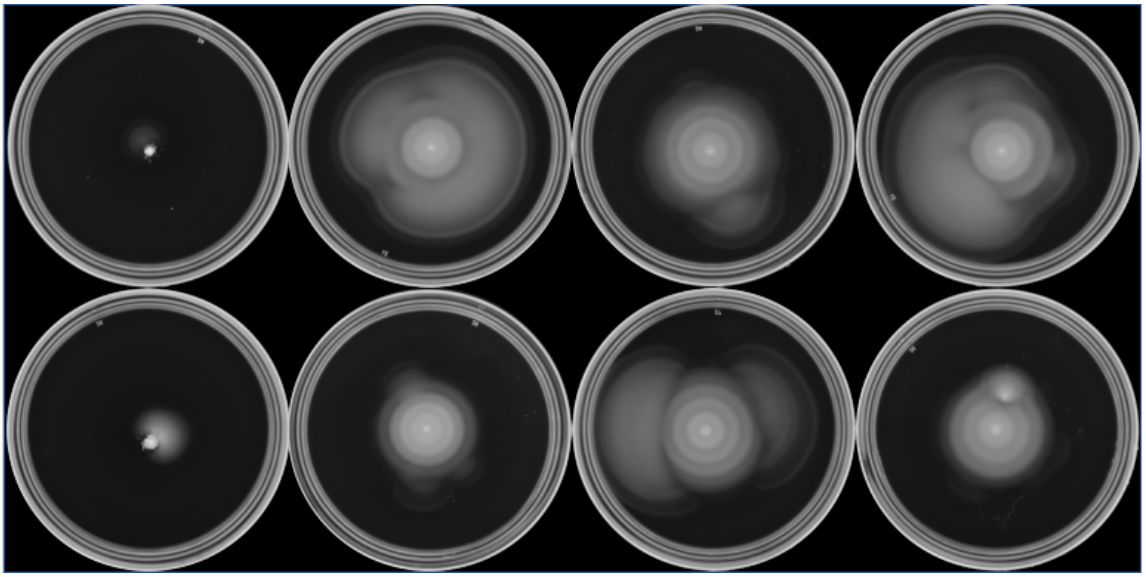

Figure 39. Deletion of trkA exhibits banding in both SMM-Na and SMM-K. Experiment preformed 2/2021.
Motility phenotypes of ion channel mutants in SMM-K and SMM-Na media. A-E) SMM-K media, I-J) SMM-Na media. A,I) KH6 (BW25113); B,G); KH318 (BW25113 Δ*lrhA*::FRT); C,H) KH10 (BW25113 Δkch::Kan ubiH^{V223G} lrhA198::IS5 A to G upstream yciV); D,I) KH1 (BW25113 ΔtrkA::FRT lrhA670::IS1 yeeJT552I cyoD104 Del); E-J) KH49 (BW25113 Δ*trkA*::Kan)

APPENDIX B

TIME-LAPSE MOVIES

The following movie files accompany the dissertation as separate files.

KH318_MOVIE_37 °C_SMM-NA

KH707_MOVIE_37°C_SMM-NA

KH1220_MOVIE_37°C_SMM-NA