Cell Growth in *Escherichia coli*: Study of Fluctuations and Asymmetry in Cell Extension

Thesis by

Manasi Subhash Gangan

In Partial Fulfilment of the Requirements

for the Degree of

Doctor of Philosophy

Indian Institute of Science Education and Research, Pune

2017

CERTIFICATE

Certified that the work incorporated in the thesis entitled "**Cell Growth in *Escherichia coli*: Study of Fluctuations and Asymmetry in Cell Extension**" submitted by Manasi S. Gangan was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

Date:

Dr. Chaitanya Athale

Associate Professor

Division of Biology, IISER Pune

DECLARATION

I declare that this written submission represents my idea in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date:

Manasi S. Gangan

20123150

So long, and thanks for all the fish

It seems that the hardest part of Ph. D. is summing up last six years. No words can truly express the importance of these years in my life. Filled with excitement, optimism and few moments of disappointments, Ph. D. has definitely resulted in a significant growth in me. Clearly, this intellectual journey was not possible without some crazy people constantly encouraging me directly or indirectly to pursue my goal. While, rest of the thesis reviews my scientific progress chronologically over the six years, in this section, I take an opportunity to express my gratitude.

I must begin thanking my parents Mr. Subhash Gangan and Mrs. Supriya Gangan for believing in my ambitions and supporting my efforts for higher education. Their patience and understanding helped me chase my target arduously.

My undergraduate years turned out to be decisive in my endeavour for graduate studies. I thank my teachers Tara ma'am, Raji ma'am and Chitra ma'am for their constant efforts to shape my scientific thoughts as well as skills.

My profound thanks are also due for Dr. Jayashree Pohnerkar (JP ma'am). I thoroughly enjoyed working with her as a dissertation student, during my years in M S university of Baroda. Just observing her working in the lab, imparted qualities like persistence and sincerity in me. Scientific discussions with her were useful to structure my thought process and sharpen my experimental acumen. I appreciate her willingness to allow me to follow my instincts while doing the research. It helped me to develop my own approach towards scientific problems. JP ma'am was everything a budding scