Comparative study of the effect of nutrients on motility and chemotaxis of *Escherichia coli* strains

Shashank Pandey

A thesis submitted for the fulfilment of the requirements for the degree of

M.Sc. (Research) Biomedical Engineering

James Watt School of Engineering

College of Science and Engineering

University of Glasgow

February 2020

ABSTRACT

This thesis evaluates two strains of *Escherichia coli* MG1655 and MDS42 for their motility in different nutrient conditions in M9 minimal medium in 2 parts. It evaluates the effect of genome deletion in the motility and also observes the heterogeneity despite sharing the same genetically encoded machinery. The first part investigates *Escherichia coli* strains' motility in 5 different medium compositions and the second part explores the chemotactic response of MG1655 to the linear gradients of different concentrations of Glucose using a single-layer membrane-based microfluidic device.

In Part 1, we study the motility of MG1655 and MDS42 in different concentrations of glucose and casamino acids in M9 minimal medium. The motility experiments conducted as a part of this study observed the average cell velocities in the range of 2.9 ± 0.5 µm/s, which are significantly less than the values recorded in literature, for the strain MG1655. The lowest motility occurs in the medium (without casamino acids) with 0M glucose, followed by 10mM Glucose and then 10µM glucose concentration. The same trend is visible in the case of both the strains MG1655 and MDS42. The presence of casamino acids did not significantly affect the motility of MG1655 in the presence or absence of Glucose. Whereas, in the case of MDS42, the casamino acids lower the motility in the presence of Glucose but tend to have no significant effect in the absence of Glucose. The two strains, however, showed no significant difference in average velocity under the same medium conditions.

In Part 2, we record and evaluate the chemotaxis of the MG1655 strain, using a single-layer membrane-based microfluidic device. The device generates a linear gradient of 10µM and 10mM glucose, to observe the chemotaxis of the MG1655 strain. The average of mean velocities for the 10µM gradient was higher than those observed in the 10mM gradient, but the difference was not significant. The higher fraction of cells (~67%) under the 10mM gradient showed almost a straight-line trajectory, unlike the cells under 10µM gradient. The cells that followed a nearly straight line path did all the more so in the case of the 10mM glucose gradient.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

Huge thanks to my supervisor and co-supervisor, Dr Huabing Yin and Dr Andrew Glidle, for guiding, helping, and supporting me in all possible ways. Special thanks to Prof. Huabing Yin for her patience and bearing with me all along.

I am also thankful to my lab research group: Louise Mason and Laura Charlton for their pep talks and their willingness to go out of their way to guide me. Yanqing Song for her patience and help with the lab techniques. Yingkai Liu, Jing Zhang, and Yuchen Fu for their precious help with the device development.

I am grateful to the PGR coordinator Julia Deans and my viva convenor Umer Ijaz, for their prompt help and guidance with the structure of my course, thesis correction and graduation.

Kudos to the Biomedical Engineering Division, and especially the supportive staff at Rankine Building for making labs accessible for long and odd hours.

A big thanks to my family for believing in me and funding my education. I also thank them for staying hopeful during the long course of my degree.

DECLARATION

Name: Shashank Pandey

Registration Number: 2405585P

I certify that the thesis presented here for examination for M.Sc. (Res.) degree of the University of Glasgow is solely my work other than where I have indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is identified in it) and that the thesis has not been edited by a third party beyond what is permitted by the University's PGR Code of Practice.

I declare that the thesis does not include work forming part of a thesis presented successfully for another degree [unless explicitly identified and as noted below].

I declare that this thesis has been produced by following the University of Glasgow's Code of Good Practice in Research.

I acknowledge that if any issues are raised regarding good research practice based on a review of the thesis, the examination may be postponed pending the outcome of any investigation of the issues.

Signature:

Date: 27 – 02 – 2020

Chapter 01 | Introduction

1.1 Motility and Chemotaxis

The complex phenomenon of life is unsustainable without the recurring interactions of the wide variety of molecules. There are interacting arrays of the molecules, which convolute to form functional networks that directly or indirectly affect the mortality of the organisms. The good examples of which could be the interactions of networking molecules that desensitize the cancerous cells to the interaction of body's signals to terminate proliferation (Hanahan 2000); the molecular interactions that instruct yeast cells to metabolize grape juice and make wine (Gonzalez 2013); the molecules that detect environmental stimulus and signal a cell to move away or towards it to survive (Partridge, Escherichia coli Remodels the Chemotaxis Pathway for Swarming 2019). The idea of the network of molecules interacting with the environment to facilitate the survival and interaction of a single cell with its environment, is what we commonly term as chemotaxis. Chemotaxis can also be seen as directed motility, as the living bacterial cells are mostly motile in a random fashion. (Keller 1971) (J. Adler 1966).

More like a common sense, we assume that the network of molecules serves only one function, i.e. whenever the cell has all the necessary biomolecules in place, they execute a certain function in an unvarying manner. But is it that simple? For now, consider the system responsible for the motility of the bacterial cells as one network of molecules which just serves the purpose of motility and then consider another network of molecules that executes the chemotactic movement in bacterial cells. However, is it a one network to one function mapping? The system is much simpler yet complex because of the cross-talks between the molecules involved in sensing the environment in the chemotaxis network and the molecules in the motility network. (Hansen 2010) (Lopes 2018)

1.1.1 What is bacterial motility and why is it important?

Antonie van Leeuwenhoek, who is acknowledged universally as the father of microbiology, in 1675 reported to the Royal Society that he had seen *Animacules* (living atoms). The *Animacules* are referred to as microbes in the current period. He also reported

that the *Animacules* seemed to move by putting forth and continuously moving two horn-like structures. This observation was the first record of observation of what we recognize as bacteria today, and this was indeed the first time when the movement (motility) was used as a significant parameter to classify the tiny particles as 'Living' (Leeuwenhoek 1677). The claims of Antonie van Leeuwenhoek were dismissed in that time and doubted until his observations were confirmed in the twentieth century as none of the scientists of his period were able to build a microscope with a resolution as good as his. (Lane 2015).

The prokaryotic organisms use a wide variety of structures for motility, archaeal flagella, type IV pili, the junctional pore, ratchet structure, and the contractile cytoskeleton being few of many. Despite such a variety, the bacterial flagellum which differs from all the structures widely has been the most deeply studied of all the motility structures in the prokaryotes (Bardy 2003). The two horn-like appendages seen by Antonie in 1675, were most probably what are known as Flagella today.

By convention, the flagella are mainly responsible for the motility and pathogenicity. More precisely, they serve the function of adhesion, adhesion assistance, and help with colonization and survival. (F. F. Liu 2017). (McCarter 2005) claimed the presence of a secondary flagellar system. The secondary flagellar system was discovered through an analysis of the genome sequence of enteroaggregative *E. coli* strain 042. This new locus of flagellar genes is known as Flag-2.

This discovery led to the possibility of the presence of a secondary flagellar system in the other strain of *E. coli.* To determine this, genomic features specific to PCN033, which is a pathogenic strain of *E. coli* and is known to have Flag – 2 loci of genes. This comparison revealed that the non – pathogenic strains of *E. coli* MG1655 (the subject of the study) and W3110, lacked the island II which is the loci for Flag – 2 also known as the lateral flagellar system. (C. Z. Liu 2015)

1.1.2 What is Chemotaxis?

Chemotaxis is the property of the living organisms to move in response to the chemicals or their gradients in their environment. These chemicals can be anything ranging from toxins to signalling molecules or nutrients, which may be crucial for the growth and survival of the organism (Neumann 2012). *Escherichia coli* being the model organism for

the prokaryotes, has been extensively studied for the bacterial chemotaxis. The cell is propelled by a set of several helical flagellar filaments that occur at arbitrary locations on its sides and extend several folds the body lengths out into the external medium. The chemotactic machinery of *E. coli* and its signalling pathway is very well understood. (Szurmant 2004). The chemotaxis is a 2-component system, which involves sensing of the environmental stimuli and one that regulates the response of the organism. (Kofoid 1988).

1.1.2.1 Chemoreceptor – Flagella Crosstalk

In the case of *Escherichia coli*, transmembrane proteins act as chemoreceptors to detect the chemical stimuli from the surroundings be it other organisms or the environment. The information about the stimulus is then passed on to the signal transduction system in the cytoplasmic region of the cell. This signal transduction system has two components, namely, CheA, which is a receptor-associated kinase, and CheY, which regulates the cellular response to the stimuli. (Szurmant 2004). As soon as the stimuli are received at the chemoreceptors, CheA histidine kinase phosphorylates automatically, and the phosphorylated histidine acts as the substrate for the CheY, which regulates the cellular response to the stimuli by taking up the phosphoryl group on its conserved aspartate. The CheY – P, in – turn controls the switch mechanism in the flagellar motor, able to change the direction and speed of the flagellar rotation. (Sagi 2003)

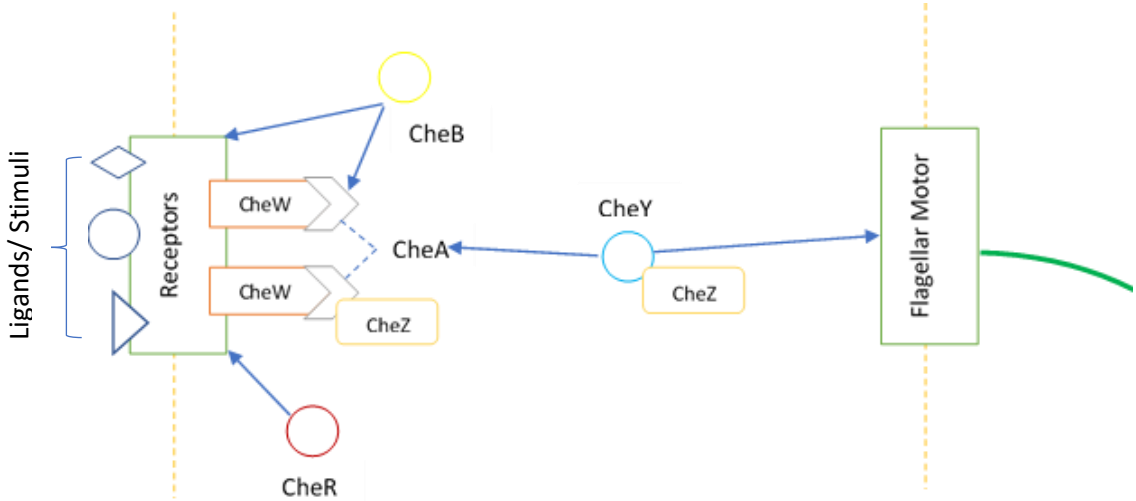


Figure 1. 1: The simplified schematic of chemoreceptor-flagella crosstalk

The organisms exhibit a chemotactic response by moving towards the molecules (Chemo – attractants) or by moving away from them; in this case, the molecules are known as

Chemo – repellents. (J. Adler, Chemotaxis in Bacteria 1966). *Escherichia coli* responds to molecules of sugars, amino acids, neurotransmitters, phenol, dipeptides, pyrimidines, quorum sensing signals, and even pH, temperature, and oxygen. (Ortega 2017). They utilize five different chemoreceptors to detect the presence of different molecules in their surroundings, but the signalling cascade followed by the chemoreceptors is common. The different chemoreceptors are spatially segregated and responsible for detecting the presence of molecules, as shown in Table 1.1. These chemoreceptors are commonly referred to as the Methyl – accepting Chemotaxis Protein molecules. (Parkinson 2015) (Hazelbauer 2008)

Chemoreceptor	**Ligand Molecule**
Trg	Ribose & Galactose
Tar	Aspartate
Tsr	Serine
Tap	Peptides (Manson 1986)
Aer	Redox by-products

Table 1. 1: The five types of Methyl-accepting Chemotaxis Protein Molecules

The five *E. coli* members of the Methyl – accepting Chemotaxis Protein family have low-abundance receptors – Tap, Trg, and Aer, which participate in signalling teams with the high-abundance receptors Tar and Tsr. (Studdert 2004).

The arrays of the chemoreceptors on the cell membrane govern the movement of the cells towards or away from the chemical or ligand molecules and their gradients. The cytoplasmic signalling cascade signals the flagellar motor to generate clockwise movements resulting in the random directional changes, also known as the tumbles. In the presence of the positive or the negative chemotaxis gradient, the tumbling frequency decreases resulting in the lesser directional changes and a more or less unidirectional. (Partridge, Escherichia coli Remodels the Chemotaxis Pathway for Swarming 2019). The tumbling occurs only when ~ 25% or more of the flagellar filaments of an organism change to clockwise direction from the conventional counter-clockwise movement. (Turner 2000) The signals for the tumbles are suppressed in case there is a chemo – effector gradient stimulus sensed by the transmembrane receptors. In this case the motor

switches back to the counterclockwise rotation, the filaments regain their normal conformation and rejoin the bundle.

The chemoreceptors form a trimer complex unit along with CheA, which is a histidine kinase and CheW, which acts as a docker molecule for CheA. The MCP – CheW – CheA altogether form a receptor signalling complex, which exhibits on and off output states. The cellular movements reflect the states exhibited by the receptor signalling complex. For instance, the presence of the attractant temporal gradient puts the signalling complex in the OFF state. This gradient presence slows down the flux of CheA phosphoryl groups to CheB and CheY, the response – regulator proteins. The phosphorylated CheY, in turn, interacts with the flagellar motor (basal body) and triggers the clockwise rotation. However, the p-CheY signal is transient due to the rapid dephosphorylation by its phosphatase CheZ. This transient nature of the signal allows the triggering of a suitable chemotactic response to the environmental triggers. (Macnab 1972) (Hazelbauer 2008)

1.1.2.2 Modern-Day Use of Chemotaxis in Bacteria

In the past years, scientists have been focusing on utilising and enhancing the sensing and chemotactic potential of the bacteria as well as other organisms such as nematodes. (Karbalaei 2018). *Escherichia coli*, which is considered to be the model prokaryotic organism to study the chemotaxis systems, has only four transmembrane proteins responsible for the detection of the environmental stimulus. These transmembrane proteins serve as the sensory receptors that are known to mediate a highly sensitive response to the presence of ligands in the environment. However, the natural occurrence of just four receptor proteins limits the spectrum of bio-sensory applications. Therefore, the scientists have been trying to design the novel chemoreceptors for *Escherichia coli* to extend their chemical sensitivity and thereby further the horizons of biosensor development. (Bi 2016)

Owing to the precision and sensitivity to low concentrations of ligands in the environment, bio-reporter devices that utilise bacterial cells have been proposed as opposed to the idea of using the chemicals to detect environmental pollutants. These bioreporter devices hold the advantage over the chemicals in several fronts from being environment friendly to short periods of detection upon exposure to the chemical targets (~30 minutes). Based on the idea of gene induction, these devices aim to quantify the

chemotaxis of the bacteria, and this was demonstrated as a proof of concept in 2017 by a group of scientists. The group quantified the movement of *Escherichia coli* towards serine, aspartate and methyl-aspartate as a function attractant concentration and exposure time. This work established that the bacterial cell-based bio-reporters might be used for developing faster bioreporter assays. (Roggo 2017)

A group of scientists came up with a microfluidic device that could prove bacterial affinity towards the cancerous cells over the normal body cells. This microfluidic device was modelled using lung cells to use E.coli to detect cancerous ones. It consisted of a central channel which had co-culture chambers on either side. The chambers contained lung cancer cells (NCI-H460) on one side and non – cancerous body cells on the other. After intensive secretome analysis and validation, it was found that the lung cancer cells produce Clusterin protein which is a chemoattractant for *Escherichia coli*. The Clusterin gradients formed in the main channel resulted in the preferential taxis of *Escherichia coli* towards lung cancer cells. This study demonstrated bacterial efficacy and usability in cancer diagnosis if not the treatment itself. (Song J. 2018)

The therapeutic use of bacteria in the medical and pharmaceutical research has shown intriguing results. The results suggest that the bacteria can be effective in the treatment of cancer. Experimental studies have found that certain strains of bacteria possess oncolytic potentials that enable them to invade and colonise the solid tumours in vivo. At the same time, there have been reports that the bacteria can be used as effective vessels for targeted drug delivery. (Song S. 2019). The treatments delivered via bacterial cells are advantageous over conventional chemotherapy which poses severe side effects upon delivery. Bacteria have experimentally proven to be effective in this aspect and have been shown to reduce tumour growth in case of animal models. It has been shown that the treatment can be delivered in 3 ways: they can be modified to approach the tumour and produce proteins to kill cells directly; they can induce apoptosis in cancerous cells via signalling pathways, or they can stimulate the immune system and induce an appropriate response against the tumour. Following up on this pathway will enable us to devise methods to target multiple cancer pathways by establishing individualised cancer medicine. (Van N. 2015)

Along with these applications of bacterial chemotaxis, there have been studies that study the chemotaxis in marine bacteria to implement their superior ability to map the spatiotemporal heterogeneity efficiently in the seascape to be able to revolutionize the pollution management and marine ecosystem productivity. (Seymour 2018) (Tout 2017).

During the same period, the inherent ability of the neutrophils to migrate to the disease sites was combined with the drug-loaded mesoporous silica nanoparticles to synthesize guided hybrid micromotors for targeted drug delivery. (Shao 2017). The efficient detection of the chemicals and their concentrations by the bacteria has also been used to develop a high throughput technique to sort the particles of the sizes of orders 10^{-6} to 10^{-9}m, with different surface properties. (Suh 2016).

On one hand there have been studies that deep dive at the genetic level, where it can be seen that the genes responsible for chemotaxis are organised in operons, which means that the transcription of several genes is dependent on a common promoter. Such an arrangement tends to result in correlations in the cell to cell variability in a way that the translated proteins may vary between individual cells, but the ration of one protein to another in a single cell does not vary a lot. (Løvdok L. 2009). It means the chemotaxis system is robust to variability in non-functional parameters sense. In simpler words, there would not be any cell that just runs and not tumble or vice versa. (Alon 1999) (Løvdok 2007)

On the other hand, there have been studies that demonstrate that even if the chemotaxis systems are robust and show no to low variability from the systems biology sense, there is plenty of cell-to-cell variability observed in genetically identical cells of a population under the same conditions. A study on bio-hybrid microswimmer system, where the bacterial swimming was integrated with some artificial components, driven by *Serratia marcescens* was designed. The chemotactic drift was observed for a large number of microswimmers under the L-Serine gradients. It was seen that the microswimmers exhibited varying speeds both towards and away from the chemoattractant. It suggested the idea that the bio-hybrid micro-swimmers can be enhanced to sense and swim better, making them a possible targeted drug delivery solution, for bioengineering applications or even as a lab-on-a-chip device. (Zhuang 2016). Later, during a study of *Escherichia coli*, the observations on single-cells in the T- maze along with a mathematical model revealed that there is a strong heterogeneity in the sensitivity and chemotactic mobility of the bacteria belonging to the genetically identical population. (Salek 2019).

1.1.2.3 Factors affecting motility and chemotaxis

Environmental stimuli may or may not induce a response in the form of bacterial motility; such cues are not well researched. However, certain sources of carbon and nitrogen are known to induce a chemotactic response in model organisms. Two of such sources are Glucose (Ambagaspitiye 2019) and Casamino Acids. (Samuels, Casamino acids slow motility and stimulate surface growth in an extreme oligotroph 2019). The laboratory experiments performed are usually nutrient-rich as compared to the natural environments in which the bacteria survive. The effects of energy-rich chemoattractants, like the glucose and casamino acids, on the motility of the oligotrophic *Escherichia coli*, is fairly unknown. In 2004 Zinser et al. (Zinser 2004) identified that *Escherichia coli* are capable of catabolizing a range of amino acids if they are maintained in the nutrient-limited conditions.

Glucose

In 1967, a study claimed that the presence of glucose prevents the synthesis of flagella in *Escherichia coli*. (J. a. Adler, The effect of environmental conditions on the motility of Escherichia coli. 1967) (Li 1993). The effect of glucose on the motility of microorganisms has majorly been studied in the form of gene expression studies. Another study in 1977, suggested that the reduced motility can be attributed to the reduced expression of the FlhD and FliA operons. (Silverman, Bacterial flagella 1977). The majority of these studies were aimed at understanding the underlying mechanism by which the presence of glucose affects the genetic machinery responsible for the motility; in other words, the biosynthesis of flagella via catabolite repression. (W. Z. Shi 1992) (W. L. Shi 1993) (Pratt 2002) (Wei 2001) and the genes responsible for the motility of bacterial cells (Inoue 2007). All these studies suggested that the bacterial cells will be non – motile, if grown in the presence of glucose.

Simultaneously, some studies aimed at determining the physical aspect of the problem, i.e. the effect of the glucose on the motility of the bacterial cells rather than the biosynthesis of flagella. In 1997, a group of scientists studied the motility of the Enterobacteria, which is the family to which *Escherichia coli* belongs, they grew the cultures on LB and MGM agar plates with varying concentrations of glucose. They concluded that the members of Enterobacteriaceae did not show a decline in the motility

with the increasing concentration of glucose from 22mM to 111mM. The *Escherichia coli* strain MC1000, MM335, and MC4100 showed inconsistent behaviour; sometimes, no motility was observed whereas other times, a completely motile population was observed (Lai 1997). Similarly, in 2008 a study on *Treponema denticola* was performed to study its motility under the influence of glucose. This experiment observed the motility using the capillary tubes, quantifying the motility by the cell numbers of *T. denticola*. To their surprise and contrary to the past observations related to the catabolite repression there was no alteration of motility nor was there a change in the flagella protein expression there was a greater number of *T. denticola* cells in the glucose-containing end of the capillaries (Ruby 2008).

During the same period the motility of the bacteria *A.hydrophila* was tested by observing its motility on the semisolid agar medium containing increasing concentrations of glucose, and all three strains of the bacteria under study showed a slow reduction in the motility with the increasing concentration of glucose. The strains ceased to be motile when the glucose concentration reached 2.5% or demonstrated highly impaired motility (Jahid 2013). Several studies have been carried out that have tried to study the motility of different bacteria strains and the effect of glucose on their motility genetically or physically, which either conclude lower gene expression of flagellar genes and how can it be altered (S. Y. Park 2019) (Rossi 2018) (S. P. Park 2016) (Delcenserie 2012) (Ling 2010) (Zhao 2007) or scarcely performed studies show inconsistent motility in the presence of glucose which does not support each other.

However, in recent years there has been a lot of work towards identifying the effect of glucose on the motility of bacteria by quantification of their motility. Wen et al. studied *Escherichia coli* BW2511 in LB broth Tris HCl agar plates to observe the chemotaxis of the bacterial cells without growth. They studied the chemotaxis of WT and ΔCheA (the mutant of CheA protein responsible for the chemotaxis) for different concentrations of glucose namely 10mM, 100mM and 1 M. It was observed that the two strains studied showed no significant chemotaxis as the bacterial circles were nearly identical. This study was rather non – specific as it did not suggest non- motile cells, but it concluded no significant chemotaxis, which means that the WT cells were motile but moved as randomly as the mutant cells (Wen 2019).

At the same time, Robert Cogger et al. studied the chemotaxis of the strain of *Escherichia coli* that is responsible for the Crohn's disease LF82 and its mutant for the protein LF82_p314, which causes defective swimming and swarming motility. In the capillary-

based, assay, it was observed that despite the popular idea of catabolite repression of the flagellar biosynthesis by glucose, the capillaries with glucose were more enriched than the ones without glucose. It suggested that the chemotaxis was intact in the wild type as well as the mutant despite the presence of glucose (Cogger-Ward 2019).

Another study that studied the fimbriae and flagella dependent motility of the uropathogenic *Escherichia coli* (UPEC) strains against the non – pathogenic (NPEC) strains attributed the difference in the motility to the presence of glucose. It was observed that the slow NPEC strains namely, W3110-LR, BW25113, AW405, and C600 did not penetrate the 0.25% agar as they did in the absence of glucose, whereas the fast NPEC strains MG1655, W3110-GSC, and RP437 swam into the 0.25% agar but slower than in the absence of glucose. On the other hand, the UPEC strains swam well irrespective of the presence of glucose, indicating a different mechanism governing the motility of the uropathogenic and non – pathogenic strains of *Escherichia coli* (Ambagaspitiye 2019).

Reyes et al. during the same period tried to study the migration bands of *Escherichia coli*, *Bacillus megaterium* and *Staphylococcus aureus* in Sugar Indole Motility medium in the presence of glucose as an attractant and alcohol as a repellent. There were no motility bands observed in the case of *S. aureus* in either the blank, glucose or alcohol case. But, *E. coli* and *B. megaterium* observed more motility bands (in the form of CFU/mL) in case of glucose as compared to the blank and the repellent, which showed lesser motility bands than blank (Reyes 2019).

Casamino Acids

As compared to the studies of the effect of glucose on the motility of the bacteria, the studies eliciting the effect of Casaminio acids on the chemotactic behaviour of the bacterial cells are even fewer. In 2000, Kohler et al. observed that the presence of Casamino acids on the swarm agar plates induced swarming motility in *Pseudomonas aeruginosa*. (Köhler 2000). Caiazza et al. also studied the swarming motility of the *Pseudomonas aeruginosa* on the M9 minimal medium in the presence of Casamino acids to identify any inhibition caused by the Rhamnolipids; the WT strain showed reduced motility in the presence of the inhibitor. (Caiazza 2005). Other similar studies on different bacterial species such as *Vibrio sp.*, *Serratia liquefacians* and *Salmonella typhimurium*,

emphasise that the addition of the casamino acids to the semi-solid and solid medium enhances the motility of the organisms (Kjelleberg 1982) (Bees 2002) (Harshey 1994).

More recently, a study which aimed at observing the effect of the casamino acids on the oligotrophic *Variovorax paradoxus.* The bacteria were grown on M9 minimal medium agar plates, which had different concentrations of agar with or without 0.1%(w/v) casamino acids. It was observed that the presence of the casamino acids significantly reduced the swimming motility of the bacterial cells and enhanced surface growth under limited nutrient conditions (Samuels, Casamino acids slow motility and stimulate surface growth in an extreme oligotroph 2019)

1.1.3 Cell to Cell Variability

Another one of the occurrences called Cell-Cell variability, has been overlooked over the past decades due to the assumption that the function performed in all the cells is similar approximately unless it is tweaked or nullified by a certain genetic mutation. However, in practical situations, the cells possessing identical genetic material do not function in the same way and have a certain level of variability.

The cellular variability is a fairly recent topic of study in biomedical science, and it is very critical in the studies dealing with the bacterial drug resistance profiles (Niepel 2009) (Weaver 2014), cancer cure and causes (Shaffer 2017) (Kessler 2014), stem cell fates and differentiation (Sheng 2012) (Cahan 2013) and formation of the biofilms (Mizan 2016). The same goes for chemotaxis, which is the movement of the cells towards or away from the chemical stimulus in the environment. The cellular heterogeneity in the area of chemotaxis is yet to be explored. The molecular and genetic system of chemotaxis has been characterized well for the prokaryotic model organism *Escherichia coli*. Although the system has been studied well, it does show differences in execution in different individuals.

The past decades have seen a great growth in industries of all sorts, which means a greater amount of chemicals are being produced than ever. This aside, there are serious cases of biomagnification, heavy metal accumulation in the food chain all over the globe (Xu 2019). Considering the alarming situations, rather than going for the chemical analysis, which will further add to the need for industries, scientist have started utilizing bacteria for the purpose. The microorganisms are easy and cheaper to harvest than the chemicals;

their smalls size also makes them favourable to be utilised as detectors inside the miniaturised portable devices. Such miniaturised portable devices are known as Bio – reporters. The outputs of the bio – reporters are simple; the bacteria can be genetically modified to produce bioluminescent proteins in response to the detection of traces of chemicals. This method is robust and requires lesser time to give results owing to the fast metabolism and short reproductive cycles of the bacteria. The concept of bio-reporters is very versatile and still being explored. The chemotactic cell studies can help us refine the two underlying principles of bio-receptors and similar point of care devices, namely: understand the repertoire of chemo-effectors a strain responds to and improve the sensitivity by selectively harvesting and culturing the sensitive individuals.

1.2 Aims and Objectives

As the first step towards contributing to the field of using bacteria for therapeutic purposes, I decided to study the motility, chemotaxis and the cell to cell variability of the bacterial strains studied previously in our lab to study the effect of genome deletion on their growth patterns under different nutrient conditions. In our lab (Yuan, Single-Cell Microfluidics to Study the Effects of Genome Deletion on Bacterial Growth Behavior 2017) studied the effects of the genome deletion on the growth of 2 strains in different medium compositions and the heterogeneity observed at an individual level. The strains used are MG1655 and MDS42. MG1655 is a strain with the genotype closer to the wild type *Escherichia coli*, whereas the strain MDS42 is the result of multiple deletion series project and has ~14% of its genome deleted. The details about the strains under study are presented in Chapter 2 Materials and Methods. I tried to observe the effects of the different medium compositions on the motility and chemotaxis of the strains discussed. The project aims to comparatively study the effect of nutrients on the chemotaxis and motility of Escherichia coli strains. The study had the following objectives:

1. To understand the basic principles of imaging, data and statistical analysis, cell culture, aseptic lab techniques.
2. To study the effects of different concentrations of Glucose and Casamino Acids on the motility of MG1655 and MDS42 strains of *Escherichia coli*
3. To evaluate heterogeneity in the motile populations in different media

4. To design, and evaluate a single-layer microfluidic device that allows the user to study cellular chemotaxis towards linear gradients of chemoattractant
5. To evaluate the effects of different concentrations of Glucose on the chemotactic behaviour of MG1655 strain of *Escherichia coli* at both the population and individual levels
6. To test the hypothesis that the absence of flagellar motility results in more unidirectional motion that is lesser changes in the direction of motion.

1.3 Thesis outline

The thesis is organised in three chapters which are structured as followed:

Chapter 2: This chapter describes the bacterial strains and common microbiology lab techniques deployed during the project. It also discusses the software used to design, sketch, code, acquire, extract and analyse the data during the experiments. The section of Statistical methods provides a basic understanding of the methods used in different cases in the following chapters.

Chapter 3: This chapter starts with an introduction to the concept of motility and the assays developed until now to study the motility in bacterial cells. It then describes the method used in the experiments to study the motility of the bacteria and to capture and extract data. It also discusses the details of the genetic machinery of MG1655 and compares it to MDS42 and discusses the results at both – population and individual level.
.

Chapter 4: This chapter starts with an introduction to the concept of chemotaxis and the basic underlying molecular and genetic system responsible for chemotaxis in prokaryotic bacteria and how the molecular cross-talk between the chemotactic system and the system responsible for the flagellar biosynthesis entails. The chapter then highlights the method used in the current study – starting from the device designs to fabrication, establishment of chemo – effector gradient, and the methods used to capture and extract data. It also discusses results at population and individual level for the MG1655 strain of *Escherichia coli*.

Chapter 2 | Materials and Methods

In this chapter, the materials, organisms, software, databases, and statistical tools & methods used to acquire data of the motility and chemotaxis of *Escherichia coli*, under different medium compositions are described in detail.

(Note: The specific procedure and setup for the motility and chemotaxis experiments is in Chapter 3 & Chapter 4 respectively)

2.1 Bacterial Culture and Growth

In continuation of the study of the effect of genome deletion on single-cell growth via microfluidics (Yuan, Single-Cell Microfluidics to Study the Effects of Genome Deletion on Bacterial Growth Behaviour 2017), the stock cultures of the wild type *Escherichia coli*, K–12 MG1655 strain and derived strain MDS42 (also referred to as Clean genome) were used for the comparative study.

The whole-genome sequence for *Escherichia coli* published in September of 1997 belonged to the strain *Escherichia coli* K-12 EMG2, which is also known as Wild – type strain (Blattner 1997). However, the MG1655 strain used in the experiments, is a second-order derivative of the ancestral strain EMG2. MG1655 does not possess the fertility factor and the λ, unlike the EMG2. (Hayashi 2006) (Yuan, Single-Cell Microfluidics to Study the Effects of Genome Deletion on Bacterial Growth Behaviour 2017)

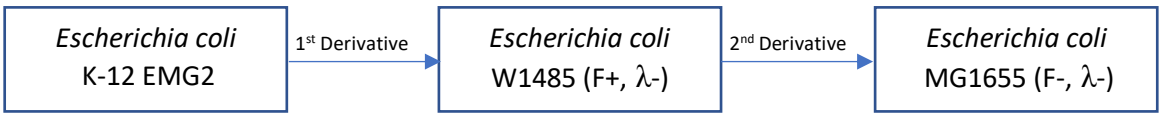


Figure 2. 1: The derivation of MG1655 from Wild-type EMG2

The MDS42 (multiple – deletion series) strain is the second strain used for the study. The MDS42 strain has undergone substantial deletion of genes deemed as "non-essential", amounting to 14.3% of the total MG1655 genome. (Pósfai, Emergent Properties of Reduced-Genome Escherichia coli 2006) (Yuan, Single-Cell Microfluidics to Study the Effects of Genome Deletion on Bacterial Growth Behaviour 2017)

2.1.1 Preparation of LB Agar Plates

Luria – Bertani broth is the most commonly used medium for culturing bacteria in labs, owing to its nutrient richness and its ability to support the quick and steady proliferation of several bacterial species. (Sezonov 2007). Luria – Bertani agar plates provide a way to obtain individual colonies, each identical to the individual ancestral cell-cultured, and can be stored for a minimum of 4 weeks (Iacoviello 2001).

2.1.1.1 Materials Required

LB agar powder (Lennox L Agar), Milli Q water, 250mL Pyrex jar, autoclave, autoclave tape, laminar hood, Petri plates, weighing balance, parafilm

2.1.1.2 Procedure *(Brent 1992)*

Weigh the correct proportion of MB Agar Powder to add to 100 ml of Milli Q water in a 250ml and add 100mL water to the Pyrex jar containing LB Agar Powder gently. Swirl the bottle to mix the contents and keep on a hot plate to boil for 1-2 minutes, to help with dissolution. Loosen the lid a bit, to allow the passage of steam and fix the lid with autoclave tape. Keep the mixture to autoclave for 20 minutes. Take out the jar and let it cool under the UV treated and ethanol wiped Laminar hood. Once the jar is cold enough to touch, pour the contents to Petri plates and let them cool down. Store the Petri plates in a 4ºC environment after covering them with parafilm and use them later.

2.1.2 Revival of Stocks

The frozen stocks of MG1655 and MDS42 were stored in -80ºC freezer.

2.1.2.1 Materials Required

Poured LB agar plates, inoculation loop, incubator, glycerol stocks of MG1655 and MDS42, 70% ethyl alcohol, absorbent paper or cotton.

2.1.2.2 Procedure

UV sterilise the laminar hood and clean the workbench with ethanol. Take out the poured plated from 4ºC refrigerator and keep them under the laminar airflow and label them as "Bacterial Strain | Date | Time.". Label one plate as "Control | Date | Time.". Bring the glycerol stocks from the -80ºC freezer on ice and keep them in the laminar hood for streaking. Do not let the glycerol stock thaw, as freezing and thawing can cause mortality in stock. Use an inoculation loop and streak on the plate gently. Use a new sterile loop for each new streak. After streaking, keep the plates with bacteria as well as the control in the incubator at 37ºC overnight. Take the plates out the next day and check for growth and single colonies or any possible contamination. Keep the cultured plates at 4ºC after covering the lid with parafilm.

2.1.3 Preparation of M9 Minimal Medium

Escherichia coli strains both MG1655 and MDS42 grow rapidly in M9 minimal medium (Yuan, Single-Cell Microfluidics to Study the Effects of Genome Deletion on Bacterial Growth Behaviour 2017). The minimal medium contains a carbon base compound glucose in case of this experiment (glucose serves both as a carbon and energy source) and M9 salts that supply nitrogen, phosphorus, and trace elements. The M9 minimal medium can be prepared by following the steps as below (Elbing 2002) (Neidhardt 1974):

Firstly, prepare the M9 minimal salt solution. This salt solution will be a 5X concentrate. Add 64g of $Na_2HPO_4.7H_2O$; 15g of KH_2PO_4; 2.5g NaCl and 5.0g NH_4Cl to 800mL of distilled water. Make up the volume to 1L by adding more distilled water. Autoclave the solution for 15 – 20 mins at 121 ºC temperature and 15 psi pressure.

Secondly, prepare the stock solutions of 1M $MgSO_4$ and 1M $CaCl_2$. Add 24.65g of MgSO4.7H2O to 100mL of distilled water. Add 147.014g of CaCl2.2H2O to 100mL of distilled water. Autoclave the solutions for 15 – 20 mins at 121 ºC temperature and 15 psi pressure. Thirdly, prepare a 100mL stock solution of 1M Glucose in distilled water and filter sterilize it. Add 18.0156g Glucose to 100mL distilled water. Prepare 1L M9 Minimal Medium by adding 200mL solution of sterilized 5X M9 salts to 800 mL distilled water to make a 1X solution. Then add 2mL of 1M $MgSO_4$ solution, 0.1mL of $CaCl_2$ solution, glucose to prepare 10mM and 10µM solutions, and 0.2% Casamino Acids.

2.1.4 Plotting the growth curves

The bacteria studied in the experiment were all studied in the log phase of their growth. To ensure that they are in the log phase, I plotted the OD_{600} graph against time to study the Optical density corresponding to the Log phase of growth. The bacteria were grown in M9 minimal medium with 10mM glucose and 0.2% casamino acids overnight at 37 ºC at 150 rpm. 20µL of the overnight culture was taken and inoculated in 10mL of fresh medium in 3 culture flasks. Another culture flask with just 10mL of minimal medium was kept as a blank. Periodic OD_{600} readings were taken for all the four culture flasks using a 96 – plate reader. **Figures 2.2** and **2.3** show the Ln(OD/OD_o) vs time plots for MG1655 and MDS42 strains. The average doubling time for both the samples of MG1655 (~42 mins) & MDS42 (~50 mins) is the same as observed previously (Yuan 2017). (Refer **Table 2.1 & 2.2**)

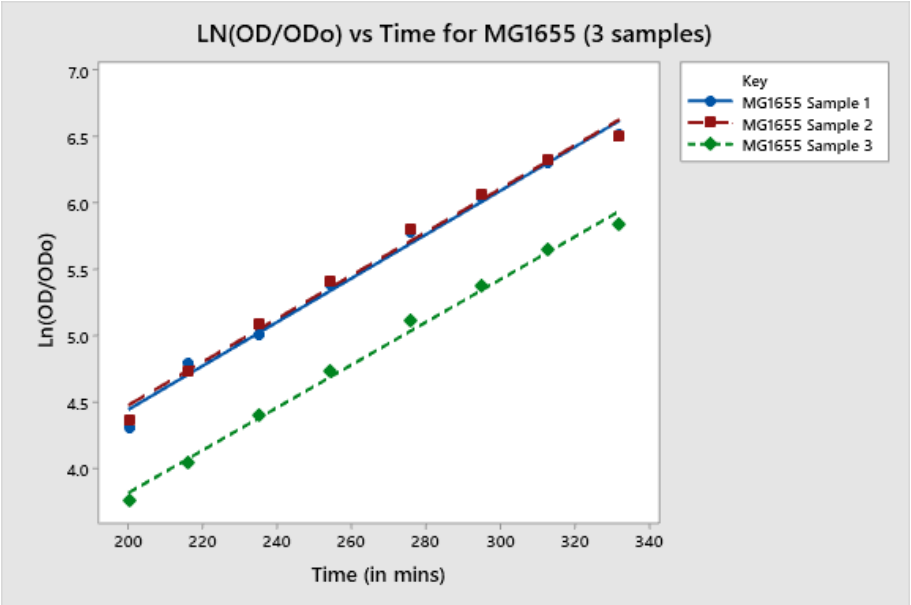


Figure 2. 2: The Ln (OD/OD0) vs time plot for three samples of MG1655 strain inoculated from the same culture into three different culture flasks and cultured under the same conditions.

Series Equations	Doubling Time	Average Doubling Time
Y = 0.0165x + 1.1493	42.01 mins	42.53 mins
Y = 0.0163x + 1.2192	42.54 mins	
Y = 0.0161x + 0.604	43.05 mins	

Table 2. 1: The line equations for Figure 2.2 (3 samples of MG1655) and their corresponding doubling times. The growth rate r was calculated using the equations obtained from the plots using the formula: r = (ln [OD2/OD1]) / (T2-T1). This growth rate r was then used to calculate the doubling time "d" given by ln2/r.

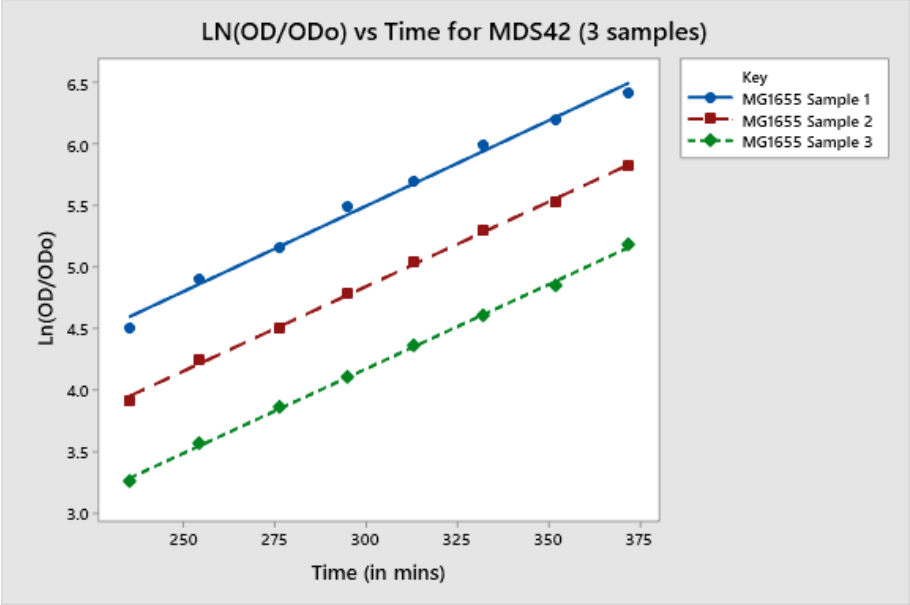


Figure 2. 3: The Ln (OD/OD0) vs time plot for three samples of MDS42 strain inoculated from the same culture into three different culture flasks and cultured under the same conditions.

Series Equations	Doubling Time	Average Doubling Time
Y = 0.0138x + 0.7180	50.23 mins	50.23 mins
Y = 0.0139x + 1.3303	49.87 mins	
Y = 0.0137x + 0.0600	50.59 mins	

Table 2. 2: The line equations for Figure 2.3 (3 samples of MDS42) and their corresponding doubling times. The growth rate r was calculated using the equations obtained from the plots using the formula: r = (ln [OD2/OD1]) / (T2-T1). This growth rate r was then used to calculate the doubling time "d" given by ln2/r.

2.2 Optical Imaging

2.2.1 Bright field microscopy

Bright-field microscopy is the simplest kind of optical microscopy. In bright-field microscopy, the sample is illuminated with the transmission of light. A contrast is then formed due to absorption of light. The dense parts appear darker on a bright background. In a bright-field microscope, the specimen is mounted on the stage. It should be made sure that the glass coverslip is not thick to focus high magnification lens through glass slip. The sample has to be positioned manually under the lens. (G. a. Wang 2012). Bright-field microscope has a simple setup, and it also enables the analysis of moving cells. (Kural 2007)

2.2.2 Fluorescence Microscopy

Fluorescence Microscopy is another type of optical microscope which creates an image using fluorescence instead of scattering or absorption of light (Tkaczyk 2010). Andor EMCCD camera was used to capture both brightfield and fluorescence images. Fluorescence imaging was used for two purposes in the project. The chemotaxis devices built for studying the effect of a nutrient on the bacteria worked on the condition of no flow and uniform gradient of chemoattractant. To determine the establishment of a gradient fluorescein was used and imaged. The no-flow condition was checked using the fluorescent microparticles. The details and images are discussed in Chapter 4.

2.2.3 Image Analysis

ImageJ, a public domain image processing and analysis program bases on Java script (Barboriak 2005), was used for imaging analysis. ImageJ can help in solving various image processing problems like live cell imaging and medical image processing (Eliceiri 2005). Along with that, it performs multiple imaging system data comparisons (Rajwa 2004).

ImageJ can perform image processing functions like sharpening, smoothing, edge detection, contrast manipulation and median filtering. It can also plot histograms.

Furthermore, it measures area and pixel value statistics, calculate angles and distances and perform geometric transformations. ImageJ provides the liberty to perform any image processing or analysis at any magnification level. Afterimage processing the data can be imported into Microsoft Excel to obtain summated set of results. Refer to Appendix 1 for details on data acquisition and extraction. Inkscape is an open-source vector graphics editor which is used to create illustrations, maps, complex drawings, diagrams and logos (Bah 2010).

2.3 Genome Analysis

2.3.1 NCBI BLAST

NCBI BLAST was used to conduct the support study of Genome comparison of MG1655 and MDS42 to validate the results of the experimental results. Basic Local Alignment Search Tool, which is also known as BLAST, is a sequence similarity search program (Johnson 2008). It is widely used to search for DNA and protein databases to find similarities in sequences (Altschul 1997). BLAST sets a threshold parameter (T) which allows a trade-off between speed and sensitivity. If the value of threshold increases, so does the speed, but the sensitivity will decrease, which in turn increase the probability of ignoring weak similarities (Altschul 1997). Refer to Appendix 2 for details of the study.

2.3.2 UNIPROT

UniProt was used in the support study of genome comparison of MG1655 and MDS42 to validate the results of the experimental results by identifying the protein substrates of the genes responsible for the flagellar motility. UniProt is an extensive database of protein sequence and their functional groups (Consortium 2015). It comprises of various biological information of proteins which is important to identify the experimental characterization of proteins. Refer to Appendix 2 for details of the study.

2.4 Statistical Methods

Several statistical methods have been used in the data analysis of the study, and this section provides a basic understanding of the method employed and the procedure to conduct certain tests using the Minitab software.

2.4.1 Minitab

Minitab is a statistical software that perform various data analysis. Data can also be analysed and summarized in MS excel, but Minitab focus more on statistical analysis and interpretation of results. It provides better visualization of results. Input of Minitab is raw data, and it then simplifies the data depending upon the statistical analysis, manipulate the datasets and identify the trends or patterns in the data.

Minitab can perform all kind of statistical tests as well as descriptive statistics. It helps to plot various statistical graphs like scatter plot, histogram etc. It enables the user to find the relationship between variables using regression. It performs analysis of variance, i.e. ANOVA test to determine the similarity and difference between the means of groups. It also creates control and time-weighted charts, and lastly, it allows us to find the best fit distribution for modelling the data.

The above features are difficult to perform by excel while with Minitab, they become easier and less time consuming (Prvan 2002).

2.4.2 Anderson Darling Test

The Anderson Darling test can be performed to check whether a sample of data follows a specific distribution. When it comes to a normal distribution, this test is the most powerful statistical tool to determine the normality of data, and it is also sensitive to the deviation of data from normality (Stephens 1974) (Nelson 1998)

Anderson Darling Test assumptions:

H_0 = the data belongs to a normal distribution

The p-value should be less than the significance level α, to reject the null hypothesis.

A p-value > α (0.05 in our case) shows that the data is normal or it follows the normal distribution.

2.4.3 Outlier Test

The presence of outliers in data can often distort the results of the analysis, so it is important that they can be detected before performing any statistical analysis on the data. Grubbs test is used to detect outliers in the univariate data set. However, Grubbs test depends on the normality of the data. That's why the data should be tested for normality first, and it should follow normal distribution before the application of Grubbs Test (Grubbs 1950)

Grubbs Test only detects one outlier at a time; therefore, the test is iterated several times until no outlier remains in the dataset. However, multiple iterations reduce the probability of detection. (Tietjen 1972)

Grubb's Test assumptions:

H_0 = the data has no outliers, or all data values belong to a normal distribution

A p-value greater than 0.05 indicates that there is no outlier in the normally distributed dataset.

2.4.4 ANOVA Test

ANOVA test perform mean comparison by analyzing the variance between the groups. ANOVA test requires the data to be normally distributed and with equal variance between the individual variables or factors. The test shows robustness to the non-normal data.

ANOVA can be one way or two ways depending upon the number of independent variables required for the test. One Way ANOVA has only one independent variable with two groups. It compares the mean of two independent groups with the help of F-distribution.

One Way ANOVA makes six assumptions, which include the dependent variable should be continuous; independent variable should have two or more categorical groups that are unrelated to each other; there should be independent of observations between the groups; if there exists a relationship between observations of groups, then the groups will not be unrelated, it will result in repeated measurements. It should be made sure that the data

fulfils above assumptions before moving to the other three assumptions (Moder 2010), which include, there should not be any presence of significant outliers as they negatively impact the performance and accuracy of One-Way ANOVA; the data for the dependent variable should approximately follow normal distribution for all groups of independent variables and the variances should be homogeneous

One Way ANOVA test has the following Null hypothesis H_0: means of normally distributed groups are equal. If the p-value is less than 0.05, then the null hypothesis would be rejected, and it can be stated that there is a statistical significance between the means of independent groups.

2.4.5 Welch's Test

To perform One – Way ANOVA the assumption of equal variances must hold. However, in situations where the variances are different, Welch's ANOVA can be used. Welch ANOVA is a form of One-Way ANOVA, and it does not assume equal variances to perform the test. Welch's test applies to all the cases with non-homogeneous variances of normally distributes data.

Welch's test can be performed without testing for equal variances as it provides the same results, even the variances are equal. It has the lowest rate of type 1 error (incorrect rejection of the null hypothesis) (Derrick 2016).

Although Welch's test is designed for unequal variances, but the assumption of normality must be maintained to perform the test. As Welch's test also compares means of different groups, therefore the hypothesis of Welch's test is similar to that of classic One-Way ANOVA. The p-value and significance interval are compared in the same way as that of classic One-Way ANOVA.

2.4.6 Games Howell Test

Games Howell test performs pairwise comparisons and provides best results for comparisons between all pairs of groups. It is an extension of Welch's test with unequal

variances. Games Howell obtains confidence interval for the mean difference between the groups by using the formula for Welch's test (Shingala 2015)

Games Howell test provides narrower confidence limits and maintains the error for groups with unequal size and variances. Moreover, Games Howell test is robust for non-normal data (Day 1989).

It is recommended for Games Howell test to have sample sizes greater than 5. Along with that, the observations of the groups should be independent of each other and should be normally distributed.

Hypothesis:

The following hypothesis should be made to perform Games Howell test (Games 1976);

Null hypothesis H_0 = mean of independent groups are equal/ all pairs of groups are equal
If the p-value is less than 0.05, then the null hypothesis will be rejected, and it can be stated that pairs or groups are significantly different.

2.4.6 Box Cox Transformation

Normality plays an important role in statistical analysis; therefore, the data must be normally distributed. In cases where data is non-normal, Box-Cox transformation can be used to stabiles the variances and transform the data into more like normal. It will help in performing a variety of statistical tests on the data.

The core of Box-Cox Transformation is an appropriate exponent known as lambda (λ) (Sakia 1992). The value of lambda lies between -5 and 5. Depending on the data, the optimal value of λ can be selected (Senvar 2016). This optimal value indicates the power to which the data should be raised to obtain the best approximation of normal distribution.

The confidence interval is used to determine if the data requires a transformation. If the confidence interval of optimal λ contains 1, then the data does not require any transformation. A value of 1 indicates to use the original data. However, if the 1 appears in the confidence interval, then the data should be transformed depending upon the selected value of optimal λ.

To transform the data, the optimal value of λ should lie in the confidence interval, and it should not contain 1.

2.4.7 Johnson Transformation

Johnson transformation was introduced as a new powerful technique to reduce the skewness of data as well as approximate the data to normality. It is also used to determine whether the original and transformed data belong to normal distribution. It can be used for both positive and negative values (Senvar 2016).

Hypothesis:

The hypothesis for Johnson transformation is similar to the Anderson Darling test, and it states that;

H_0 = Data is normally distributed

If the p-value is less than 0.05, then the original and transformed data is not normal, and the null hypothesis will be rejected. However, if the p-value is greater than 0.05, then it can be stated that there is no significant evidence to reject the null hypothesis, and the data follows a normal distribution.

MINITAB then forms normal probability plots and displays p-value for the original and transformed data. For normally distributed data, the points on the probability plot will follow the straight line and p-value will be greater than alpha, i.e. 0.05. If the data is non-normal, the points will not fall along the fitted normal distribution line, and p-value will be less than 0.05.

If the original data follows a normal distribution, then MINITAB will only display one probability plot and will not perform Johnson transformation.

2.4.8 Goodness – to – fit Test

Goodness of fit test is performed to determine if the sample data fits any defined distribution. It means that the test is used to select the best fit and appropriate distribution for the analysis of data. It is done by measuring the distance between the desired distribution and sample. This distance is then compared with a threshold value. In statistical terms, the distance is referred to as the test statistic, while the threshold value is known as the critical value.

A test statistic is calculated from the sample data and is used to reject and accept the null hypothesis (Casella 2001). Test statistic is also used to calculate the p-value. If data has evidence against the null hypothesis, then the value of test statistic either become too small or too large owing to the alternative hypothesis. And this change in test statistic decreases the p-value enough so the null hypothesis can be rejected.

However, critical value defines a point on test distribution which is compared with test statistics to test the hypothesis. Test statistics are related to the significance level (alpha). Therefore, their values remain fixed from the beginning of the test.

Hypothesis:

Goodness of fit test holds the following hypothesis;

H_0 = the distribution fits best for the data

Either test statistics is compared with critical value, or p-value is compared with significance interval, to test the hypothesis. If the test statistic is less than a critical value, then it can be stated that the distribution is a good fit for the data. However, if the test statistic greater than the critical value, it can be deduced that the chosen distribution is not a good fit for the analysis of the data.

Similarly, if the p-value is greater than alpha, i.e. 0.05 then it provides evidence to accept the null hypothesis otherwise lower p-value indicates evidence to reject the null hypothesis.

2.4.9 Kruskal Wallis Test

Kruskal-Wallis test is used to compare medians between two or more groups. It is a distribution-free test and is used when the assumptions of one-way ANOVA are not satisfied (McKight 2010) (Kruskal 1952). For one-way ANOVA, the assumptions of normality and equal variances must be made while for Kruskal-Wallis test; no such assumptions are made.

One of the assumptions that must be made to perform Kruskal-Wallis test is that the distributions of data for all groups have a similar shape, i.e. they do not vary in skewness (Vargha 1998).

Hypothesis:

The following null hypothesis must be considered to perform the Kruskal-Wallis test

H_0: medians among the groups are similar.

If the p-value is less than 0.05, then the null hypothesis would be rejected, and it can be stated that there is a statistical significance between the means of independent groups.

2.4.10 Levene's Test

A Levene's test is used to determine if the variances are equal or not (Gastwirth 2009). The Levene's test used in Minitab, is based on the modifications by Brown and Forsythe to Levene's procedure. It considers the distances of the data points from their sample median rather than their sample mean. Using the sample median rather than the sample mean makes the test more robust for smaller samples (M. a. Brown 1974). Levene's test is absolute for any continuous distribution. For skewness heavy distributions as well, this method is more reliable than Bonett's and F – test method. (Schultz 1985).

If the results from the test conclude that the data sets have homogeneous variances, an ANOVA F-test can be performed but, it the results indicate otherwise, the Welch modification test is better suited. (Gastwirth 2009)

Chapter 3 | Comparative Studies of Motility

3.1. Motility studies on *Escherichia coli*

Among the prokaryotes, *Escherichia coli* K-12 is one of the model organisms for the genetic, biomolecular and physiological research, as it has been understood and analysed thoroughly. It is also used for the commercial synthesis of biomolecules, hormones and enzymes of medical. therapeutic & industrial significance. (Pósfai, Emergent Properties of Reduced - Genome Escherichia coli 2006). The genome of MG1655 which is closely related to the K-12 Strain, as stated in Chapter 2, has also been completely sequenced, and nearly 87% of its genes have defined functions as well. (Serres 2004) (Riley 2006) (Blattner 1997).

MDS42, the second strain under consideration in this study, a result of Multiple Deletion Series project. It displayed numerous advantages for the biotechnological applications over the MG1655, absence of cryptic virulence elements and mobile elements being the most desirable among them (Karcagi, Indispensability of Horizontally Transferred Genes and Its Impact on Bacterial Genome Streamlining 2016). Although the strain proved competent when it came to Growth yield, cell size, acid stress tolerance but the genome reduction took a toll on nutrient utilization ability (Price 2004), and the stress tolerance of the MDS42 strain (Bochner 2001). Since, the idea of Multiple Deletion Series Project was to get rid of the mobile and cryptic virulence elements primarily. Does that mean the motility related genes were affected too?

Escherichia coli are peritrichous organisms and usually move with a run and tumble strategy in their environment. It has also been found in certain studies that their trajectories tend to become loop like near the surface (Lauga 2006) (Percival 2014). The bacterial flagellum has three major components – long helical filament, a hook and a basal body integrated in the cell membrane. All the proteins responsible for the flagellar components are synthesized in the cytoplasm but vary in their final place of localization. The final destinations can be as different as the cytoplasm, the inner, outer and peripheral cell membrane, the periplasm, and the cell exterior. All the protein components are capable of self – assembly once the prerequisite subunits are available.

The motility studies date back to 1880s, where two groups led by Engelmann (1881) and Pfeffer (1884) tried to observe the motion of the bacteria with the help of the newly

developed microscopes. (J. Adler 1966) (Drews 2005). However, the topic of motility and chemotaxis boomed after the 1950s. Two of the prominent experimental studies took place in the 1970s, which are still fundamental to current motility studies. In these preliminary studies the live cells were attached to the surface of the microscope slides through flagella, and the information like speed and direction of the rotation of the flagellar motor was recorded along with the events of stop and the reversal of the direction of rotation. (Silverman, Flagellar rotation and the mechanism of bacterial 1974) (Berry 2000)

Another series of experimental methods was started by Berg, where the motile cells were tracked and studied under the microscope on a movable stage. The stage was controlled by the light intensity of the images via an electronic feedback loop. (H. C. Berg 1971) (H. a. Berg 1972) The light intensity also corresponded to the position of the bacterial cells in the 3-D space. Several similar techniques have been developed by different researcher groups to investigate bacterial motility (D. a. Brown 1974) (Duffy 1997)

Later decade observed a similar study in which the cells were attached to the microscope slides via the cell body, unlike the previous studies where they were attached through the flagellum. The flagellum was sheared, and latex biomarker (bead) was attached to it. This marker facilitated the observation under microscope easier as the movement of sheared flagellum moved the bead, and the motion was recorded. (Ryu 2000) (Reid 2006) (Sowa 2005)

Further ahead in time, the scientists added the automation to the tracking process of the bacterial cells. They tried to study a large number of cells (~100 – 200 cells) in a fixed field of view using a cell tracking software (Postlethwaite 2013). Independent studies carried out by Duffy and Ford (Duffy 1997), Frymier et al. (Frymier 1995), Vigeant and Ford (Vigeant 1997) and Berg and Brown (D. a. Brown 1974), were high throughput studies with greater cell motility readout.

This chapter aims at studying the differences in the motility behaviour of the two strains, namely MG1655 and MDS42 and their responses to the two chemo-attractants: Glucose and Casamino Acids. This study is a combination of the studies mentioned above. In this study, the cells are grown to the exponential phase to ensure maximum metabolic activity in different culture medium compositions. The cells are then studied using the microscope in a fixed field of view, and they are tracked one by one using the manual tracking software to ensure accuracy.

3.2 Materials and Methods

3.2.1. Experimental Setup

Prepare the experimental setup using the given procedure to study the motility of the strains MG1655 and MDS42, as shown in **Figure 3.2**. The setup is assembled using a simple sterile microscope slide, coverslips and parafilm sheet. Fix the coverslips on the microscope slide in a way that they form a few millimetres wide channel. The ends of the channel must be covered with parafilm to avoid overflow of the liquid inoculums. The bacteria strains, both MG1655 and MDS42, must be grown in the different medium compositions, as mentioned in **Table 3.1**, to study their motility.

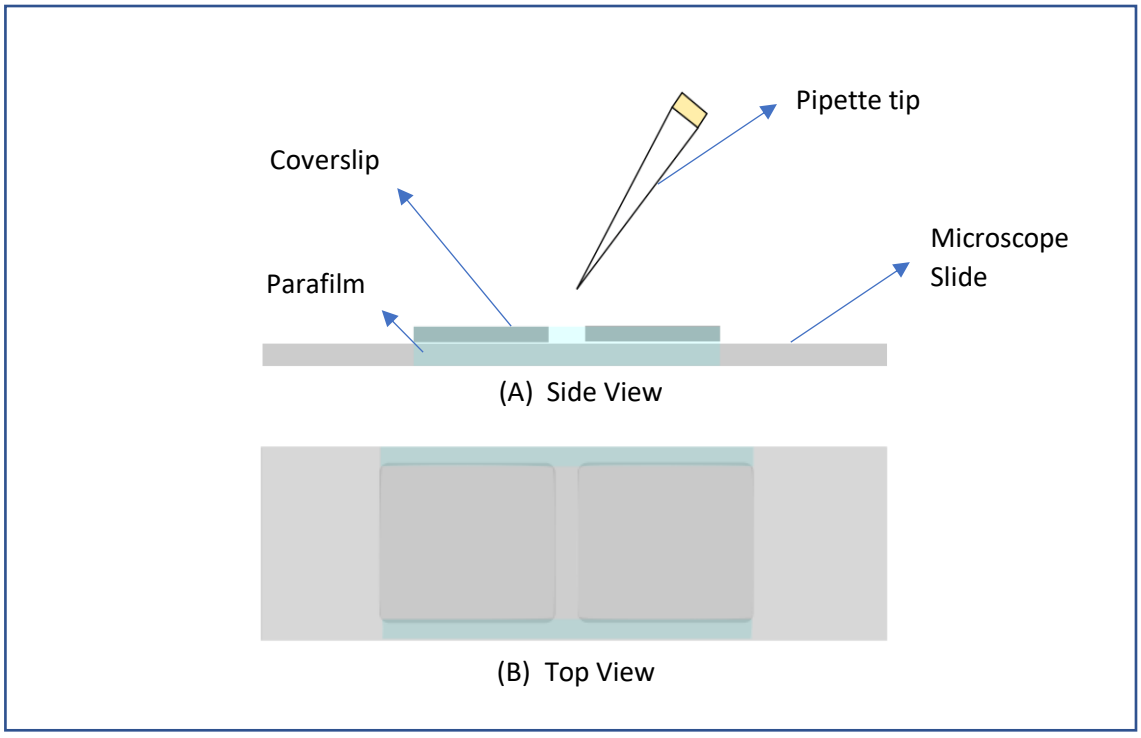


Figure 3. 1: The experimental setup to study the motility of the 2 strains of Escherichia coli - MDS42 & MG1655. (A) The side view showing all the components including the cavity below pipette tip for the sample. (B)The top view to show the channel-like cavity

The prepared apparatus is put in transparent cover under UV light, until the bacterial samples were ready. *Escherichia coli* cells are cultured in 10mL of different medium compositions as mentioned in **Table 3.2**, overnight at 37ºC, until they reach stationary phase. Take 20µL of the overnight culture was inoculated into 10mL of the same fresh medium for 2.5 hours (unless specified otherwise) at 37ºC at 150 rpm. Once the desired Optical Density of the bacterial culture is achieved (as mentioned in **Table 3.2**), use a 200µL pipette to inoculate 100µL of bacterial culture suspension into the channel. Place the slide with the sample on the centre of your microscope stage. To reduce the amount

of light passing through the sample on stage, lower the condenser lens to the lowest position.

Once the sample is on the stage, use the 4X objective, focus on the edge of the slide. Switch to the 60X objective and refocus on edge. Now slowly move the specimen away from the edge and look for the presence of motility. It may be necessary to decrease the intensity of light passing through the specimen to increase the contrast. After manually focussing the microscope with a 60X objective, record the videos using the Andor EMCCD camera and Andor Solis software. Refer to **Appendix 1** for details on data acquisition and extraction.

To validate and support the results obtained in the experiments, a genome comparison study between MG1655 and MDS42 was performed. The details of the study and the motive are present in **Appendix 2,** and the results have been used to build arguments in the results section of Chapter 3.

	M9 SALTS				
M9 SALTS			**GLUCOSE**		
			10mM	**10µM**	**0M**
	CASAMINO ACIDS	**0.2%**	**YES**	**NO**	**YES**
		0%	**YES**	**YES**	**YES**

Table 3. 1: The matrix showing the compositions of the M9 minimal medium (labelled in green) used for the motility study

Medium Composition	**MG1655 (OD_{600})**	**MDS42 (OD_{600})**
M9 Salts + 10mM Glucose + 0.2% Casamino Acids	~0.1	~0.1
M9 Salts + 10mM Glucose + No Casamino Acids	~0.1	~0.1
M9 Salts + 10µM Glucose + No Casamino Acids	~0.05	~0.05
M9 Salts + No Glucose + 0.2% Casamino Acids	~0.1	~0.1
M9 Salts + No Glucose + No Casamino Acids	~0.04	~0.04

Table 3. 2: The list of the Optical densities at 600nm for the two strains of Escherichia coli in different medium compositions at the time of the experiment

Medium Composition	**Label**
M9 salts + 10mM Glucose + 0.2% Casamino Acids	A
M9 salts + 10mM Glucose + No Casamino Acids	B
M9 salts + 10µM Glucose + No Casamino Acids	C
M9 salts + No Glucose + No Casamino Acids	D
M9 salts + No Glucose + 0.2% Casamino Acids	E

Table 3. 3: Labels assigned to medium composition

3.3 Statistical Analysis: Procedure

Using the Data Acquisition and Extraction Methodology mentioned in **Appendix 3**, the cell movement data was collected for ~50 cells of both strains, in the mentioned different medium compositions (refer to **Table 3.2**). The data of each of the cells was collected for 100 frames. It gave us parameters such as – "Angular change per frame" – this was the angular change in the trajectory of the cells with each frame, "Overall Displacement" – the distance between the initial and final position of the cell, "Framewise Velocity" – the absolute value of velocity calculated based on the displacement of the cell from one frame to another, and "Average Velocity" – the average of the framewise velocities of a cell in all the frames. The first set of results is based on the average velocity measurements of the specimen, studied collectively for the populations in different medium composition and different strains. The procedure followed for statistical analysis of the data is described below.

Perform a Goodness – of – fit Test to identify the data distribution. It helps in recognising if the data followed any other distribution more suitably than normal. However, for the experimental data, the A-values and p-values obtained, indicated more inclination towards normal distribution than any other probability distribution. The A-squared statistic values were found to be the smallest with p-values higher than 0.05, which indicated that the data is more likely to follow a normal distribution.

Generate a data and graphical summary. For further data analysis, a graphical data summary was generated for each strain in different medium compositions. The data summary is necessary to understand the centre of the data collected, determine to a confidence interval of mean, median & standard deviation and to assess the spread, shape and the distribution the data follows. The data summary helps in deciding whether parametric or empirical tests will be suitable for the study. The graphs, as shown in **Figure 3.2 (A)** and **(B)** were obtained to summarise the data for each of the medium composition data for both the strains. The tables related to data summary are present in **Appendix 3**. From the data, the statistic of primary importance is the one corresponding to the Anderson – Darling Test, i.e. A – squared and p-value. It turned out that 3 out of 10 cases deviated a little from the normal distribution.

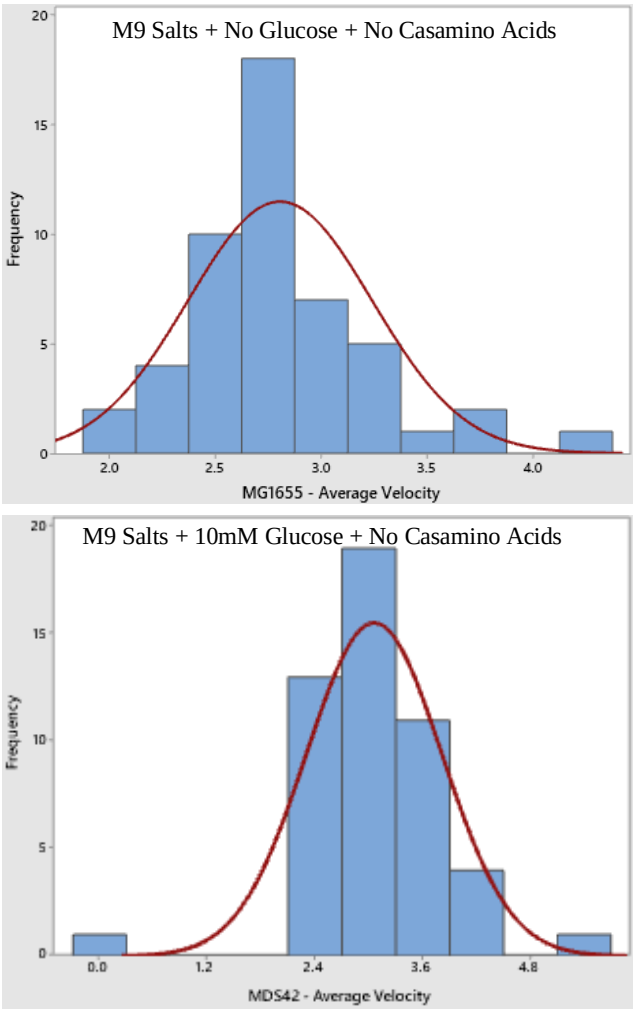


Figure 3. 2: The histograms and the corresponding normal distribution fits of MG1655 and MDS42 Average Velocity data in different media compositions. The data was checked for normal distribution using Anderson Darling Test and the non – fitting data was transformed

Perform necessary data transformations. If all the datasets follow the similar probability distribution, proceed to the next step but if they do not follow, use an appropriate data transformation method. Notably, for the strain MDS42 in medium compositions labels **A** and **B** (**Table 3.3),** the values of A – squared statistic from Anderson Darling are high, and the p-value is also low – [1.49, <0.005] and [1.02, 0.010] respectively. Also, the strain MG1655 in the medium composition label **E**, the values of A- squared statistic from the Anderson Darling Test and the p-value is [1.127, 0.005]. It necessitated the need to perform data transformation to bring the data distribution closer to normal distribution. The data transformations performed for this purpose did not result in homogeneous variances.

Check for the deviations such as skewness and kurtosis. The data collected for the two strains in different medium compositions other than the above three cases of data transformation, show skewness and kurtosis. The presence of skewness indicates the

asymmetry in the data distribution, and kurtosis implies the deviation from the normal distribution data peak (Kim 2013). However, all the values of skewness and kurtosis are contained within [-2,2]. The range of [-2,2] indicates that the data distribution can be considered relatively normal. (Muzaffar 2016) (George 2010).

Perform Grubb's Test and Levene's Test. Grubb's test was performed to identify the datasets with outliers for each strain in different medium compositions (a representative example is shown in **Figure 3.3**). The tabular summary of the Grubb's Test can be found in **Appendix 3**. Levene's test was performed to assess the homogeneity of the variances. The Levene's test was performed for pairs to check if the homogeneity holds for them. The groups selected for this test were the ones that deviated largely from the normal distribution and observed the outliers. The test results are summarised in **Appendix 3**. The Levene's used is based on the modifications by Brown and Forsythe to Levene's procedure. It considers the distances of the data points from their sample median rather than their sample mean. Using the sample median rather than the sample mean makes the test more robust for smaller samples (M. a. Brown 1974). Levene's test is absolute for any continuous distribution. For skewness heavy distributions as well, this method is more reliable than Bonett's and F – test method. (Schultz 1985)

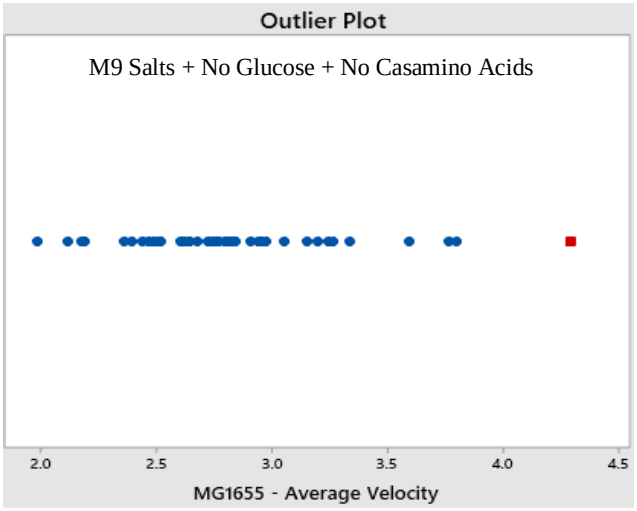


Figure 3. 3: Outlier Identification using the Individual value plot by Grubb's Test. The red dot represents the individual average velocity, which is identified as outliers.

Perform One Way ANOVA. If from the above steps, it is evident that the datasets are normally distributed, free from outliers and have a homogeneous variance, perform the One-way ANOVA test.

However, in the data obtained in the experiments, it has certain groups which contain outliers, and at last, the data collected does not have Homogeneous Variances. These results tend to violate the assumptions for the data to be analysed by One-Way ANOVA Test (refer to the Statistical Methods section in Chapter 2)

Perform Welch's Test and Games Howell Test. The presence of outliers and heterogenous variances render the ANOVA inappropriate, and the F-test also yields unreliable results when the group sizes are not equal (Jan 2014). Although the assumptions of One-Way ANOVA are not completely satisfied by the data sets, an alternative test called Welch's test is a way to analyse such data. The Welch-Satterthwaite equation used in Welch's t-test tackles the shortcomings of the aforementioned tests. Welch's test is capable of dealing with unequal sample sizes, heterogeneous variances (Moser 1992)and it also maintains a control over the Type 1 errors just like the One-Way ANOVA test (Jan 2014).

Welch's test, just like the One-Way ANOVA test, determines if the null hypothesis of equal means can be accepted. Therefore, we just get to know if the means are equal or not. But the real need of the analysis is in determining which pair of means are equal and which ones are not at the same time. We apply the Games–Howell test for this purpose. The Games Howell ranks and groups the data sets based on the equality or inequality of the means. The Games Howell test works well in the case of unequal variances and group sizes (Shingala 2015) (Refer to Statistical Methods Section of Chapter 2). The summary tables from the tests are present in **Appendix 3**. The results are discussed in the next section.

Perform Non-Parametric Test (Kruskal Wallis). To confirm the results obtained via Welch's test were meaningful, I did a non-parametric test called the Kruskal Wallis Test. It was done to confirm the results of Welch's and Games Howell tests and to make sure that the data transformed to treat non-normality did not show erroneous results. Appendix 3 holds the information about the medians for each strain in every medium composition.

3.4 Results and Discussion

3.4.1 Experimental Motility of the two strains

Several independent studies have been carried out at different points in the past to determine the speed with which the wild type strain of *Escherichia coli* moves. In 1972, Berg and Brown conducted a study where they tried to study the differences in the motility of the wild type, uncoordinated mutant strain and a non-chemotactic strain, while swimming up the gradient of serine in a capillary tube. It was found that the wild type moved with an average speed of 14.2 ± 3.4 µm/s up the gradient. (H. a. Berg, Chemotaxis in Escherichia coli analysed by Three-dimensional Tracking 1972).

In 1996, a group studied the velocity of *Escherichia coli* clusters in extremely narrow (3 – 6 µm) capillary tubes and observed that the cluster velocities were observed to be ~18 µm/s. (Mitchell 2006). And in one of the latest studies carried out in 2019 by Johnathan et al. it was seen that the bacteria grown in a liquid medium had a lower motility speed as compared to the ones grown on the solid surfaces.

For *Escherichia coli*, it was observed that the ones grown in liquid medium showed average speeds of ~21 µm/s while the ones grown on the solid surfaces had average velocities as high as ~25 µm/s. This difference is because the signalling protein CheZ has a higher stability in the bacteria grown on solid surfaces, thereby reducing the tumble bias. (Partridge, Escherichia coli Remodels the Chemotaxis Pathway for Swarming 2019),

3.4.1.1 MG1655

The medium compositions with labels B (i.e. M9 salts + 10mM Glucose + No Casamino Acids), C (i.e. M9 salts + 10µM Glucose + No Casamino Acids) and D (M9 salts + No Glucose + No Casamino Acids) differ only in the concentration of glucose 10mM, 10µM and 0M respectively in the absence of Casamino Acids. The strain MG1655 exhibited the mean average velocities of 2.9091 ± 0.3516 µm/s (B), 3.1811 ± 0.3703 µm/s (C) and 2.8056 ± 0.4346 µm/s (D) respectively (Refer **Table A3.8** in **Appendix 3**).

The medium compositions A and E on the other hand, differ only in the presence of glucose and MG1655 exhibited the mean average velocities of 3.0552 ± 0.5559 µm/s and

2.6426 ± 0.2379 µm/s in A (i.e. M9 salts + 10mM Glucose + 0.2% Casamino Acids) and E (i.e. M9 salts + No Glucose + 0.2% Casamino Acids) respectively. Thus, the strain MG1655 showed the highest average velocity in the medium containing 10µM glucose and no casamino acids, followed by medium label A, B, D and E (**Figure 3.4 (A)**).

According to the results **Table A3.10 & A3.11** of Welch's Test in **Appendix 3** and **Figure 3.4 (B)**, the average velocity observed in medium composition A was not significantly different from B and C. The pairs to be paid attention to are the ones belonging to the pairs B-C; C-D; B-D; D-E and A-E (marked with * in **Figure 3.4(B)**). These results suggest that the medium composition of 10 µM Glucose observes the highest motility, and the motility decreases when the concentration is increased to 10mM and decreased to 0M glucose without casamino acids. The comparison of average velocities between pairs E-A and D-C prompt that the motility in the presence is higher than the cases where glucose is absent.

When we consider the pairs A-B and D-E, where we observe the effect of the presence and absence of the casamino acids on the average velocities of the groups when the glucose concentrations are same in both the medium preparations of the pair, no significant effect of Casamino acids on the motility was observed.

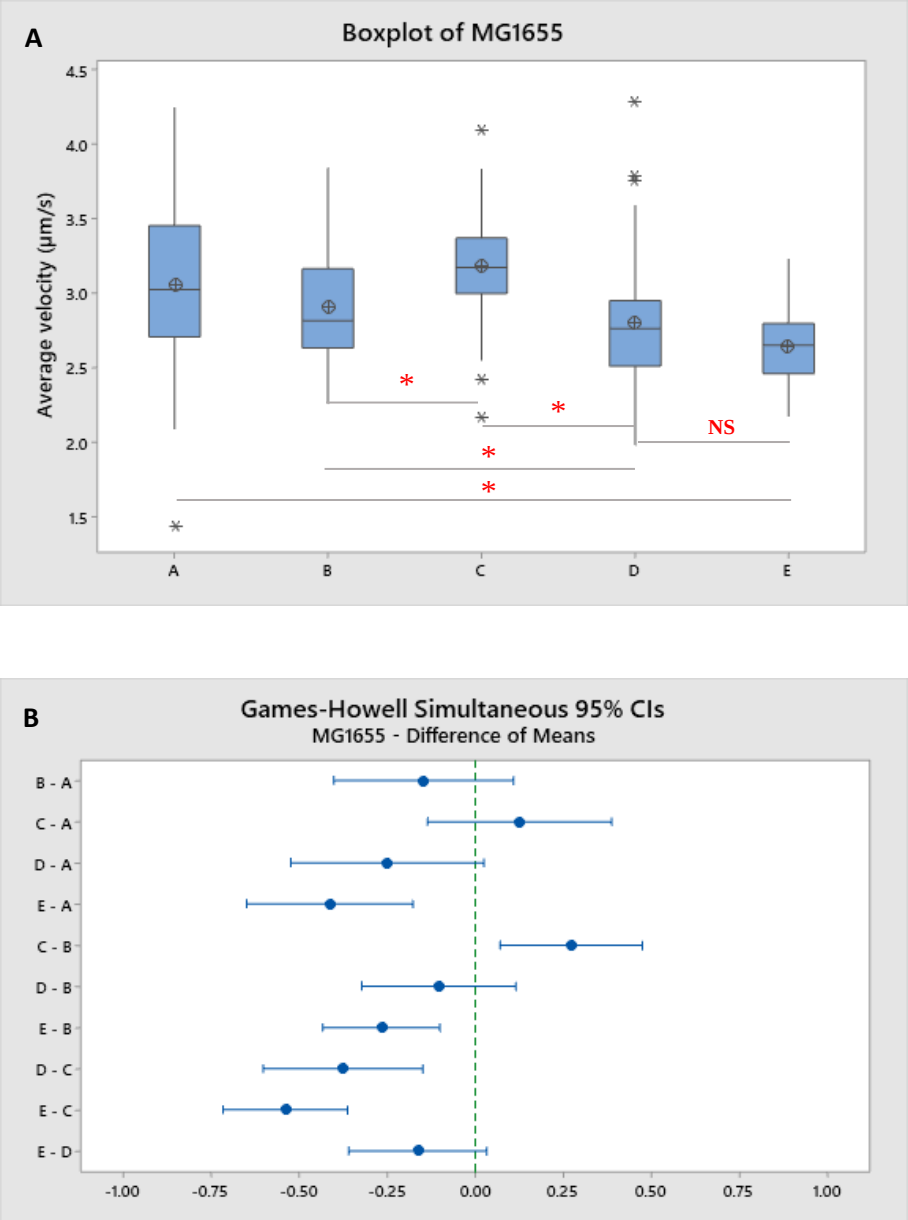


*Figure 3. 4: (A) The Box plot of MG1655's average velocity in all the medium compositions is shown. The * mark denotes the sets which were not significantly different, and NS denotes that the means are not significantly different (B) The Games Howell plot of mean differences. If the pair for the interval contains 0, i.e. if the pair mean difference intervals intersect the dotted green line, then they are not significantly different*

3.4.1.2 MDS42

The medium compositions with labels B (i.e. M9 salts + 10mM Glucose + No Casamino Acids), C (i.e. M9 salts + 10µM Glucose + No Casamino Acids) and D (M9 salts + No Glucose + No Casamino Acids) differ only in the concentration of glucose 10mM, 10µM and 0M respectively in the absence of Casamino Acids. The strain MDS42 exhibited the mean average velocities of 3.059 ± 0.755 µm/s (B), 3.066 ± 0.3664 µm/s (C) and 2.5799 ± 0.3121µm/s (D) respectively (Refer **Table A3.8** in **Appendix 3**). The medium compositions A and E on the other hand, differ only in the presence of glucose and MDS42 exhibited the mean average velocities of 2.8138 ± 0.5729 µm/s and 2.6992 ± 0.2576 µm/s in A (i.e. M9 salts + 10mM Glucose + 0.2% Casamino Acids) and E (i.e. M9 salts + No Glucose + 0.2% Casamino Acids) respectively. Thus, the strain MDS42 showed the highest average velocity in the medium containing 10µM glucose and no casamino acids, followed by medium label B, A, E and D (**Figure 3.5 (A)**).

According to the results **Table A3.12** & **A3.13** of Welch's Test in **Appendix 3** and **Figure 3.5 (B)**, the average velocities observed in medium composition A, B and C were not significantly different from each other. Similarly, the average velocities observed in the medium compositions A, D and E were not significantly different from each other. The pairs to be paid attention to are the ones belonging to the pairs B-C; C-D; B-D; D-E and A-E (marked with * in Figure 3.5(B)). These results suggest that the medium composition of 10 µM glucose (C) observes the highest motility and the motility decreases when the concentration is increased to 10mM (B) and decreased to 0M glucose (D) without casamino acids. The comparison of average velocities between pairs E-D shows no significant difference even though the average velocity observed in medium label E is higher than D, indicating the presence of casamino acids does not affect the average velocity significantly. We observe the effect of the presence and absence of the casamino acids (consider the pairs A-B and D-E) on the average velocities of the groups when the glucose concentrations are same in both the compositions of the pair. The presence of Casamino does not result in a significant difference in the average velocities.

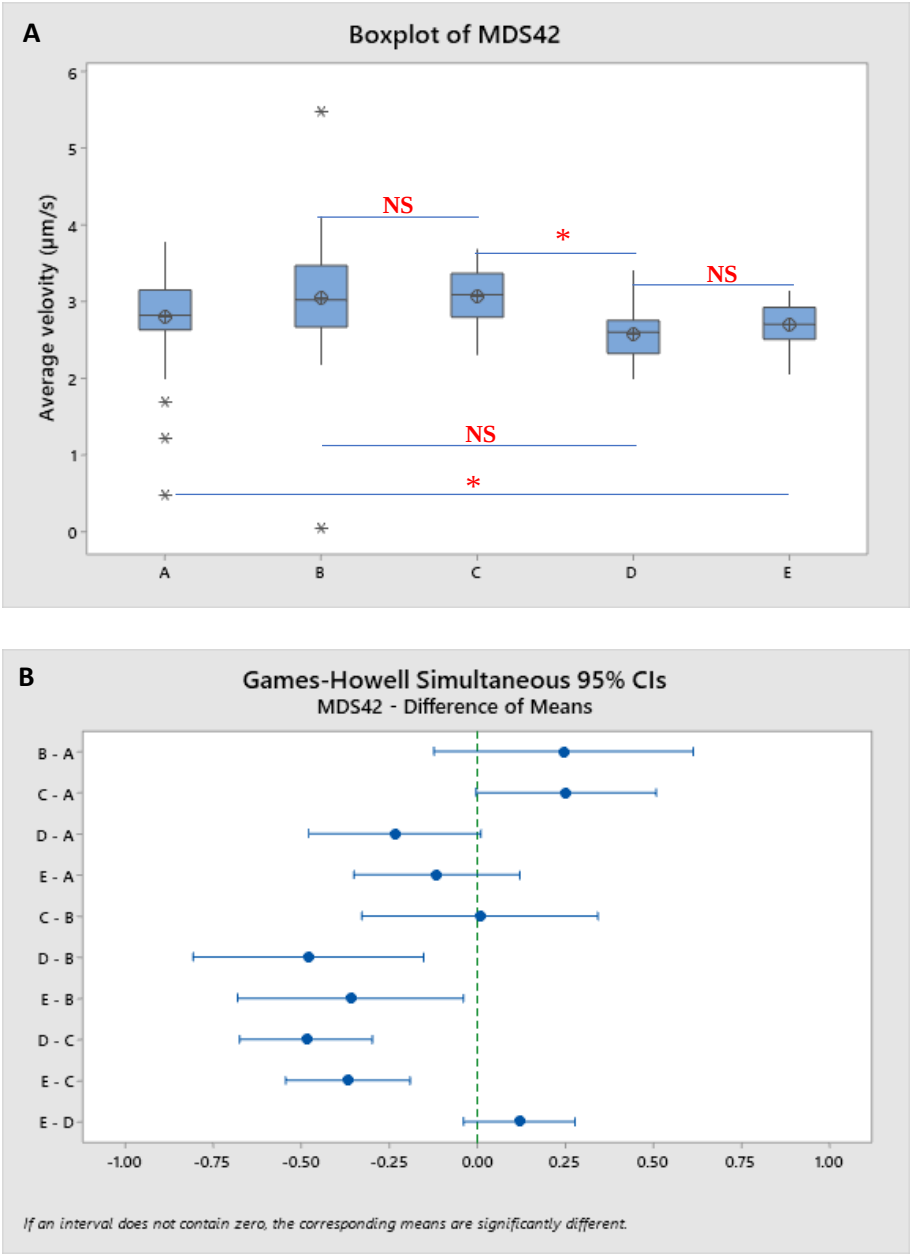


*Figure 3. 5: (A). The Box plot of MDS42's average velocity in all the media compositions. The * mark denotes the sets which were not significantly different, and NS denotes that the means are not significantly different (B) The Games Howell plot of mean differences. If the pair for the interval contains 0 i.e. if the pair mean difference intervals intersect the dotted green line, then they are not significantly different*

3.4.1.3 MG1655 vs MDS42

After comparing the effects of the concentration of glucose and casamino acids on the motility of individual strain, I tried to observe and compare the strain motilities in the same medium compositions. For the medium labels – A, C and D MG1655 showed higher average velocities as compared to MDS42 as opposed to the case of medium labels B and E where MDS42 showed higher average velocities as compared to MG1655. (Refer to **Table 3.4**)

From Welch's test (Refer **Table A3.14** and **A3.15** in **Appendix 3**), it could be seen that the motility sample grown in medium composition there was no significant difference between the average velocities of MG1655 and MDS42, in any of the five medium compositions.

Medium Label	**Strain**	**Average velocity (μm/s)**
A	MG1655	3.0552 ± 0.5559
	MDS42	2.8138 ± 0.5729
B	MG1655	2.9091 ± 0.3516
	MDS42	3.059 ± 0.755
C	MG1655	3.1811 ± 0.3703
	MDS42	3.066 ± 0.3664
D	MG1655	2.8056 ± 0.4346
	MDS42	2.5799 ± 0.3121
E	MG1655	2.6426 ± 0.2379
	MDS42	2.6992 ± 0.2576

Table 3. 4: The average velocity measures for MG1655 and MDS42 strains in different medium compositions

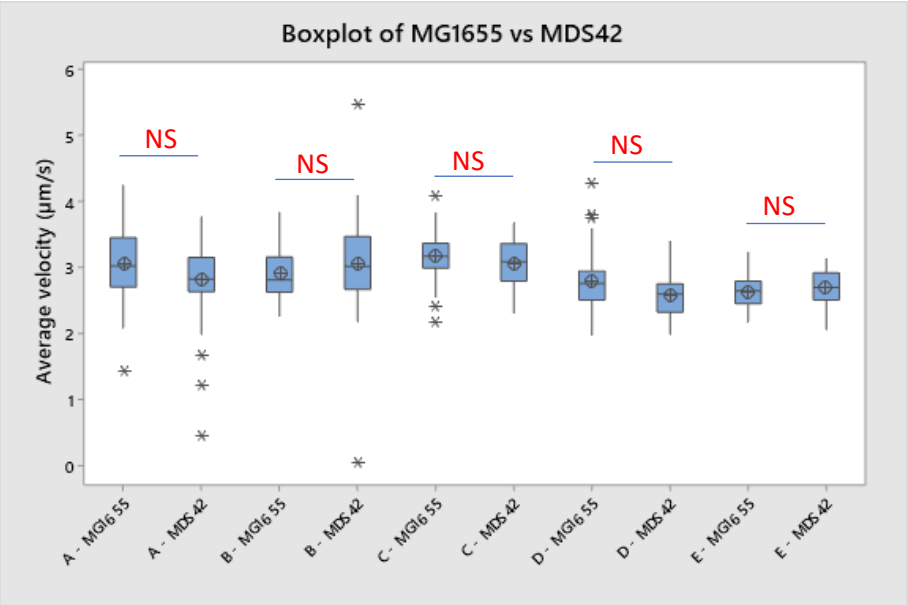


*Figure 3. 6: The Box plot comparison of MG1655 and MDS42 average velocity in all the media compositions. The * mark denotes the sets which were not significantly different, and NS denotes that the means are not significantly different*

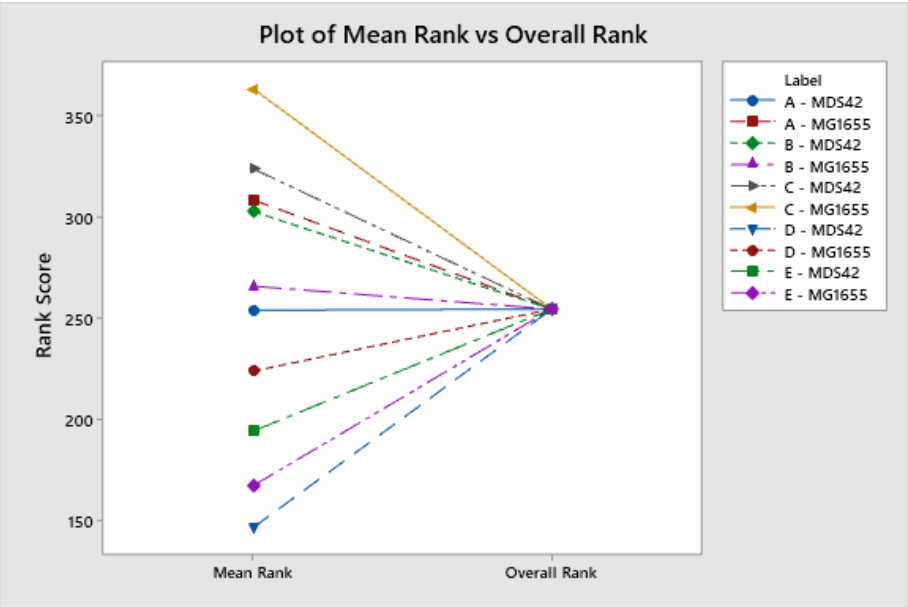


Figure 3. 7: The plot of strain medium ranks vs the overall ranks. Kruskal Wallis is a median based method. The higher Rank score indicates the particular label has a higher number of individual cells with high average velocities. Based on the scoring for each value of average velocity, the medium label of a strain is assigned a Mean rank. Overall rank indicates the overall score obtained by weighing all the individual values of average velocity irrespective of medium composition.

The z-values given against the group labels are a more relaxed way of making sense of how the mean rank of the group compares to the mean rank of the overall observations. They can be interpreted for each of the groups as follows: A greater absolute Z-value suggests that the group ranks further above or below from the overall average rank. A

negative Z- value means the group ranks below the overall mean rank, and a positive value indicates otherwise.

These results are in line with the parametric results obtained after performing the Welch's and Games Howell Tests. The maximum motility is observed in the case of MG1655 in the medium composition containing 10µM glucose and no casamino acids, followed by MDS42 strain in the same medium composition (**Figure 3.7**). The results of the parametric tests were confirmed and were in line with the observations of the non – parametric test.

3.4.2 Difference in the motility associated genes for the two strains

From the motility experiments conducted as a part of this study observed the average individual velocities in the range of 2.9 ± 0.5 µm/s (**Table 3.4**), which is significantly less than what was observed in most of the previously conducted experiments – 14.2 ± 3.4 µm/s (H. a. Berg, Chemotaxis in Escherichia coli analysed by Three-dimensional Tracking 1972), ~18 µm/s (Mitchell 2006) and 21-25 µm/s (Partridge, Escherichia coli Remodels the Chemotaxis Pathway for Swarming 2019) respectively. It goes for both the strains of *Escherichia coli* under study. Let us go through the case of both the strains one by one.

3.4.2.1 MDS42

Upon performing the Protein BLAST as mentioned in Chapter 3 (Refer to **Appendix 2** for details on the support study), it was found that the strain MDS42 lacks the crucial gene FliA, lack of which is responsible for the downregulation of the flagellar biosynthesis. (Karcagi 2016). It is not just the lack of FliA, but the strain MDS42 lacks 33 out of 35 other genes which were checked for in the genome comparison in this study (refer to section Genome Comparison: Methodology, in Chapter 3). The observed lack of matches for the genes responsible for flagellar biosynthesis makes sense because MDS42 was a result of the Multiple Deletion Series project on *Escherichia coli*. The main aim of the project was to remove all the mobile and cryptic virulence elements from the genome. The observed motility could possibly be observed due to the twitching due to the presence of Curli Fimbriae in MDS42. (Nuccio 2007) The evidence of the presence of curli fimbriae in MDS42 can be found in the RefSeq database[1]. (Fredens 2019). This database

[1] RefSeq DB Link:https://www.genome.jp/dbget-bin/www_bget?refseq:NC_020518 | The corresponding genes for Curli Fimbriae are present in MDS42

also confirms the absence of flagellar genes in MDS42 as found after the protein BLAST in Genome comparison section of Chapter 3 (Refer to **Appendix 2**). Therefore, the minimal motility of MDS42 might possibly be due to the presence of appendages called curli fimbriae.

3.4.2.2 MG1655:

Despite having the genes responsible for Flagellar biosynthesis, the strain MG1655 demonstrates the motility comparable to MDS42, which lacks the flagella or has downregulated flagellar biosynthesis. The possible reasons could either be that the flagella are not formed in MG1655 just like MDS42 strain, or MG1655 has an environmental stimulus that slows down the flagellar motor. The environment in which both the strains were kept or grown during the experiments was the M9 minimal medium in the absence of glucose and/or casamino acids or the presence of either or both of them, at pH ~ 7 and temperature ranging between room temperature and 37ºC. Therefore, if we look at the environmental factors one by one, we might find a reason for the low motility.

Glucose

In 1967, Adler conducted a series of experiments, in which he observed that the *Escherichia coli* cultured in glucose in absence, as well as the presence of amino acids, lost their motility and upon further investigation, they found that the bacteria that were grown in the presence of amino acids with no trace of glucose had flagella while the ones grown in the presence of glucose or grown in the presence of only glucose and no amino acids had no flagella, sometimes a short flagellum or in 0-0.1% cases typical flagella. (J. a. Adler, The Effect of Environmental Conditions on the motility of Escherichia coli 1967). In 1993, another set of experiments that were designed to discover the adverse conditions responsible for the loss of flagella in *Escherichia coli* revealed that the presence of the glucose in the culture medium is considered as a favourable condition for the growth of *E. coli* and this causes the catabolite repression of flhD operon. It stops the flagellar biosynthesis. (W. L. Shi, Mechanism of Adverse Conditions Causing Lack of Flagella in Escherichia coli 1993).

The presence of glucose in the absence of nitrogen or poor nitrogen sources tends to inhibit the synthesis of the signalling molecule cAMP, which is responsible for the activation of the global transcription factor CRP. The activation of CRP further affects

the activation of the further cascade and primarily FlhDC, the master operon. It happens because cAMP forms a complex with CAP (cAMP-CAP), which directly interacts with the RNA polymerase and promote the activation of FlhDC. This activation leads to the expression of the further cascade in flagellar biosynthesis, primarily fliA. fliA is the second level of the flagellar biosynthesis cascade and encodes for the flagellar sigma factor, thereby controlling the expression of flagellar biosynthesis genes present on level 3 of the cascade. fliC also known as flagellin structural gene, is present in the third level of the cascade. (O. K.-W. Soutourina 1999) (Bren 2016) (O. a. Soutourina, Regulation cascade of flagellar expression in Gram-negative bacteria 2003).

The glucose thus affects the flagella and, as a result, the motility of *Escherichia coli* in two ways. Firstly, glucose lowers the cAMP levels, which prevent the activation of the global transcription factor CRP. The lack of activation of which furthers the lack of activation of FlhDC, which lies of the level 1 of cascade thereby disabling the expression of genes like fliA and fliC present on level 2 and 3 of the cascades, respectively. It leads to the complete shutdown of the flagellar biosynthesis despite the presence of flagellar genes. Secondly, the lowering of cAMP by glucose leads to the lesser cAMP – CAP complex. The lower concentration of cAMP – CAP leads to less or no interaction with RNA polymerase, failing to promote the activation of FlhDC. The rest of the cascade falls apart as the dominos effect. Therefore, the presence of glucose in the growth medium leading to the loss of flagella causes low motility in Escherichia coli.

Casamino Acids

Previously, there have been studies which advocate that the addition of Casamino acids promote motility in a solid medium (Köhler 2000) (Caiazza 2005). (Kjelleberg 1982) (Bees 2002) (Harshey 1994). However, the presence of Casamino acids in the nutrient-limited medium is known to cause low motility (Samuels, Casamino acids slow motility and stimulate surface growth in an extreme oligotroph 2019). In this study, the presence of casamino acids on motility was studied by observing the colony size of the oligotroph. The cell colony size underwent significant reduction upon the addition of Casamino acids.

Cationic Inhibition

The presence of metal ions is known to cause poor motility on *Escherichia coli* in the absence of a chelating agent in the growth and/or motility medium. This case also applies to using the M9 minimal medium, which contains $CaCl_2$ and $MgSO_4$. The inhibition of

motility starts at higher concentration, observed to be ~10^{-3} M for Ca^{2+} and Mg^{2+} and 10^{-2} M for NH_4^+ (J. a. Adler, The Effect of Environmental Conditions on the motility of Escherichia coli 1967). As for the composition of M9 minimal medium for growth of the bacterial strains, we have the presence of $CaCl_2$, $MgSO_4$ and NH_4Cl salts as primary constituents. Stoichiometric calculations showed that the Ca^{2+} ions were 10^{-4} M, Mg^{2+} ions were 2 X 10^{-3} M, and NH_4^+ ions were 18 X 10^{-6} M. As it can be seen that the Ca^{2+} and NH_4^+ ions are way below the concentration at which they have been known to inhibit motility, while Mg^{2+} ions lie in the order of concentration that is known to inhibit the motility. Therefore, one of the possible reasons for the lower or inhibited motility observed could be the presence of Mg^{2+} ions in the growth medium.

Anionic Inhibition

The presence of Cl^- and SO_4^{2-} also inhibits the motility of the bacteria. SO_4^{2-} is known to cause more inhibition than Cl^-, which is known to cause inhibition at 0.1M concentration. We have both the Cl^- and SO_4^{2-} ions in the growth medium but the concentration of Cl^- is way below the inhibitory concentration. The specific inhibitory concentration for the latter (SO_4^{2-}) is not known. Therefore, it might be one of the possible reasons that cause reduced motility observed in the current study. (J. a. Adler, The Effect of Environmental Conditions on the motility of Escherichia coli 1967)

3.4.3 Effect of concentration

During the study, the motility of the bacterial population was studied in different medium compositions to understand the effect of different concentrations of glucose as well as the effect of absence or presence of Casamino Acids. The reduced motility as a result of the presence of glucose was reflected in the data recorded for the different concentrations of glucose in the medium – 10mM and 10µM, respectively. Let us see how the concentration of glucose and the presence of Casamino acids reflected in the motility readings of the two strains MG1655 and MDS42.

3.4.3.1 Glucose

To understand the effect of the change of glucose concentration, I mainly focused on the Medium compositions that did not have casamino acids and had different concentrations of glucose, namely: 0M, 10µM and 10mM. As it can be seen from the result **Tables A3.10**

– **A3.13** of Games Howell in **Appendix 3** that despite the reduced motility there is a difference in the recorded motility for the strain in the medium compositions with different glucose concentrations (Medium compositions B, C and D). The lowest motility occurs in the medium with 0M glucose, followed by 10mM glucose and then 10µM glucose concentration (refer to **Table 3.4**). The same trend is visible in case of both the strains MG1655 and MDS42.

Although glucose inhibits the motility, contrary to the fact 0M glucose medium composition observed the lowest motility possibly because glucose is the only available Carbon or energy source, therefore its absence causes cells to cease or lower their activity mainly for the reason of conserving energy and survival. On the other hand, 10mM glucose depletes cAMP more compared to 10uM glucose medium, leading to enhanced repression of CRP and FlhDC and further lower possibility of cAMP – CAP complexes. Therefore, the higher motility observed in case of 10uM glucose media can be justified.

The framewise velocity at which the individual bacteria travel between two frames for all the media compositions, were plotted (**Figure 3.8**). From the graphs, it can be seen that although the individuals demonstrate average velocities much lower than the usual as observed by other groups in their experiments, some individuals in specific frames achieve speeds as high as 15 – 30 µm/s.

The variation and the inconsistency in the speed can be attributed to the regulation of the rotation of the flagellar motor. *Escherichia coli* is known to fine-tune its swimming speed by using the protein YcgR. YcgR is also known as the molecular brake, and it is known to regulate the motor speed by binding with the cyclic di – GMP (Boehm, Second messenger-mediated adjustment of bacterial swimming velocity 2010) (Ryjenkov 2006). The condition of low glucose concentration results in slower rates of increase of cyclic di – GMP concentrations,

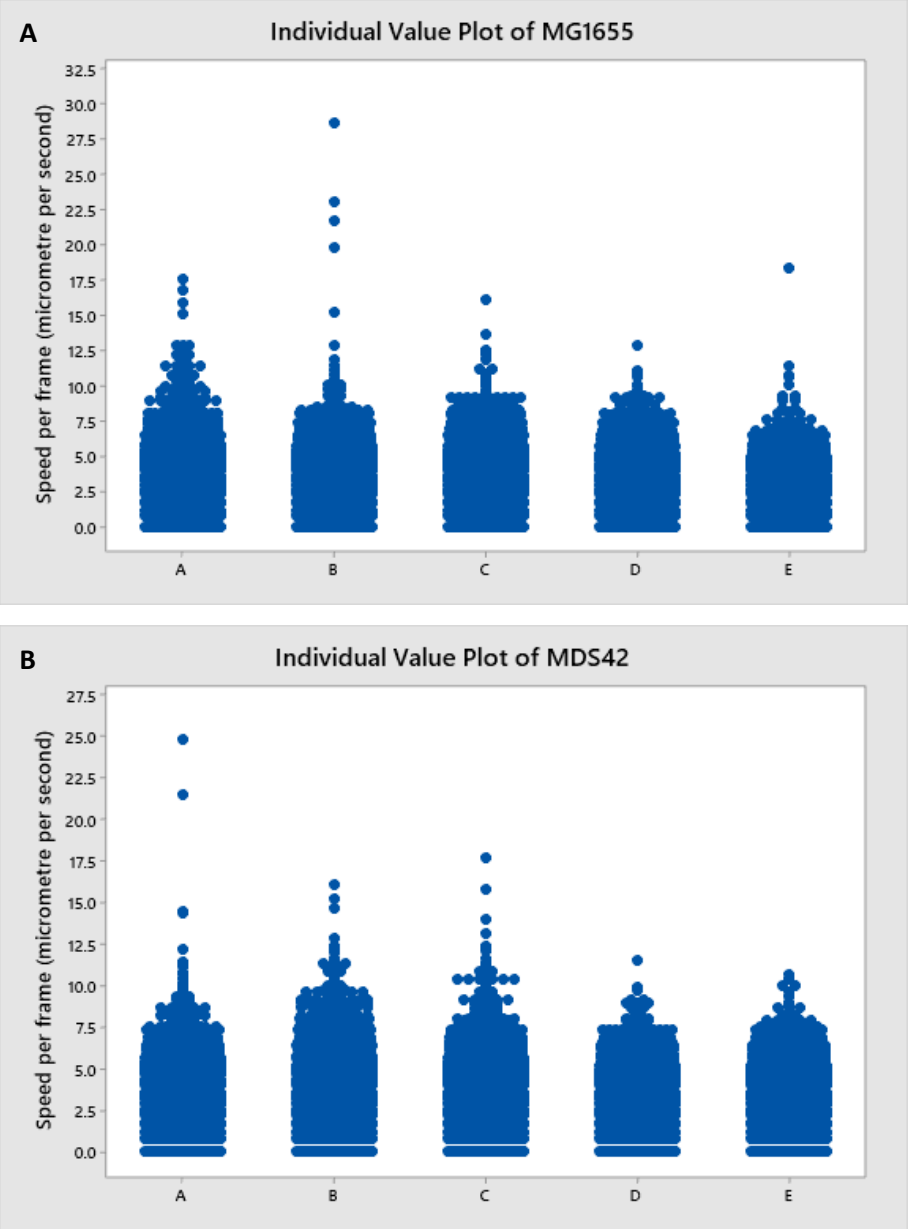


Figure 3. 8: Framewise speed profile of MG1655 (A) and MDS42 (B) in all 5 media composition. Under the label of a medium, each dot corresponds to the individual value of the velocity calculated between two subsequent frames (termed as framewise velocity) for all the bacteria sampled in the medium, in a total of 100 frames. Despite the presence of high velocities (15 – 30 µm/s) as observed in few frames, most of the framewise velocities range from 0.0 – 7.5 µm/s. Each composition has M9 minimal media and varied concentrations of Glucose and Casamino Acids, with A: 10mM glucose and 0.2% Casamino Acids; B: 10mM Glucose and No casamino Acids; C: 10uM glucose and No Casamino Acids; D: No Glucose and No Casamino Acids and E: 0M Glucose and 0.2% Casamino Acids. No significant differences were observed.

whereas the increasing levels of GMP cause reduced motility. Consequently, the YcgR – GMP complex binds to the flagellar motor leading it to slow down. The synthesis of c di – GMP dimer depends on the individual organisms which are dependent on the glucose utilization capability and the presence of adenylate cyclase, which is necessary for the synthesis of GMP. This can be the reason behind the occasional high but overall lower

framewise speed of the individuals in the medium compositions which contain different concentrations of glucose. (Shibuya 1977) (Nieto 2019).

3.4.3.2 Casamino Acids

To understand the effect of the presence of Casamino Acids, I mainly focused on the Medium compositions that did not have casamino acids and had 0.2% Casamino Acids, in the presence of 10mM and 0M Glucose in the medium. As it can be seen from the result **Tables A3.10 – A3.13** of Games Howell Test in **Appendix 3**, that despite the reduced motility there is a difference in the recorded motility for the strain in the medium compositions with different glucose concentrations with or without Casamino Acids (Medium compositions A, B, D and E).

The lowest motility occurs in the medium with 0M glucose with and without casamino acids in case of both the strains MG1655 and MDS42 (refer **Table 3.4**). By observing the average velocity, we can see that in the medium labels A and B (with 10mM glucose in presence and absence of casamino acids respectively) MG1655 has a higher motility in the presence of casamino acids (medium label A). However, the average velocity in medium A and B do not have a significant difference. Whereas, MDS42 shows a higher average velocity in the absence of casamino acids (medium label B) and the average velocity observed in A and B differ significantly.

In the medium labels D and E (with 0M glucose in absence and presence of casamino acids respectively), MG1655 has a higher motility in the absence of casamino acids (medium label D). However, the average velocity in medium D and E do not have a significant difference. Whereas MDS42 shows a higher average velocity in the presence of casamino acids (medium label E), but just like MG1655, the average velocity observed in D and E did not differ significantly in case of MDS42.

Conclusively, it can be said that the presence of casamino acids did not significantly affect the motility of MG1655 in the presence or absence of glucose. Whereas, in the case of MDS42, the casamino acids lower the motility in the presence of glucose but tend to have no significant effect in the absence of glucose.

3.4.4 Quantifying Changes in Direction

As a part of the motility study, the change in the direction of motion of the bacterial cells in the population was also observed as illustrated in **Figure 3.9**. From the position co-ordinates of the bacterial cell in each of the frame, the change in the trajectory angles was calculated. The data of direction change in degrees was recorded for 100 frames for each bacteria in every medium composition. This framewise change in direction was then aggregated and plotted as a histogram for all the bacteria in a media composition in intervals of 10 degrees.

The MATLAB code written to fetch the changes in angles in each frame can be referred to in **Appendix 5 (2)**. The observed angular changes were both positive and negative, when calculated. However, the direction was of lesser importance as compared to the magnitude of angular change. As a result, the histograms which were nearly symmetric about the peak at 0º, are represented in **Figure 3.10** and **3.11**.

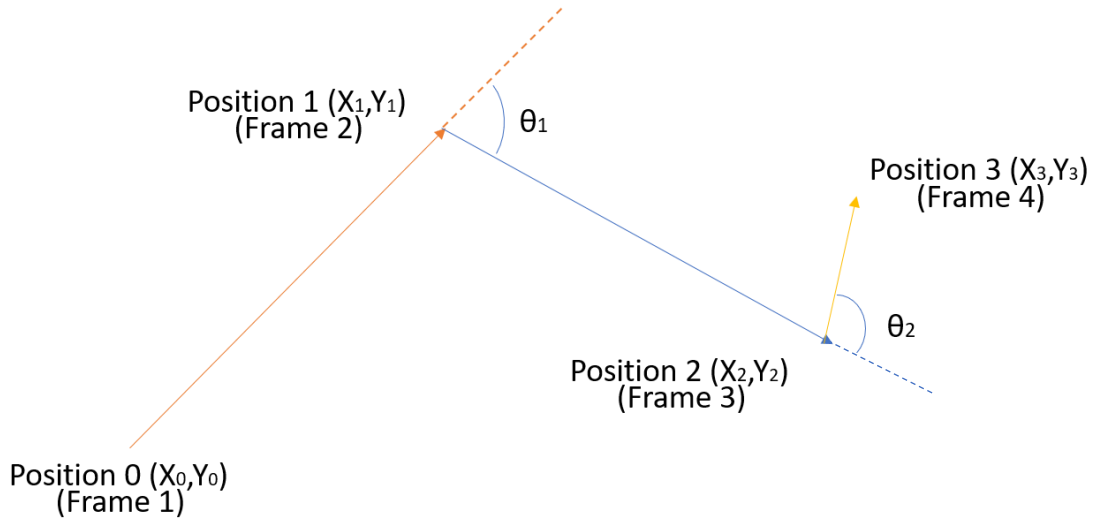


Figure 3. 9:Visual Illustration of how the angular changes in the trajectory of individual bacteria cells. Each frame records the position coordinates of the bacteria, which change with each frame in case of motile bacteria. The angles between the line of motion in each frame is referred to as Angular change or Directional Change in motion. Unit of directional change is Degrees.

For MG1655, in all the medium compositions, 38.325 ± 1.441% of the frames observed no change (0º-10º) in the angle, i.e. a unidirectional trajectory. I had expected more significant angular changes in the trajectory in case of medium compositions with lesser concentrations of Glucose or no casamino acids. However, the samples observed 21.967 ± 1.114% portion of the total frames in which the trajectory changed by 60º-100º. The

angular change data showed a uniform trend across different medium compositions. The graphical representation of the data is given in **Figure: 3.10 (A)** and **(B)**.

It can be seen that the histograms for the different medium compositions show the same shape/trendline. In **Figure 3.10 (B)**, the box plots also show that the sample data has consistent mean and median data with the mean, median and the quartile values slightly higher in case of Medium composition C. The mean angular change was significantly different from the means observed in medium label D and E as shown in **Figure 3.10 (B)**.

Medium Composition	Percent (60^{o}-100^{o} change)
A	22.03019
B	22.5285
C	23.49225
D	20.8207
E	20.96419

Table 3. 5: The percentage frames that observed 60-100 degrees of angular change in the trajectory of MG1655 individuals

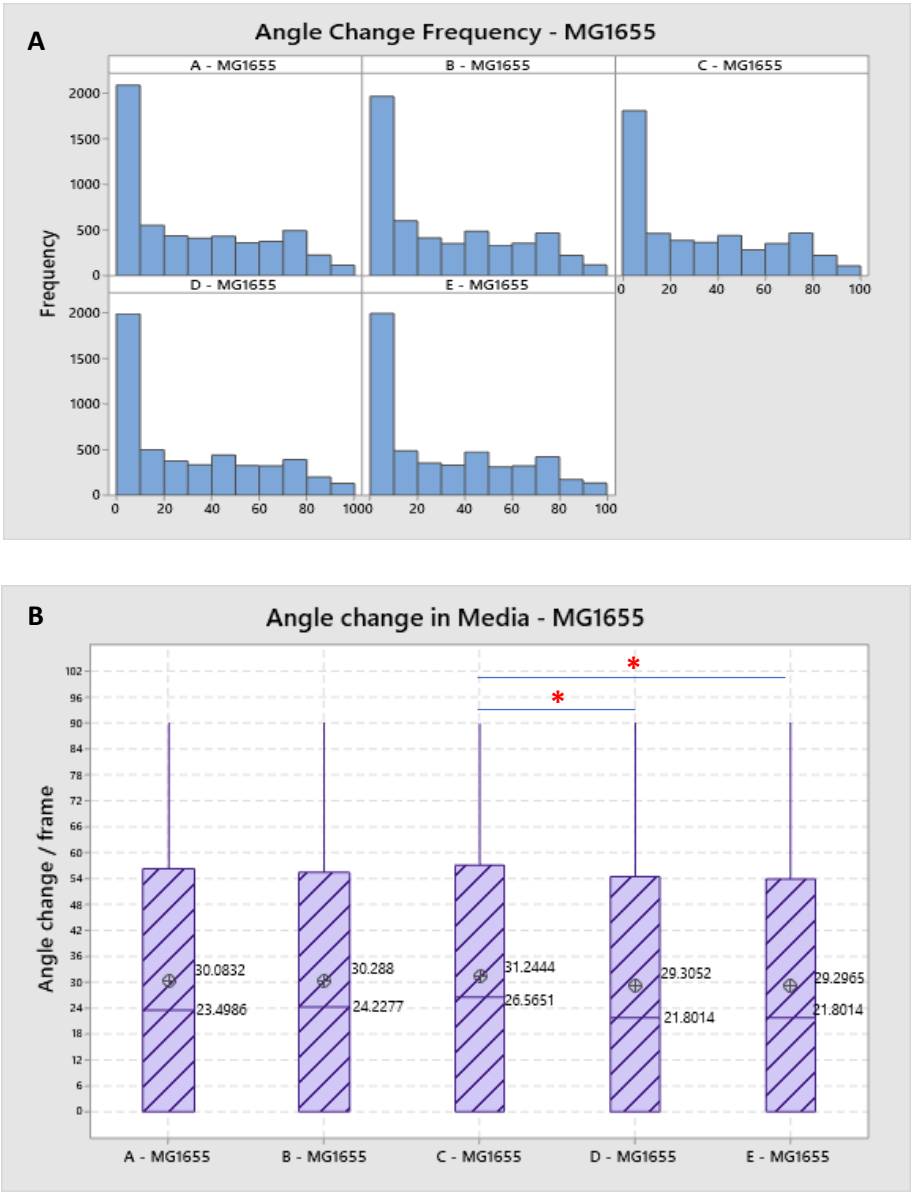


*Figure 3. 10: (A) The data of direction change in degrees was recorded for 100 frames for each bacterium in every medium composition. This framewise change in direction was then aggregated and plotted as a histogram for all the bacteria in a media composition in intervals of 10 degrees.Histogram of modulus of angular change observed in a frame for strain MG1655 in different media compositions. (B) Comparative box plots of angle change per frame in different media compositions for MG1655. The * indicates the pairs are significantly different from each other.*

For MDS42, in all the medium compositions, 38.902 ± 2.457% of the frames observed no change (0º-10º) in the angle, i.e. a unidirectional trajectory. I had expected more significant angular changes in the trajectory in case of medium compositions with lesser concentrations of Glucose or no casamino acids. However, the samples observed 22.012 ± 1.474% portion of the total frames in which the trajectory changed by 60º-100º. It was also a hypothesis that the strain MDS42 will show lesser angular changes since it has lesser functional flagellar motility because of the downregulated flagellar biosynthesis,

but the angular changes observed were at par with that of MG1655. In case of MDS42, no angular changes >90º were observed.

The angular change data showed a uniform trend across different medium compositions. The graphical representation of the data is given in **Figure: 3.11 (A)** and **(B)**. In the **Figure 3.11 (B)**, the box plots show that the mean values slightly higher in case of medium composition C. The mean angular change was significantly different from the means observed in all the other medium labels, i.e. A, B, D and E as shown in **Figure 3.11 (B)**. The angular changes in A & D and B&D were not significantly different.

The trajectories of individual cells further demonstrate the angular changes of cell motion. **Figure 3.12** represents the generalised trajectory, followed by most of the organisms. They tend to stay near their initial positions moving in a run and tumble fashion in a roughly circular trajectory. The trajectories observed for MDS42 are similar to that of MG1655. The unique thing about their trajectories were that their run lengths were longer or equal to the run lengths observed for MG1655.

Medium Composition	Percent (60º-100º change)
A	21.03623
B	22.11502
C	24.15373
D	20.29723
E	22.46

Table 3. 6: The percentage frames that observed 60-100 degrees of angular change in the trajectory of MDS42 individuals

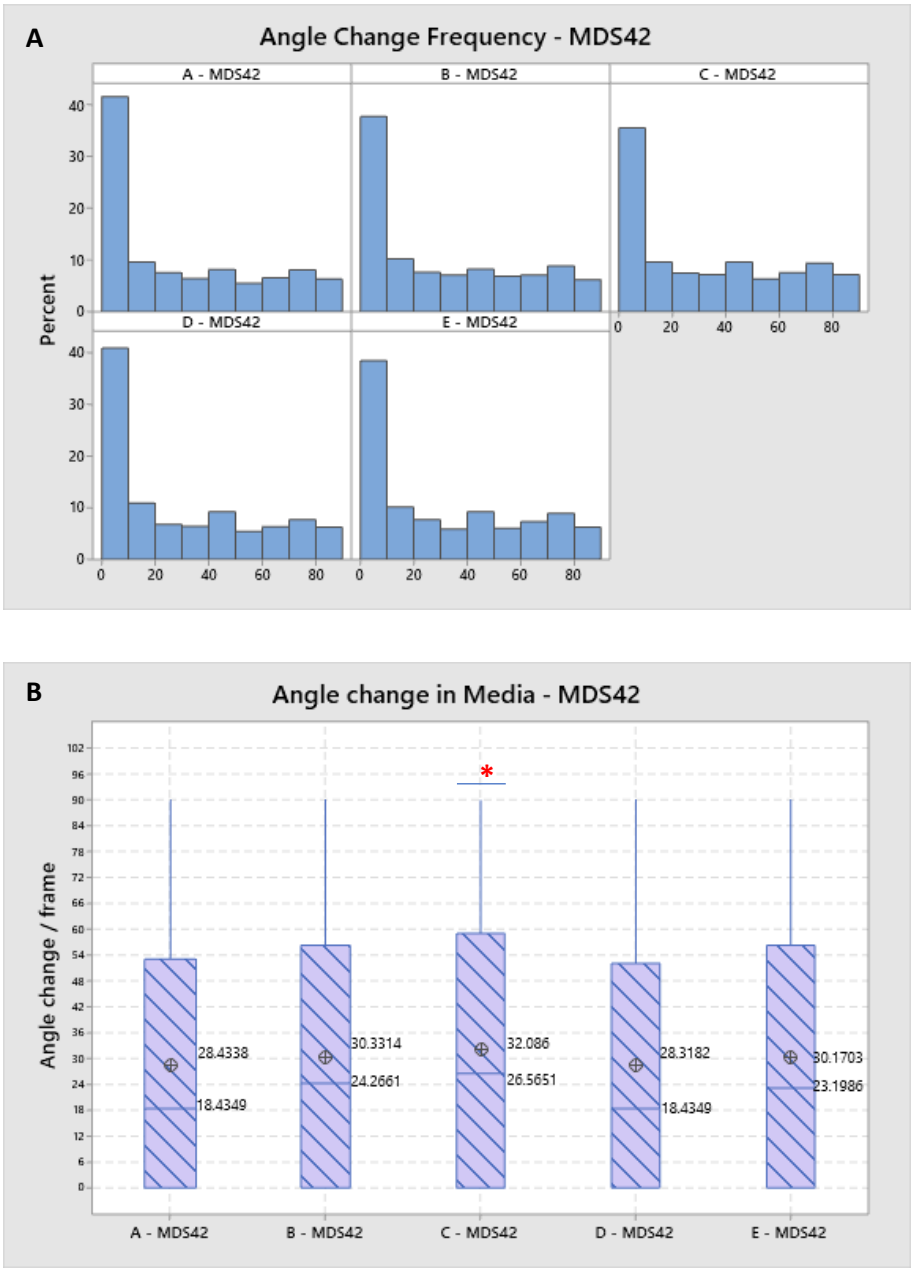


*Figure 3. 11: (A) The data of direction change in degrees was recorded for 100 frames for each bacterium in every medium composition. This framewise change in direction was then aggregated and plotted as a histogram for all the bacteria in a media composition in intervals of 10 degrees. Histogram of modulus of angular change observed in a frame for strain MDS42 in different media compositions. (B) Comparative box plots of angle change per frame in different media compositions for MDS42. The * over C- MDS42 indicates that the angular change per frame observed in this case is significantly different from all the other conditions.*

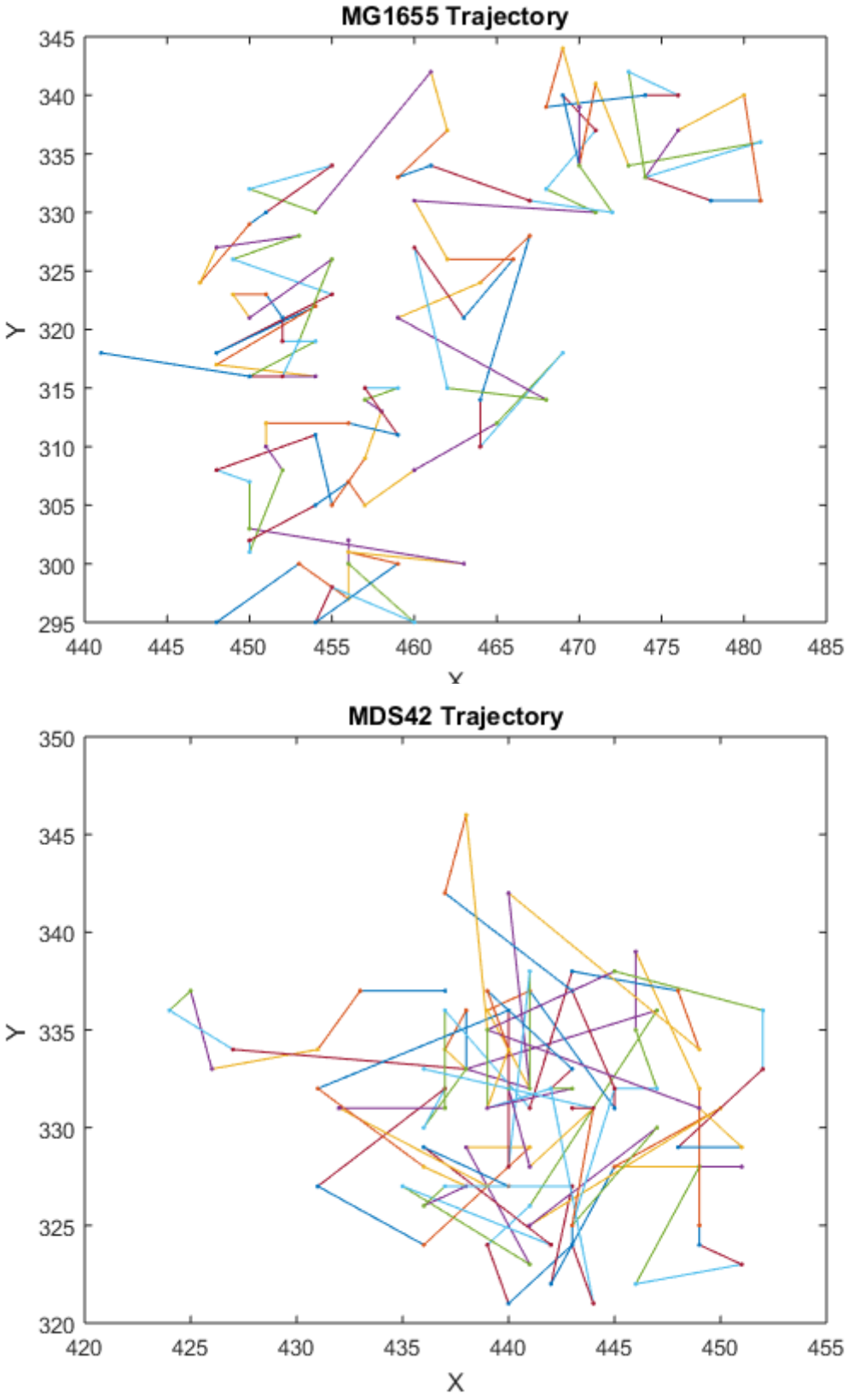


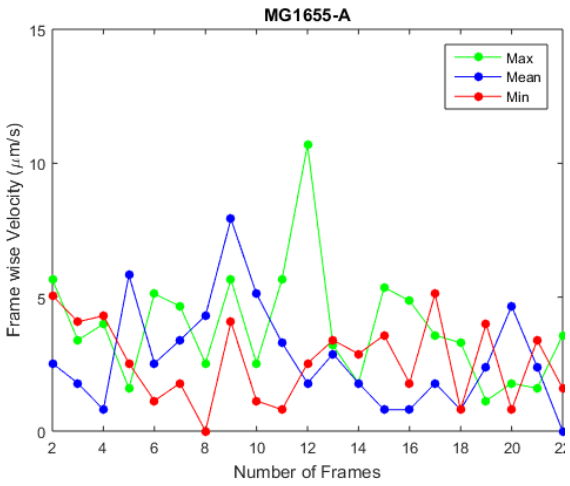
Figure 3. 12: Two Representative trajectories of both the strains - MG1655 and MDS42. X and Y axis are the cartesian coordinated of the cell recorded in each frame. The trajectory presented was recorded in 100 frames. The unit is pixels.

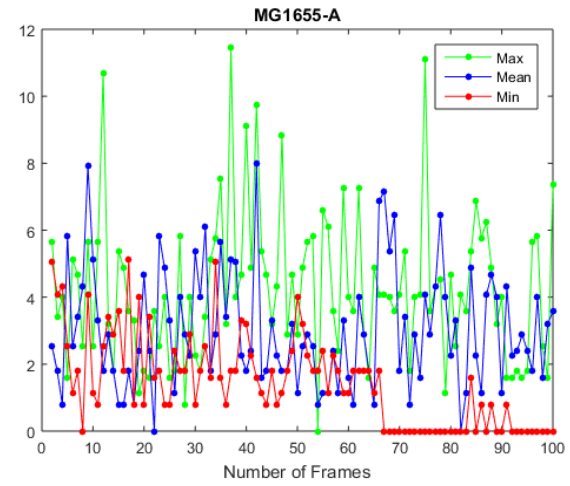
3.4.5 Framewise Velocities of Individuals

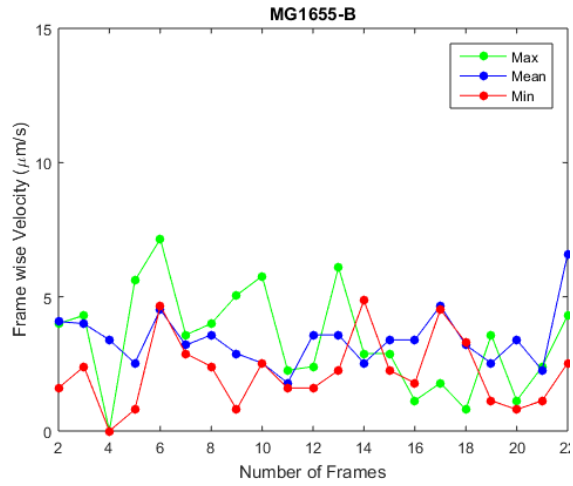
The framewise velocities of the individuals were analysed from the recorded video. Three individuals from each of the medium composition were selected based on the following criteria:
(1) the one with the highest average velocity (labelled as Max, represented by Green), (2) the one with the lowest average velocity (labelled as Min, represented by Red) and (3) the one with the average velocity nearly equal to the mean velocity of the population (labelled as Mean, represented by Blue). This data was difficult to read in 100 frames in some cases, so I represented the snippet in a shorter window of 20 frames.

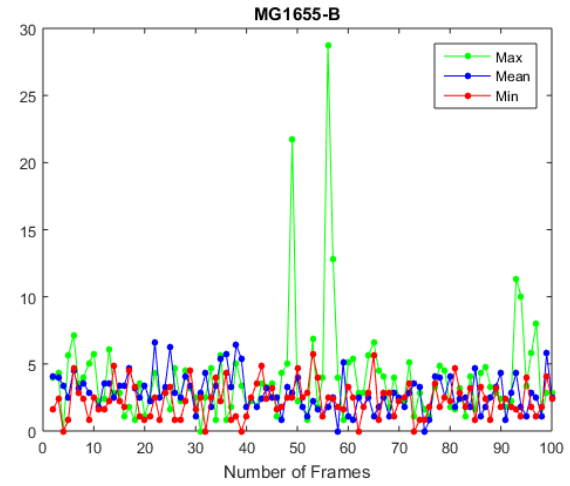
The framewise velocities for MG1655 in all five medium compositions are shown in **Figure 3.13**. It can be seen that there is a general pattern of a peak followed by a downfall in the velocity per frame graphs. It usually shows that the bacterium moves and then slows down, maybe prior to tumbling or changing direction. Another feature that is noticeable in the graphs is that the velocity increases or decreases further into another peak or a downfall respectively but with a different slope. It suggests the presence of acceleration or deceleration depending upon the immediate conditions. For MG1655 (Figure 3.11), it can be seen that the bacterium demonstrating the highest average velocity is not far from the mean velocity of the population. However, upon observing the framewise velocity of the bacterium in different medium compositions, there were velocities as high as 11µm/s (case of A); 29µm/s (case of B); 12µm/s (case of C); 13µm/s (case of D) and 7.5µm/s (case of E). But the bacteria have been unable to demonstrate such velocities consistently.

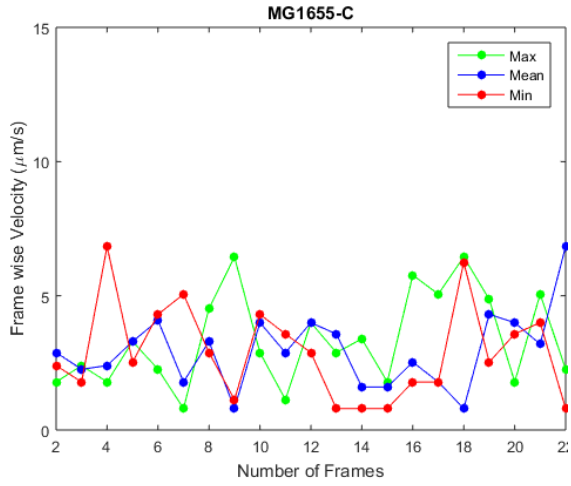
The framewise velocities for MDS42 in all five medium compositions are shown in **Figure 3.14**. It can be seen that there is no visibly evident pattern of a peak followed by a downfall in the velocity per frame graphs. It might be due to the down-regulated genes for flagellar biosynthesis leading to abnormally functioning flagella due to the absence of FliA operon as discussed in the Genome comparison section. Another feature that is noticeable in the graphs is that the velocity increases or decreases further into another peak or a downfall respectively but with a different slope. It suggests the presence of acceleration or deceleration depending upon the immediate conditions just like the MG1655 strain. In MDS42, similar to MG1655 velocities as high as 11µm/s (case of A); 15µm/s (case of B); 18µm/s (case of C); 9µm/s (case of D) and 10µm/s (case of E). However, the bacteria have been unable to demonstrate such velocities consistently, and they have also not shown velocities >~20 – 30µm/s.

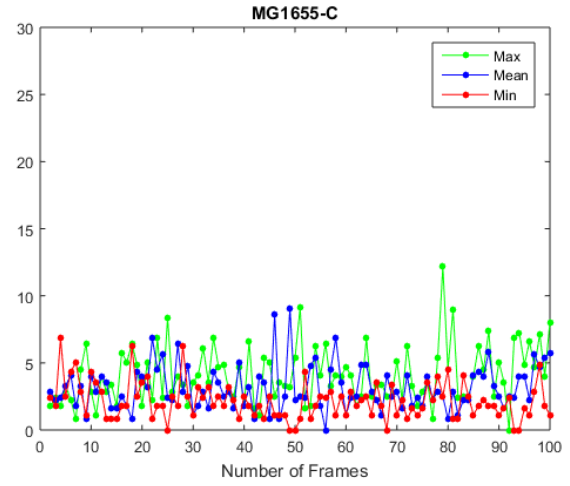

MG1655-A
Max
Mean
Min
Frame wise Velocity (μm/s)
Number of Frames


MG1655-A
Max
Mean
Min
Number of Frames


MG1655-B
Max
Mean
Min
Frame wise Velocity (μm/s)
Number of Frames


MG1655-B
Max
Mean
Min
Number of Frames


MG1655-C
Max
Mean
Min
Frame wise Velocity (μm/s)
Number of Frames


MG1655-C
Max
Mean
Min
Number of Frames

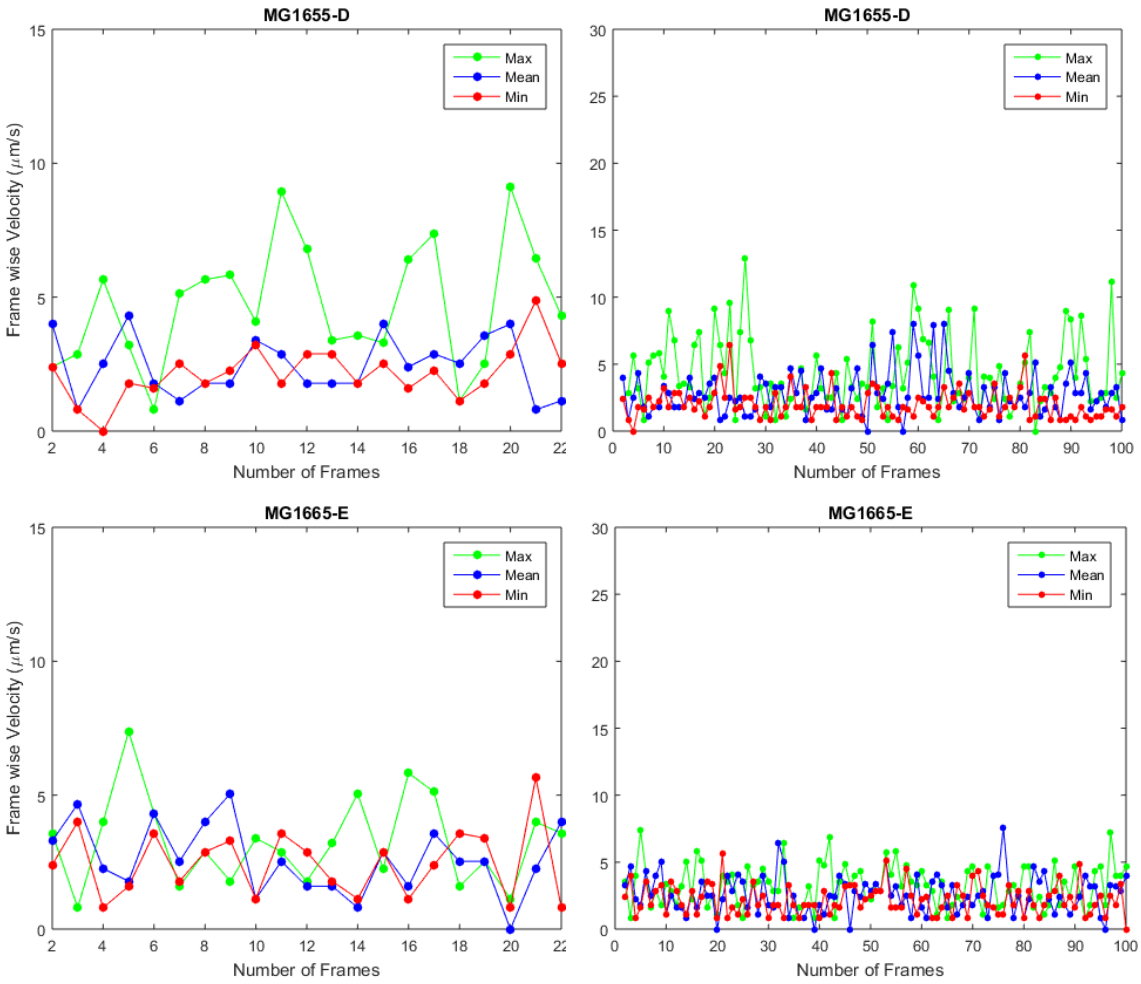
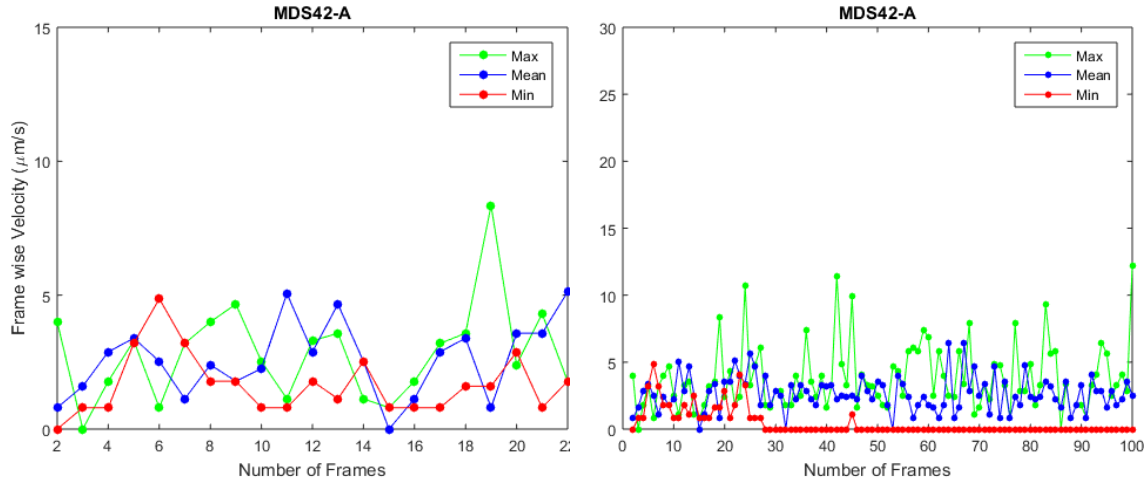


Figure 3. 13: Framewise velocity plots of MG1655 in different media compositions A, B, C, D, & E. The plots marked with same letter correspond respective media composition and depict the framewise velocity plot for 100 frames (right) and 20 frames (left). In each of the media 3 bacteria were selected to represent the framewise velocity profiles. These bacteria were selected based on the average velocity profiles. Red: The bacterium showing the lowest average velocity in the given medium. Blue: The bacterium showing the average velocity close to the mean of average velocities observed for all the bacteria in the given media. Green: The bacterium showing the highest average velocity in the given medium.



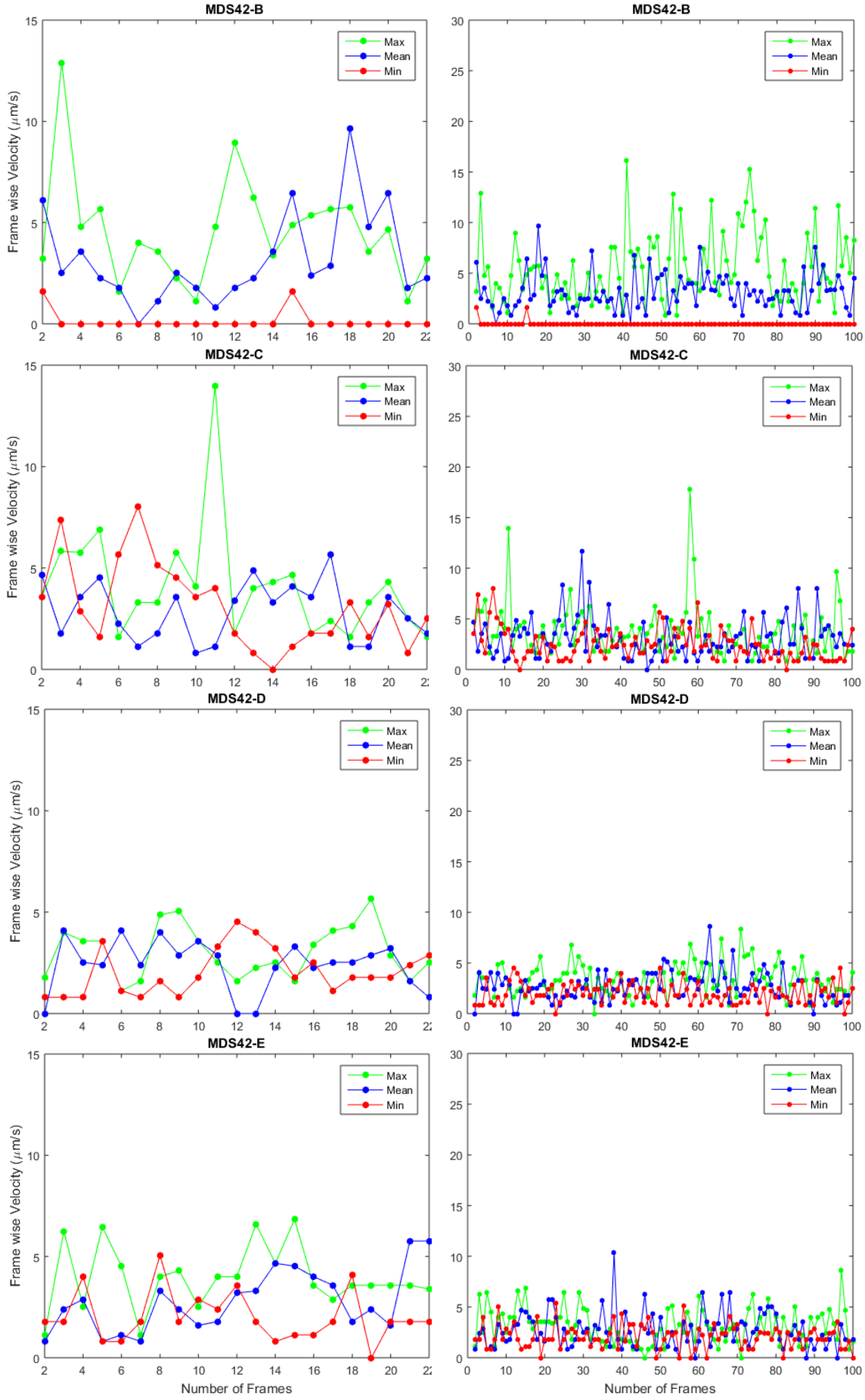


Figure 3. 14: Framewise velocity plots of MDS42 in different media compositions A, B, C, D, & E. The plots marked with same letter correspond respective media composition and depict the framewise velocity plot for 100 frames (right) and 20 frames (left). In each of the media 3 bacteria were selected to represent the framewise velocity profiles. These bacteria were selected based on the average velocity profiles. Red: The bacterium showing the lowest average velocity in the given medium. Blue: The bacterium showing the average velocity close to the mean of average velocities observed for all the bacteria in the given media. Green: The bacterium showing the highest average velocity in the given medium.

3.5 Conclusion

Cell movement and growth are the drivers of cell and tissue behaviour. Their mechanisms and responses to specific conditions are fundamental to understanding the biology of cells, diseases & pharmaceutical industry and its improvement. This led to the idea of understanding the effect of glucose and casamino acids on the motility of Escherichia coli, studied via two different strains MG1655 and MDS42, which vary mainly in their genetic make – up. The project was based upon studying the motility and cellular chemotaxis by creating a simple membrane-based microfluidic device that could be used to observe the real-time cell migration under the influence of linear gradients of varying concentrations of chemoattractant.

For *Escherichia coli*, it was observed that the ones grown in liquid medium showed average speeds of ~21 µm/s while the ones grown on the solid surfaces had average velocities as high as ~25 µm/s. However, from the motility experiments conducted as a part of this study observed the average individual velocities in the range of 2.9 ± 0.5 µm/s, which is significantly less than what was observed in most of the previously conducted experiments for the strain MG1655.

The strain MDS42 lacks the crucial gene FliA, lack of which is responsible for the downregulation of the flagellar biosynthesis. (Karcagi 2016). It is not just the lack of FliA, but the strain MDS42 lacks 33 out of 35 other genes. The absence of flagella might be the possible reason for the lack of motility in MDS42. The strain MG1655, despite having the genes responsible for Flagellar biosynthesis demonstrates the motility comparable to MDS42 possibly die to the catabolite repression by glucose and other cationic and anionic inhibitors in the growth medium like Ca^{2+} and Mg^{2+}.

In case of the motility experiments on MG1655 and MDS42, the medium composition of 10µM Glucose observes the highest motility, and the motility decreases when the concentration is increased to 10mM and decreased to 0M glucose without casamino acids. The motility of MG1655 to was observed to be higher in 10mM glucose medium with 0.2% casamino acids than the medium with no glucose and 0.2% casamino acids, which indicates that glucose is essential for motility but inhibits it at the same time. Similarly, the motility in the absence of casamino acids showed that the cells would be more motile in the presence of glucose than no glucose at all. MDS42 strain, despite the lack of the

flagellar system, showed the same trend. However, in one specific case, it showed more motility in the absence of glucose than the medium with 10mM glucose in the absence of casamino acids, unlike MG1655 which showed the least motility in no glucose medium in the absence of casamino acids.

The framewise velocity of the MG1655 in different medium compositions shot up to velocities as high as 11µm/s (medium A); 29µm/s (medium B); 12µm/s (medium C); 13µm/s (medium D) and 7.5µm/s (medium E). Framewise velocities being 11µm/s (medium A); 15µm/s (medium B); 18µm/s (medium C); 9µm/s (medium D) and 10µm/s (medium E) in case of MDS42. The interesting observation is that although such high velocities exist but not in a sustained manner, the presence of the inhibiting factors might be causing inconsistent performance of motility system of MG1655 and MDS42 irrespective of the fact it is flagella driven or not.

The presence of casamino acids did not significantly affect the motility of MG1655 in the presence or absence of glucose. Whereas, in the case of MDS42, the casamino acids lower the motility in the presence of glucose but tend to have no significant effect in the absence of glucose.

Chapter 4 | Chemotaxis Studies

4.1 Introduction: Experimental methods for analysing chemotaxis

Not just the molecular techniques were used to study the chemotactic pathways of bacteria; there were several experimental techniques employed to study the chemotaxis of bacteria under different environmental conditions. I have tried to list down the critical assays employed to study the chemotaxis in bacteria, since the earliest record.

In 1966, the first-ever record of assay to study chemotaxis was documented. This was the Capillary assay, in which a capillary was filled with the motility buffer (free from chemoattractant), and the bacterial accumulation in the capillary was taken as a measure of the motility. The gradients of the chemoattractant or the chemo – repellent, was developed in the capillary using diffusion, to study the positive and negative chemotaxis, respectively. The capillary containing the motility buffer was immersed in the chemoattractant or chemo – repellent to form their gradients via diffusion. The bacteria accumulating in the capillary were counted under the microscope to understand the movement towards the gradients, and it was also observed that the bacteria did not respond to the high concentration gradients. (J. Adler, Chemotaxis in Bacteria 1966) (J. Adler, Effect of Amino Acids and Oxygen on Chemotaxis in Escherichia coli 1966 (2)) (Julius Adler 1973).

During the same period, in another experimental setup involving a capillary tube and two bottles acting as the reservoirs of bacteria and buffer. The chemo – effectors, L – Serine and L – Histidine were put in the bottles to form the positive/negative gradients in different concentrations. Microscopy imaging was used to observe the average velocities as the measure of the chemotaxis exhibited by the bacterial cells in different concentration temporal gradients. (Macnab 1972)

In 1989, the Ring Forming assay was designed. It is also known as the Swarming Plate assay. In this assay, the bacteria were placed at a point in a container filled with semisolid agar with a chemo – effector gradient. The bacteria use up the local chemo – effector and move along the gradient leading to the expansion of the ring. The expansion of the spot to a ring was used as a quantification of the chemotaxis up the gradient. This procedure had a limitation; it could only be performed for the chemoattractants which were

metabolizable, there was no control on how the gradient forms and the response also depends on how the bacteria metabolise and grow on the medium. (Wolfe 1989)

In the 1990s, there was a significant breakthrough in the ways to study the chemotaxis. The stopped-flow chamber diffusion assay was designed in 1991. It was used to study the migration of bacterial cells in response to the gradients. There were two chambers containing the bacterial populations grown in different chemoattractant concentrations. The chambers had continuous flows which did not allow the bacterial suspensions to interact. However, as soon as the flow stopped, the chemoattractant gradient is formed by diffusion, allowing the bacteria to migrate. The bacterial population density was measured using the light scattering and was used as a quantitative measure of chemotaxis. (Ford 1991)

A group developed a Diffusion gradient chamber assay, in which a central arena was surrounded by the reservoirs separated from the arena by semi-permeable membranes (Widman 1992). The mediums with and without chemo – effectors, were pumped from the source and sink reservoirs respectively to help form the gradient. Once the stable gradients were established the cells were inoculated at the centre of the arena, and their movement up the gradient was studied. The chemoattractant used was Aspartate, and the levels of glucose and oxygen in the system were monitored using microsensors. During the same period, a simple setup of 2 reservoirs joined by a microcapillary was designed. A linear gradient of the chemo – effector, was obtained by filling one of the reservoirs with motility buffer and the other with the chemo – effector. (Z. a. Liu 1996)

All the assays devised till now had a few things in common. All of them consisted of the primary components such as – reservoirs for inlet and/or outlet of the medium & chemoattractant; channel or a capillary, which served as an area to observe the chemotaxis; cell inoculation area and pumps/membranes to help form the gradients. However, these large setups had certain limitations and drawbacks such as they could not be used to develop molecular gradient to study the single-cell characteristics (F. W. Wang 2008), they require large quantities of the reagents thereby increasing the experimental cost (Y. B. Liu 2012); it is harder to control and challenging to generate the stable gradients viable for longer times (VanDersarl 2011), it is incompatible for direct use in various biochemical applications as a lot of devices assembled with materials that might be reactive or interrupt the assay. As a result, the micro-scale microfluidic devices are preferred over them as they provide better control, resolution, stability, ability to build dynamic gradients, usability and applicability in biochemical applications.

4.1.1 Microfluidic Assays

The microfluidic chemotaxis assays mainly differ based on how the microfluidic device develops the chemo – effector gradients. There are several ways to generate the gradient, each having its own advantages and limitations. There are tree-shaped pattern devices that are easy to design but have the disadvantage of possible leakage and blockage. The cells might also suffer shear force in such devices.

The tree designs can be altered to suit the different requirements. The Y – shaped devices have two inlets from where different fluid streams enter a common channel to form a gradient perpendicular to the channel length. They are easy to design but hard to deal when it comes to calculations. The cells are also under the threat of shearing. The membrane-based devices are relatively simple in all the aspects and do not pose the risk of shearing the cells. However, the speed of the gradient generation is slower in this case.

There is another type of devices, which require excellent pressure control. These devices have faster gradient generation speeds compared to the membrane-based devices. (X. L. Wang 2017). The present-day devices are basically based on one or more of the types mentioned above. For instance, Ahmed and Stocker built a device that had the main and side channels perpendicular to each other. The gradient was established by diffusion, and the cells migrated towards the main channel as they were injected through the side channel with the motility buffer at a fixed flow rate. The main channel contained the fluorophore and the chemoattractant. (T. a. Ahmed 2008).

In 2009, Kalinin et al. designed a three parallel channel device. This device had the channels separated by membranes formed by agarose gel. The top channel, also known as the source channel had the chemo – effector along with buffer flowing through it and the bottom channel had motility buffer flowing through it. The central channel, which had a linear gradient developed due to the diffusion from the two parallel flows in the top and bottom channels, was then inoculated with the bacterial cells and their trajectories were studied. (Kalinin 2009)

The current study aims at studying the differences in the chemotaxis behaviour of the MG1655 strain under the linear gradients of the different concentrations of the chemo-attractant – Glucose. In this study, the cells are grown to exponential phase to ensure maximum metabolic activity in different culture medium compositions and studied in a membrane-based microfluidic device that has two reservoirs and a single channel The

cells are then time-lapse recorded in a fixed field of view, and they are tracked one by one using the manual tracking software to ensure accuracy.

4.2 Materials and methods

4.2.1 Making silicon masters

A Mask was formed by transferring a computer-aided design on a silicon or glass substrate. The mask is then used to transfer the pattern onto the required material. L-Edit – the CAD editor was used to create layouts and patterns for lithography. The process of pattern transferring onto the substrate is known as lithography (Bahreyni 2009). Pattern transfer can be done by either using an electron beam or ultraviolet light source. Photolithography is a process in which pattern transfer takes place using UV light. The process of Photolithography was done by Yingkai Liu in Huabing Yin's Lab group.

Before photolithography, the silicone substrate/wafer is first cleaned in opticlear, acetone and propan-2-ol (A. G. Arafat 2007). Each process is applied for 5 mins under sonication (A. S. Arafat 2004). Any moisture on the substrate can affect the photoresist adhesion; therefore; the substrate is dried in the oven at 180^{o}C for 30 min.

After cleaning the silicon substrate, the photoresist is deposited on it. A uniform layer of negative photoresist (SU-8) is applied to the substrate using a spin coater, and the mask is aligned on the substrate. The pattern is then transferred on the substrate by exposure of UV light of proper wavelength. Due to the use of negative photoresist, the areas that are exposed to light will remain on the substrate while the rest will be washed off.

After the exposure, the substrate is etched either by chemical or plasma. The etching is used to remove the substrate from the areas that are not covered with photoresist. At the end photoresist stripper is used to remove the photoresist of the substrate. Photolithography yields the desired pattern mould known as master, which can be used further for soft lithography.

4.2.2 Soft lithography

Soft lithography can be used to obtain PDMS (Polydimethylsiloxane) microfluidic channels. It is an extension of photolithography and is well suited for polymers and gels. Before pouring PDMS to the master, the silicon substrate should be coated with silane.

SU-8 can stick to PDMS and detaching the PDMS from the master will destroy the surface of the substrate.

To perform silane coating, a single drop of silane is placed next to the substrate in a petri dish. The petri dish, along with the substrate, is then kept in a desiccator, which allows the silane to evaporate and then settle and cover the surface of the substrate.

PDMS (Slygard 184) is comprised of two parts - base elastomer and curing agent. At first 10:1 of two parts is mixed thoroughly. The continuous agitation creates air bubbles which are not desirable for microfluidic devices. To remove the air bubbles, PDMS is placed in a moisture-free environment of vacuum desiccator.

The de-gasified PDMS is poured in the silicon master, and a constant thickness of PDMS should be maintained. It is then allowed to cure either for 1-2 hrs at 80^{o}C or overnight in an oven at 75^{o}C. The temperature and time for curing can be changes as per requirements. When PDMS is cured, it is peeled from the surface of the master and bonded to the glass/silicon slab.

4.2.3 Plasma Bonding

The surface of the substrate and PDMS should be modified to enable the bonding between them. Moreover, the hydrophobic nature of PDMS limits its use in various applications (Tan 2010). Therefore, oxygen plasma treatment can be done to both glass slab and PDMS device.

At first, both the PDMS and glass are cleaned with distilled water, acetone and propan-2-ol under sonication for 5 min each and then dried out with nitrogen gun. Both PDMS and substrate are exposed to air plasma for 30-60 sec. Both the PDMS and substrate are bonded to produce a microfluidic device. To ensure a proper seal, the device is pacing in the oven at 60^{o}C for 20-30 min.

4.2.4 3D Printing

For a different set of designs and to avoid inconvenience due to the inaccessibility of the cleanroom, I tried to use 3-D printing to print the moulds for the desired microfluidic devices. 3D printing is a manufacturing process which creates 3D models by depositing layers of material onto one another (Schubert 2014) (Bhushan 2017). 3D printing creates the model of complex objects or biomedical devices with the help of computer-aided designs (Chia 2015). These models comprise of the geometric properties of the object. To

3D print, the CAD file should be converted into STL, which includes the details of each layer (Jasveer 2018). The STL file is then sent to the 3D printer.

Miicraft 125 series printer was used to print these moulds. Miicraft 125 series printer can create objects with dimensions between 150x80 mm 57x32 mm. The full XY resolution of Miicraft 125 series printer ranges from 78µm to 30µm. It can develop structures as thin as 0.005 mm.[2] The printer uses the light source of wavelength 365/405nm to cure resins and to form clear prints.

4.3 Microfluidic Device Designs

I started with one of the simple microfluidic device, developed by Yanqing Song in Huabing Yin's Lab group. The device was a membrane-based, single-channel and 2 – reservoir device, as shown in **Figure 4.1**. The reservoir 1 was meant for the introduction of bacteria into the channel, which connects it to reservoir 2 through a channel of height 10µm. The reservoir 2 facing the end of the channel has pillars. The purpose of the pillars is to stop the PBS buffer in the channel and form a meniscus. The reservoir 2 serves the purpose of setting the agar gel membrane near the pillars at the channel end. This gel allows diffusion of the chemo-effector, into the channel and enables the formation of a linear gradient.

Once the gradient is formed, the bacteria are introduced in the reservoir 1, and their chemotaxis is observed. However, the device posed the following issues – first one being, the agarose gel tends to solidify when delivered through the syringe pump system, prior to reaching the site of gel fixing. When the gel is delivered manually using a syringe to the site of interest, the pressure applied varies with the user. As the slightest change in pressure drives the agarose in the channel, the small dimensions of the device make it difficult to achieve consistent gradients and thus results. Some failures and the fallacies in the gel formation are presented in **Figure 4.1 (A) (B) (C)** and **(D)**.

[2] (https://www.think3d.in/young-optics-launches-miicraft-125-after-the-success-of-miicraft/

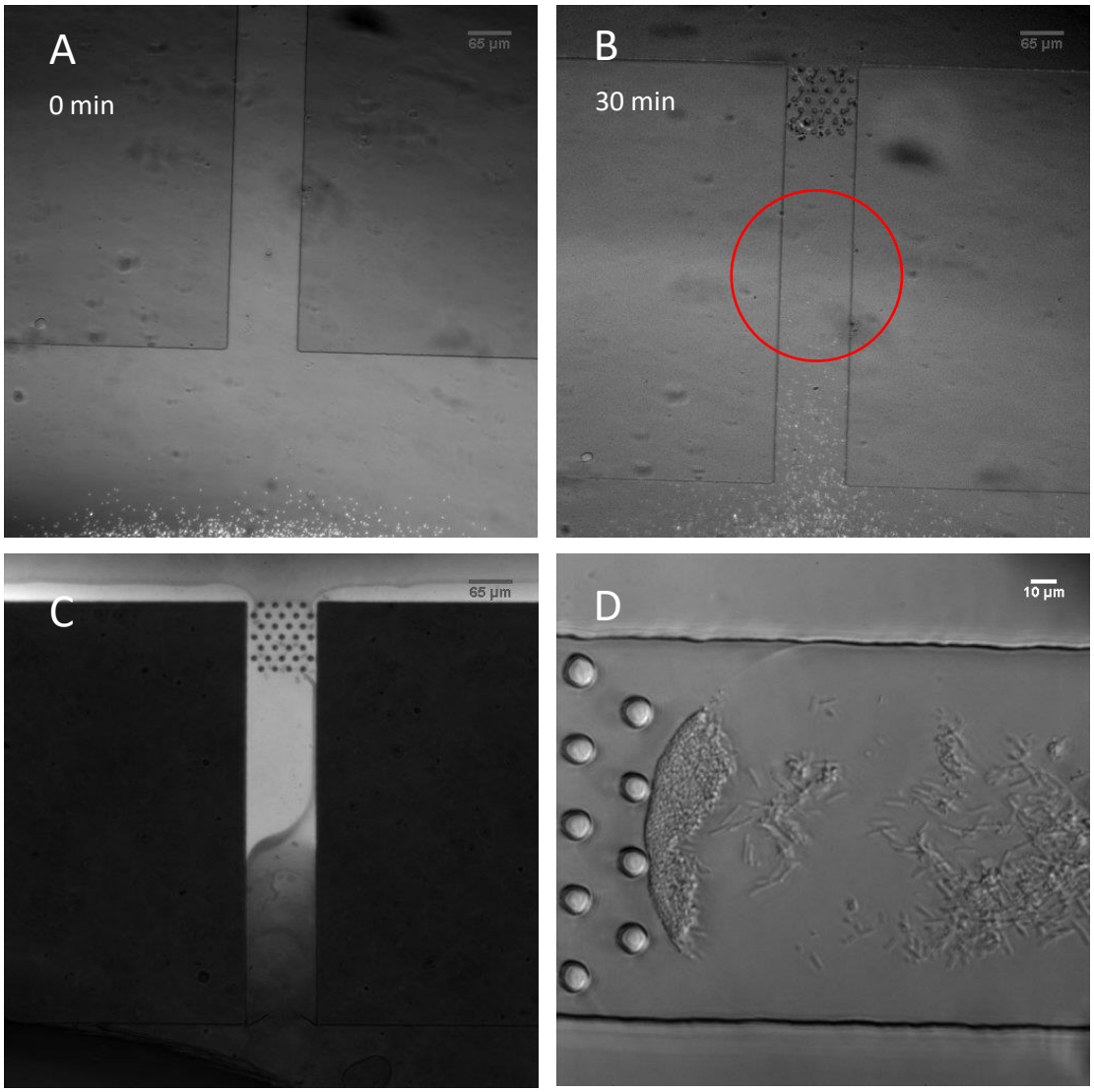


Figure 4. 1: (A) & (B) The images of Device 1 taken at o min and 30 min apart after adding the 1 µm fluorescent microbeads. The no flow condition could not be achieved after fixing the gel and the beads can be seen flowing from the reservoir to the channel. (C) & (D) Demonstrate the inconsistency in fixing the gel due to user variability and manual application of pressure. In (C) the gel has fixed mid – way in the channel and in (D) although the gel seems to have fixed near the pillars, there is a failure to achieve no flow condition. (The respective scales are on the images)

Since the first device could not be used for the experiments, I tried to use 3D printing to make the moulds with a different design. **Figure 4.2 (A)** shows the device design. This device design as well consisted of 2 reservoirs and a single channel. The design improvement that could improve the consistency of the device was that it had 2 more reservoirs connected to each other with a channel higher than the main channel. These 2 reservoirs were meant for the agarose gel. Reservoir 3 was meant for the inlet of the gel into the secondary channel, and the reservoir 4 was meant for the outlet. The main channel was perpendicular to the secondary (agarose) channel. The main channel with reservoir 1 and 2 served the same purpose as the previous device.

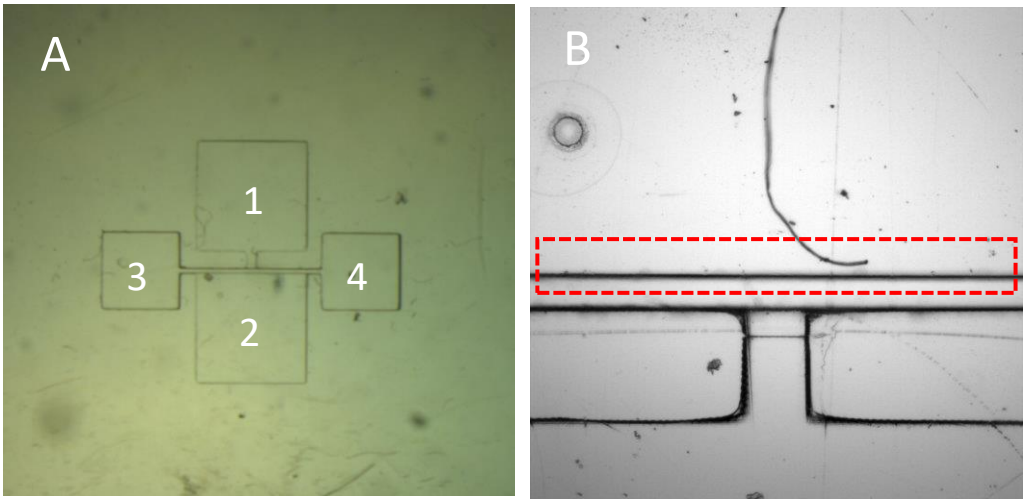


Figure 4. 2: (A)The 3D-printed mould for the chemotaxis device using the MiiCraft printer. (B) Due to the resolution constrains the features smaller than 50µm could not be printed distinctly. The gap between reservoir 2 and the agarose side channel was 50µm, but the printed mould had the walls merged.

The device designs were first prepared using Solidworks sketch and assembly features. The assembly files were then converted to **.stl* extension for the 3D printer. However, the dimension 50µm could not be printed with the Miicrafr 125 series 3D printer, as visible in **Figure 4.2 (B)**. The ink used to 3D print the moulds also poisoned the polymerization of PDMS, as a result of which the resulting devices had the unpolymerized sticky surface, which was in contact with the mould. **Figure 4.3** shows the effect of poisoned polymerization of PDMS on the microfluidic device. As a result, the device was designed by preparing a silicone mould using Photolithography by Yingkai Liu, in Huabing Yin's lab group.

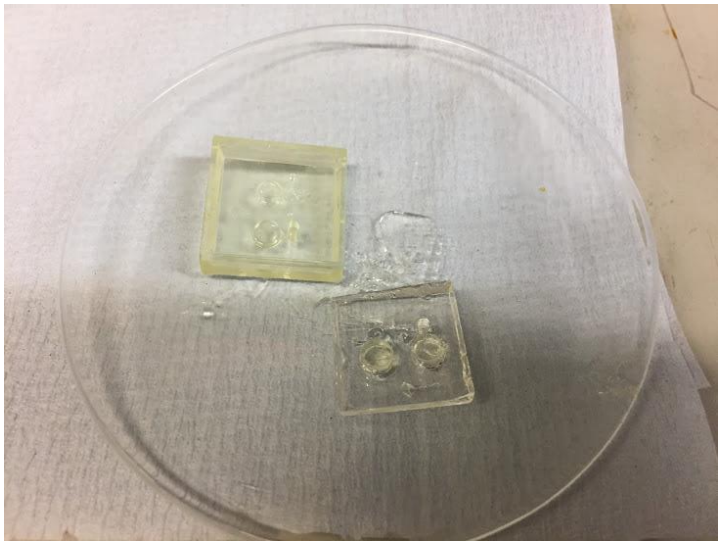


Figure 4. 3: The ink used in the MiiCraft printer poisons the PDMS polymerization near the mould substrate, making it difficult to obtain usable devices with clean surface and distinct features

Two fanned outflank channels were introduced between the reservoir two and pillar end of the main channel, to prevent overflowing of agarose into the main channel (**Figure 4.4 (A) & (B)**. The idea behind introducing the fan – like side channels was to allow the agarose to spread sideways due to lesser pressure on the sides as compared to the main channel. This way, the agarose does not enter the main channel and helps to form a uniform gradient of chemo – effector. This device design was finalised to carry out the Chemotaxis experiment.

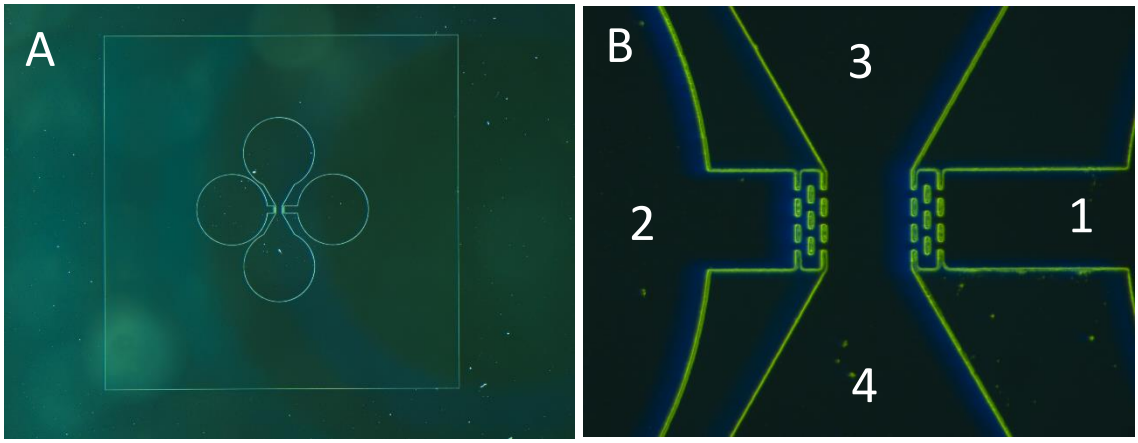


Figure 4. 4: (A) The mask of Device 3 (device used in the chemotaxis experiments). (B) 5X magnification of the mask shows Reservoir 2 on the left (to introduce chemo attractant) connected to a fanned – out agarose channel which is in turn connected to the main channel followed by reservoir 1

4.4 How the device works?

Once the devices prepared using PDMS were bonded to the glass substrate, they were taken through Plasma surface treatment so as to make the channels and reservoirs hydrophilic prior to commencing the experiment. The devices were ashed for 20 seconds at 20% power, if the experiment was to be conducted the same day and the power and time settings of 100% and 50 seconds was used, if the experiment was scheduled for the next day.

The experiment is set up by preparing 2% Agarose. For the experiments, 30mL of 2% agarose was prepared every time by using 0.6g of Agarose in 30mL distilled water. The agar must be kept at 60 – 70ºC in a water – bath. After this, take 0.5µL of PBS / Distilled water / M9 Salts to the Reservoir 1 of the microfluidic device and let it flow to the channel and form a meniscus. Along with the agar, keep a 2mL syringe and 0.8 mm blunt needle soaked in 60 – 70 ºC water – bath. Use the syringe and needle to deliver the 2% agar drop by drop in reservoir 2, until it reached the meniscus formed by the PBS / Water / M9 Salts (in this case) at the pillars at the channel end. Observe under the microscope to ensure the right delivery.

After adding agar to the device, keep the device on the ice – pack to speed up the gel fixing at the pillars. After 10 – 15 minutes of adding the agarose gel membrane to the device, add Chemo – effector of desired concentration.

The bacteria should be prepared for the experiment simultaneously. Grow MG1655 overnight at 150rpm and 37 ºC, in “M9 salts + 10mM Glucose + No Casamino Acids” and “M9 salts + 10µM Glucose + No Casamino Acids.” Culture 100µL of the inoculate from the overnight culture in fresh medium for 2 hours. Centrifuge the culture at 4000 – 5000 rpm for 5 – 10 minutes and resuspend them in the same medium as in the microfluidic device (M9 salts in this case). After 12-15 minutes of adding the chemoattractant, add the bacteria in Reservoir 1. Wait for the bacteria to move from the reservoir to the channel and start recording the data.

4.5 Statistical Analysis: Procedure

The Chemotaxis data was collected for ~50 cells of MG1655 strain, in the two different medium compositions, namely – M9 salts + 10mM Glucose + No Casamino Acids (A) and M9 salts + 10µM Glucose + No Casamino Acids (B). The data of each of the cell

was collected for a minimum of 100 frames or till the time it crossed the field of vision of the microscope. This gave us parameters such as – "Displacement per frame", "Overall Displacement", "Framewise Velocity" and "Average Velocity". The first set of results is based on the Average Velocity measurements of the specimen, studied collectively for the populations in different medium composition for MG1655.

The data was processed using the same statistical procedure, followed in Chapter 3 for the motility data. The Graphical summary of the Chemotaxis data collected in a different medium is presented in the Tabular format in **Appendix 4**. As evident from **Figure 4.8**, the data deviates from the normal distribution. I tried to transform the data to make it follow the normal distribution, but both the Box-Cox transformation and Johnson Transformation could not make the data normal, as shown in **Appendix 4**. The Games Howell Test was used to compare and contrast the data obtained for two gradient conditions.

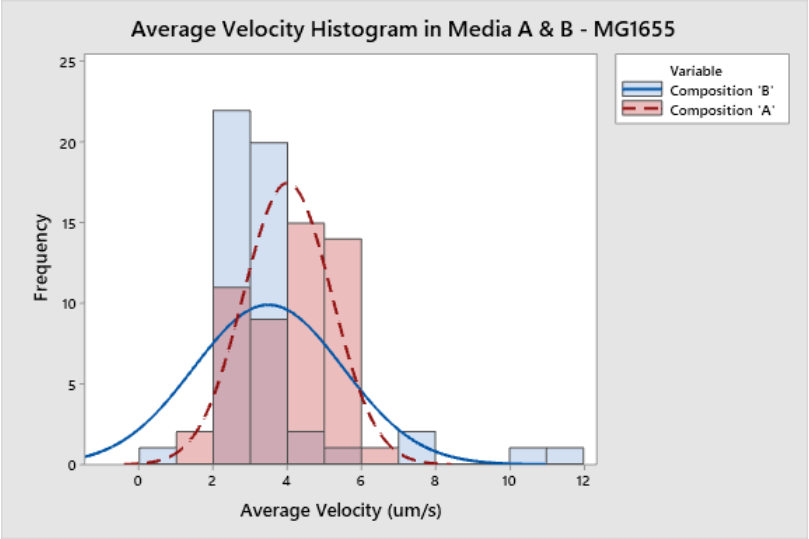


Figure 4. 5:The histograms and the corresponding normal distribution fits of MG1655 Average Velocity data in different media compositions. The data was checked for normal distribution using Anderson Darling Test and the non – fitting data was transformed to see if it can follow normal distribution.

4.6 Results and Discussion

4.6.1 Linear Gradient formation

The time required to form the chemo – effector gradient is decided by the diffusivity of the chemo – effector through the medium in the channel from one reservoir to another. The time characteristic is roughly given by the expression: $T = L^2/(2D)$, where T is the time taken to form the gradient, L is the length of the channel or the distance the chemo – effector must travel to form the gradient and D is the diffusion coefficient of the chemo – effector in the medium (T. S. Ahmed 2010). As per the device design, the attractant must diffuse 400um in Agarose gel and 600um in PBS. The diffusion coefficients of the glucose and fluorescein in Agarose are: 6.46×10^{-10} m^2/s and 2.5×10^{-10} m^2/s respectively, while in the PBS their diffusion coefficients are: 6.78×10^{-10} m^2/s and 2.6×10^{-10} m^2/s respectively. (T. a. Zhang 2005) (Hishikawa 2016)

Upon calculations, the time taken by Fluorescein to form the gradient is 16 – 20 mins theoretically, but during the experiment, it took ~30 minutes to form the gradient. Similarly, the theoretical time for glucose gradient is ~8 mins but, I allowed 15 – 20 mins for the glucose gradient to form during the experiments, as the gradient stays stable once it is formed.

4.6.1.2 Forming the Fluorescein Gradients

To understand the gradient formation in the main channel, I studied the fluorescence intensity of the fluorescein gradient. The gradient ranged from 100µM-0µM. The plots shown in **Figure 4.6** show the Gradient at 30 min after adding the fluorescein left to right. The microfluidic channel is 600µm in length. The gradient was stable for the duration of the experiment.

The fluorescein intensity profile was plotted along three parallel lines along and perpendicular to the channel. The illustrative lines for which intensity profile has been studied are shown in **Figure 4.6 (A)**. In **Figure 4.6: (B)** The x-axis starts from 100 pixels because, from 0-100 pixels, the intensity was constant with slope 0. The linear gradient along the length of the channel is reasonably stable, as is evident from the plots. As can be seen from **Figure 4.6 (C),** the horizontal profile also shows a gradient. The gradient closest to the fluorescein input reservoir is a bit steep as compared to the other two positions. The other two positions show a relatively stable gradient which is not too steep.

The bacteria were seen to move obliquely due to the presence of the horizontal gradient rather than going straight up the vertical gradient (refer **Figure 4.13** for the bacteria trajectory).

From the intensity plots of fluorescein along the channel length and across the channel width, the linear gradients were obtained as recorded in **Table 4.2**. The linear equations have been plotted for the x values corresponding to the beginning and end of the channel length and channel width. For the plots along the channel length $0 \leq x \leq 1500$ (as 2.5 pixels = 1μm) and for the plots along the channel width $0 \leq x \leq 500$ pixels respectively. The corresponding plots are also shown in **Figure A5.1 in Appendix 5.** I used MATLAB to plot these linear equations; the corresponding code can be found in **Appendix 5**.

Line Label	**Equations**	**Slopes of Vertical Gradient**
Right	Y = -0.2002x + 1277.4	-0.2002
Mid	Y = -0.241x + 1307	-0.241
Left	Y = -0.294x + 1333	-0.294
Line Label	**Equations**	**Slopes of Horizontal Gradient**
Line 1	Y = -0.1759x + 1277.9	-0.1759
Line 2	Y = -0.2244x + 1209.1	-0.2244
Line 3	Y = -0.3043x + 1298.6	-0.3043

Table 4. 1: The table of equations of gradient plots along the channel length (vertical) and across the channel width (horizontal)

There is a slight gradient along the width of the channel. This is probably due to the inconsistency caused due to the manual application of the agarose gel membrane, that could lead to the unequal diffusion path for the chemoattractant. The chemotaxis was observed for the two linear gradients 0 – 10mM glucose and 0 – 10μM glucose, defined by the linear functions given in **Table: 4.1**.

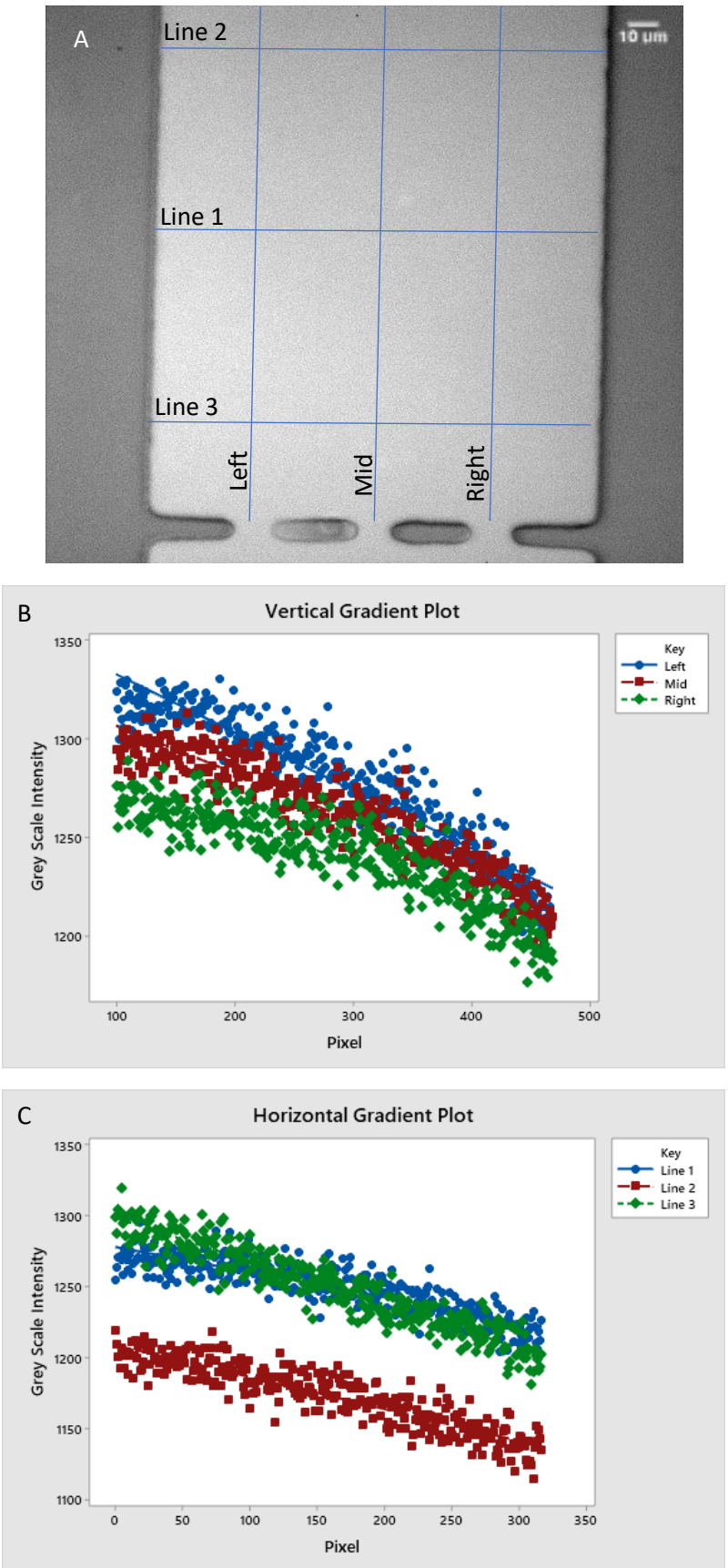


Figure 4. 6: The Gradient plots using fluorescent intensity. (A) The Main channel of device with Fluorescein gradient and illustrative lines along which the intensity profile has been recorded. (B) The Vertical Gradient Plot shows the fluorescein intensity profile along 3 parallel lines along the channel. (C) The Horizontal gradient plots show the fluorescein intensity plots along 3 parallel lines along the width of the channel (perpendicular to the channel)

4.6.2 Chemotaxis Parameters

The distance travelled by the bacteria from one frame to other when added for all the 50 frames is called the Total distance, while the shortest distance between the initial (position in the Frame 1) and the final position (position in frame 50) is called the Total Displacement (**Figure 4.7**). The data recorded for distance travelled, and displacement from the initial position in the first 50 frames can be utilised to see the organisms which actively followed the gradient or were more sensitive to the gradient. The closer are the values of displacement and distance will be the straighter will be the path followed by the bacteria. Therefore, the individuals who have a displacement to distance ratio closer to 1 will be the individuals of interest.

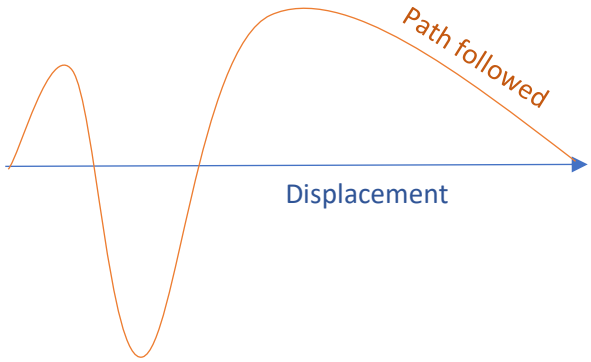


Figure 4. 7: Illustration to demonstrate the concept of displacement to distance travelled. The "Path Followed" refers to the Total distance.

In order to make sense of the distance and displacement data obtained for each of the individuals, the ratio of displacement to distance travelled in the first 50 frames by the cells was calculated under each gradient condition. It is a fact that the ratio cannot be more than "1" and the ratio of "1" indicated an exact straight-line path of the individual. The displacement to distance ratio of 0.7 and above was considered to approximate a nearly straight-line motion, t. In the individual value plot shown in **Figure 4.8**, it can be observed that the number of bacteria showing the displacement to distance ratio of 0.7 and higher are more significant in the gradient of 10 mM glucose as compared to the 10µM glucose. This suggests that the gradient established using the microfluidic device for 10mM glucose was sufficiently detected by the individuals resulting in ~67% of them following a nearly straight line trajectory along the length of the channel.

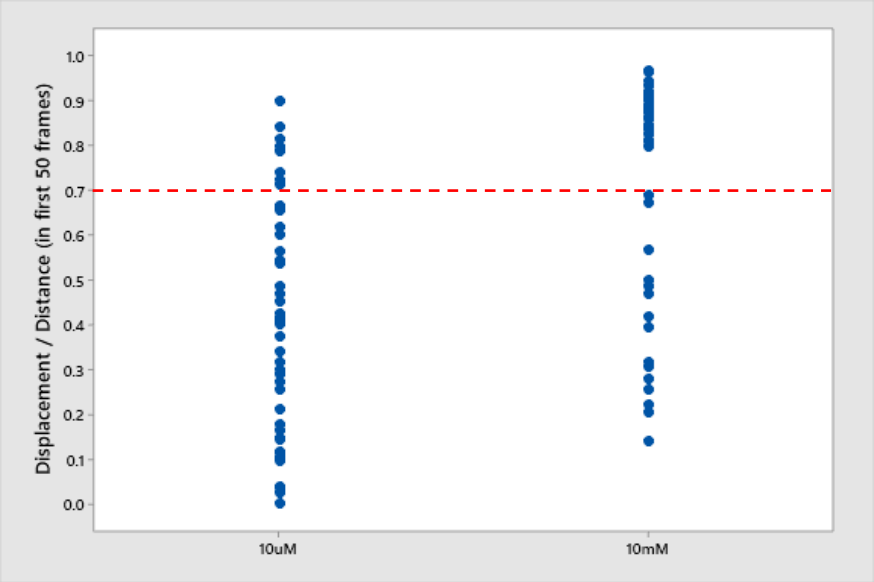


Figure 4. 8: Individual value plots of the ratio of displacement to distance to observe individuals that followed nearly straight-line path under the gradients of 10mM and 10µM glucose.

4.6.3 Total Distance

The average distance travelled in 10µM gradient (110 ± 90.1µm) is greater than the distance travelled in 10mM gradient (105.85 ± 42.03 µm), the values are not significantly different. (**Figure 4.9(C)**)

While looking at the total distance traversed by the individuals of MG1655 strain in the 10µM glucose gradient, ~ 68% of the sample individuals travel 25-100 µm in the first 50 frames. However, there were 32% of the sample population that were observed to move distances as 100-500µm, but only 25% of that 32 % could travel more than 200µm. On the other hand, for the chemotaxis experiment with 10mM gradient of glucose, only ~ 40% of the sampled individuals were recorded to cover a distance of 60 – 100 µm in the first 50 frames. Out of the remaining individuals who covered distances greater than 100 µm, only one individual could travel more than 200 µm. Upon comparing the average distance travelled by the sample individuals in both the gradients, the average turned out to be higher for the 10µM gradient. This was possibly due to some individuals that travelled more than 200 µm in 50 frames; otherwise, 10µM gradient experiment recorded more individuals than the other group that travelled more than 100µm. (refer to **Figure 4.9**). However, the distance travelled by the cells under the two concentration gradients was not significantly different.

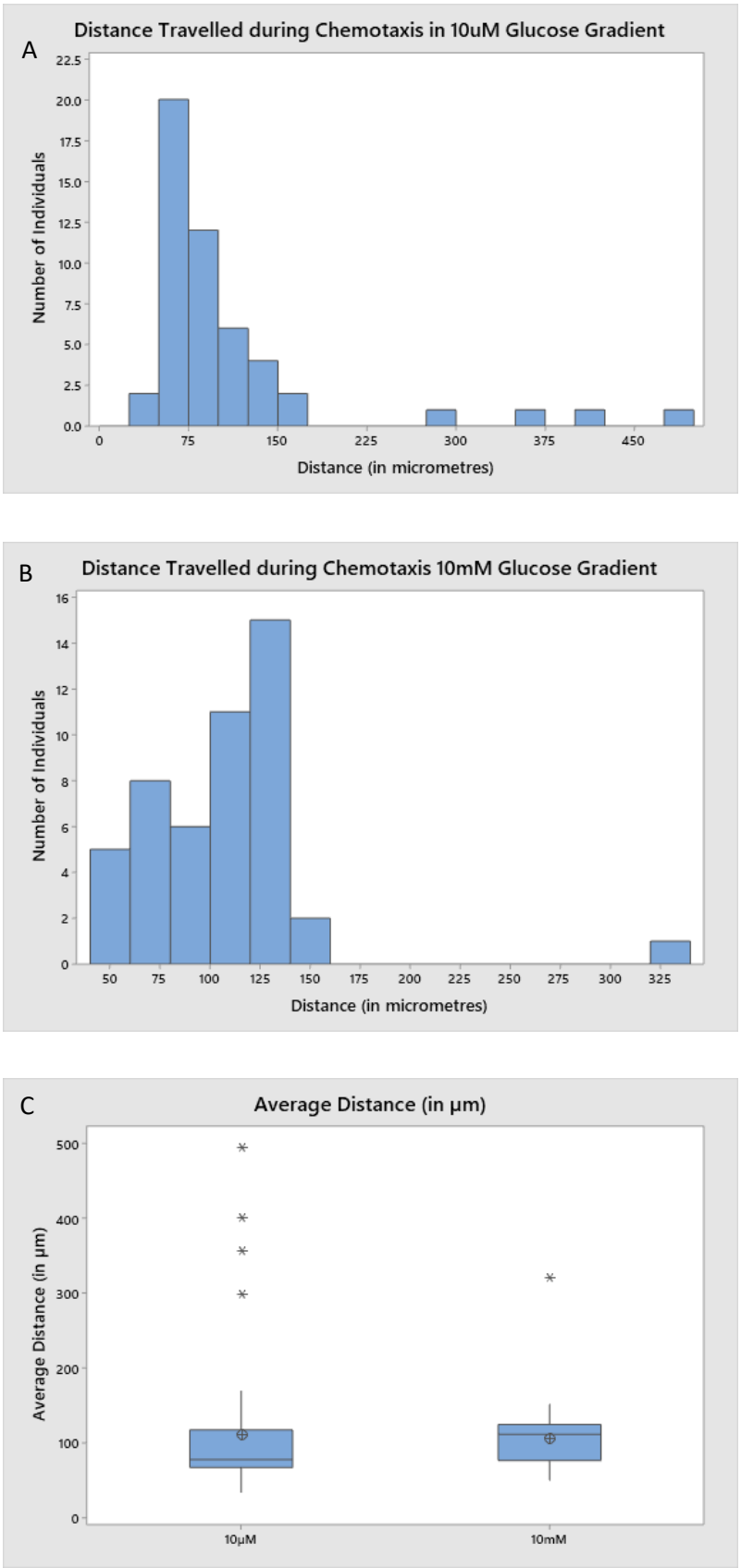


*Figure 4. 9: (A) & (B) show histograms of the distances travelled by individuals of MG1655 strain in 10uM and 10mM gradient respectively. (C) compares the average distance travelled in 10µM gradient to 10mM gradient, although the average distance travelled in 10µM gradient (110 ± 90.1µm) is greater than the distance travelled in 10mM gradient (105.85 ± 42.03 µm), the values are not significantly different. The * here represent the outlier values.*

4.6.4 Total Displacement

The average displacement from the initial position in 10µM gradient (37.45 ± 23.78 µm) is lesser than the displacement in 10mM gradient (80.29 ± 38.57 µm), the values are significantly different in 10µM gradient from 10mM gradient (**Figure 4.10 (C)).**

While looking at the total displacement of the individuals of MG1655 strain in the 10µM glucose gradient, it can be seen that on average, the individuals show a displacement ~38um in the first 50 frames. However, there were nearly 36% of the individuals that were observed to have displaced from 40 – 110 µm during these 50 frames of observation. There were five individuals that demonstrated almost double the average displacement observed in 10 µm gradient of glucose.

On the other hand, for the chemotaxis experiment with 10mM gradient of glucose, it was seen that on average the individuals show a displacement ~80 µm in the first 50 frames, which is more than double of the average displacement observed in the case of 10µM gradient. In fact, there were nearly 63% of the individuals in the sample that were observed to have displaced from 80-150 µm during these 50 frames of observation. However, just looking at displacement does not make sense unless we consider the displacement respective to the length and width of the channel. (refer **Figure 4.10**)

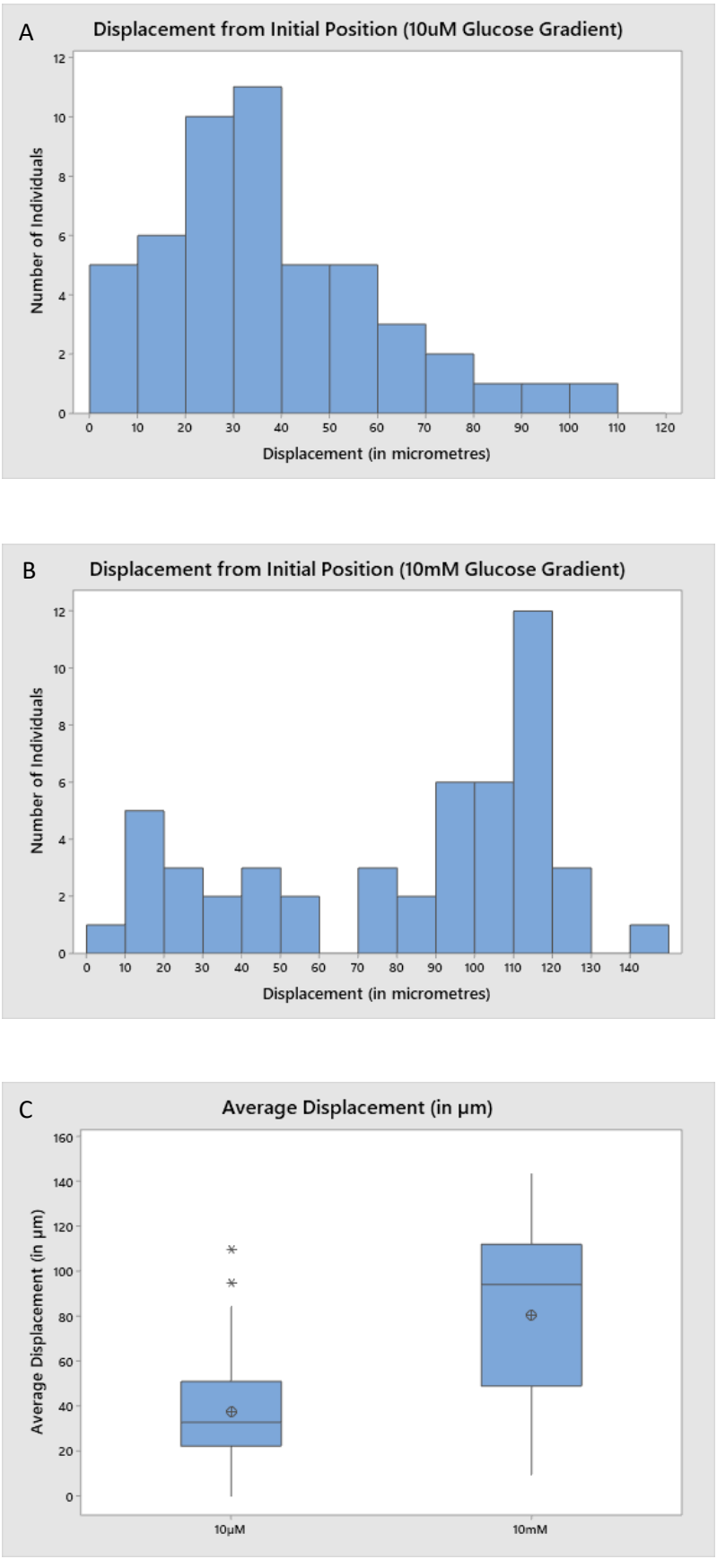


*Figure 4. 10: (A) & (B) show histograms of the displacements of the individuals of MG1655 strain in 10uM and 10mM gradient respectively. (C) The average displacement in 10µM gradient (37.45 ± 23.78 µm) is lesser than the displacement in 10mM gradient (80.29 ± 38.57 µm), the values are significantly different.in 10uM gradient from 10mM gradient. The * here represent the outlier values.*

4.6.5 Lateral and Axial Displacement

4.6.5.1 Lateral Direction

While looking at the total displacement of the individuals of MG1655 strain in the 10µM glucose gradient, we will first consider their displacement in the lateral direction, i.e. along the width of the channel for the sake of simplicity.

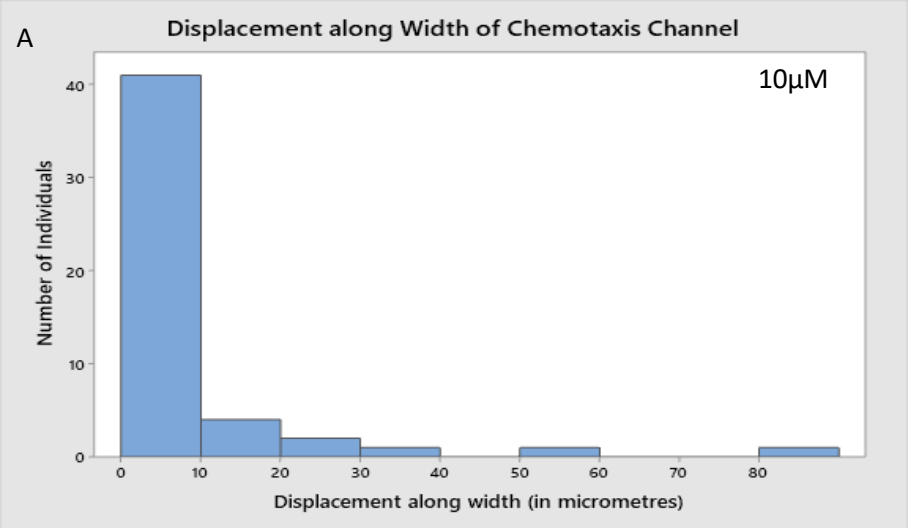
The average displacement along the channel width in the 10µM gradient (9.37 ± 14.96 µm) is lesser than the displacement in 10mM gradient (10.44 ± 6.52 µm), but the values are not significantly different.in 10uM gradient from 10mM gradient (**Figure 4.11(C)**)

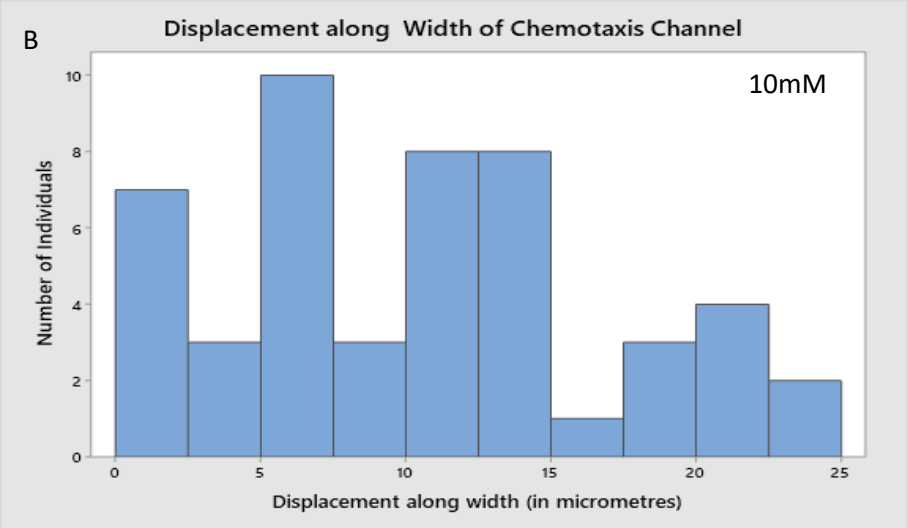
It can be seen that on average, the individuals show a displacement ~9.37 µm in the first 50 frames. Nearly 82% of the individuals were observed to have displaced from 0-10µm, during the 50 frames along the minute gradient across the width of the channel. There were two individuals that demonstrated almost five times the average displacement observed in 10 µm gradient of glucose along the X-direction.

On the other hand, for the chemotaxis experiment with 10mM gradient of glucose, it was seen that on average the individuals show a displacement ~10.5 µm in the first 50 frames, which is almost comparable to the average displacement observed in the case of 10 µM gradient. However, it was also observed that none of the individuals showed displaced more than 25 µm from their initial position along the width of the channel. (refer **Figure 4.11 (A) & (B)**)

This happened probably because the vertical and horizontal gradients in case of 10 µM Glucose were both weak, but in case of 10 mM glucose the concentration difference along the vertical direction was too high to go undetected, and therefore bacteria chose to migrate lesser towards the X direction.

In both the cases, we can say that the concentration difference along the width of the channel was so small, that the bacteria chose to displace no more than 25 µm in most of the case in both the gradients. A few individuals did show more significant displacements along the X direction; it could mean that they were capable of detecting smaller changes in the concentrations around them or it could just be attributed to the event of chance as we do not have any reading to prove it.





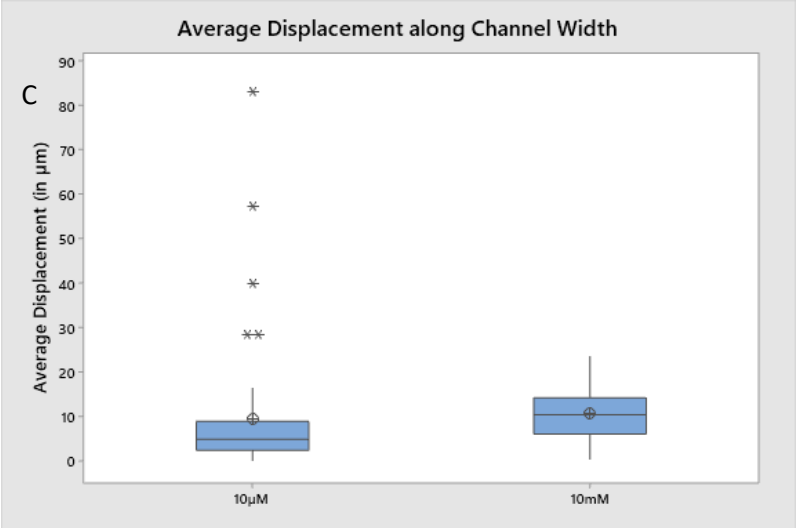


*Figure 4. 11: (A) & (B) show histograms of the displacements along the width of the chemotaxis channel for the individuals of MG1655 strain in 10uM and 10mM gradients respectively. (C) The average displacement along the channel width in 10µM gradient (9.37 ± 14.96 µm) is lesser than the displacement in 10mM gradient (10.44 ± 6.52 µm), but the values are not significantly different.in 10uM gradient from 10mM gradient. The * here represent the outlier values.*

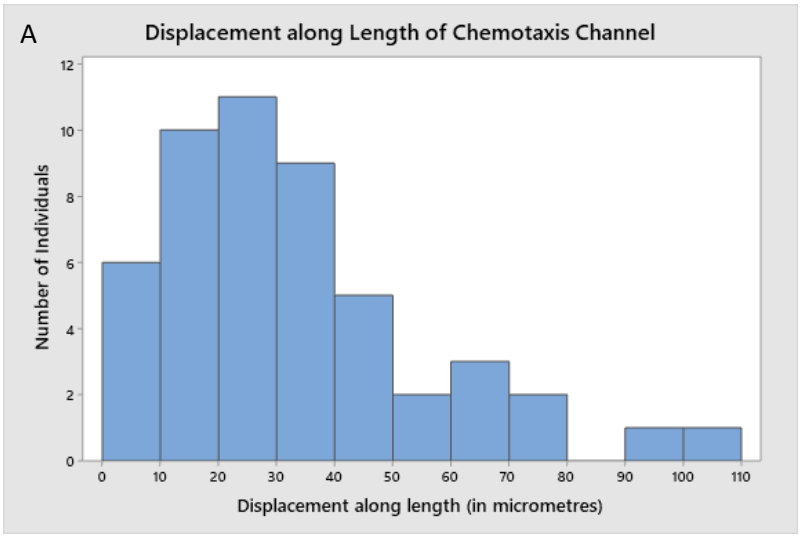
4.6.5.2 Axial Direction

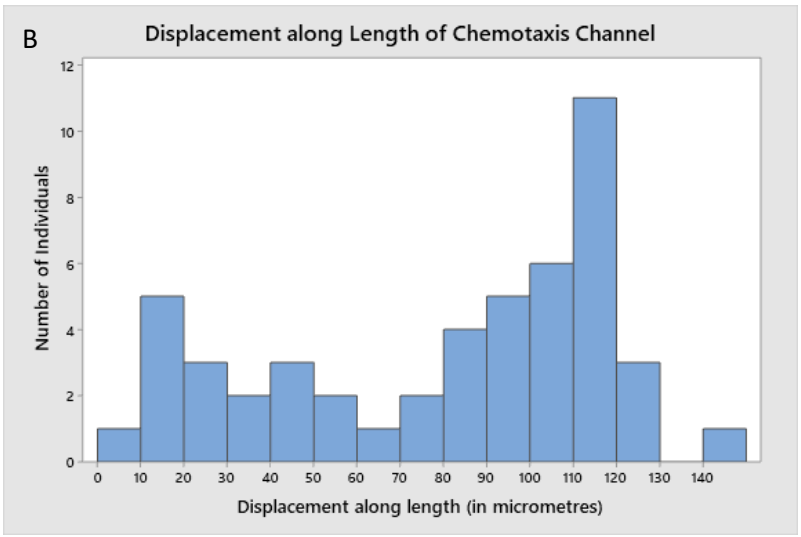
After looking at the total displacement of the individuals of MG1655 strain in the lateral direction (along the width of the channel), we will now observe the displacement of the individuals in the axial direction, i.e. along the length of the channel. The average displacement along the channel length in 10µM gradient (32.71 ± 23.23 µm) is lesser than the displacement in 10mM gradient (79.40 ± 38.47 µm), the values are significantly different in 10 µM gradient from 10mM gradient. (**Figure 4.12(C)**)

For the case of 10 µM glucose gradient, it can be seen that on average the individuals show a displacement ~33 µm in the first 50 frames. Nearly 28% of the individuals were observed to have displaced from 40 – 110 µm during these 50 frames along the gradient along the length of the channel. There were two individuals that demonstrated the average displacement of the range [90, 110] µm along the 10µM gradient of glucose.

On the other hand, for the chemotaxis experiment with 10mM gradient of glucose, it was seen that on average the individuals show a displacement ~80 µm in the first 50 frames, which is very large as compared to the average displacement observed in the case of 10 µM gradient.

As, a matter of fact, the average displacement along axial direction also happens to be nearly equal to the average total displacement observed in Section 4.7.3 for 10µM glucose gradient. This suggests that the individuals in the sample travel in a nearly straight line along the channel length with nearly negligible displacement along the width of the channel. Moreover, nearly 62.5 % of individuals from the sample displaced more than 80 µm along the length of the channel with one individual reaching as high as 150 µm from their initial position along the length of the channel. This happened because probably the gradient of 10mM Glucose was stable and distinctly detectable by the individuals of strain MG1655. (Refer to **Figure: 4.12**)





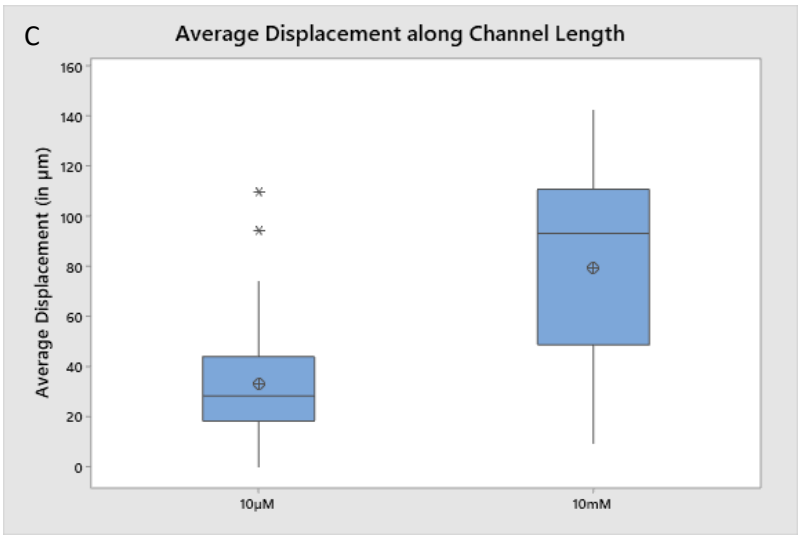


*Figure 4. 12: :(A) & (B) show histograms of the displacements along the length of the chemotaxis channel for the individuals of MG1655 strain in 10uM and 10mM gradients respectively. (C) The average displacement along the channel length in 10µM gradient (32.71 ± 23.23 µm) is lesser than the displacement in 10mM gradient (79.40 ± 38.47 µm), the values are significantly different.in 10uM gradient from 10mM gradient. The * here represent the outlier values.*

4.6.6 Individual Trajectories: MG1655 in 10mM (A) and 10µM (B) Glucose gradient

The Figure 4.13 represents the commonly observed trajectories during the chemotaxis experiments. Figure 4.13(A) shows the oblique trajectory. Some bacteria moved obliquely rather than moving up the channel gradient. This characteristic movement is caused by the presence of a slight horizontal gradient. Figure 4.13(B) & (C) represent the loopy and zig-zag trajectory, respectively. These trajectories were followed by some bacteria in the 10µM glucose gradient; these cells did not respond via movement towards the gradient.

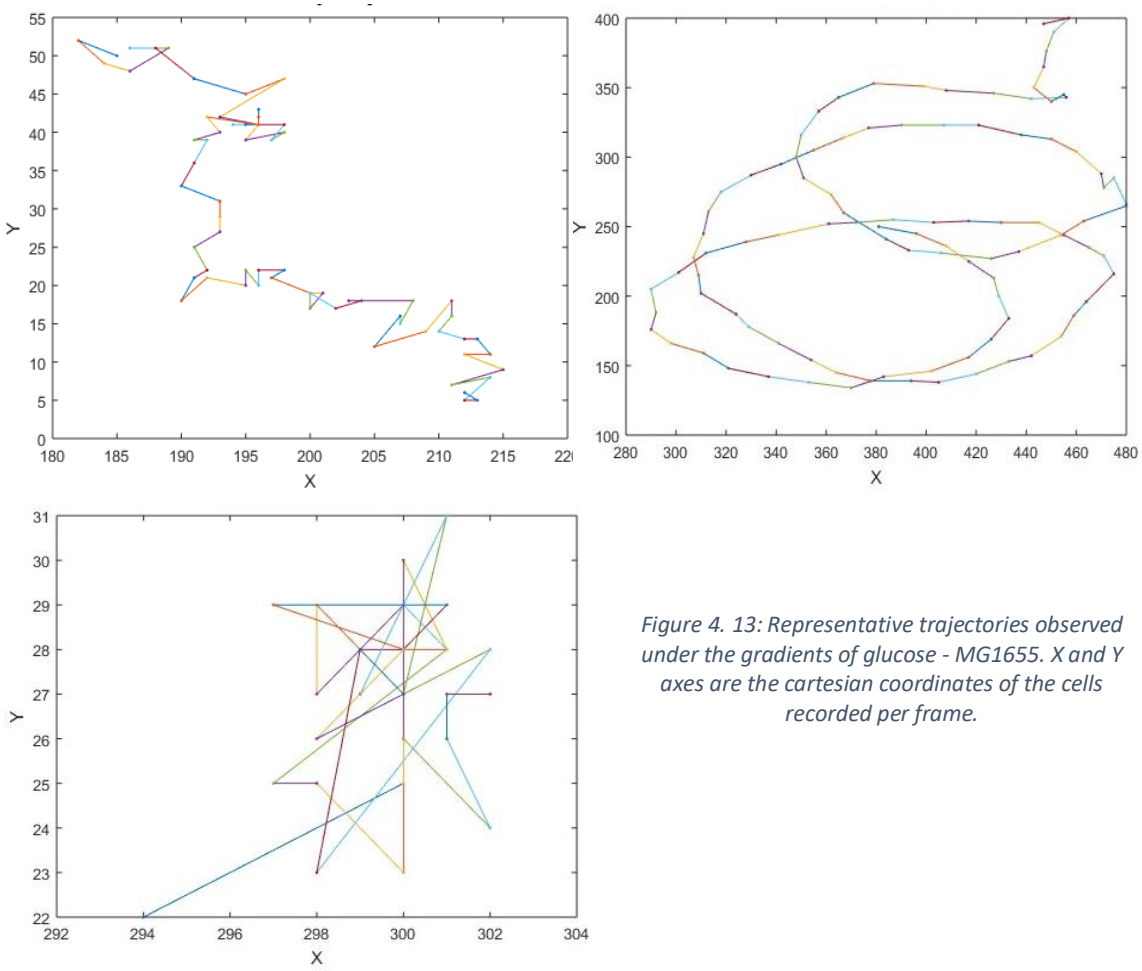


Figure 4. 13: Representative trajectories observed under the gradients of glucose - MG1655. X and Y axes are the cartesian coordinates of the cells recorded per frame.

4.7 Conclusion and Future Scope

4.7.1 Conclusion

The experiments performed to study the motility and chemotaxis of MG1655 and MDS43 in different medium compositions, involved the use of various techniques such as imaging, cell culture, data collection, data extraction and cleaning, and statistical analysis.

From the motility study, it has been observed that the medium composition of 10µM Glucose observes the highest motility and the motility decreases when the concentration is increased to 10mM and decreased to 0M glucose without casamino acids. The motility of MG1655 to was observed to be higher in 10mM glucose medium with 0.2% casamino acids than the medium with no glucose and 0.2% casamino acids, which indicates that glucose is essential for motility but inhibits it at the same time. Similarly, the motility in the absence of casamino acids showed that the cells would be more motile in the presence of glucose than no glucose at all. MDS42 strain, despite the lack of the flagellar system, showed the same trend. However, in one specific case, it showed more motility in the absence of glucose than the medium with 10mM glucose in the absence of casamino acids, unlike MG1655 which showed the least motility in no glucose medium in the absence of casamino acids. This fulfils one of the objectives of the study to observe the effects of different concentrations of Glucose and Casamino Acids on the motility of MG1655 and MDS42 strains of *Escherichia coli.*

Despite being the genetically identical population, the individuals demonstrated heterogeneity which could be observed in the form of their movement patterns, average velocities and their choice to stay stationary and multiply. This was observed in the case of both MG1655 and MDS42. Although it was easier to observe the framewise velocity profiles and average velocities of the individual bacteria, it is difficult to quantify heterogeneity using qualitative parameters such as trajectory pattern and even the quantitative parameters involving the velocity and directional changes.

As a part of the study, a hypothesis “The absence of flagellar motility results in more unidirectional motion that is lesser changes in the direction of motion” was put to the test

by studying the directional change in motion of MG1655 and MDS42 individuals in different medium compositions. It was observed that despite the lack of genes for flagellar assembly, there was no significant difference in the angular changes during motion of MDS42 from MG1655. Therefore, as per the study, the absence of flagellar motility does not necessarily result in more unidirectional motion that is lesser changes in the direction of motion.

As the second part of the study, the chemotaxis of MG1655 was studied using a single layer membrane-based microfluidic device. The device design allowed the formation of a linear gradient of chemoattractant at different concentrations, and the movement of individual bacterial cells could be easily recorded and studied. The objective to design and evaluate a single-layer microfluidic device that allows the user to study cellular chemotaxis towards linear gradients of chemoattractant was successfully achieved.

In the chemotaxis, the response of MG1655 was studied in response to the 10µM and 10mM linear gradient of Glucose. The average distance travelled in 10µM gradient (110 ± 90.1µm) is greater than the distance travelled in 10mM gradient (105.85 ± 42.03 µm), the values are not significantly different. This was in agreement with the motility readings as in case of motility experiments the average velocity was observed to be higher in case of medium containing 10µM glucose than the medium containing 10mM glucose, but the difference was not significant.

In terms of the total distance travelled in the first 50 frames, it was seen that in the 10 µM glucose gradient ~ 68% of the sample individuals travel a distance ranging from 25 – 100 µm, while in 10 mM gradient of glucose gradient it was seen that only ~ 40% of the sampled individuals were recorded to cover a distance of 60 – 100 µm in the first 50 frames rest of the individuals travelled greater than 100 µm.

Although, there was no significant difference in the framewise velocity and angular change, the chemotaxis path traversed by the strain individuals in the two cases varied widely. Under the influence of 10mM gradient, more individuals showed a ratio of displacement to distance ≥ 0.7, nearly 67% of the individuals as compared to the individuals under the 10µM gradient. The average vertical displacement in the case of 10mM glucose gradient was roughly equal to average total displacement under the same gradient, indicating that the cells followed an approximately straight trajectory. There

were no significant differences observed in terms of angle change per frame and the framewise velocities. This fulfils the final objective of the study to evaluate the effects of different concentrations of Glucose on the chemotactic behaviour of MG1655 strain of *Escherichia coli* at both the population and individual levels.

4.7.2 Scope of Future Work

The present work can be further improved and extended to the following domains:

1. A more efficient flow – based gradient generating microfluidic device can be used to study the effect of chemo-attractants on the chemotaxis, and it can be integrated to a single cell capture device which can help in selecting and culturing the cells exhibiting the cellular variability or physiological robustness.

2. The population-based responses of mutants and wildtype cells can help in better understanding of the cascade of gene translation and inter-relation of the genes comprising the chemotactic and motility systems. It can also help in mapping how certain genetic or environmental situations translate into specific morphological and physiological responses.

3. We can calibrate a microfluidic device to the response of chemotactic strain of *Escherichia coli* to a natural chemoattractant like Mannose and use it to detect the presence of heavy metal radicals in the natural samples of water. Owing to the precision and sensitivity to low concentrations of ligands in the environment, bio-reporter devices that utilise bacterial cells have been proposed as opposed to the idea of using the chemicals to detect environmental pollutants.

Appendix 1

Data Acquisition

After manually focusing the microscope with 60X objective using the steps mentioned in Chapter 3, the mode of the image acquisition was changed to Camera. We used Andor camera customised microscope for the purpose of image/video acquisition using the Andor Solis software interface. Following steps were followed to acquire the data:

1. Open the Andor Solis software.
2. Click on 'Acquisition Setup' and then go to 'Camera Setup'.
3. Change the parameters as shown in Table A1.1
4. Click 'OK' and then 'Take Signal'
5. Save the image series with an appropriate name for analysis later.

Parameters	**Values**
Acquisition Mode	Kinetic
Triggering	Internal
Readout Mode	Image
Exposure Time (in seconds)	0.1
Number of Accumulations	1
Kinetic Series Length	100
Kinetic Cycle Time (in seconds)	0.50006
Number of Prescans	0
Shift Speed (in microseconds)	[1.7]
Vertical Clock Voltage Amplitude	Normal
Readout Rate	3MHz at 14 bit
Pre-Amplifier Gain	1x
Output Amplitude	Conventional

Table A1. 1:The parametric values of the Acquisition setup in Andor Solis prior to data acquisition

Data Extraction

The image-series data captured using Andor Solis software, is primarily videos with no quantitative meaning as such. ImageJ – an open-source Java-based tool was used to quantify the image data. ImageJ has plugin packages that can read '.sif.' files generated using Andor Solis. These files were then quantified using the 'Manual Tracking' plugin. Following steps were followed to extract the data from image series files:

1. Open ImageJ and click on ‘PLUGINS’ option on the toolbar.

2. Click on ’READ SIF’; this will open a dialogue box from where you can choose the’.sif’ extension file from which you wish to extract the data.

3. Click on ‘MULTIPOINT’ tool; this tool helps in keeping a manual count of the cells on the screen.

4. Carefully count and label the cells as “Total Cells”, “Motile Cells”, “Non-motile Cells” and “Dividing Non-motile Cells”.

5. After capturing the cell count data, cells track data needed to be quantified from the image series obtained for each medium composition and each strain.

6. Click on ‘PLUGINS’ and select ‘MANUAL TRACKING.’

7. In the “TRACKING” dialogue box that opens, fill in the parameters as given in Table 3.5. The time interval is the duration of time between two consecutive frames. X-Y calibration refers to the number of micrometres in 1 pixel.

Parameter	**Value**
Time Interval (in seconds)	0.5
x/y calibration (in micrometres)	0.4
Z calibration	0

Table A1. 2: The parametric values in the Manual Tracking plugin of ImageJ prior to data extraction from image series.

8. Leave the other values as “Default.”

9. Now click on “ADD TRACK.”

10. Click on the bacterial cell of interest and choose an end of the cell.

11. Click on the same end of the cell on each frame (till 100 frames)

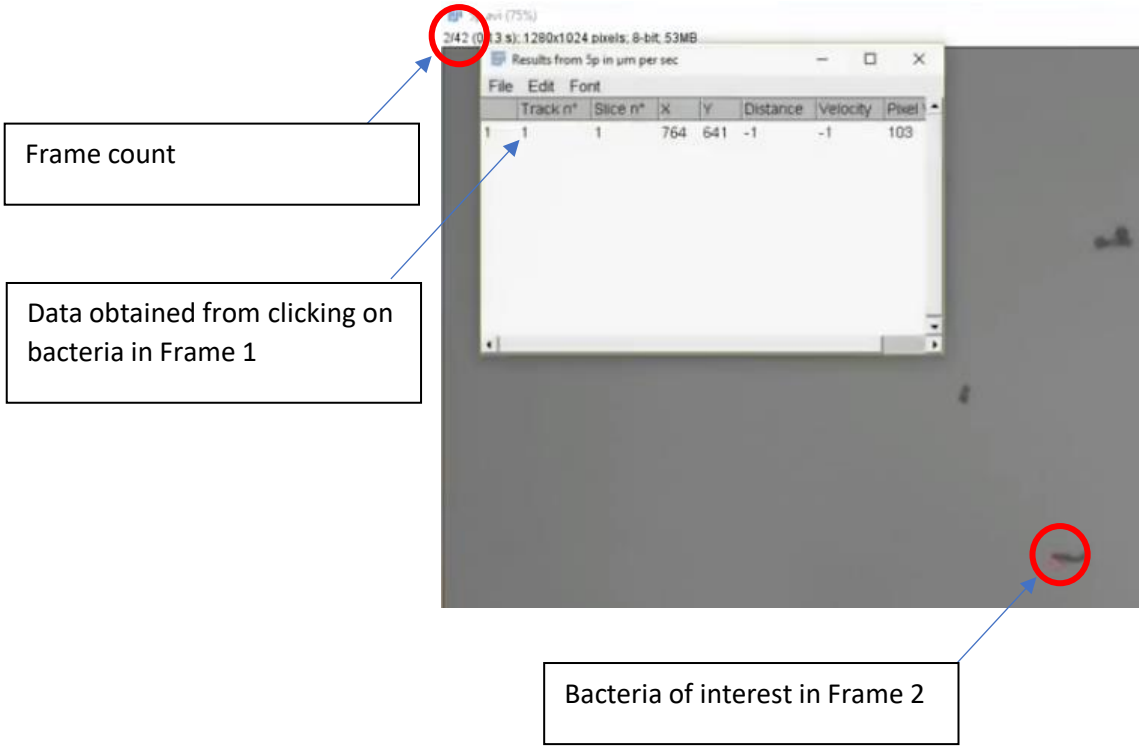
12. Click on “END TRACK” and save the obtained data in the format as shown in Table A.3

Track No.	Slice No.	X	Y	Distance	Velocity	Pixel Value

Table A1. 3: The headers of the data collected in the output files after the Manual Tracking procedure

In the above table, “Track No.” refers to the bacteria under observation, “Slice No.” denotes the x of 100 frames captured, “X” & ”Y” are the co-ordinates of the bacteria in each of the frames, “Distance” is the displacement of the bacteria in the present frame from the last frame, “Velocity” is the frame-wise velocity of the bacteria (calculated by the software based on the time interval and calibration data input initially), “Pixel Value” is the pixel number which the software uses for calculating distance and velocity based on the parameters specified in the Table A1.3.

Given below is a 3-frame illustration of manually tracking cells. Frame 1 is the first time we click on the cell; as a result, the initial position gets recorded as X and Y, and since the distance travelled and velocity of travel cannot be calculated, they are given a default value of -1. Figure A1.1 (A), (B) and (C) demonstrate how the data of bacteria movement is recorded frame by frame in ImageJ manual tracking plugin.



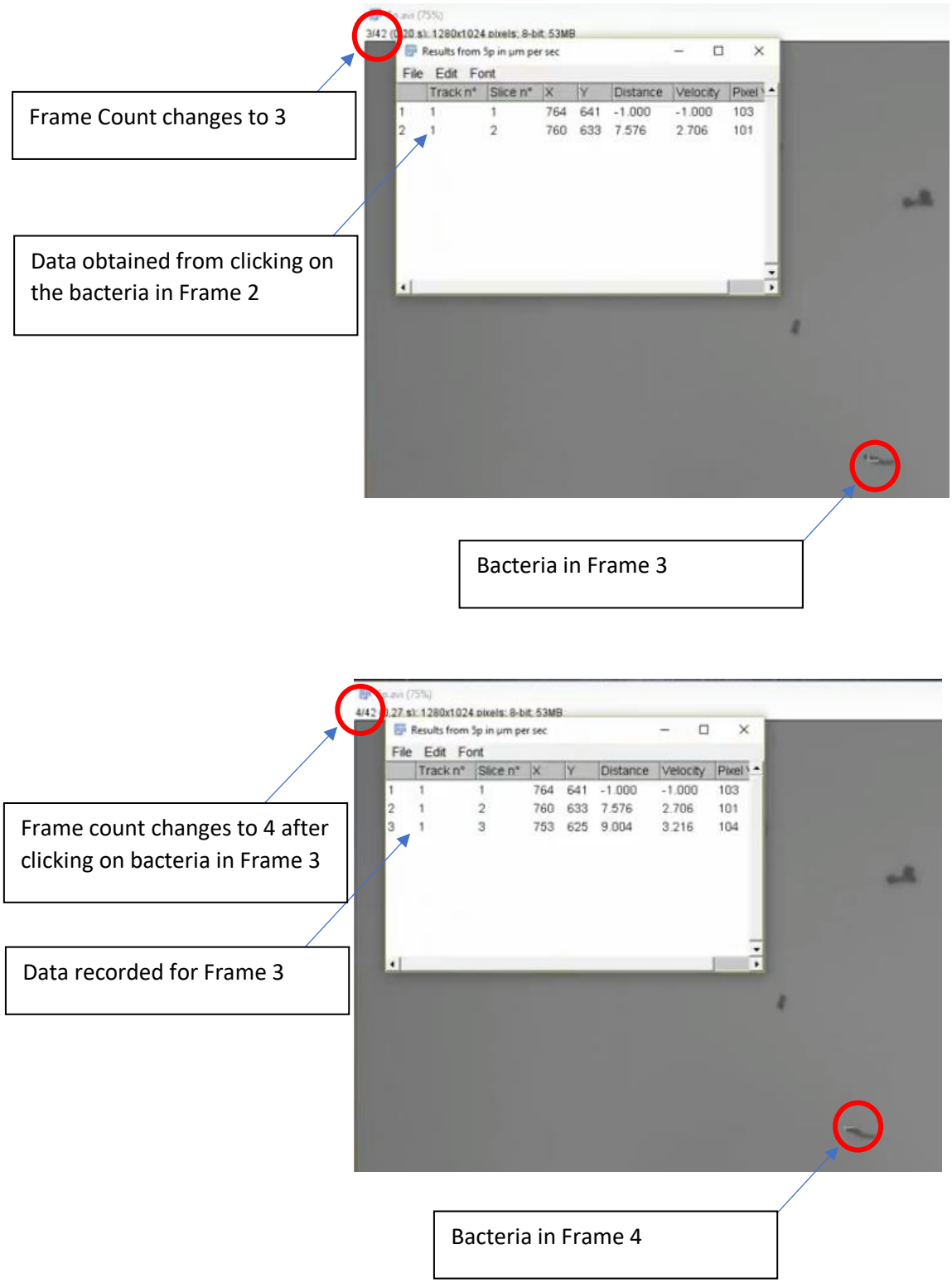


Figure A1. 1: The 3-frame illustration of data extraction from recorded video using Manual Tracking plugin of ImageJ

Appendix 2

Genome Comparison: Support Study

Among the prokaryotes, *Escherichia coli* K-12 is one of the model organisms for the genetic, biomolecular and physiological research, as it has been understood and analysed thoroughly. It is also used for the commercial synthesis of biomolecules, hormones and enzymes of medical. therapeutic & industrial significance. (Pósfai, Emergent Properties of Reduced - Genome Escherichia coli 2006). The genome of MG1655 which is closely related to the K-12 Strain, as stated in Chapter 2, has also been completely sequenced, and nearly 87% of its genes have defined functions as well. (Serres 2004) (Riley 2006) (Blattner 1997)

Once the genome of K-12 and closely related MG1655 was sequenced entirely and defined functionally, questions on the efficiency of genetic machinery rose and gave rise to deletion/reduction projects. The curious questions like “What is the smallest set of genes necessary for sustaining in a specific environment?”; “Can the genome be reduced to the point where the remaining genes can be easily categorised?”; “How far can the reduction simplify the genome?”; “Which genes are unnecessary for specific applications or even counterproductive?”, became the motivation behind the Genome Reduction Projects, which were fuelled by the academic and commercial interests. The aim of these significant – scale genome reductions was to construct the strains with the minimal genome that enhances their cellular functions, makes them metabolically robust and upgrades the sustainability. (Fehér 2007) (Pósfai, Emergent Properties of Reduced-Genome Escherichia coli 2006) (Karcagi 2016). In 2005, Hashimoto et al. prepared medium to large-sized sets of deletions and arranged them together such that they maximize the deletions of DNA regions without compromising with the functionality of housekeeping genes of MG1655 strain. (Hashimoto 2005)

Followed by this project, a group led by Posfai in 2006 started a Multiple Deletion Series project. During the project, several strains were developed beginning from MDS12 strain with 12 deletions (Kolisnychenko 2002) to MDS43 strain which had a total of 43 deletions resulting in the elimination of 743 genes, that make up to a total of 15.3% of the total genome size (Pósfai, Emergent Properties of Reduced - Genome Escherichia coli 2006). The Multiple Deletion Series project was aimed at deleting all the sequences corresponding to the mobile elements in the genome, with a motive to increase the

genomic stability and make it more robust to mutations. This stability was achieved through the removal of elements of multiple sequence repeats, insertion elements, transposes and prophages which are mainly responsible for the genetic derangements and rearrangements like inversions, deletions, transpositions and horizontal gene transfer (Hacker 2003).

MDS42, the second strain under consideration in this study, a result of Multiple Deletion Series project. It displayed numerous advantages for the biotechnological applications over the MG1655, absence of cryptic virulence elements and mobile elements being the most desirable among them (Karcagi, Indispensability of Horizontally Transferred Genes and Its Impact on Bacterial Genome Streamlining 2016). Although the strain proved competent when it came to Growth yield, cell size, acid stress tolerance but the genome reduction took a toll on nutrient utilization ability (Price 2004) and the stress tolerance of the MDS42 strain (Bochner 2001). Since, the idea of Multiple Deletion Series Project was to get rid of the mobile and cryptic virulence elements primarily. Does that mean the motility related genes were affected too?

As for the primary role, motility enables a bacterium to acclimatise to the environment and its changing conditions, to evade adverse conditions like changing pH; competition; scarce resources and chances of predation. This is accomplished with the help of timely gene expression in response to a corresponding external stimulus. (Duan 2012) (O. a. Soutourina, Regulation cascade of flagellar expression in Gram-negative bacteria 2003) (Fenchel 2002). The flagellar motility is well studied in microbiology, and the question “Is flagella related to virulence?”; “Is flagellar gene set a participant or contributor to pathogenicity?”, are of primary importance for this study as they might answer if the genes responsible for flagellar biosynthesis or the factors responsible for its up- or down-regulation were deleted in the Multiple Deletion Series project and specifically MDS42.

Flagellar motility is the most extensively studied among the modes of motility in bacterial life forms. And the question about the relation of pathogenicity with flagellar gene expression is being considered, and the results have been inconsistent (Caldwell 1985) (Nachamkin 1993). The idea that the presence of flagella or just the corresponding gene set is responsible for virulence is still under debate. So, the more natural way to check if MDS42 really had the genetic code for the flagellar biosynthesis and related mechanisms or was motile is to compare the genetic code of MDS42 with that of MG1655.

Genome Comparison: Methodology

The complete genome sequences of MG1655 and MDS42 are available in the NCBI genome database, which can be accessed using the Taxonomy IDs, as mentioned in Table A2.1. Based on the literature review and protein names available in the NCBI database, a table of the loci associated with flagella biosynthesis and biosynthesis-related – proteins were fetched for MG1655 from the website: https://www.ncbi.nlm.nih.gov/genome/proteins

Taxonomy ID	**Current NCBI Name**	**NCBI BLAST Name**
511145	*Escherichia coli* str. K – 12 substr. MG1655	Enterobacteria
1110693	*Escherichia coli* str. K – 12 substr. MDS42	Enterobacteria

Table A2. 1:The NCBI Identifiers of the two strains of Escherichia coli (MG1655 & MDS42) used for the experiments

Once the list of the proteins/genes associated with the flagellar biosynthesis was prepared, the following steps were followed:

1. Fetch the Gene ID of the flagellar biosynthesis associated proteins

2. Go to the website: https://www.ncbi.nlm.nih.gov/gene and input the GENE ID and click 'SEARCH.'

3. The Gene summary, Genomic context and the corresponding NUCLEOTIDE SEQUENCE appears

4. Save the nucleotide sequence in FASTA format and name the file based on the GENE ID

5. Go to the website: https://ncbi.nlm.nih.gov/Blast.cgi

6. Click on the NUCLEOTIDE BLAST (Z. S. Zhang 2000) to access the **blastn suite**

7. Under the Enter Query Sequence, upload the FASTA sequence corresponding to the GENE ID (Refer to the FASTA Sequences of Flagellar genes at the end of Appendix 2)

8. Under Choose Search Set:
    a. Select **Standard Databases (nr etc.)** option under databases header
    b. Type the Taxonomy ID corresponding to MDS42 that is TaxID: 1110693

9. Hit BLAST and see if a corresponding match can be found in the MDS42 genome sequence

10. Make a note of all the matches found along with the corresponding identity and e-value

11. All the GENE IDs can also be checked at once for the no identity

12. Collect the FASTA sequences in a .txt file in the following format:

FASTA Format Elements	Explanation
>	Angular Bracket, followed by no space
NC_000913.3:	Accession of the Organism, followed by a colon
2015868-2016554	Nucleotide numbers from the beginning to end of the sequence, if the corresponding strand is positive. Otherwise, put a 'c' and mention the end to start nucleotide numbers
Escherichia coli str. K-12 substr. MG1655,	Current NCBI name of the Taxonomy ID followed by a comma and a space
complete genome	Type of genome
ATTGCCATGA…	Nucleotide sequence corresponding to the GENE ID

Table A2. 2: The FASTA files created for the BLAST must be in this format prior to checking for identity

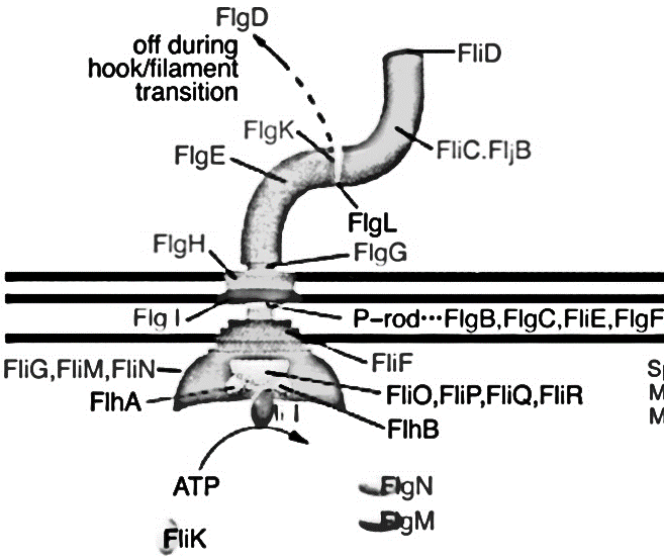


Figure A2. 1:The schematic of flagellum denoting the proteins corresponding to the morphological components

(Blocker 2003)

Gene ID	Locus	Locus tag	Protein product	Protein name
945634	flgN	b1070	NP_415588.1	flagellar biosynthesis protein FlgN
946300	flgA	b1072	NP_415590.1	flagellar basal body P-ring formation protein FlgA
945678	flgB	b1073	NP_415591.1	flagellar basal-body rod protein FlgB
946687	flgC	b1074	NP_415592.1	flagellar basal-body rod protein FlgC
945813	flgD	b1075	NP_415593.1	flagellar biosynthesis, initiation of hook assembly
945636	flgE	b1076	NP_415594.1	flagellar hook protein FlgE
945639	flgF	b1077	NP_415595.1	flagellar basal-body rod protein FlgF
945647	flgG	b1078	NP_415596.1	flagellar basal-body rod protein FlgG
946996	flgH	b1079	NP_415597.1	flagellar L-ring protein
947534	flgI	b1080	NP_415598.3	flagellar P-ring protein
945648	flgK	b1082	NP_415600.1	flagellar hook-filament junction protein 1
945646	flgL	b1083	NP_415601.1	flagellar hook-filament junction protein 2
947609	ycgR	b1194	NP_415712.1	flagellar brake protein YcgR
946094	flhE	b1878	NP_416392.1	flagellar protein
946390	flhA	b1879	NP_416393.1	flagellar biosynthesis protein FlhA
946391	flhB	b1880	NP_416394.1	flagellar biosynthesis protein FlhB
949101	fliC	b1923	NP_416433.1	flagellar filament structural protein
946428	fliD	b1924	NP_416434.1	flagellar filament capping protein
946429	fliS	b1925	NP_416435.1	flagellar biosynthesis protein FliS
946433	fliT	b1926	NP_416436.1	flagellar biosynthesis protein FliT
946446	fliE	b1937	NP_416447.1	flagellar basal-body protein FliE
946448	fliF	b1938	NP_416448.1	flagellar basal-body MS-ring and collar protein
946451	fliG	b1939	NP_416449.1	flagellar motor switch protein FliG
946456	fliH	b1940	NP_416450.2	flagellar biosynthesis protein FliH
946457	fliI	b1941	NP_416451.1	flagellum-specific ATP synthase FliI
946454	fliJ	b1942	NP_416452.1	flagellar biosynthesis protein FliJ
946449	fliK	b1943	NP_416453.1	flagellar hook-length control protein
946443	fliL	b1944	NP_416454.1	flagellar protein FliL
946442	fliM	b1945	NP_416455.1	flagellar motor switch protein FliM
946423	fliN	b1946	NP_416456.1	flagellar motor switch protein FliN
946458	fliO	b1947	NP_416457.4	flagellar biosynthesis protein FliO
946462	fliP	b1948	NP_416458.1	flagellar biosynthesis protein FliP
946463	fliQ	b1949	NP_416459.1	flagellar biosynthesis protein FliQ
946464	fliR	b1950	NP_416460.1	flagellar biosynthesis protein FliR
946776	flk	b2321	NP_416824.1	putative flagella assembly protein
948824	fliA	b1922	NP_416432.3	RNA polymerase, sigma 28 (sigma F) factor

Table A2. 3: The summary table of functional genes involved in the flagellar biosynthesis in MG1655 and their corresponding protein products. This table has the gene IDs which were used to identify the relevant genes for flagellar biosynthesis in MDS42

Query Sequence	Match Description	Query Cover	e- value	Percent Identity	Accession Number
NC_000913.3:c1244527-1243793 *Escherichia coli* str. K-12 substr. MG1655, complete genome	*Escherichia coli* str. K-12 substr. MDS42 DNA, complete genome	100%	0	100%	AP012306.1
NC_000913.3:2437950-2438945 *Escherichia coli* str. K-12 substr. MG1655, complete genome	*Escherichia coli* str. K-12 substr. MDS42 DNA, complete genome	100%	0	100%	AP012306.1

Table A2. 4: The BLAST of flagellar genes listed for MG1655 against the Genome of MDS42 gave a match for two genes with minimum e-value and 100% identity. There were no matches found for the rest of the genes.

Upon performing the BLAST against the MDS42 complete genome, it turned out that all the GENE IDs except for 946776 and 947609 showed no significant match. To understand the functionality of the genes for which 100 % identity was found, the UniProt database was accessed. UniProt is an open-source platform that provides the high – grade, detailed information about the protein sequences and their respective functions. The following steps were performed to obtain the functional knowledge of the Proteins corresponding to the Genes with 100% identity:

1. Go to the website: https://www.uniprot.org/
2. Click on the Retrieve / ID Mapping
3. Under the header –" Provide your identifiers", mention the GENE ID for the gene of interest
4. Under the Select Options header, select "GENE ID (Entrez Gene)" in the "From" drop-down and "UniProtKB" from the "To" drop-down menu.
5. Click on the SUBMIT button

Gene ID	UniProtKB ID	Entry Name	Protein Name	Gene Name
946776	P15286	FLK_ECOLI	Flagellar Regulator flk	flk div, b2321, JW2318
947609	P76010	YCGR_ECOLI	Flagellar Brake Protein YcgR	ycgR b1194, JW1183

Table A2. 5: The GeneIDs retrieved from BLAST of MG1655 flagellar genes were used to retrieve the UniProtKB IDs to understand the function of the encoded proteins

Using the functional information available on UniProt it was found that The Gene ID 946776 corresponds to the gene responsible for the putative flagella assembly protein and this nucleotide sequence regulates the gene expression of flagella by adjusting the level of protein corresponding to FlgM. FlgM is an anti-sigma factor. The protein level regulation takes place by sensing the ring completion or hook elongation (Detailed description of Flagellar structure and genetic makeup is present in Chapter 1: Introduction). There is a possibility that Flk can inhibit the secretion of FlgM, where it acts like a braking system for a more complex T3S – Type III Secretion system associated with Flagella. It prevents the interaction between FlhB and FliK to probably downregulate the switching of the flagellar Type-III Secretion system specificity to filament-type substrates before the Hook – Basal – Body is complete. The default specificity is towards the rod and hook – type substrates. This function has been deduced based on the identity and belongs to the Swiss – Prot section of UniProtKB. This means that even though the entry has been manually updated, it has been reviewed by the UniProtKB curators.

The Gene ID 947609 corresponds to the gene responsible for the flagellar brake protein YcgR and this nucleotide acts as a flagellar brake. This protein regulates the swimming and swarming modes of motility via a bis-(3'-5') cyclic diguanylic acid (c-di-GMP)-dependent mechanism. YcgR binds to the flagellar motor, mainly MotA and also to FliG and FliM, when it (YcgR) is bound to c-di-GMP, causing the flagellar motor to slow down. It attaches to FliM in the absence of c-di-GMP as well, resulting in the same fate of slowing down. It was also concluded that the motility could be decreased by increasing the concentration of c-di-GMP. (Ko 2000) (Ryjenkov 2006) (Boehm, Second messenger-mediated adjustment of bacterial swimming velocity 2010) (Paul 2010)

Another vital gene that is of significance while comparing the flagella mediated motility of MG1655, and MDS42 strains is Gene locus of FliA, corresponding to the Gene ID 948824. This gene codes for the sigma factors; these elements are the initiation factors that assist the attachment of RNA polymerase to specific initiation sites and are then released. This sigma factor controls the expression of flagella-related genes (X. a. Liu 1995). This gene is found in MG1655 and has been removed in the MDS42 strain.

(Karcagi 2016). As a result, the corresponding flagellar gene expression is downregulated in MDS42 as compared to that in the MG1655 strain.

FASTA Sequences of the flagellar genes of MG1655

Below are the FASTA-sequences of the identified flagellar genes of MG1655. These sequences were obtained from the MG1655 genome data at NCBI and used to run BLAST for Flagellar genes in MDS42.

> NC_000913.3:c1129830-1129414 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGACACGTCTTGCAGAGATCCTCGACCAGATGTCCGCTGTGCTTAACGATCTCAAAACGGT
AATGGATCAAGAGCAGCAACATCTCTCTATGGGGCAGATCAACGGCAGCCAGTTGCAATGG
ATTACAGAACAAAAAAGCTCACTGCTGGCGACGCTGGATTACCTCGAACAGTTACGCAGGA
AAGAACCCAATACAGCAAATAGCGTTGATATTAGTCAACGCTGGCAGGAAATTACTGTGAA
AACGCAGCAACTACGCCAAATGAATCAACATAACGGCTGGTTACTGGAAGGACAGATTGAG
CGCAATCAACAGGCGCTGGAAATGTTGAAACCGCATCAGGAACCGACGCTATATGGGGCGA
ACGGTCAGACCTCAACAACCCATCGCGGCGGTAAAAAGATTTCGATCTGA

> NC_000913.3:c1130863-1130204 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGCTGATAATAAAACGTAGCGTGGCGATCATCGCGATACTGTTCAGTCCGTTAAGTACGGC
GAGCAATCTCACATCGCAATTGCACAACTTTTTTAGCGCCCAACTCGCGGGGGTAAGTGATG
AGGTTCGTGTTTCTATTCGTACAGCGCCCAATCTACTACCGCCATGCGAGCAGCCATTGCTTT
CGATGAGCAATAATTCCCGCCTGTGGGGCAATGTGAATGTGTTGGCACGCTGCGGTAACGAC
AAACGATATTTACAGGTTAATGTACAGGCCACAGGAAATTATGTGGTTGCCGCGATGCCCAT
TGCGCGGGGAGGAAAGCTGGAAGCTGGCAATGTCAAACTGAAACGCGGACGGCTGGATACC
CTGCCACCGCGTACGGTGCTGGATATCAATCAACTTGTTGATGCCATTAGCCTGCGCGATCT
ATCACCCGATCAACCTATCCAGTTAACCCAGTTTCGCCAGGCATGGCGGGTAAAAGCGGGAC
AACGCGTCAATGTGATCGCCAGCGGTGATGGGTTTAGCGCCAACGCAGAAGGTCAGGCGCT
GAACAATGCAGCCGTCGCACAGAATGCGCGGGTGCGCATGGTATCGGGACAGGTAGTCAGC
GGCGTTGTTGATGCAGATGGGAATATTCTTATAAACCTGTAA

> NC_000913.3:1131018-1131434 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGCTCGATAAGCTCGACGCCGCCTTACGTTTTCAACAAGAGGCGCTCAATCTGCGCGCCCA
GCGTCAGGAAGTGCTGGCAGCAAACATCGCCAATGCCGATACCCCTGGTTATCAGGCGCGC
GATATCGATTTTGCCAGTGAACTTAAAAAAGTCATGCAACGTGGACGGGATGCAACCAGTGT
GGTTGCACTGACGATGACCTCAACGCAACACATTCCGGCGCAGGCGCTGACGCCTCCTACCG
CAGAACTGCAATACCGTATTCCGGACCAGCCTTCGCTTGACGGTAATACCGTCGATATGGAT
CGCGAACGCACCCAGTTTGCCGATAACAGCCTGCAATACCAGATGAGCCTTAGCGCGTTGAG
CGGGCAAATCAAAGGCATGATGAACGTTTTACAGAGCGGAAATTAA

> NC_000913.3:1131438-1131842 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGGCACTGCTGAATATTTTTGATATCGCCGGGTCGGCGTTAACTGCCCAGTCCCAGCGCCT
GAACGTGGCGGCCAGTAATCTGGCGAATGCTGATAGCGTGACCGGTCCCGATGGACAGCCA
TATCGGGCAAAACAGGTGGTATTCCAGGTTAACGCTGCACCAGGTGCTGCGACAGGCGGCG
TAAAGGTTGCCGATGTTATAGAAAGTCAGGCCCCGGACAAACTGGTTTATGAACCGGGTAAT
CCGCTGGCAGATGCAAAGGGCTACGTAAAAATGCCGAACGTTGATGTTGTCGGAGAGATGG
TTAACACCATGTCGGCGTCACGCAGCTATCAGGCCAATGTTGAAGTGCTCAACACGGTGAAA
AGCATGATGCTGAAAACCCTTACGCTCGGTCAATAA

- NC_000913.3:1131854-1132549 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGTCCATTGCGGTAACCACCACCGATCCGACAAATACCGGCGTCAGTACCACCAGCAGTAG
TTCGCTCACGGGCAGCAACGCCGCAGATTTACAAAGCAGTTTTCTGACTTTGCTGGTGGCGC
AGCTGAAAAACCAGGACCCGACCAATCCAATGGAAAACAACGAGCTGACGTCGCAATTGGC
ACAAATCAGCACGGTCAGCGGGATTGAAAAACTCAATACCACGCTCGGATCTATTTCCGGAC
AGATTGATAACAGCCAGTCGTTACAGGCCAGTAACCTGATCGGTCACGGCGTGATGATCCCC
GGCACCACTGTTCTTGCGGGAACCGGCAGTGAAGAAGGGGCTGTGACCACGACCACGCCGT
TTGGTGTTGAGCTGCAACAGGCGGCAGACAAAGTTACGGCCACCATCACCGATAAAAATGG
CGCGGTTGTGCGCACCATTGATATTGGTGAACTGACCGCCGGAGTTCACAGTTTCACCTGGG
ACGGTACGTTGACTGATGGCAGCACTGCGCCGAACGGTTCTTACAATGTAGCGATTAGCGCC
AGTAACGGTGGTACACAACTGGTTGCCCAGCCGCTGCAGTTTGCTCTGGTGCAGGGTGTGAT
CCGCGGCAACAGCGGTAATACGCTGGATCTCGGCACTTACGGCACCACCACCCTCGACGAA
GTACGGCAGATAATTTAA

- NC_000913.3:1132574-1133782 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGGCCTTTTCTCAAGCGGTTAGCGGATTAAACGCTGCCGCCACCAACCTCGATGTTATTGG
CAACAATATCGCCAACTCCGCCACCTACGGCTTTAAATCAGGCACGGCCTCTTTTGCCGATA
TGTTTGCCGGTTCGAAAGTGGGACTGGGGGTAAAAGTTGCCGGTATCACTCAGGACTTTACC
GATGGCACGACCACCAACACCGGGCGAGGTCTGGACGTTGCTATCAGCCAGAACGGTTTTTT
CCGTCTGGTAGACAGCAACGGTTCGGTGTTCTACAGCCGTAACGGACAATTTAAGCTGGATG
AAAACCGTAACCTGGTGAATATGCAAGGTTTACAGCTGACGGGTTACCCGGCAACCGGTAC
GCCGCCGACTATTCAGCAAGGGGCGAATCCGACCAATATTTCGATCCCGAATACCCTGATGG
CAGCGAAAACTACCACCACGGCATCGATGCAGATCAACCTGAATTCCAGTGATCCGCTTCCT
ACTGTTACGCCATTCAGCGCCAGCAATGCGGATAGCTATAACAAAAAAGGTTCGGTGACTGT
TTTCGACAGTCAGGGTAATGCTCATGACATGAGCGTCTACTTTGTGAAGACCGGGGATAATA
ACTGGCAGGTCTACACCCAGGATAGCAGTGATCCAAACAGCATTGCGAAGACAGCGACAAC
ACTGGAATTTAATGCTAATGGCACATTAGTGGATGGTGCGATGGCGAATAATATCGCAACCG
GCGCAATTAACGGTGCAGAACCCGCCACGTTTAGTCTGAGCTTCCTCAACTCCATGCAGCAA
AATACCGGCGCTAACAATATTGTGGCAACCACCCAGAACGGCTACAAACCGGGCGATCTGG
TGAGTTATCAAATCAATGATGACGGTACGGTTGTCGGCAACTATTCCAACGAACAAACCCAA
CTGCTGGGGCAGATTGTACTGGCGAACTTTGCCAACAACGAAGGTCTGGCATCCGAAGGCG
ACAACGTCTGGTCTGCGACGCAATCTTCTGGCGTGGCGCTGTTGGGGACAGCCGGGACGGG
AAACTTTGGCACCCTGACCAACGGTGCGCTGGAAGCGTCCAACGTCGATCTCAGTAAAGAA

CTGGTCAATATGATCGTTGCCCAGCGTAACTATCAGTCTAACGCCCAGACCATCAAAACCCA
GGACCAGATCCTCAACACGCTGGTTAACTTACGCTAA

- NC_000913.3:1133802-1134557 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGGATCACGCAATTTATACCGCGATGGGAGCAGCCAGCCAGACACTGAATCAACAGGCGG
TAACCGCCAGTAATCTGGCCAATGCCTCAACGCCCGGTTTTCGCGCGCAGTTGAATGCTTTA
CGCGCGGTGCCAGTGGAAGGGCTTTCTCTGCCCACGCGCACGTTGGTCACGGCGTCAACGCC
GGGCGCAGATATGACGCCCGGCAAAATGGATTACACCTCGCGCCCGCTGGACGTCGCGTTG
CAGCAGGATGGCTGGCTGGCCGTGCAGACCGCTGACGGCAGCGAAGGGTATACGCGTAATG
GCAGCATTCAGGTTGATCCCACCGGGCAACTGACAATTCAGGGGCATCCGGTGATAGGCGA
GGCTGGGCCAATTGCTGTGCCGGAAGGGGCGGAAATCACTATTGCTGCCGATGGCACAATCT
CGGCGCTCAATCCGGGCGATCCGGCAAATACGGTTGCGCCAGTAGGGCGTCTTAAACTGGTG
AAAGCCACGGGCAGCGAAGTGCAGCGCGGTGACGACGGCATTTTTCGTTTAAGCGCAGAAA
CCCAGGCCACGCGTGGGCCGGTACTGCAGGCAGATCCAACCTTGCGTGTGATGTCGGGGGTT
CTGGAAGGCAGTAACGTCAATGCCGTTGCGGCAATGAGCGACATGATTGCCAGCGCGCGGC
GTTTTGAAATGCAGATGAAGGTGATCAGCAGCGTCGATGATAACGCAGGCCGTGCCAACCA
ACTGCTGTCGATGAGTTAA

- NC_000913.3:1134729-1135511 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGATCAGTTCATTATGGATCGCCAAAACGGGCCTTGACGCCCAGCAAACCAATATGGACGT
CATTGCCAACAACCTGGCAAACGTCAGTACTAACGGTTTTAAGCGTCAGCGCGCGGTGTTTG
AAGATCTGCTTTATCAAACCATTCGCCAGCCGGGGGCACAGTCTTCCGAACAAACCACCTTA
CCCTCCGGATTACAAATCGGCACGGGGGTACGCCCGGTCGCCACTGAACGCTTACACAGCCA
GGGAAACCTGTCGCAGACCAACAACAGCAAAGATGTCGCGATTAAAGGGCAGGGCTTTTTC
CAGGTGATGTTGCCAGATGGTTCATCAGCCTATACCCGTGACGGCTCTTTCCAGGTGGATCA
GAACGGGCAGCTGGTGACGGCTGGTGGTTTTCAGGTGCAGCCAGCGATCACCATTCCGGCG
AATGCGTTAAGTATCACCATCGGTCGTGATGGCGTGGTCAGCGTAACCCAACAAGGCCAGG
CAGCTCCGGTTCAGGTTGGGCAGCTCAATCTCACCACCTTTATGAATGACACCGGGCTGGAG
AGCATTGGCGAAAACCTCTACACCGAAACGCAATCCTCTGGTGCACCGAACGAAAGCACGC
CGGGCCTGAACGGCGCGGGACTGCTGTATCAAGGGTATGTTGAAACGTCTAACGTCAACGT
GGCGGAAGAACTGGTCAATATGATTCAGGTGCAACGCGCTTACGAAATCAACAGTAAAGCG
GTGTCCACCACCGATCAGATGCTGCAAAAACTGACGCAACTCTAA

- NC_000913.3:1135564-1136262 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGCAAAAAAACGCTGCGCATACTTATGCCATTTCCAGCTTGTTGGTGCTTTCACTAACCGG
CTGCGCCTGGATACCCTCCACGCCGCTGGTGCAGGGGGCGACCAGTGCACAACCGGTTCCCG
GTCCGACGCCCGTCGCCAACGGTTCTATTTTCCAGTCTGCTCAGCCGATTAACTATGGCTATC
AACCGCTGTTTGAAGATCGTCGACCACGCAATATTGGCGATACGCTGACCATCGTGTTGCAG
GAGAACGTCAGCGCCAGCAAAAGCTCCTCTGCGAATGCCAGCCGTGACGGTAAAACTAATT
TTGGCTTTGATACTGTGCCGCGCTATTTGCAGGGGCTGTTTGGTAACGCTCGTGCCGATGTCG
AAGCCTCCGGTGGTAACACGTTCAACGGAAAGGGCGGGGCCAATGCCAGCAATACCTTTAG
CGGCACGTTGACGGTGACGGTTGACCAGGTACTGGTCAACGGCAACCTGCATGTGGTGGGT

GAAAAACAGATTGCCATTAATCAGGGTACCGAATTTATTCGCTTCTCTGGCGTGGTTAATCC
ACGCACTATCAGCGGCAGCAATACCGTACCGTCTACTCAGGTGGCGGATGCGCGCATTGAAT
ACGTAGGCAATGGCTACATTAACGAAGCGCAAAATATGGGCTGGTTGCAGCGTTTCTTCCTT
AACCTGTCGCCAATGTAA

- NC_000913.3:1136274-1137371 Escherichia coli str. K-12 substr. MG1655, complete genome

GTGATTAAATTTCTCTCTGCATTAATTCTTCTACTGGTCACGACGGCGGCTCAGGCTGAGCGT
ATTCGCGATCTCACCAGTGTTCAGGGGGTAAGGCAAAACTCACTGATTGGCTATGGTCTGGT
GGTGGGGCTGGATGGCACCGGTGACCAGACAACCCAGACGCCGTTTACCACACAAACGCTT
AATAACATGCTCTCACAGCTGGGAATTACCGTTCCGACGGGCACCAATATGCAGCTAAAAA
ACGTCGCTGCGGTAATGGTGACAGCGTCACTTCCTCCGTTTGGACGTCAGGGGCAAACCATC
GATGTGGTGGTTTCTTCCATGGGAAATGCCAAAAGCTTGCGTGGAGGTACGTTGTTGATGAC
ACCGCTTAAGGGCGTTGACAGTCAGGTGTATGCGCTGGCGCAGGGCAATATTCTGGTTGGCG
GCGCAGGAGCCTCCGCTGGCGGTAGCAGTGTTCAGGTTAACCAACTGAACGGTGGACGGAT
CACCAATGGTGCGGTTATTGAACGTGAATTGCCCAGCCAGTTTGGCGTCGGGAATACCCTTA
ATTTGCAACTTAACGACGAAGATTTCAGCATGGCGCAGCAAATCGCTGACACCATCAACCGC
GTGCGTGGATATGGCAGCGCCACCGCGTTAGATGCGCGGACTATTCAGGTGCGCGTACCGA
GTGGCAACAGTTCCCAGGTCCGCTTCCTTGCCGATATTCAGAATATGCAGGTTAATGTCACC
CCGCAGGACGCTAAAGTAGTGATTAACTCGCGCACCGGTTCGGTGGTGATGAATCGCGAAG
TGACCCTCGACAGCTGCGCGGTAGCGCAGGGGAATCTCTCAGTAACAGTTAATCGTCAGGCC
AATGTCAGCCAGCCAGATACACCGTTTGGTGGTGGACAGACTGTGGTTACTCCACAAACGCA
GATCGATTTACGCCAGAGCGGCGGTTCGCTGCAAAGCGTACGTTCCAGCGCCAGCCTCAATA
ACGTGGTGCGCGCGCTCAATGCGCTGGGCGCTACGCCGATGGATCTGATGTCCATACTGCAA
TCAATGCAAAGTGCGGGATGTCTGCGGGCAAAACTGGAAATCATCTGA

- NC_000913.3:1138378-1140021 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGTCCAGCTTGATTAATAACGCCATGAGCGGACTGAACGCGGCCCAGGCGGCGTTAAATA
CGGCAAGTAATAATATCTCCAGCTATAACGTTGCCGGATATACCCGCCAAACCACTATTATG
GCGCAGGCCAATAGCACGTTGGGCGCTGGCGGCTGGGTTGGCAATGGTGTCTACGTTTCTGG
TGTGCAGCGTGAGTATGATGCGTTTATTACCAACCAGTTACGTGCGGCGCAGACGCAAAGTA
GCGGTCTGACTGCCCGCTATGAGCAGATGTCGAAAATCGACAATATGCTCTCCACCAGTACC
TCTTCGCTGGCAACACAGATGCAGGATTTCTTCACCAGCCTGCAAACGCTGGTGAGTAACGC
GGAAGACCCGGCAGCGCGCCAGGCGCTGATTGGGAAATCAGAAGGATTGGTGAATCAGTTT
AAAACCACCGATCAATATCTGCGCGACCAGGACAAACAGGTCAATATCGCGATAGGTGCCA
GCGTTGATCAGATCAACAACTACGCTAAACAAATTGCCAGCCTGAACGATCAAATCTCGCGC
CTGACAGGCGTGGGGGCAGGGGCGTCACCTAACAATCTGCTGGATCAACGCGATCAACTGG
TGAGCGAATTAAACCAGATTGTTGGTGTAGAAGTCAGCGTTCAGGATGGCGGCACTTATAAC
ATCACGATGGCCAATGGTTACTCACTGGTTCAGGGAAGTACGGCGCGGCAACTGGCGGCAG
TTCCTTCCAGCGCTGACCCTTCTCGTACGACTGTCGCTTATGTTGATGGGACGGCAGGCAATA
TTGAGATCCCGGAGAAATTACTGAATACCGGGTCGCTGGGCGGCATTCTGACATTCCGTTCT
CAGGATCTGGACCAGACGCGTAATACGCTTGGACAACTGGCGCTGGCATTTGCCGAGGCTTT
CAACACCCAACACAAAGCCGGATTTGATGCTAACGGCGATGCCGGTGAAGATTTCTTTGCTA

TCGGTAAGCCCGCGGTTCTGCAAAACACGAAAAACAAAGGTGACGTTGCGATCGGTGCCAC
GGTAACTGATGCCTCCGCGGTACTGGCGACAGATTACAAAATCTCGTTCGATAATAATCAGT
GGCAGGTCACCCGCCTTGCCAGCAATACCACTTTTACGGTGACGCCGGATGCCAACGGTAAA
GTGGCATTTGATGGTCTGGAGTTGACGTTTACAGGAACGCCTGCCGTTAACGACAGCTTCAC
GCTGAAACCAGTAAGTGACGCCATCGTCAACATGGATGTATTAATCACCGACGAAGCGAAA
ATAGCGATGGCGAGCGAAGAAGATGCGGGTGATAGCGATAACCGCAACGGTCAGGCCCTGC
TGGATCTGCAAAGCAACAGTAAAACGGTGGGCGGTGCGAAATCCTTTAACGACGCTTATGC
CTCGTTAGTGAGTGATATCGGTAATAAAACCGCGACGTTGAAAACCAGTAGCGCCACGCAA
GGTAATGTGGTGACGCAGCTTTCCAATCAGCAGCAGTCGATTTCCGGTGTCAATCTCGATGA
GGAGTACGGAAATCTGCAACGTTTTCAGCAGTATTACCTGGCGAATGCGCAGGTTCTGCAGA
CGGCAAACGCGATTTTTGATGCGCTGATTAACATTCGCTAA

- NC_000913.3:1140033-1140986 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGCGTTTCAGTACACAGATGATGTACCAGCAAAACATGCGTGGTATCACCAATTCTCAGGC
AGAATGGATGAAGTACGGCGAACAGATGTCGACGGGTAAGCGAGTCGTTAACCCTTCTGAC
GATCCCATTGCTGCATCACAAGCCGTAGTTCTCTCCCAGGCACAGGCGCAAAACAGCCAGTA
CACGCTGGCGCGTACTTTCGCCACTCAAAAAGTGTCACTGGAAGAGAGTGTACTTAGCCAGG
TCACCACTGCTATCCAGAATGCTCAGGAAAAAATTGTCTACGCCAGCAATGGCACCTTGAGT
GACGATGACCGGGCCTCGCTGGCTACGGATATTCAGGGGCTTCGTGACCAGTTGCTGAATCT
GGCAAACACCACTGACGGTAACGGGCGCTACATTTTTGCCGGTTATAAAACAGAGACTGCG
CCGTTTAGCGAAGAGAAAGGGAAATACGTCGGTGGAGCAGAAAGTATTAAACAACAGGTCG
ATGCTTCGCGTTCGATGGTGATAGGGCACACGGGTGACAAAATTTTCGACAGTATTACCAGC
AACGCGGTAGCGGAACCAGACGGTAGCGCTTCTGAAACCAATCTTTTTGCCATGCTGGATAG
TGCCATCGCAGCCCTGAAAACGCCGGTCGCGGATAGCGAAGCGGATAAAGAAACCGCCGCT
GCGGCGTTAGATAAAACCAACCGCGGACTGAAAAACTCGCTGAACAATGTGCTGACTGTTC
GCGCGGAATTAGGCACGCAGCTGAACGAACTGGAGTCGCTGGATTCATTAGGTAGCGATCG
CGCTTTAGGGCAAACGCAGCAGATGAGCGATCTGGTTGATGTGGACTGGAATGCAACTATTT
CATCTTACATCATGCAGCAAACGGCATTGCAGGCATCGTATAAAGCATTTACCGATATGCAG
GGATTGTCGCTCTTCCAGCTCAGCAAATAA

- NC_000913.3:c1244527-1243793 Escherichia coli str. K-12 substr. MG1655, complete genome

GTGAGTCATTACCATGAGCAGTTCCTGAAACAAAATCCGTTAGCCGTCCTGGGCGTGTTACG
CGATTTGCACAAAGCCGCAATTCCTTTGCGTCTCAGTTGGAATGGCGGGCAGCTGATCAGCA
AATTACTGGCAATAACCCCGGATAAACTGGTGCTGGATTTCGGCAGTCAAGCCGAAGACAA
CATCGCCGTGCTAAAGGCACAGCACATTACCATTACCGCCGAAACTCAGGGTGCGAAAGTC
GAGTTTACTGTTGAACAACTACAGCAGAGTGAATACTTGCAGCTTCCGGCATTTATTACCGT
ACCGCCTCCCACCTTATGGTTTGTACAACGACGCCGATATTTCCGCATCTCCGCCCCACTCCA
TCCGCCTTATTTTTGCCAGACCAAACTGGCGGATAACAGTACGTTACGTTTCCGCCTGTATGA
TTTGTCGTTAGGCGGCATGGGCGCATTACTGGAAACAGCAAAGCCTGCCGAATTACAAGAA
GGCATGCGCTTCGCTCAGATTGAAGTCAACATGGGGCAATGGGGTGTTTTTCACTTTGACGC
CCAGTTAATCTCCATCAGCGAGCGCAAAGTGATTGATGGCAAGAATGAAACCATCACCACTC

CCCGTCTGAGCTTCCGTTTTCTTAACGTCAGCCCGACGGTGGAGCGGCAATTACAGCGGATT
ATTTTCTCTCTCGAGCGAGAAGCCCGGGAAAAAGCGGACAAAGTGCGCGACTGA

- NC_000913.3:c1962972-1962580 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGAGAACCTTATTAGCAATATTATTGTTTCCGCTGCTGGTGCAAGCCGCCGGGGAGGGGAT
GTGGCAGGCAAGTAGTGTAGGAATTACGCTAAATCATCGCGGTGAGTCGATGTCGTCTGCGC
CTCTTTCTACGCGACAACCTGCTTCAGGATTGATGACGCTGGTAGCGTGGCGTTATCAGCTTA
TCGGCCCGACACCTTCAGGACTGCGGGTTCGCTTGTGTTCGCAATCTCGTTGTGTCGAATTAG
AGGGGCAGAGCGGAACCACCGTGGCCTTTTCCGGTATAGCGGCAGCAGAACCGTTGCGATT
TATCTGGGAAGTGCCAGGCGGTGGGCGGTTAATTCCACCGCTAAAGGTACAACGTAATGAA
GTGATTGTGAATTATCGCTGA

- NC_000913.3:c1965050-1962972 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGAGTAATCTGGCCGCGATGCTGCGCCTGCCCGCAAACCTGAAATCGACACAATGGCAGA
TCCTTGCCGGACCGATTTTGATCCTGTTGATCTTGTCGATGATGGTGCTGCCACTGCCCGCAT
TCATACTCGACCTGTTGTTTACCTTCAATATTGCCTTGTCGATCATGGTGTTGCTGGTGGCGA
TGTTTACCCAGCGCACGCTTGAGTTTGCTGCGTTTCCGACCATTCTGTTGTTTACCACGCTGT
TGCGTCTGGCACTTAACGTGGCTTCAACCCGTATCATTTTAATGGAAGGGCATACCGGCGCG
GCGGCGGCAGGGAAGGTGGTCGAAGCGTTCGGTCACTTCCTCGTTGGTGGCAATTTCGCTAT
CGGTATCGTGGTGTTTGTCATTCTCGTGATCATCAACTTTATGGTCATTACCAAAGGTGCCGG
GCGTATCGCAGAAGTGGGTGCGCGCTTTGTTCTCGATGGTATGCCGGGTAAGCAGATGGCGA
TTGACGCCGACCTTAACGCCGGATTGATTGGTGAAGATGAGGCGAAAAAACGCCGCTCCGA
AGTGACTCAGGAAGCCGATTTTTACGGCTCAATGGACGGGGCAAGTAAGTTTGTTCGCGGCG
ATGCCATCGCCGGGATCCTCATCATGGTCATTAACATTGTCGGCGGGTTGCTGGTCGGCGTG
CTGCAACATGGCATGAGCATGGGACACGCGGCGGAAAGTTATACGCTATTGACCATTGGCG
ACGGTCTGGTGGCACAAATTCCGGCGCTGGTGATTTCTACCGCCGCGGGGGTCATCGTTACG
CGTGTCAGCACCGATCAGGATGTTGGCGAGCAGATGGTGAATCAGCTTTTCAGTAACCCAAG
CGTTATGTTGTTAAGCGCCGCCGTGCTCGGTTTACTCGGCCTGGTGCCTGGAATGCCGAACCT
GGTATTTTTGCTGTTCACTGCCGGATTGCTCGGGCTGGCCTGGTGGATACGCGGACGCGAAC
AAAAAGCGCCTGCCGAACCCAAACCGGTAAAAATGGCAGAGAATAATACCGTTGTCGAAGC
GACGTGGAACGATGTACAACTGGAAGATTCTCTGGGAATGGAAGTGGGTTATCGACTGATC
CCGATGGTCGATTTCCAGCAGGATGGTGAGTTGTTGGGCCGTATACGCAGTATCCGCAAGAA
ATTTGCCCAGGAGATGGGATTTCTGCCGCCAGTGGTGCACATTCGCGACAATATGGATCTGC
AACCTGCCCGCTATCGCATTTTGATGAAAGGCGTGGAGATTGGCAGTGGTGATGCTTATCCG
GGGCGCTGGCTGGCGATTAACCCTGGAACCGCTGCCGGGACGTTACCTGGTGAGGCGACCG
TCGATCCGGCATTTGGCCTGAATGCTATCTGGATTGAAAGTGCGCTAAAAGAACAGGCGCAG
ATTCAGGGGTACACAGTGGTTGAGGCCAGCACGGTGGTAGCAACGCATCTTAACCACCTCAT
TAGCCAGCATGCCGCAGAGCTGTTTGGTCGTCAGGAGGCGCAACAGCTGTTGGATCGCGTCG
CCCAGGAGATGCCAAAGCTGACGGAAGATCTCGTTCCTGGCGTCGTCACGCTCACCACACTG
CATAAAGTGCTGCAAAATCTCCTCGATGAAAAAGTACCGATTCGCGATATGCGCACCATTCT
CGAAACGCTGGCGGAACATGCGCCCATCCAAAGCGATCCACATGAATTAACCGCCGTCGTG
CGCGTGGCGTTGGGACGGGCGATTACCCAGCAGTGGTTTCCTGGCAAAGATGAAGTCCATGT

TATTGGCCTCGATACACCGCTGGAACGTTTGTTACTACAGGCGCTGCAGGGCGGGGGAGGAC
TGGAGCCAGGGCTGGCGGATCGTTTACTGGCGCAAACTCAGGAAGCGCTATCCCGTCAGGA
GATGCTGGGTGCGCCGCCAGTATTGTTGGTGAACCACGCGCTGCGACCATTATTGTCTCGCT
TCCTGCGCCGCAGCTTGCCGCAGTTAGTGGTCCTGTCGAATCTGGAACTGTCTGATAACCGA
CATATCCGCATGACGGCGACAATTGGCGGCAAATAA

- NC_000913.3:c1966191-1965043 Escherichia coli str. K-12 substr. MG1655, complete genome

GTGTCTGACGAGAGCGACGACAAAACAGAAGCCCCCACACCTCACCGACTAGAAAAAGCGC
GGGAAGAGGGGCAAATCCCGCGTTCCCGTGAACTGACCTCACTGCTGATTTTGCTGGTGGGC
GTTAGTGTTATCTGGTTTGGCGGTGTGTCGCTGGCCCGTCGATTGTCGGGCATGCTCTCCGCT
GGGCTGCATTTTGATCACAGTATTATCAATGACCCGAATCTGATCCTCGGGCAGATTATTCTG
CTGATCAGAGAAGCCATGCTGGCGCTGCTGCCGCTGATTAGCGGCGTGGTGCTGGTGGCGCT
CATTTCTCCGGTCATGCTGGGAGGGCTGGTATTTAGCGGCAAATCCTTGCAGCCGAAGTTTT
CCAAACTCAACCCGCTACCGGGCATTAAACGGATGTTCTCGGCTCAGACTGGCGCGGAGTTG
CTTAAAGCAATTTTGAAAACCATCCTGGTTGGCAGCGTGACGGGGTTTTTTCTCTGGCATCAC
TGGCCGCAGATGATGCGCTTGATGGCCGAGTCTCCGATTACCGCCATGGGTAATGCGATGGA
TTTGGTAGGGCTATGCGCACTGCTGGTGGTGCTTGGTGTCATTCCAATGGTGGGATTTGACGT
CTTTTTCCAAATATTCAGCCACCTGAAAAAGCTGCGTATGTCACGGCAGGATATTCGTGATG
AGTTCAAACAAAGCGAAGGTGACCCTCATGTTAAAGGGCGGATCCGTCAGATGCAGCGAGC
TGCTGCACGGCGTCGGATGATGGCCGATGTGCCGAAAGCGGATGTCATTGTCAATAACCCGA
CCCACTATTCGGTAGCGTTGCAGTATGACGAAAACAAAATGAGCGCACCGAAAGTGGTCGC
TAAAGGTGCAGGGCTGGTCGCGCTGCGCATTCGTGAAATTGGCGCTGAAAATAACGTCCCG
ACGCTTGAAGCGCCGCCGCTGGCGCGAGCGCTGTATCGACATGCGGAGATTGGTCAACAAA
TCCCGGGTCAACTGTACGCCGCGGTGGCGGAAGTGCTGGCCTGGGTCTGGCAACTGAAACG
CTGGCGTCTGGCTGGTGGACAGCGCCCTGTACAACCTACTCATCTTCCGGTGCCGGAAGCCC
TGGATTTTATTAACGAGAAACCGACCCATGAGTAA

- NC_000913.3:c2003606-2002110 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATAATATCAACAAGAA
CCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAA
GGATGACGCAGCGGGTCAGGCGATTGCTAACCGTTTCACCTCTAACATTAAAGGCCTGACTC
AGGCGGCCCGTAACGCCAACGACGGTATCTCCGTTGCGCAGACCACCGAAGGCGCGCTGTC
CGAAATCAACAACAACTTACAGCGTGTGCGTGAACTGACGGTACAGGCCACTACCGGTACT
AACTCTGAGTCTGATCTGTCTTCTATCCAGGACGAAATTAAATCCCGTCTGGATGAAATTGA
CCGCGTATCTGGTCAGACCCAGTTCAACGGCGTGAACGTGCTGGCAAAAAATGGCTCCATGA
AAATCCAGGTTGGCGCAAATGATAACCAGACTATCACTATCGATCTGAAGCAGATTGATGCT
AAAACTCTTGGCCTTGATGGTTTTAGCGTTAAAAATAACGATACAGTTACCACTAGTGCTCC
AGTAACTGCTTTTGGTGCTACCACCACAAACAATATTAAACTTACTGGAATTACCCTTTCTAC
GGAAGCAGCCACTGATACTGGCGGAACTAACCCAGCTTCAATTGAGGGTGTTTATACTGATA
ATGGTAATGATTACTATGCGAAAATCACCGGTGGTGATAACGATGGGAAGTATTACGCAGT
AACAGTTGCTAATGATGGTACAGTGACAATGGCGACTGGAGCAACGGCAAATGCAACTGTA
ACTGATGCAAATACTACTAAAGCTACAACTATCACTTCAGGCGGTACACCTGTTCAGATTGA

TAATACTGCAGGTTCCGCAACTGCCAACCTTGGTGCTGTTAGCTTAGTAAAACTGCAGGATT
CCAAGGGTAATGATACCGATACATATGCGCTTAAAGATACAAATGGCAATCTTTACGCTGCG
GATGTGAATGAAACTACTGGTGCTGTTTCTGTTAAAACTATTACCTATACTGACTCTTCCGGT
GCCGCCAGTTCTCCAACCGCGGTCAAACTGGGCGGAGATGATGGCAAAACAGAAGTGGTCG
ATATTGATGGTAAAACATACGATTCTGCCGATTTAAATGGCGGTAATCTGCAAACAGGTTTG
ACTGCTGGTGGTGAGGCTCTGACTGCTGTTGCAAATGGTAAAACCACGGATCCGCTGAAAGC
GCTGGACGATGCTATCGCATCTGTAGACAAATTCCGTTCTTCCCTCGGTGCGGTGCAAAACC
GTCTGGATTCCGCGGTTACCAACCTGAACAACACCACTACCAACCTGTCTGAAGCGCAGTCC
CGTATTCAGGACGCCGACTATGCGACCGAAGTGTCCAATATGTCGAAAGCGCAGATCATCCA
GCAGGCCGGTAACTCCGTGTTGGCAAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCTCTGC
TGCAGGGTTAA

- NC_000913.3:2003872-2005278 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGGCAAGTATTTCATCGCTGGGAGTCGGGTCAGGTCTGGATTTAAGTTCCATCCTTGATAG
CCTCACCGCCGCGCAAAAAGCGACGCTAACCCCCATTTCAAATCAGCAATCGTCGTTTACCG
CTAAACTTAGCGCCTACGGTACGCTGAAAAGCGCGCTGACGACTTTCCAGACCGCCAATACT
GCATTGTCTAAAGCCGATCTTTTTTCCGCCACCAGCACCACCAGCAGCACCACCGCGTTCAG
TGCCACCACTGCGGGTAACGCCATCGCCGGGAAATACACCATCAGCGTCACCCATCTGGCGC
AGGCGCAAACCCTGACCACGCGCACCACCAGAGACGATACGAAAACGGCGATCGCCACCAG
CGACAGTAAACTCACCATTCAACAAGGCGGCGACAAAGATCCGATTACCATTGATATCAGC
GCGGCTAACTCATCGTTAAGCGGGATCCGTGATGCCATCAACAACGCAAAAGCAGGCGTAA
GCGCAAGCATCATTAACGTGGGTAACGGTGAATATCGTCTGTCAGTCACATCAAATGACACC
GGCCTTGATAATGCGATGACACTCTCGGTCAGCGGTGATGATGCGCTACAAAGTTTTATGGG
CTATGACGCCAGTGCCAGCAGCAACGGTATGGAGGTCTCGGTTGCCGCCCAGAATGCGCAG
CTGACAGTCAACAACGTCGCCATCGAGAACAGCAGCAACACCATCAGCGACGCGCTGGAAA
ACATCACCCTGAACCTGAACGATGTCACCACGGGCAACCAGACGCTAACCATCACTCAGGA
CACCTCCAAAGCGCAAACGGCGATTAAAGACTGGGTGAATGCCTACAACTCGCTAATAGAT
ACCTTCAGCAGCCTGACCAAATACACCGCCGTAGATGCGGGAGCTGATAGCCAGAGTTCTA
GCAATGGTGCACTGCTCGGCGACTCCACGCTGCGGACGATTCAGACGCAGTTGAAATCGATG
CTGAGTAATACCGTCAGTTCTTCCAGCTATAAAACGTTGGCGCAGATTGGTATCACGACCGA
TCCCAGCGATGGCAAACTGGAACTGGATGCCGACAAACTCACCGCTGCACTGAAAAAAGAT
GCCAGCGGCGTAGGTGCATTGATTGTTGGCGATGGTAAAAAAACCGGCATCACGACCACCA
TCGGCAGCAACCTGACCAGTTGGCTTTCGACAACGGGCATTATTAAAGCCGCTACCGATGGC
GTTAGTAAGACCCTGAATAAATTAACTAAAGACTACAACGCCGCCAGCGATCGCATTGATGC
GCAGGTCGCTCGCTACAAAGAACAATTTACCCAACTGGACGTTTTAATGACCTCGTTAAACA
GCACCAGCAGCTACTTAACGCAGCAGTTCGAAAACAACAGTAATTCCAAGTAA

- NC_000913.3:2005303-2005713 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGTACGCGGCAAAAGGCACCCAGGCCTATGCACAAATTGGCGTCGAAAGCGCCGTAATGA
GCGCCAGCCAGCAGCAGCTGGTCACCATGCTATTTGATGGAGTGCTGAGCGCACTGGTTAGA
GCGAGCCTGTTTATGCAGGACAACAATCAGCAAGGCAAAGGCGTCTCTTTGTCAAAAGCGA
TCAACATCATTGAGAACGGACTGCGGGTGAGTCTTGATGAAGAGAGCAAAGACGAACTAAC

CCAAAACTTGATTGCTCTTTATAGCTATATGGTCAGGCGCTTGCTGCAAGCCAATTTACGCA
ACGATGTCTCCGCAGTCGAAGAAGTGGAAGCATTAATGCGCAATATTGCCGATGCCTGGAA
AGAGTCGTTACTCTCCCCTTCTTTGATTCAGGACCCAGTCTGA

- NC_000913.3:2005713-2006078 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGAACCATGCACCGCATTTATATTTCGCCTGGCAACAACTCGTCGAAAAAAGCCAGCTCAT
GTTACGCCTGGCAACGGAAGAACAATGGGACGAACTCATCGCCAGCGAAATGGCGTATGTG
AATGCGGTGCAGGAGATTGCACATTTGACTGAAGAGGTTGACCCGTCCACCACGATGCAGG
AGCAGCTCCGCCCGATGCTGCGCCTGATTCTCGACAACGAAAGCAAGGTAAAGCAGTTATTA
CAGATTCGGATGGATGAACTGGCGAAACTGGTCGGTCAGTCATCGGTGCAAAAATCGGTGTT
AAGTGCCTATGGCGATCAGGGCGGCTTTGTGCTGGCTCCGCAGGATAACCTCTTTTGA

- NC_000913.3:c2013014-2012700 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGTCAGCGATACAGGGGATTGAAGGGGTTATCAGCCAGTTACAGGCTACGGCGATGAGTG
CGCGTGCGCAGGAATCACTGCCGCAACCGACCATTAGTTTTGCCGGGCAGCTGCACGCCGCG
CTCGATCGCATTAGTGATACACAAACAGCTGCCCGCACGCAGGCAGAAAAATTCACTCTCGG
TGAACCCGGCGTGGCGTTAAACGATGTGATGACCGATATGCAAAAAGCCTCAGTTTCTATGC
AAATGGGGATTCAGGTGCGTAATAAGCTGGTGGCGGCGTATCAGGAAGTGATGAGCATGCA
GGTGTAG

- NC_000913.3:2013229-2014887 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGAATGCGACTGCAGCCCAGACAAAATCTCTTGAGTGGCTTAATCGCCTGCGTGCGAATCC
GAAAATTCCATTGATTGTTGCCGGTTCCGCGGCAGTGGCGGTCATGGTCGCACTGATCCTGT
GGGCGAAAGCCCCCGACTACCGCACATTATTCAGCAATCTTTCCGATCAGGATGGTGGCGCA
ATTGTCAGCCAACTGACGCAAATGAATATTCCTTACCGCTTCAGCGAAGCCAGCGGCGCTAT
TGAAGTTCCGGCAGATAAAGTTCACGAACTGCGTCTGCGCCTGGCACAACAAGGTTTGCCAA
AAGGCGGCGCGGTCGGTTTCGAACTGCTTGATCAGGAAAAGTTTGGTATCAGCCAGTTCAGC
GAACAGGTGAATTATCAGCGGGCGCTGGAAGGCGAGCTTTCTCGTACCATCGAAACTATCG
GCCCGGTAAAAGGGGCGCGCGTACATCTGGCAATGCCGAAACCGTCTTTATTCGTCCGTGAA
CAAAAATCCCCTTCTGCATCGGTGACGGTAAATCTGTTACCCGGCCGCGCACTCGATGAAGG
GCAAATTAGCGCCATTGTGCATCTGGTTTCCAGCGCCGTTGCTGGTCTGCCGCCGGGAAACG
TCACGCTGGTGGATCAGGGCGGACATCTGTTAACCCAGTCCAATACCAGCGGGCGCGATCTT
AATGACGCTCAGTTGAAATATGCCAGCGATGTCGAAGGCCGTATTCAGCGGCGTATTGAAGC
GATCCTGTCGCCTATTGTTGGTAACGGTAATATTCACGCCCAGGTTACGGCGCAGCTGGACT
TCGCCAGTAAAGAACAAACGGAAGAACAGTATCGCCCTAACGGTGATGAATCTCATGCGGC
GCTTCGTTCACGCCAGCTTAATGAGAGCGAGCAAAGCGGTTCCGGTTATCCGGGCGGCGTAC
CGGGGGCGTTGTCGAATCAACCGGCACCTGCGAATAACGCGCCAATCAGCACGCCTCCGGC
AAATCAAAATAACCGCCAGCAGCAGGCGAGCACCACCAGCAATAGTGGGCCGCGTAGCACA
CAGCGGAATGAAACCAGTAACTACGAAGTCGATCGCACCATTCGTCATACCAAAATGAACG
TGGGCGATGTGCAACGTCTGTCAGTCGCGGTCGTGGTGAATTACAAAACCTTGCCAGATGGC
AAACCGTTGCCTCTCAGCAACGAACAGATGAAGCAAATTGAAGATCTGACCCGCGAGGCGA
TGGGCTTTTCTGAAAAACGCGGTGACTCGCTCAATGTCGTTAACTCGCCGTTCAATAGCAGT

GACGAAAGCGGCGGAGAACTGCCATTCTGGCAACAGCAAGCGTTTATCGATCAGTTACTTGC
TGCCGGTCGCTGGTTGCTGGTACTGCTGGTGGCGTGGCTGCTGTGGCGGAAAGCGGTACGTC
CGCAGCTAACACGTCGCGCTGAGGCGATGAAAGCTGTACAGCAACAGGCGCAGGCCCGCGA
GGAAGTGGAAGATGCGGTGGAAGTCCGCCTGAGCAAAGACGAACAACTACAACAACGGCG
CGCTAACCAACGTCTGGGGGCAGAAGTCATGAGCCAGCGTATCCGTGAAATGTCTGATAAC
GATCCGCGCGTGGTGGCGCTGGTCATTCGCCAGTGGATAAATAACGATCATGAGTAA

- NC_000913.3:2014880-2015875 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGAGTAACCTGACAGGCACCGATAAAAGCGTCATCCTGCTGATGACCATTGGCGAAGACC
GGGCGGCAGAGGTGTTCAAGCACCTCTCCCAGCGTGAAGTACAAACCCTGAGCGCTGCAAT
GGCGAACGTCACGCAGATCTCCAACAAGCAGCTAACCGATGTGCTGGCGGAGTTTGAGCAA
GAAGCTGAACAGTTTGCCGCACTGAATATCAACGCCAACGATTATCTGCGCTCGGTATTGGT
CAAAGCTCTGGGTGAAGAACGTGCCGCCAGCCTGCTGGAAGATATTCTCGAAACTCGCGAT
ACCGCCAGCGGTATTGAAACGCTCAACTTTATGGAGCCACAGAGCGCCGCCGATCTGATTCG
CGATGAGCATCCGCAAATTATCGCCACCATTCTGGTGCATCTGAAGCGCGCCCAAGCCGCCG
ATATTCTGGCGTTGTTCGATGAACGTCTGCGCCACGACGTGATGTTGCGTATCGCCACCTTTG
GCGGCGTGCAGCCAGCCGCGCTGGCGGAGCTGACCGAAGTACTGAATGGCTTGCTCGACGG
TCAGAATCTCAAGCGCAGCAAAATGGGCGGCGTGAGAACGGCAGCCGAAATTATCAACCTG
ATGAAAACTCAGCAGGAAGAAGCCGTTATTACCGCCGTGCGTGAATTCGACGGCGAGCTGG
CGCAGAAAATCATCGACGAGATGTTCCTGTTCGAGAATCTGGTGGATGTCGACGATCGCAGC
ATTCAGCGTCTGTTGCAGGAAGTGGATTCCGAATCGCTGTTGATCGCGCTGAAAGGAGCCGA
GCAGCCACTGCGCGAGAAATTCTTGCGCAATATGTCGCAGCGTGCCGCCGATATTCTGCGCG
ACGATCTCGCCAACCGTGGTCCGGTGCGTCTGTCGCAGGTGGAAAACGAACAGAAAGCGAT
TCTGCTGATTGTGCGCCGCCTTGCCGAAACTGGCGAGATGGTAATTGGCAGCGGCGAGGATA
CCTATGTCTGA

- NC_000913.3:2015868-2016554 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGTCTGATAATCTGCCGTGGAAAACCTGGACGCCGGACGATCTCGCGCCACCACAGGCAG
AGTTTGTGCCCATAGTCGAGCCGGAAGAAACCATCATTGAAGAGGCTGAACCCAGCCTTGA
GCAGCAACTGGCGCAACTGCAAATGCAGGCCCATGAGCAAGGTTATCAGGCGGGTATTGCC
GAAGGTCGCCAGCAAGGTCATAAGCAGGGCTATCAGGAAGGACTGGCCCAGGGGCTGGAGC
AAGGTCTGGCAGAGGCGAAGTCTCAACAAGCGCCAATTCATGCCCGGATGCAGCAACTGGT
CAGCGAATTTCAAACTACCCTTGATGCACTTGATAGTGTGATAGCGTCGCGCCTGATGCAGA
TGGCGCTGGAGGCGGCACGTCAGGTCATCGGTCAGACGCCAACGGTGGATAACTCGGCACT
GATCAAACAGATCCAACAGTTGTTGCAGCAAGAACCGTTATTCAGCGGTAAACCACAGCTG
CGCGTGCACCCGGATGATCTGCAACGTGTGGATGATATGCTCGGCGCTACCTTAAGTTTGCA
TGGCTGGCGCTTGCGGGGCGATCCCACCCTCCATCCTGGCGGCTGTAAAGTCTCCGCCGATG
AAGGCGATCTCGACGCCAGTGTCGCCACTCGCTGGCAAGAACTCTGCCGTCTGGCAGCACCA
GGAGTGGTGTAA

- NC_000913.3:2016554-2017927 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGACCACGCGCCTGACTCGCTGGCTAACCACGCTGGATAACTTTGAAGCCAAAATGGCGCA
GTTGCCTGCGGTACGTCGCTACGGGCGATTAACCCGCGCTACCGGGCTGGTGCTGGAAGCCA
CCGGATTACAATTGCCGCTCGGCGCAACCTGTGTCATTGAGCGCCAGAACGGCAGCGAAAC
GCACGAAGTAGAAAGCGAAGTCGTTGGCTTTAACGGTCAACGGCTGTTTTTAATGCCGCTGG
AGGAAGTCGAAGGTGTCCTGCCCGGCGCGCGTGTTTATGCCAAAAACATTTCGGCAGAAGG
GCTGCAAAGCGGCAAGCAGTTGCCGCTCGGTCCGGCGTTATTAGGTCGCGTTCTGGACGGCA
GCGGTAAACCGCTCGATGGCCTGCCCTCCCCCGATACGACGGAAACCGGTGCGCTGATTACC
CCGCCATTTAACCCGTTGCAACGTACACCGATTGAACATGTGCTGGACACCGGCGTGCGCCC
AATCAATGCCCTGCTTACCGTTGGGCGTGGGCAGCGTATGGGGCTGTTTGCCGGGTCCGGCG
TTGGTAAAAGTGTGCTGCTGGGGATGATGGCACGTTACACCCGCGCCGATGTCATTGTCGTG
GGTTTGATTGGTGAACGTGGGCGCGAAGTAAAAGATTTTATTGAGAACATCCTCGGTGCCGA
AGGGCGTGCACGCTCAGTGGTGATTGCCGCTCCGGCGGATGTTTCTCCGCTCCTGCGAATGC
AGGGTGCCGCCTATGCCACGCGCATTGCCGAAGATTTTCGCGATCGTGGTCAGCATGTGTTG
CTGATTATGGACTCCCTCACCCGCTACGCGATGGCCCAGCGTGAAATTGCGCTGGCGATTGG
CGAACCCCCCGCCACCAAAGGTTATCCACCGTCGGTGTTTGCCAAATTACCGGCACTGGTCG
AGCGTGCCGGAAATGGCATTAGCGGCGGCGGCTCGATTACCGCGTTTTATACCGTGCTCACT
GAAGGCGATGACCAGCAGGATCCGATTGCCGACTCCGCGCGGGCGATCCTCGACGGTCACA
TTGTGCTGTCTCGCCGACTGGCGGAAGCCGGGCACTATCCGGCTATCGATATTGAAGCGTCG
ATCAGCCGCGCAATGACGGCGTTGATCAGTGAGCAACATTACGCGCGAGTGCGCACCTTCA
AACAGCTGTTGTCGAGTTTTCAGCGTAACCGCGATCTGGTTAGCGTCGGCGCGTATGCCAAA
GGCAGCGATCCGATGCTCGATAAAGCCATCGCCCTGTGGCCGCAGCTGGAGGGCTATTTGCA
ACAAGGCATTTTTGAACGCGCGGACTGGGAAGCGTCTCTCCAGGGGCTGGAGCGTATTTTCC
CGACAGTGTCATAA

- NC_000913.3:2017946-2018389 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGGCAGAACATGGTGCGCTGGCGACCCTGAAAGATCTGGCAGAAAAAGAGGTAGAGGATG
CCGCGCGCCTGCTGGGTGAAATGCGTCGCGGATGTCAGCAGGCGGAAGAACAGCTCAAAAT
GCTGATTGATTATCAGAATGAATATCGCAATAACCTCAACAGCGATATGAGTGCCGGGATAA
CCAGCAACCGCTGGATCAACTATCAGCAGTTTATCCAGACGCTGGAAAAAGCCATTACTCAG
CATCGCCAGCAACTTAATCAGTGGACGCAGAAAGTTGACATTGCCCTGAACAGTTGGCGAG
AAAAAAAACAACGTTTGCAGGCCTGGCAGACACTGCAGGAACGGCAATCCACGGCGGCACT
GCTTGCAGAAAACCGCCTCGATCAGAAAAAGATGGATGAGTTCGCCCAGCGCGCCGCCATG
AGGAAACCTGAATGA

- NC_000913.3:2018386-2019513 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGATTCGCTTAGCGCCCTTGATTACCGCCGACGTTGACACCACCACATTGCCTGGCGGCAA
AGCCAGCGATGCTGCACAAGATTTTCTCGCGTTGTTGAGCGAAGCATTAGCAGGCGAGACA
ACTACCGACAAAGCGGCCCCCCAGTTGCTGGTGGCAACAGATAAGCCCACGACAAAAGGCG
AGCCGCTGATCAGCGATATTGTTTCCGACGCGCAACAAGCTAATTTACTGATCCCTGTGGAT
GAAACACCGCCTGTCATCAACGACGAACAATCCACATCAACACCGTTAACCACCGCTCAGA
CGATGGCGTTGGCTGCGGTGGCTGACAAAAATACGACAAAAGACGAAAAAGCGGATGATCT
GAATGAAGACGTCACCGCAAGCCTGAGCGCCCTTTTTGCGATGTTGCCGGGTTTTGACAATA

CGCCCAAAGTGACTGATGCGCCGTCAACCGTGTTACCGACAGAGAAACCAACGCTCTTCACA
AAACTGACTTCTGAGCAACTCACAACAGCACAGCCTGATGACGCCCCCGGCACACCAGCTC
AGCCATTAACACCGCTGGTAGCAGAAGCCCAGAGTAAAGCGGAAGTCATCAGCACACCTTC
ACCGGTGACCGCTGCCGCCAGCCCGCTAATCACTCCACACCAGACACAGCCACTGCCCACCG
TCGCCGCACCTGTTTTGAGTGCACCGCTGGGTTCTCACGAATGGCAACAATCATTAAGCCAG
CATATTTCGCTGTTCACCCGCCAGGGGCAACAAAGTGCAGAGTTGCGTCTGCACCCGCAGGA
TTTAGGTGAAGTGCAAATCTCCCTCAAAGTGGATGATAACCAGGCGCAAATCCAGATGGTTT
CACCGCATCAGCATGTACGCGCCGCCCTGGAAGCAGCGCTGCCGGTACTGCGCACGCAGCT
GGCCGAAAGTGGCATTCAGTTAGGGCAAAGCAACATCAGTGGCGAAAGCTTTAGTGGTCAG
CAGCAGGCCGCTTCCCAGCAACAGCAAAGCCAACGCACAGCAAACCATGAACCTCTGGCGG
GGGAAGACGACGATACGCTTCCGGTTCCCGTCTCTTTACAAGGGCGTGTAACAGGCAACAGC
GGCGTTGATATTTTCGCCTAA

- NC_000913.3:2019618-2020082 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGACTGATTACGCGATAAGCAAGAAAAGCAAGCGATCGCTTTGGATCCCGATTCTGGTATT
CATTACCCTCGCGGCCTGTGCCAGCGCAGGTTACAGCTACTGGCATTCGCATCAGGTTGCCG
CTGACGACAAAGCGCAGCAACGCGTCGTGCCCTCACCGGTCTTCTACGCGCTGGATACCTTC
ACGGTCAATTTGGGCGATGCGGATCGCGTACTTTATATCGGCATAACCCTGCGCCTGAAAGA
TGAAGCTACCCGCTCGCGGCTGAGTGAGTATTTGCCGGAAGTCCGTAGTCGCTTGCTGTTAC
TGTTTTCGCGTCAGGATGCTGCCGTACTGGCGACAGAAGAAGGCAAGAAAAACCTGATTGC
CGAGATTAAAACCACACTTTCCACCCCGCTTGTTGCCGGGCAACCGAAACAGGATGTCACCG
ACGTGCTGTATACCGCTTTTATTCTGCGATAA

- NC_000913.3:2020087-2021091 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGGGCGATAGTATTCTTTCTCAAGCTGAAATTGATGCGCTGTTGAATGGTGACAGCGAAGT
CAAAGACGAACCGACAGCCAGTGTTAGCGGCGAAAGTGACATTCGTCCGTACGATCCGAAT
ACCCAACGACGGGTTGTGCGCGAACGTTTGCAGGCGCTGGAAATCATTAATGAGCGCTTTGC
CCGCCATTTTCGTATGGGGCTGTTCAACCTGCTGCGTCGTAGCCCGGATATAACCGTCGGGG
CCATCCGCATTCAGCCGTACCATGAATTTGCCCGCAACCTGCCGGTGCCGACCAACCTGAAC
CTTATCCATCTGAAACCGCTGCGCGGCACTGGGCTGGTGGTGTTCTCACCGAGTCTGGTGTTT
ATCGCCGTGGATAACCTGTTTGGCGGCGATGGACGCTTCCCGACCAAAGTGGAAGGTCGCG
AGTTTACCCATACCGAACAGCGCGTCATCAACCGCATGTTGAAACTGGCGCTTGAAGGCTAT
AGCGACGCCTGGAAGGCGATTAATCCGCTGGAAGTTGAGTACGTGCGTTCGGAAATGCAGG
TGAAATTTACCAATATCACCACCTCGCCGAACGACATTGTGGTTAACACGCCGTTCCATGTG
GAGATTGGCAACCTGACCGGCGAATTTAATATCTGCCTGCCATTCAGCATGATCGAGCCGCT
ACGGGAATTGTTGGTTAACCCGCCGCTGGAAAACTCGCGTAATGAAGATCAGAACTGGCGC
GATAACCTGGTGCGCCAGGTGCAGCATTCACAGCTGGAGCTGGTCGCCAACTTTGCCGATAT
CTCGCTACGCCTGTCGCAGATTTTAAAACTGAACCCCGGCGACGTCCTGCCGATAGAAAAAC
CCGATCGCATCATCGCCCATGTTGACGGCGTCCCGGTGCTGACCAGTCAGTATGGCACCCTC
AACGGTCAGTATGCGTTACGGATAGAACATTTGATTAACCCGATTTTAAATTCTCTGAACGA
GGAACAGCCCAAATGA

- NC_000913.3:2021088-2021501 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGAGTGACATGAATAATCCGGCCGATGACAACAACGGCGCAATGGACGATCTGTGGGCTG
AAGCGTTGAGCGAACAAAAATCAACCAGCAGCAAAAGCGCTGCCGAGACGGTGTTCCAGCA
ATTTGGCGGTGGTGATGTCAGCGGAACGTTGCAGGATATCGACCTGATTATGGATATTCCGG
TCAAGCTGACCGTCGAGCTGGGCCGTACGCGGATGACCATCAAAGAGCTGTTGCGTCTGACG
CAAGGGTCCGTCGTGGCGCTGGACGGTCTGGCGGGCGAACCACTGGATATTCTGATCAACG
GTTATTTAATCGCCCAGGGCGAAGTGGTGGTCGTTGCCGATAAATATGGCGTGCGGATCACC
GATATCATTACTCCGTCTGAGCGAATGCGCCGCCTGAGCCGTTAG

- NC_000913.3:2021504-2021869 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGAATAACCACGCTACTGTGCAATCTTCCGCGCCGGTTTCTGCTGCGCCACTGCTGCAGGT
GAGCGGCGCACTCATCGCCATTATTGCCCTGATCCTCGCTGCTGCCTGGCTGGTAAAACGGT
TGGGATTTGCCCCTAAACGCACTGGCGTTAACGGTCTGAAAATTAGCGCCAGTGCTTCACTG
GGCGCGCGTGAAAGGGTTGTGGTGGTCGATGTGGAAGATGCACGGCTGGTGCTCGGCGTTA
CCGCAGGTCAAATCAATCTGCTGCATAAACTTCCCCCTTCTGCACCAACGGAAGAGATACCG
CAGACCGATTTTCAGTCGGTCATGAAAAATTTGCTTAAGCGTAGCGGGAGATCCTGA

- NC_000913.3:2021869-2022606 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGCGTCGTTTATTGTCTGTCGCACCTGTCCTTCTCTGGCTGATTACGCCCCTCGCCTTCGCGC
AACTGCCGGGTATCACCAGCCAGCCGCTGCCTGGCGGTGGACAAAGCTGGTCGCTCCCGGTG
CAGACGCTGGTGTTCATCACCTCGTTGACGTTTATTCCGGCAATTTTACTGATGATGACCAGT
TTCACCCGCATCATCATTGTTTTTGGTTTATTGCGTAACGCGCTGGGAACACCCTCCGCGCCA
CCTAACCAGGTATTGCTGGGGCTGGCACTGTTTTTGACCTTTTTTATTATGTCACCGGTGATC
GACAAAATTTATGTAGATGCGTACCAGCCATTCAGCGAAGAGAAAATATCAATGCAGGAGG
CGCTGGAAAAAGGGGCGCAGCCGCTGCGTGAGTTTATGCTGCGTCAGACCCGTGAGGCAGA
TTTAGGGTTGTTTGCCAGACTGGCGAATACCGGCCCGTTGCAGGGACCTGAAGCCGTGCCGA
TGCGCATTTTGCTCCCGGCCTACGTGACCAGCGAGTTGAAAACCGCATTTCAGATAGGCTTC
ACGATTTTCATCCCTTTTTTGATTATCGACCTGGTGATAGCCAGCGTGTTGATGGCATTGGGG
ATGATGATGGTTCCCCCAGCCACCATTGCTCTGCCCTTTAAACTGATGCTGTTTGTACTGGTG
GATGGCTGGCAATTGCTGGTCGGTTCGCTGGCGCAGAGCTTTTACAGCTAG

- NC_000913.3:2022616-2022885 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGACACCTGAATCGGTCATGATGATGGGGACTGAAGCGATGAAAGTCGCGCTGGCACTGG
CTGCCCCGCTATTGTTGGTAGCGTTGGTCACGGGCCTTATCATCAGTATTTTGCAGGCCGCCA
CGCAGATTAACGAAATGACGCTGTCGTTTATTCCGAAAATCATCGCCGTATTTATCGCCATT
ATTATTGCCGGACCGTGGATGCTCAATCTGTTGCTGGATTACGTCCGCACCTTGTTCACTAAC
CTGCCGTATATCATCGGGTAG

- NC_000913.3:2022893-2023678 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGTTGCAGGTGACAAGCGAACAATGGCTATCCTGGTTAAACCTGTACTTCTGGCCGTTACT
GCGCGTGCTGGCGCTGATCTCCACCGCGCCGATTCTGAGCGAACGCAGCGTACCGAAACGG
GTAAAACTGGGTCTGGCAATGATGATCACGTTCGCCATTGCCCCATCATTACCTGCCAACGA
TGTTCCTGTTTTTTCGTTCTTTGCTCTGTGGCTGGCCGTGCAGCAGATCCTGATCGGCATTGC

GCTTGGTTTTACCATGCAATTTGCCTTTGCCGCTGTGCGAACCGCTGGCGAAATTATCGGTCT
GCAAATGGGGCTGTCATTTGCGACGTTTGTCGATCCGGCCAGCCATCTTAATATGCCCGTTTT
AGCGCGTATCATGGATATGCTGGCGTTACTGCTGTTCCTGACATTTAACGGTCATTTATGGTT
GATTTCACTGCTGGTCGATACCTTTCACACCCTGCCGATTGGTGGCGAACCGTTGAACAGCA
ATGCGTTTCTGGCACTCACCAAAGCAGGGAGTTTGATTTTCCTTAACGGGCTGATGCTGGCG
TTACCGCTCATTACTCTGCTGCTGACACTGAATCTGGCATTAGGTTTACTTAATCGTATGGCC
CCGCAATTATCCATTTTTGTTATTGGATTTCCATTAACTCTGACTGTCGGCATCTCTTTAATGG
CGGCATTAATGCCGTTAATTGCACCTTTTTGCGAACATTTATTCAGTGAAATTTTTAATTTGC
TGGCTGATATTATTAGTGAATTGCCATTAATATAA

- NC_000913.3:2437950-2438945 Escherichia coli str. K-12 substr. MG1655, complete genome

ATGATACAACCTATTTCCGGCCCTCCTCCTGGGCAACCACCAGGTCAGGGAGATAATCTGCC
GTCTGGCACGGGCAATCAGCCTTTATCCAGTCAGCAACGTACTTCGCTGGAAAGCTTAATGA
CGAAAGTGACCTCACTGACGCAACAGCAAAGAGCAGAACTGTGGGCGGGTATCAGGCACGA
TATTGGTCTGTCGGGAGATTCACCGCTGCTTTCGCGTCACTTCCCTGCCGCTGAGCATAATCT
GGCGCAACGTCTGCTGGCCGCGCAAAAAAGCCATTCTGCCCGCCAGCTTTTAGCGCAATTAG
GGGAGTATTTACGTCTGGGGAATAATCGTCAGGCGGTCACGGATTATATCCGTCATAACTTT
GGTCAGACGCCGCTGAATCAGCTCTCACCGGAGCAATTAAAAACCATTCTCACCCTGTTGCA
GGAAGGGAAGATGGTTATTCCGCAACCACAGCAGCGCGAGGCGACCGACCGTCCTTTATTA
CCGGCGGAGCACAATGCGCTAAAACAGCTGGTGACCAAACTTGCGGCGGCAACGGGGGAAC
CCAGCAAACAGATCTGGCAATCGATGCTGGAACTTTCCGGGGTGAAAGATGGCGAGTTAAT
TCCAGCGAAACTGTTTAACCATCTGGTGACCTGGCTACAGGCGCGTCAGACGCTAAGCCAGC
AAAATACGCCGACGCTGGAATCACTACAGATGACGCTAAAACAACCTTTAGATGCCAGTGA
ACTGGCGGCGTTATCGGCATATATCCAGCAAAAATATGGTCTTTCTGCGCAATCATCGCTTTC
TTCTGCCCAGGCCGAGGATATTCTTAATCAGCTTTATCAACGGCGGGTTAAAGGGATTGATC
CGCGTGTTATGCAACCGCTGCTTAATCCTTTTCCACCGATGATGGACACGTTGCAAAATATG
GCAACGCGTCCCGCGCTGTGGATACTGTTAGTCGCGATTATCCTGATGCTGGTCTGGCTGGTT
CGTTAA

Appendix 3

Statistical Summary and Result Tables for Tests

Statistical Summary of the Collected Data

Below are the tables that contain the data summary of the motility data collected for MG1655 and MDS42 in different medium compositions

Medium Composition	**Statistic**	**MG1655**	**MDS42**
M9 Salts + 10mM Glucose + 0.2% Casamino Acids	Anderson Darling Normality Test		
	A – squared	0.20	1.49
	P – value	0.867	<0.005
	General Statistics		
	Mean (µm/s)	3.0552	2.8138
	Standard Deviation	0.5559	0.5729
	Variance	0.3090	0.3282
	Skewness	-0.124491	-1.55632
	Kurtosis	0.542216	4.81442
	N	52	57
	Box – Plot		
	Minimum	1.4339	0.4593
	1st Quartile	2.7072	2.6356
	Median	3.0253	2.8270
	3rd Quartile	3.4551	3.1557
	Maximum	4.2487	3.7740

Table A3. 1: The data summary of the average velocity data for MG1655 and MDS42 in Medium composition A

Medium Composition	Statistic	MG1655	MDS42
M9 Salts + 10mM Glucose + No Casamino Acids	Anderson Darling Normality Test		
	A – squared	0.69	1.02
	P – value	0.067	0.010
	General Statistics		
	Mean (μm/s)	2.9091	3.0586
	Standard Deviation	0.3516	0.7552
	Variance	0.1236	0.5703
	Skewness	0.563930	-0.55639
	Kurtosis	-0.204304	5.95109
	N	51	49
	Box – Plot		
	Minimum	2.2642	0.0323
	1st Quartile	2.6345	2.6714
	Median	2.8162	3.0229
	3rd Quartile	3.1656	3.4739
	Maximum	3.8418	5.4933

Table A3. 2: The data summary of the average velocity data for MG1655 and MDS42 in Medium composition B

Medium Composition	Statistic	MG1655	MDS42
M9 Salts + 10uM Glucose + No Casamino Acids	Anderson Darling Normality Test		
	A – squared	0.36	0.30
	P – value	0.431	0.566
	General Statistics		
	Mean (μm/s)	3.1811	3.0658
	Standard Deviation	0.3703	0.3664
	Variance	0.1371	0.1343
	Skewness	-0.114165	-0.163381
	Kurtosis	0.778590	-0.804069
	N	48	50
	Box – Plot		
	Minimum	2.1653	2.3172
	1st Quartile	2.9989	2.7995
	Median	3.1747	3.0944
	3rd Quartile	3.3740	3.3691
	Maximum	4.0977	3.6817

Table A3. 3: The data summary of the average velocity data for MG1655 and MDS42 in Medium composition C

Medium Composition	Statistic	MG1655	MDS42
M9 Salts + No Glucose + No Casamino Acids	Anderson Darling Normality Test		
	A – squared	1.13	0.32
	P – value	0.005	0.525
	General Statistics		
	Mean (μm/s)	2.8056	2.5799
	Standard Deviation	0.4346	0.3121
	Variance	0.1888	0.0974
	Skewness	1.09989	0.439402
	Kurtosis	2.31833	0.288866
	N	50	51
	Box – Plot		
	Minimum	1.9812	1.9949
	1st Quartile	2.5122	2.3247
	Median	2.7653	2.6050
	3rd Quartile	2.9518	2.7566
	Maximum	4.2967	3.4004

Table A3. 4: The data summary of the average velocity data for MG1655 and MDS42 in Medium composition D

Medium Composition	Statistic	MG1655	MDS42
M9 Salts + No Glucose + 0.2% Casamino Acids	Anderson Darling Normality Test		
	A – squared	0.26	0.51
	P – value	0.700	0.193
	General Statistics		
	Mean (µm/s)	2.6426	2.6992
	Standard Deviation	0.2379	0.2576
	Variance	0.0566	0.0664
	Skewness	0.057146	-0.463546
	Kurtosis	-0.267845	-0.468343
	N	50	50
	Box – Plot		
	Minimum	2.1757	2.0636
	1st Quartile	2.4615	2.5129
	Median	2.6525	2.7046
	3rd Quartile	2.7980	2.9254
	Maximum	3.2329	3.1398

Table A3. 5: The data summary of the average velocity data for MG1655 and MDS42 in Medium composition E

Medium Composition	Strain	p-value	Outlier
M9 salts + 10mM Glucose + 0.2% Casamino Acids	MDS42	0.0004429406	0.459333
	MG1655	0.124	No
M9 salts + 10mM Glucose + No Casamino Acids	MDS42	0.0005089303	0.0323232
	MG1655	0.311	No
M9 salts + 10uM Glucose + No Casamino Acids	MDS42	1.000	No
	MG1655	0.209	No
M9 salts + No Glucose + No Casamino Acids	MDS42	0.338	No
	MG1655	0.013	4.29673
M9 salts + No Glucose + 0.2% Casamino Acids	MDS42	0.558	No
	MG1655	0.533	No

Table A3.6: The summary table of Grubb's test on the two strain samples in different medium compositions.

Groups 1	Group 2	Test Statistic	p-value
MDS42 – A	MDS42 – B	20.07	0.0000146144
MDS42 – A	MDS42 – C	43.20	0.0000000020
MG1655 – A	MG1655 – D	58.49	0.0000000000129012

Table A3.7: Levene's Test for Homogeneity of Variance

Label	Number of Observations	Mean	Standard Deviations	95% Confidence Interval
A	52	3.0552	0.5559	(2.9004, 3.2100)
B	51	2.9091	0.3516	(2.8103, 3.0080)
C	48	3.1811	0.3703	(3.0736, 3.2886)
D	50	2.8056	0.4346	(0.9785, 1.0627)
E	50	2.6426	0.2379	(2.5750, 2.7102)

Table A3.8: Means Table showing sample mean, and population mean in a 95% confidence interval, for MG1655 in all medium compositions. This test was performed for the average velocity values, and the units are µm/s

Label (MDS42)	Number of Observations	Mean	Standard Deviations	95% Confidence Interval
A	57	2.8138	0.5729	(2.245, 3.385)
B	49	3.059	0.755	(2.304, 3.814)
C	50	3.0658	0.3664	(2.9616, 3.1699)
D	51	2.5799	0.3121	(2.4921, 2.6677)
E	50	2.6992	0.2576	(2.6260, 2.7724)

Table A3.9: Means Table showing sample mean, and population mean in a 95% confidence interval, for MDS42 in all medium compositions. This test was performed for the average velocity values, and the units are µm/s

Factor	N	Mean	Grouping
C	48	3.1811	A
A	52	3.0552	A B
B	51	2.9091	B
E	50	2.6426	C
D	50	1.0206	B C

Table A3. 60: Games Howell's mean – based grouping of the samples in different medium composition for MG1655. This test was performed for the average velocity values, and the units are µm/s

Difference of Levels	Difference of Means	95% CI	Adjusted P-Value
B-A	-0.1461	(-0.4009, 0.1088)	0.503
C-A	0.1259	(-0.1354, 0.3873)	0.666
D-A	-0.2496	(-0.5235, 0.0244)	0.092
E-A	-0.4126	(-0.6481, -0.1770)	0.000
C-B	0.2720	(0.0700, 0.4739)	0.003
D-B	-0.1035	(-0.3223, 0.1153)	0.683
E-B	-0.2665	(-0.4326, -0.1004)	0.000
D-C	-0.3755	(-0.6018, -0.1492)	0.000
E-C	-0.5385	(-0.7149, -0.3621)	0.000
E-D	-0.1630	(-0.3587, 0.0327)	0.148

Table A3. 71: Pairwise comparison of the MG1655 samples in different medium composition to observe the significance of the difference in their means. This test was performed for the average velocity values.

Factor	N	Mean	Grouping
A	57	2.8138	A B
C	50	3.0658	A
E	50	2.6992	B
D	51	2.5799	B
B	49	3.059	A

Table A3. 82: Games Howell's mean – based grouping of the samples in different medium composition for MDS42. This test was performed for the average velocity values, and the units are µm/s

Difference of Levels	Difference of Means	95% CI	Adjusted P-Value
B-A	0.245	(-0.123, 0.612)	0.349
C-A	0.2520	(-0.0034, 0.5073)	0.055
D-A	-0.2339	(-0.4778, 0.0101)	0.067
E-A	-0.1146	(-0.3497, 0.1205)	0.654
C-B	0.007	(-0.328, 0.342)	1.000
D-B	-0.479	(-0.805, -0.152)	0.001
E-B	-0.359	(-0.680, -0.039)	0.020
D-C	-0.4859	(-0.6742, -0.2975)	0.000
E-C	-0.3666	(-0.5430, -0.1901)	0.000
E-D	0.1193	(-0.0388, 0.2774)	0.230

Table A3. 93: Pairwise comparison of the MDS42 samples in different medium composition to observe the significance of the difference in their means

Factor	N	Mean	Grouping					
A-MDS42	57	2.8138		B	C	D		
C-MG1655	48	3.1811	A					
C-MDS42	50	3.0658	A	B				
A-MG1655	52	3.0552	A	B				
B-MG1655	51	2.9091		B				
E-MDS42	50	2.6992			C	D		
E-MG1655	50	2.6426				D		
D-MDS42	51	2.5799				D		
D-MG1655	50	2.8056		B	C	D		
B-MDS42	49	3.059	A	B	C			F

Table A3. 104: Games Howell's mean–based grouping of the samples in selected medium composition for MG1655 & MDS42. This test was performed for the average velocity values, and the units are µm/s

Difference of Levels	Difference of Means	95% CI	Adjusted P-Value
A-MDS42-A-MG1655	-0.241	(-0.591, 0.108)	0.442
B-MDS42-B-MG1655	0.149	(-0.239, 0.538)	0.959
C-MDS42-C-MG1655	-0.1153	(-0.3564, 0.1257)	0.86831163
D-MDS42-D-MG1655	-0.2257	(-0.4705, 0.0191)	0.097
E-MDS42-E-MG1655	0.0566	(-0.1040, 0.2171)	0.97904382

Table A3. 115: Pairwise comparison of the MG1655 and MDS42 average velocity in the same medium composition to observe the significance of the difference in their means

Label	N	Median	Mean Rank	Z-Value
A-MDS42	57	2.82699	254.0	-0.03
A-MG1655	52	3.02530	308.6	2.81
B-MDS42	49	3.02289	302.8	2.42
B-MG1655	51	2.81616	265.9	0.59
C-MDS42	50	3.09436	324.1	3.53
C-MG1655	48	3.17468	363.6	5.41
D-MDS42	51	2.60501	146.4	-5.55
D-MG1655	50	2.76527	223.8	-1.56
E-MDS42	50	2.70456	194.0	-3.07
E-MG1655	50	2.65252	166.9	-4.45
Overall	508		254.5	

Table A3. 126: Strain-Medium Composition pairs ranked according to the cumulative individual ranks (Kruskal Wallis Test). This test was performed for the average velocity values, and the units are µm/s

Appendix 4

Statistical Summary and Test results

Statistical Summary of the Collected Data (MG1655)

Below id the statistical summary of the chemotaxis average velocity data collected for MG1655 under 10µM and 10mM gradients.

Statistic	M9 Salts + 10mM Glucose + No Casamino Acids	M9 Salts + 10mM Glucose + No Casamino Acids
Anderson Darling Normality Test		
A – squared	1.51	7.04
P – value	<0.005	<0.005
General Statistics		
Mean (µm/s)	4.0096	3.4923
Standard Deviation	1.1865	2.0117
Variance	1.4079	4.0467
Skewness	-0.36520	2.87297
Kurtosis	-1.20033	8.78564
N	52	50
Box – Plot		
Minimum	1.6886	0.7169
1st Quartile	2.8127	2.6116
Median	4.3590	3.0172
3rd Quartile	5.0519	3.4385
Maximum	6.0375	11.8941

Table A4. 1: Statistical summary of the average velocity data of MG1655 under 10uM and 10mM gradients

Goodness – of – Fit Test MG1655

Goodness of Fit Test (M9 salts + 10uM Glucose + No Casamino Acids)-A

Distribution	AD	P	LRT P
Normal	7.041	<0.005	
Box-Cox Transformation	3.495	<0.005	
Lognormal	3.495	<0.005	
3-Parameter Lognormal	3.468	*	0.972409
Exponential	10.367	<0.003	
2-Parameter Exponential	8.225	<0.010	0.000005
Weibull	5.753	<0.010	
3-Parameter Weibull	5.119	<0.005	0.007212
Smallest Extreme Value	9.456	<0.010	
Largest Extreme Value	3.376	<0.010	
Gamma	4.486	<0.005	
3-Parameter Gamma	5.154	*	1.000000
Logistic	4.214	<0.005	
Loglogistic	1.986	<0.005	
3-Parameter Loglogistic	1.932	*	0.639420
Johnson Transformation	0.390	0.37094658	

Box-Cox transformation: $\lambda = 0$
Johnson transformation function: -0.524651 + 0.678146 × Asinh((X-2.67208) / 0.247142)

Goodness of Fit Test (M9 salts + 10mM Glucose + No Casamino Acids)-B

Distribution	AD	P	LRT P
Normal	1.506	<0.005	
Box-Cox Transformation	1.506	<0.005	
Lognormal	2.096	<0.005	
3-Parameter Lognormal	1.553	*	0.005795
Exponential	11.758	<0.003	
2-Parameter Exponential	5.288	<0.010	0.000000
Weibull	1.537	<0.010	
3-Parameter Weibull	1.149	<0.005	0.319603
Smallest Extreme Value	1.055	<0.010	
Largest Extreme Value	1.963	<0.010	
Gamma	1.908	<0.005	
3-Parameter Gamma	1.747	*	0.135559
Logistic	1.524	<0.005	
Loglogistic	1.997	<0.005	
3-Parameter Loglogistic	1.526	*	0.015100

Box-Cox transformation: $\lambda = 1$
Johnson transformation function: Not Possible

Label	N	Median	Mean Rank	Z-Value
A–MG1655	52	4.35897	60.6	3.17
B-MG1655	50	3.01722	42.0	-3.17
Overall	102		51.5	

Table A4. 2:Strain-Medium Composition pairs ranked according to the cumulative individual ranks (Kruskal Wallis Test). This was done for the average velocity values recorded in different conditions. The units of median are $\mu m/s$

Appendix 5

MATLAB Codes

1. Velocity Plot Code – MATLAB

```
A = xlsread ('Framewise_All.xlsx', 'A1:M9112');

T = a(:,1);

s = a(:,2);

x= a(:,3);

y = a(:,4);

D = a(:,5);

V = a(:,6);

z = [s,V];

figure (1)

for I = 303:400

hi = plot ([z(i-1,1) z(i,1)], [z(i-1,2) z(i,2)], 'g.-')

hold on

end

hold on

for i = 203: 300

me = plot ([z(i-1,1) z(i,1)], [z(i-1,2) z(i,2)], 'b.-')

hold on

end

hold on

for i= 703: 800

lo = plot ([z(i-1,1) z(i,1)], [z(i-1,2) z(i,2)], 'r.-')
```

```
hold on

end

title ('MDS42-E')

xlabel ('Number of Frames')

ylabel ('Velocity')

legend ([hi,me,lo],'Max', 'Mean', 'Min')
```

2. Angle Change Code – MATLAB

```
a = xlsread ('MDS42_NoGCA.xlsx', 'A1:R5001');

T = a(:,1);

s = a(:,2);

x = a(:,3);

y = a(:,4);

z = [x,y];

for i = 2 : 4999

m1 = ((z(i,2)-z(i-1,2))/(z(i,1)-z(i-1,1)));

m2 = ((z(i+1,2)-z(i,2))/(z(i+1,1)-z(i,1)));

thed = atand ((m2-m1)/(1+ (m1*m2)));

arrd(i) = reshape (thed,[],1);

hold on

end

arrd= arrd'
```

3. Trajectory Plot Code – MATLAB

```
a = xlsread ('Framewise_All.xlsx', 'A1:M9112');

T = a(:,1);

s = a(:,2);

x = a(:,3);

y = a(:,4);

z = [x,y];

figure (1)

for i = 3102: 3199

plot ([z(i-1,1) z(i,1)], [z(i-1,2) z(i,2)], '.-');

hold on

end

title ('MG1655 Trajectory_ 01')

xlabel ('X')

ylabel ('Y')

figure (2)

for i = 8514: 8611

plot ([z(i-1,1) z(i,1)], [z(i-1,2) z(i,2)], '.-');

hold on

end

title ('MG1655 Trajectory_ 02')

xlabel ('X')

ylabel ('Y')

axis ([301 310 325 337])
```

4. Gradient Curve Code – MATLAB

```
X= 0:1500

Y = -0.2002.*X + 1277.4

Y = plot (X,Y)

hold on

Y1 = -0.241.*X + 1307

y1 = plot (X,Y1)

hold on

Y2 = -0.294.*X + 1333

y2 = plot (X,Y2)


title ('Fluorescein intensity across channel length')

xlabel ('X')

ylabel ('Y')

legend ([y,y1,y2],'-0.2002.*X + 1277.4', '-0.241.*X + 1307', '-0.294.*X + 1333')
```

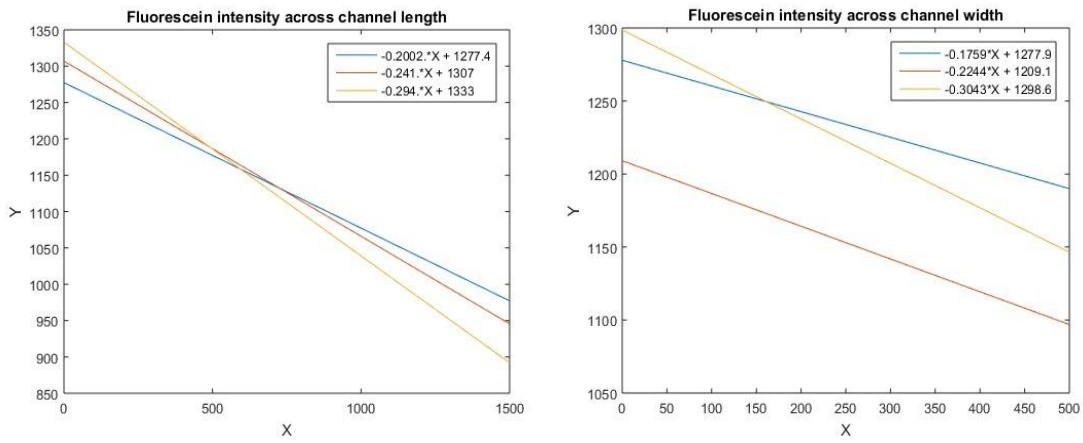


Figure A5. 1: The plots of the linear gradients recorded in Table 4.2 along the channel length and across the channel width. The x – axis scale is in pixels. 1 pixel - 0.4µm.

LIST OF REFERENCES

Adler, J. and Templeton, B. 1967. “The Effect of Environmental Conditions on the motility of Escherichia coli.” *Journal of General Microbiology* 46: 175 - 184.

Adler, J. 1966. “Chemotaxis in Bacteria.” *Science* 153 (3737): 708 - 716.

Adler, J. 1966 (2). “Effect of Amino Acids and Oxygen on Chemotaxis in Escherichia coli.” *Journal of Bacteriology* 92 (1): 121 - 129.

Ahmed, T. and Stocker, R. 2008. “Experimental Verification of the Behavioral Foundation of Bacterial Transport Parameters Using Microfluidics.” *Biophysical journal* 95 (9): 4481 - 4493.

Ahmed, T., Shimizu, T.S. and Stocker, R. 2010. “Bacterial chemotaxis in linear and nonlinear steady microfluidic gradients.” *Nano Letters* 10 (9): 3379 - 3385.

Al-Maliki, J.Q. and Al-Maliki, A.J.Q. 2015. “The Processes and Technologies of 3D Printing.” *International Journal of Advances in Computer Science and Technology* 4 (10): 161 - 165.

Alon, U., Surette, M.G., Barkai, N. and Leibler, S. 1999. “Robustness in bacterial chemotaxis.” *Nature* 397 (6715): 168-171.

Altissimo, M. 2010. “E-beam lithography for micro-/nanofabrication.” *Biomicrofluidics* 4: 026503-1 - 6.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. 1997. “Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.” *Nucleic Acids Research* 25 (17): 3389 - 3402.

Ambagaspitiye, S., Sudarshan, S., Hogins, J., McDill, P., De Nisco, N.J., Zimmern, P.E. and Reitzer, L. 2019. “FIMBRIAE AND FLAGELLA MEDIATED SURFACE MOTILITY AND THE EFFECT OF GLUCOSE ON NONPATHOGENIC AND UROPATHOGENIC ESCHERICHIA COLI.” *Biorxiv.*

Arafat, A., Giesbers, M., Rosso, M., Sudhölter, E.J., Schroën, K., White, R.G., Yang, L., Linford, M.R. and Zuilhof, H. 2007. “Covalent biofunctionalization of silicon nitride surfaces.” *Langmuir* 23 (11): 6233-44.

Arafat, A., Schroën, K., de Smet, L.C., Sudhölter, E.J. and Zuilhof, H. 2004. “Tailor-made functionalization of silicon nitride surfaces.” *Journal of the American Chemical Society* 126 (28): 8600-1.

Bah, T. 2010. *Inkscape Guide to a Vector Drawing Program.* 3. Boston: Pearson.

Bahreyni, B. 2009. “Chapter 2 - Microfabrication.” In *Fabrication and Design of Resonant Microdevices*, 9 - 46. William Andrew Applied Science Publishers.

Barboriak, D.P., Padua, A.O., York, G.E. and MacFall, J.R. 2005. “Creation of DICOM—Aware Applications Using ImageJ.” *Journal of Digital Imaging* 18 (2): 91 - 99.

Bardy, S.L., Ng, S.Y. and Jarrell, K.F. 2003. “Prokaryotic motility structures.” *Microbiology* 149 (2): 295 - 304.

Bees, M.A., Andresen, P., Mosekilde, E. and Givskov, M. 2002. “Quantitative effects of medium hardness and nutrient availability on the swarming motility of Serratia liquefaciens.” *Bulletin of mathematical Biology* 64: 565-587.

Berg, H. C. 1971. “How to track bacteria.” *Review of Scientific Instruments* 42: 868 - 871.

Berg, H.C. and Brown, D.A. 1972. “Chemotaxis in Escherichia coli analysed by Three-dimensional Tracking.” *Nature* 239: 500 - 504.

Berg, H.C. and Brown, D.A. 1972. “Chemotaxis in Rhodobacter sphaeroides analysed by three - dimensional tracking.” *Nature* 239: 500 - 504.

Berry, R.M. and Armitage, J.P. 2000. “Response kinetics of tethered Rhodobacter sphaeroides to changes in light intensity.” *Biophysical Journal* 78: 1207 - 1215.

Bhushan, B. and Caspers, M. 2017. “An overview of additive manufacturing (3D printing) for microfabrication.” *Microsystem Technologies* 23 (4): 1117 - 1124.

Bi, S., Pollard, A.M., Yang, Y., Jin, F. and Sourjik, V. 2016. “Engineering Hybrid Chemotaxis Receptors in Bacteria.” *ACS Synthetic Biology* 5 (9): 989-1001.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F. and Gregor, J. 1997. “The Complete Genome Sequence of Escherichia coli K-12.” *Science* 227 (5331): 1453-1462.

Blocker, A., Komoriya, K. and Aizawa, S.I. 2003. “Type III secretion systems and bacterial flagella: Insights into their function from structural similarities.” *PNAS* 100 (6): 3027 - 3030.

Bochner, B.R., Gadzinski, P. and Panomitros, E. 2001. “Phenotype microarrays for high-throughput phenotypic testing and assay of gene function.” *Genome Research* 11 (7): 1246-1255.

Boehm, A., Kaiser, M., Li, H., Spangler, C., Kasper, C.A., Ackermann, M., Kaever, V., Sourjik, V., Roth, V. and Jenal, U. 2010. “Second messenger-mediated adjustment of bacterial swimming velocity.” *Cell* 141: 107-116.

Boehm, A., Kaiser, M., Li, H., Spangler, C., Kasper, C.A., Ackermann, M., Kaever, V., Sourjik, V., Roth, V. and Jenal, U. 2010. “Second messenger-mediated adjustment of bacterial swimming velocity.” *Cell* 141 (1): 107 - 116.

Bren, A., Park, J.O., Towbin, B.D., Dekel, E., Rabinowitz, J.D. and Alon, U. 2016. “Glucose becomes one of the worst carbon sources for E.coli on poor nitrogen sources due to suboptimal levels of cAMP.” *Scientific Reports* 6 (24834): 1 - 10.

Brent, R. 1992. “Section 1: Escherichia coli.” In *Short Protocols in Molecular Biology*, 1.1-1.52. New York: John Wiley & Sons.

Brown, D.A. and Berg, H.C. 1974. “Temporal stimulation of chemotaxis in Rhodobacter sphaeroides.” *PNAS* 71: 1388 - 1392.

Brown, M.B. and Forsythe, A.B. 1974. “Robust Tests for the Equality of Variances.” *Journal of the American Statistical Association* 69 (346): 364 - 367.

Cahan, P. and Daley, G.Q. 2013. “Origins and implications of pluripotent stem cell variability and heterogeneity.” *Nature Reviews Molecular Cell Biology* 14: 357-368.

Caiazza, N.C., Shanks, R.M. and O'toole, G.A. 2005. “Rhamnolipids Modulate Swarming Motility Patterns of Pseudomonas aeruginosa.” *Journal of Bacteriology* 187 (21): 7351-7361.

Caldwell, M.B., Guerry, P., Lee, E.C., Burans, J.P. and Walker, R.I. 1985. “Reversible expression of flagella in Campylobacter jejuni.” *Infection and Immunity* 50 (3): 941-943.

Casella, G. and Berger, R.L. 2001. *Statistical Inference.* 2. Duxbury Press.

Chia, H.N. and Wu, B.M. 2015. “Recent advances in 3D printing of biomaterials.” *Journal of Biological Engineering* 9 (4): 1-14.

Consortium, U. 2015. “UniProt: a hub for protein information.” *Nucleic Acids Research* 43 (D1): D204 - D212.

Day, R.W. and Quinn, G.P. 1989. “Comparisons Of Treatments After An Analysis Of Variance In Ecology.” *Ecological Monographs* 59 (4): 433-463.

Delcenserie, V., LaPointe, G., Charaslertrangsi, T., Rabalski, A. and Griffiths, M.W. 2012. “Glucose Decreases Virulence Gene Expression of Escherichia coli O157:H7.” *Journal of Food Protection* 75 (4): 748 - 752.

Derrick, B., Toher, D. and White, P. 2016. “Why Welchs test is Type I error robust?” *The quantitative methods for Psychology* 12 (1): 30-38.

Drews, G. 2005. “Contributions of Theodor Wilhelm Engelmann on phototaxis, chemotaxis and photosynthesis.” *Photosynthesis research* 83: 25 - 34.

Duan, Q., Zhou, M., Zhu, L. and Zhu, G. 2012. “Flagella and bacterial pathogenicity.” *Journal of Basic Microbiology* 52: 1-8.

Duffy, K.J. and Ford, R.M. 1997. “Turn angle and run time distributions characterize swimming behavior for Pseudomonas putida.” *Journal of Bacteriology* 179: 1428 - 1430.

Elbing, K. and Brent, R. 2002. “Media Preparation and Bacteriological Tools.” *Current Protocols in Molecular Biology* 1.1.1 - 1.1.7.

Eliceiri, K.W. and Rueden, C. 2005. “Tools for Visualizing Multidimensional Images from Living Specimens.” *Photochemistry and Photobiology* 81 (5): 1116 - 1122.

Fehér, T., Papp, B., Pál, C. and Pósfai, G. 2007. “Systematic Genome Reductions: Theoretical and Experimental Approaches.” *Chemical Reviews* 107 (8): 3498 - 3513.

Fenchel, T. 2002. “Microbial behavior in a heterogeneous world.” *Science* 296 (5570): 1068-1071.

Fligner, M.A. and Rust, S.W. 1982. “A modification of Mood’s median test for the generalized Behrens-Fisher problem.” *Biometrika* 69 (1): 221 - 226.

Ford, R.M., Phillips, B.R., Quinn, J.A. and Lauffenburger, D.A. 1991. “Stopped-flow chamber and image analysis system for quantitative characterization of bacterial population migration: Motility and chemotaxis ofEscherichia coli K12 to fucose.” *Microbial Ecology* 22 (1): 127 - 138.

Fredens, J., Wang, K., de la Torre, D., Funke, L.F., Robertson, W.E., Christova, Y., Chia, T., Schmied, W.H., Dunkelmann, D.L., Beránek, V. and Uttamapinant, C. 2019. “Total synthesis of Escherichia coli with a recoded genome.” *Nature* 569 (7757): 514 - 518.

Frymier, P.D., Ford, R.M., Berg, H.C. and Cummings, P.T. 1995. “Three-dimensional tracking of motile bacteria near a solid planar surface.” *PNAS* 92: 6195 - 6199.

Games, P.A. and Howell, J.F. 1976. “Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances: A Monte Carlo Study.” *Journal of Educational Statistics* 1 (2): 113 - 125.

George, D. and Mallery, P. 2010. *SPSS for Windows Step by Step: A Simple Guide and Reference.* 10. Boston: Pearson.

Gonzalez, R., Quirós, M. and Morales, P. 2013. “Yeast respiration of sugars by non-Saccharomyces yeast species: a promising and barely explored approach to lowering alcohol content of wines.” *Trends in Food Science and Technology* 29: 55 - 61.

Grubbs, F.E. 1950. “Sample Criteria for Testing Outlying Observations.” *The Annals of Mathematical Statistics* 21 (1): 27 - 59.

Hacker, J., Hentschel, U. and Dobrindt, U. 2003. “Prokaryotic Chromosomes and Disease.” *Science* 301 (5634): 790-793.

Hanahan, D. and Weinberg, R.A. 2000. “The hallmarks of Cancer.” *Cell* 100 (1): 57 - 70.

Hansen, C.H., Sourjik, V. and Wingreen, N.S. 2010. “A dynamic-signaling-team model for chemotaxis receptors in Escherichia coli.” *PNAS* 107 (40): 17170-17175.

Harshey, R.M. and Matsuyama, T. 1994. “Dimorphic transition in Escherichia coli and Salmonellatyphimurium: surface-induced differentiation into hyperflagellate swarmer cells.” *Proceedings of the National Academy of Sciences* 91: 8631-8635.

Hashimoto, M., Ichimura, T., Mizoguchi, H., Tanaka, K., Fujimitsu, K., Keyamura, K., Ote, T., Yamakawa, T., Yamazaki, Y., Mori, H. and Katayama, T. 2005. “Cell size and nucleoid organization of engineered Escherichia coli cells with a reduced genome.” *Molecular Microbiology* 55 (1): 137 - 149.

Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., Ohtsubo, E., Baba, T., Wanner, B.L., Mori, H. and Horiuchi, T. 2006. “Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110.” *Molecular Systems Biology* 2 (1).

Hazelbauer, G.L., Falke, J.J. and Parkinson, J.S. 2008. “Bacterial chemoreceptors: High performance signalling in networked arrays.” *Trends in Biochemical Science* 33 (1): 9 - 19.

Hishikawa, Y., Kakino, Y., Tsukamoto, H., Tahara, K., Onodera, R. and Takeuchi, H. 2016. “Control of Drug Diffusion Behavior of Xanthan and Locust Bean Gum Gel by Agar Gel.” *Chemical & Pharmaceutical Bulletin* 64 (10): 1450 - 1457.

Iacoviello, M.P. and Rubin, S.A. 2001. “Sterile Preparation of Antibiotic selective LB Agar plates using a Microwave oven.” *BioTechniques* 30: 962-965.

Inoue, T., Shingaki, R., Hirose, S., Waki, K., Mori, H. and Fukui, K. 2007. “Genome-Wide Screening of Genes Required for Swarming Motility in Escherichia coli K-12.” *Journal of Bacteriology* 189 (3): 950-957.

Jahid, I.K., Lee, N.Y., Kim, A. and Ha, S.D. 2013. “Influence of Glucose Concentrations on Biofilm Formation,Motility, Exoprotease Production, and Quorum Sensing in Aeromonas hydrophila.” *Journal of Food Protection* 76 (2): 239-247.

Jan, S.L. and Shieh, G. 2014. “Sample size determinations for Welch's test in one-way heteroscedastic ANOVA.” *The Britishh Journal of Mathematical and Statistical Psychology* 67 (1): 72-93.

Jasveer, S. and Jianbin, X. 2018. “Comparison of Different Types of 3D Printing Technologies.” *International Journal of Scientific and Research Publications* 8 (4).

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. and Madden, T.L. 2008. “NCBI BLAST: a better web interface.” *Nucleic Acids Research* 36 (2): W5 - W9.

Julius Adler, Gerald L. Hazelbauer, and M. M. Dahl. 1973. “Chemotaxis Toward Sugars in Escherichia coli.” *Journal of Bacteriology* 115 (3): 824 - 847.

Kalinin, Y.V., Jiang, L., Tu, Y. and Wu, M. 2009. “Logarithmic Sensing in Escherichia coli Bacterial Chemotaxis.” *Biophysical Journal* 96: 42439 - 42448.

Karbalaei, A. and Cho, H.J. 2018. “Microfluidic Devices Developed for and Inspired by Thermotaxis and Chemotaxis.” *Micromachines* 9 (149): 1-28.

Karcagi, I., Draskovits, G., Umenhoffer, K., Fekete, G., Kovács, K., Méhi, O., Balikó, G., Szappanos, B., Györfy, Z., Fehér, T. and Bogos, B. 2016. “Indispensability of Horizontally Transferred Genes and Its Impact on Bacterial Genome Streamlining.” *Molecular Biology and Evolution* 33 (5): 1257 - 1269.

Karcagi, I., Draskovits, G., Umenhoffer, K., Fekete, G., Kovács, K., Méhi, O., Balikó, G., Szappanos, B., Györfy, Z., Fehér, T. and Bogos, B. 2016. “Indispensability of Horizontally Transferred Genes and Its Impact on Bacterial Genome Streamlining.” *Molecular Biology and Evolution : Oxford Journals* 33 (5): 1257 - 1269.

Keller, E.F. and Segel, L.A. 1971. “Model for Chemotaxis.” *J. theor. Biol.* 30: 225-234.

Kessler, D.A., Austin, R.H. and Levine, H. 2014. “Resistance to Chemotherapy: Patient Variability and Cellular Heterogeneity.” *Cancer Research* 74 (17): 4663-4670.

Kim, H.Y. 2013. “Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis.” *Restorative Dentistry and Endodontics* 38 (1): 52-54.

Kjelleberg, S., Humphrey, B.A. and Marshall, K.C. 1982. “Effect of interfaces on small, starved marine bacteria.” *Applied Environmental Microbiology* 43: 1166-1172.

Ko, M. and Park, C. 2000. “Two novel flagellar components and H-NS are involved in the motor function of Escherichia coli.” *Journal of Molecular Biology* 303 (3): 371-382.

Kofoid, E.C. and Parkinson, J.S. 1988. “Transmitter and receiver modules in bacterial signaling proteins.” *Proceedings of the National Academy of Sciences of the United States of America* 85: 4981 - 4985.

Köhler, T., Curty, L.K., Barja, F., Van Delden, C. and Pechère, J.C. 2000. “Swarming of Pseudomonas aeruginosa Is Dependent on Cell-to-Cell Signaling and Requires Flagella and Pili.” *Journal of bacteriology* 182 (21): 5990-5996.

Kolisnychenko, V., Plunkett, G., Herring, C.D., Fehér, T., Pósfai, J., Blattner, F.R. and Pósfai, G. 2002. “Engineering a Reduced Escherichia coli Genome.” *Genome Research* 12 (4): 640-647.

Kruskal, W.H. and Wallis, W.A. 1952. “Use of ranks in one criterion variance analysis.” *Journal of American Statistical Association* 47: 583 - 621.

Kural, C., Serpinskaya, A.S., Chou, Y.H., Goldman, R.D., Gelfand, V.I. and Selvin, P.R. 2007. "Tracking melanosomes inside a cell to study molecular motors and their interaction." *PNAS* 104 (13): 5378 - 5382.

Lai, H.C., Shu, J.C., Ang, S., Lai, M.J., Fruta, B., Lin, S., Lu, K.T. and Ho, S.W. 1997. "Effect of Glucose Concentration on Swimming Motility in Enterobacteria." *Biochemical and Biophysical Research Communications* 231 (3): 692-695.

Lane, N. 2015. "The unseen world: reflections on Leeuwenhoek (1677) 'Concerning little animals'." *Philosophical Transactions of the Royal Society Biological Sciences* 370 (1666).

Lauga, E., DiLuzio, W.R., Whitesides, G.M. and Stone, H.A. 2006. "Swimming in circles: motion of bacteria near solid boundaries." *Biophysical Journal* 90: 400 - 412.

Leeuwenhoek, A. 1677. "Observations, communicated to the publisher by Mr. Antony van Leewenhoeck, in a dutch letter of the 9th Octob. 1676. here English'd: concerning little animals by him observed in rain-well-sea- and snow water; as also in water wherein pepper had lain infus." *The Royal Society - Philosophical Transactions* 12 (133).

Li, C.O.N.G.Y.I., Louise, C.J., Shi, W.E.N.Y.U.A.N. and Adler, J. 1993. "Adverse Conditions Which Cause Lack of Flagella in Escherichia coli." *Journal of Bacteriology* 175 (8): 2229-2235.

Ling, H., Kang, A., Tan, M.H., Qi, X. and Chang, M.W. 2010. "The absence of the luxS gene increases swimming motility and flagella synthesis in Escherichia coli K12." *Biochemical and Biophysical Research Communications* 410 (4): 521-526.

Liu, C., Zheng, H., Yang, M., Xu, Z., Wang, X., Wei, L., Tang, B., Liu, F., Zhang, Y., Ding, Y. and Tang, X. 2015. "Genome analysis and in vivo virulence of porcine extraintestinal pathogenic Escherichia coli strain PCN033." *BMC Genomics* 16 (717): 1890 - 99.

Liu, F., Fu, J., Liu, C., Chen, J., Sun, M., Chen, H., Tan, C. and Wang, X. 2017. "Characterization and distinction of two flagellar systems in extraintestinal pathogenic Escherichia coli PCN033." *Mcrobiological Research* 196: 69 -79.

Liu, X. and Matsumura, P. 1995. "An alternative sigma factor controls transcription of flagellar class-III operons in Escherichia coli: gene sequence, overproduction, purification and characterization." *Gene* 164 (1): 81-84.

Liu, Y., Butler, W.B. and Pappas, D. 2012. "Spatially selective reagent delivery into cancer cells using a two-layer microfluidic culture system." *Anal Chim Acta* 743: 125 - 130.

Liu, Z. and Papadopoulos, K.D. 1996. "A method for measuring bacterial chemotaxis parameters in a microcapillary." *Biotechnology and Bioengineering* 51: 120 - 125.

Lopes, J.G. and Sourjik, V. 2018. "Chemotaxis of Escherichia coli to major hormones and polyamines present in human gut." *The ISME Journal* 12: 2736 - 2747.

Løvdok, L., Bentele, K., Vladimirov, N., Müller, A., Pop, F.S., Lebiedz, D., Kollmann, M. and Sourjik, V. 2009. "Role of Translational Coupling in Robustness of Bacterial Chemotaxis Pathway." *PLoS Biology* 7 (8): e1000171.

Løvdok, L., Kollmann, M. and Sourjik, V. 2007. "Co-expression of signaling proteins improves robustness of the bacterial chemotaxis pathway." *Hournal of Biotechnology* 129 (2): 173-180.

Macnab, R.M. and Koshland, D.E. 1972. “The Gradient-Sensing Mechanism in Bacterial Chemotaxis.” *Proceedings of the National Academy of Sciences of the United States of America* 69 (9): 2509 - 2512.

Manson, M.D., Blank, V., Brade, G. and Higgins, C.F. 1986. “Peptide Chemotaxis in E.coli involves the Tap Signal transducer and the dipeptide permease.” *Nature* 321: 253 - 256.

McCarter, L.L. 2005. “Multiple Modes of Motility: a Second Flagellar System in Escherichia coli.” *Journal of Bacteriolgy* 187 (4): 1207 - 1209.

McKight, P.E. and Najab, J. 2010. “Kruskal-Wallis Test.” In *The Corsini Encyclopedia of Psychology*, edited by Irving B. Weiner and W. Edward Craighead. John Wiley & Sons, Inc.

Mitchell, J.G. and Kogure, K. 2006. “Bacterial motility: links to the environment and a driving force for microbial physics.” *FEMS Microbiology Ecology* 55 (1): 3 - 16.

Mizan, M.F.R., Jahid, I.K., Kim, M., Lee, K.H., Kim, T.J. and Ha, S.D. 2016. “Variability in biofilm formation correlates with hydrophobicity and quorum sensing among Vibrio parahaemolyticus isolates from food contact surfaces and the distribution of the genes involved in biofilm formation.” *Biofouling* 32 (4): 497-509.

Moder, K. 2010. “Alternatives to F-Test in One Way ANOVA in case of heterogeneity of variances (a simulation study).” *Psychological Test and Assessment Modeling* 52 (4): 343-353.

Moser, B.K. and Stevens, G.R. 1992. “Homogeneity of Variance in the Two-Sample Means Test.” *The American Statistician* 46 (1): 19-21.

Muzaffar, B. 2016. “The Development and Validation of a Scale to Measure Training Culture: The TC Scale.” *Journal of Culture, Society and Development* 23: 49 - 58.

Nachamkin, I.R.V.I.N.G., Yang, X.H. and Stern, N.J. 1993. “Role of Campylobacter jejuni flagella as colonization factors for threeday-old chicks: analysis with flagellar mutants.” *Applied and Environmental Microbiology* 59 (5): 1269 - 1273.

Neidhardt, F.C., Bloch, P.L. and Smith, D.F. 1974. “Culture Medium for Enterobacteria.” *Journal of Bacteriology* 119 (3): 736 - 747.

Nelson, L.S. 1998. “The Anderson Darling Test for Normality.” *Journal Of Quality Technology* 30 (3): 298-299.

Neumann, S., Grosse, K. and Sourjik, V. 2012. “Chemotactic signaling via carbohydrate phosphotransferase systems in Escherichia coli.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (30): 12159 - 12164.

Niepel, M., Spencer, S.L. and Sorger, P.K. 2009. “Non-genetic cell-to-cell variability and the consequences for pharmacology.” *Current Opinion in Chemical Biology* 13 (5-6): 556 - 561.

Nieto, V., Partridge, J.D., Severin, G.B., Lai, R.Z., Waters, C.M., Parkinson, J.S. and Harshey, R.M. 2019. “Under Elevated c-di-GMP in Escherichia coli, YcgR Alters Flagellar Motor Bias and Speed Sequentially, with Additional Negative Control of the Flagellar Regulon via the Adaptor Protein RssB.” *Journal of Bacteriology* 202 (1): 1 - 14.

Noether, G. E. 1967. *Elements of Nonparametric Statistics.* Vol. 133, 94 - 97. Wiley & Sons.

Nuccio, S.P. and Bäumler, A.J. 2007. “Evolution of the Chaperone/Usher Assembly Pathway: Fimbrial Classification Goes Greek.” *Microbiology and Molecular Biology Reviews* 71 (4): 551 - 575.

Ortega, Á., Zhulin, I.B. and Krell, T. 2017. “Sensory Repertoire of Bacterial Chemoreceptors.” *Microbiology and Molecular Biology Reviews* 81 (4): 1 - 28.

Park, S., Park, Y.H., Lee, C.R., Kim, Y.R. and Seok, Y.J. 2016. “Glucose induces delocalization of a flagellar biosynthesis protein from the flagellated pole.” *Molecular Microbiology* 101 (5): 795-808.

Park, S., Yoon, J., Lee, C.R., Lee, J.Y., Kim, Y.R., Jang, K.S., Lee, K.H. and Seok, Y.J. 2019. “Polar landmark protein HubP recruits flagella assembly protein FapA under glucose limitation in Vibrio vulnificus.” *Molecular Microbiology* 112 (1): 266-279.

Parkinson, J.S., Hazelbauer, G.L. and Falke, J.J. 2015. “Signaling and sensory adaptation in Escherichia coli chemoreceptors: 2015 update.” *Trends in Microbiology* 23 (5): 257 - 266.

Partridge, J.D., Nhu, N.T., Dufour, Y.S. and Harshey, R.M. 2019. “Escherichia coli Remodels the Chemotaxis Pathway for Swarming.” *Molecular Biology and Physiology* 10 (2): 316 - 319.

Partridge, J.D., Nhu, N.T., Dufour, Y.S. and Harshey, R.M. 2019. “Escherichia coli Remodels the Chemotaxis Pathway for Swarming.” *Molecular Biology and Physiology* 10 (2): 316 - 19.

Partridge, J.D., Nhu, N.T., Dufour, Y.S. and Harshey, R.M. 2019. “Escherichia coli Remodels the Chemotaxis Pathway for Swarming.” *Molecular Biology and Physiology* 10 (2): 00316 - 19.

Paul, K., Nieto, V., Carlquist, W.C., Blair, D.F. and Harshey, R.M. 2010. “The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a 'backstop brake' mechanism.” *Molecular Cell* 38 (1): 128-139.

Percival, S.L. and Williams, D.W. 2014. “Chapter Six - Escherichia coli.” In *Microbiology of Waterborne Diseases (Second Edition)*, edited by Marylynn V. Yates, Nicholas F. Gray, David W. Williams, Rachel M. Chalmers Steven L. Percival, 89 - 117. Academic Press.

Pósfai, G., Plunkett, G., Fehér, T., Frisch, D., Keil, G.M., Umenhoffer, K., Kolisnychenko, V., Stahl, B., Sharma, S.S., De Arruda, M. and Burland, V. 2006. “Emergent Properties of Reduced - Genome Escherichia coli.” *Science* 312: 1044-1046.

Pósfai, G., Plunkett, G., Fehér, T., Frisch, D., Keil, G.M., Umenhoffer, K., Kolisnychenko, V., Stahl, B., Sharma, S.S., De Arruda, M. and Burland, V. 2006. “Emergent Properties of Reduced-Genome Escherichia coli.” *Science* 312 (5776): 1044-1046.

Postlethwaite, C.M., Brown, P. and Dennis, T.E. 2013. “A new multi-scale measure for analysing animal movement data.” *Journal of Theoretical Biology* 317: 175 - 185.

Pratt, L.A. and Kolter, R. 2002. “Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili.” *Journal of Molecular Microbiology* 30 (2): 285-293.

Price, N.D., Reed, J.L. and Palsson, B.Ø. 2004. “Genome-scale models of microbial cells: evaluating the consequences of constraints.” *Nature Reciews Microbiology* 2 (11): 886-897.

Prvan, T., Reid, A. and Petocz, P. 2002. “Statistical Laboratories Using Minitab, SPSS and Excel: A Practical Comparison.” *Teaching Statistics* 24 (2): 68 - 75.

Rajwa, B., Mcnally, H.A., Varadharajan, P., Sturgis, J. and Robinson, J.P. 2004. “AFM/CLSM data visualization and comparison using an open-source toolkit.” *Microscopy Research and Technique* 64 (2): 176 - 184.

Reid, S.W., Leake, M.C., Chandler, J.H., Lo, C.J., Armitage, J.P. and Berry, R.M. 2006. “The maximum number of torque-generating units in the flagellar motor of Rhodobacter sphaeroides is at least 11.” *PNAS* 103: 8066 - 8071.

Reyes, A.T., Ambita, D., Batalon, J.L., Aba, B.L., Cortes, A., Macabecha, C.G. and Montecillo, A. 2019. “Motility and chemotaxis of Escherichia coli in medium with attractant and repellent.” *Journal of Drug Delivery and Therapeutics* 9 (5): 15-18.

Riley, M., Abe, T., Arnaud, M.B., Berlyn, M.K., Blattner, F.R., Chaudhuri, R.R., Glasner, J.D., Horiuchi, T., Keseler, I.M., Kosuge, T. and Mori, H. 2006. “Escherichia coli K-12: a cooperatively developed annotation snapshot--2005.” *Nucleic Acids Research* 1 (34): 1-9.

Robert Cogger-Ward, Adam Collins, Denise McLean, Jacob Dehinsilu, Alan Huett. 2019. “A conserved protein, BcmA, mediates motility, biofilm formation, and host colonisation in Adherent Invasive Escherichia coli.” *bioRxiv.*

Roggo, C., Picioreanu, C., Richard, X., Mazza, C., Van Lintel, H. and Van Der Meer, J.R. 2017. “Quantitative chemical biosensing by bacterial chemotaxis in microfluidic chips.” *Environmental Microbiology* 20 (1): 241-258.

Rossi, E., Paroni, M. and Landini, P. 2018. “Biofilm and motility in response to environmental and host-related signals in Gram negative opportunistic pathogens.” *Journal of Applied Microbiology* 125 (6): 1587-1602.

Ruby, J.D., Lux, R., Shi, W., Charon, N.W. and Dasanayake, A. 2008. “Effect of glucose on Treponema denticola cell behavior.” *Oral Microbiology and Immunology* 23 (3): 234-238.

Ryjenkov, D.A., Simm, R., Römling, U. and Gomelsky, M. 2006. “The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria.” *Journal of Biological Chemistry* 281 (303): 310-314.

Ryjenkov, D.A., Simm, R., Römling, U. and Gomelsky, M. 2006. “The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria.” *Journal of Biological Chemistry* 281 (41): 30310 - 30314.

Ryu, W.S., Berry, R.M. and Berg, H.C. 2000. “Torque-generating units of the flagellar motor of Escherichia coli have a high duty ratio.” *Nature* 403: 444 - 447.

Sagi, Y., Khan, S. and Eisenbach, M. 2003. “Binding of the Chemotaxis Response Regulator CheY to the Isolated, Intact Switch Complex of the Bacterial Flagellar Motor.” *Journal of Biological Chemistry* 278 (28): 25867 - 25871.

Sakia, R. M. 1992. “The Box-Cox Transformation Technique: A Review.” *The Statistician* 41 (2): 169 - 178.

Salek, M.M., Carrara, F., Fernandez, V., Guasto, J.S. and Stocker, R. 2019. “Bacterial chemotaxis in a microfluidic T-maze reveals strong phenotypic heterogeneity in chemotactic sensitivity.” *Nature Commuications* 10 (1877): 1-11.

Samuels, T., Pybus, D. and Cockell, C.S. 2019. “Casamino acids slow motility and stimulate surface growth in an extreme oligotroph.” *Environmental and Microbiology Reports* 12 (1): 63 - 69.

Schubert, C., Van Langeveld, M.C. and Donoso, L.A. 2014. “Innovations in 3D printing: a 3D overview from optics to organs.” *British Journal of Opthalmology* 98 (2): 159 - 161.

Schultz, B.B. 1985. “Levene's Test for Relative Variation.” *Systematic Zoology* 34 (4): 449 - 456.

Senvar, O. and Sennaroglu, B. 2016. “Comparing Performances of Clements, Box-Cox, Johnson Methods with Weinbull Distributions for Assessing Process Capability.” *Journal of Industrial Engineering and Management* 9 (3): 634 - 656.

Serres, M.H., Goswami, S. and Riley, M. 2004. “GenProtEC: an updated and improved analysis of functions of Escherichia coli K-12 proteins.” *Nucleic Acids Research* (32): 300 - 302.

Seymour, J.R. and Raina, J.B. 2018. “Swimming in the sea: chemotaxis by marine bacteria.” *Microbiology Australia* 10: 12-16.

Sezonov, G., Joseleau-Petit, D. and d'Ari, R. 2007. “Escherichia coli Physiology in Luria-Bertani Broth.” *Journal of Bacteriology* 189 (23): 8746 - 8749.

Shaffer, S.M., Dunagin, M.C., Torborg, S.R., Torre, E.A., Emert, B., Krepler, C., Beqiri, M., Sproesser, K., Brafford, P.A., Xiao, M. and Eggan, E. 2017. “Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance.” *Nature* 546: 431-435.

Shao, J., Xuan, M., Zhang, H., Lin, X., Wu, Z. and He, Q. 2017. “Chemotaxis-Guided Hybrid Neutrophil Micromotors for Targeted Drug Transport.” *Angewantde chemie* 56 (42): 12935-12939.

Sheng, X., Reppel, M., Nguemo, F., Mohammad, F.I., Kuzmenkin, A., Hescheler, J. and Pfannkuche, K. 2012. “Human Pluripotent Stem Cell-Derived Cardiomyocytes: Response to TTX and Lidocain Reveals Strong Cell to Cell Variability.” *PloS One* 7 (9): 13-17.

Shi, W., Zhou, Y., Wild, J., Adler, J. and Gross, C.A. 1992. “DnaK, DnaJ, and GrpE Are Required for Flagellum Synthesis in Escherichia coli.” *Journal of Bacteriology* 174 (19): 6256-6263.

Shi, W.E.N.Y.U.A.N., Li, C.O.N.G.Y.I., Louise, C.J. and Adler, J. 1993. “Mechanism of Adverse Conditions Causing Lack of Flagella in Eschenichia coli.” *Journal of Bacteriology* 175 (8): 2236-2240.

Shibuya, M., Takebe, Y. and Kaziro, Y. 1977. “A Possible Involvement of cya Gene in the Synthesis of Cyclic Guanosine 3' :5' Monophosphate in E.coli.” *Cell* 12: 521 - 528.

Shingala, M.C. and Rajyaguru, A. 2015. “Comparison of Post Hoc Tests for Unequal Variance.” *International Journal of New Technologies in Science and Engineering* 2 (5): 22-33.

Silverman, M. and Simon, M. 1977. “Bacterial flagella.” *Annual Review of Microbiology* 31: 397-419.

Silverman, M. and Simon, M. 1974. “Flagellar rotation and the mechanism of bacterial.” *Nature* 249: 73 - 74.

Song, J., Zhang, Y., Zhang, C., Du, X., Guo, Z., Kuang, Y., Wang, Y., Wu, P., Zou, K., Zou, L. and Lv, J. 2018. “A microfluidic device for studying chemotaxis mechanism of bacterial cancer targeting.” *Scientific Reports* 8 (6394).

Song, S., Vuai, M.S. and Zhong, M. 2019. “The role of bacteria in cancer therapy – enemies in the past, but allies at present.” *Infectious Agents and Cancer* 13 (9): 1 - 7.

Soutourina, O., Kolb, A., Krin, E., Laurent-Winter, C., Rimsky, S., Danchin, A. and Bertin, P. 1999. “Multiple Control of Flagellum Biosynthesis in Escherichia coli: Role ofH-NS Protein and the Cyclic AMP-Catabolite Activator Protein Complex in Transcription of the flhDC Master Operon.” *Journal of Bacteriology* 181 (24): 7500 - 7508.

Soutourina, O.A. and Bertin, P.N. 2003. “Regulation cascade of flagellar expression in Gram-negative bacteria.” *FEMS Microbiology Reviews* 27 (4): 505 - 523.

Soutourina, O.A. and Bertin, P.N. 2003. “Regulation cascade of flagellar expression in Gram-negative bacteria.” *FEMS Microbiology Review* 27 (4): 505-523.

Sowa, Y., Rowe, A.D., Leake, M.C., Yakushi, T., Homma, M., Ishijima, A. and Berry, R.M. 2005. “Direct observation of steps in rotation of the bacterial flagellar motor.” *Nature* 437: 916 - 919.

Stephens, M.A. 1974. “EDF Statistics for Goodness of Fit and Some Comparisons.” *Journal of American Statistical Association* 69 (347): 730 - 737.

Studdert, C.A. and Parkinson, J.S. 2004. “Crosslinking snapshots of bacterial chemoreceptor squads.” *Proceedings of the National Academy of Sciences of the United States of America* 101: 2117 - 2122.

Suh, S., Traore, M.A. and Behkam, B. 2016. “Bacterial chemotaxis-enabled autonomous sorting of nanoparticles of comparable sizes.” *Lab Chip* 16: 1254-1260.

Szurmant, H. and Ordal, G.W. 2004. “Diversity in Chemotaxis Mechanisms among the Bacteria and Archaea.” *Microbiology and Molecular Biology Reviews* 68 (2): 301 - 319.

Tan, S.H., Nguyen, N.T., Chua, Y.C. and Kang, T.G. 2010. “Oxygen plasma treatment for reducing hydrophobicity of a sealed polydimethylsiloxane microchannel.” *Biomicrofluidics* 4 (3): 322041-8.

Tietjen, G.L. and Moore, R.H. 1972. “Some Grubbs-Type Statistics for the Detection of Several Outliers.” *Journal Technometrics* 14 (3): 583 - 597.

Tkaczyk, T.S. 2010. “Specialised Techniques - Fluorescence.” In *Field Guide to Microscopy*, edited by John E. Greivenkamp, 90 - 93. Bellingham, Washington USA: SPIE Press.

Tout, J., Astudillo-García, C., Taylor, M.W., Tyson, G.W., Stocker, R., Ralph, P.J., Seymour, J.R. and Webster, N.S. 2017. “Redefining the sponge-symbiont acquisition paradigm: sponge microbes exhibit chemotaxis towards host-derived compounds.” *Environmental Microbiology Reports* 9: 750 - 755.

Turner, L., Ryu, W.S. and Berg, H.C. 2000. “Real-Time Imaging of Fluorescent Flagellar Filaments.” *Journal of Bacteriology* 182 (10): 2793 - 2801.

Van Dessel, N., Swofford, C.A. and Forbes, N.S. 2015. “Potent and tumor specific: arming bacteria with therapeutic proteins.” *Therapeutic Delivery* 6 (3): 385 - 399.

VanDersarl, J.J., Xu, A.M. and Melosh, N.A. 2011. “Rapid spatial and temporal controlled signal delivery over large cell culture areas.” *Lab Chip* 11 (18): 3057 - 3063.

Vargha, A. and Delaney, H.D. 1998. “The Kruskal Wallis Test and Stochastic Homogeneity.” *Journal of Educational And Behavioural Statistics* 23 (2): 170-192.

Vigeant, M.A. and Ford, R.M. 1997. “Interactions between motile Escherichia coli and glass in media with various ionic strengths, as observed with a three-dimensional-tracking microscope.” *Applied Environmental Microbiology* 63: 3474 - 3479.

Wang, F., Wang, H., Wang, J., Wang, H.Y., Rummel, P.L., Garimella, S.V. and Lu, C. 2008. “Microfluidic delivery of small molecules into mammalian cells based on hydrodynamic focusing.” *Biotechnology and Bioengineering* 100 (1): 150 -158.

Wang, G. and Fang, N. 2012. “Detecting and Tracking Nonfluorescent Nanoparticle Probes in Live Cells.” In *Methods in Enzymology*, 84 - 102. Elsevier.

Wang, X., Liu, Z. and Pang, Y. 2017. “Concentration gradient generation methods based on microfluidic systems.” *Royal Society of Chemistry* 7: 29966 - 29984.

Weaver, W.M., Tseng, P., Kunze, A., Masaeli, M., Chung, A.J., Dudani, J.S., Kittur, H., Kulkarni, R.P. and Di Carlo, D. 2014. “Advances in high-throughput single-cell microtechnologies.” *Current Opinion in Biotechnology* 25: 114-123.

Wei, B.L., Brun-Zinkernagel, A.M., Simecka, J.W., Prüß, B.M., Babitzke, P. and Romeo, T. 2001. “Positive regulation of motility and flhDC expression by the RNA-binding protein CsrA of Escherichia coli.” *Molecular Microbiology* 40 (1): 245-256.

Wen, A. and Rakab, E. 2019. “Developing a qualitative chemotaxis-sensitive assay for bacteriophage-host interactions on E. coli BW2511.” *The Meducator* 23-25.

Widman, M.T., Emerson, D., Chiu, C.C. and Worden, R.M. 1992. “Modeling microbial chemotaxis in a diffusion gradient chamber.” *Biotechnology and Bioengineering* 55 (1): 191 - 205.

Wolfe, A.J. and Berg, H.C. 1989. “Migration of bacteria in semisolid agar.” *Proceedings of the National Academy of Sciences of the United States of America* 86 (18): 6973 - 6977.

Xu, X., Bryan, A.L., Mills, G.L. and Korotasz, A.M. 2019. “Mercury speciation, bioavailability, and biomagnification in contaminated streams on the Savannah River Site (SC, USA).” *Science of the Total Environment* 668: 261-270.

Yuan, X., Couto, J.M., Glidle, A., Song, Y., Sloan, W. and Yin, H. 2017. “Single-Cell Microfluidics to Study the Effects of Genome Deletion on Bacterial Growth Behavior.” *ACS Synthetic Biology* 6 (12): 2219 - 2227.

Zhang, T. and Fang, H.H.P. 2005. “Effective diffusion coefficients of glucose in artificial biofilms.” *Environmental Technology* 26 (2): 155 - 160.

Zhang, Z., Schwartz, S., Wagner, L. and Miller, W. 2000. “A greedy algorithm for aligning DNA sequences.” *Journal of Computational Biology* 7 (1-2): 201-214.

Zhao, K., Liu, M. and Burgess, R.R. 2007. “Adaptation in bacterial flagellar and motility systems: from regulon members to ‘foraging’-like behavior in E. coli.” *Nucleic Acids Research* 35 (13): 4441-4452.

Zhuang, J. and Sitti, M. 2016. “Chemotaxis of bio-hybrid multiple bacteria-driven microswimmers.” *Scientific Reports* 6 (32315): 1-10.

Zinser, E.R. and Kolter, R. 2004. “Escherichia coli evolution during stationary phase.” *Research in Microbiology* 155 (5): 328 - 336.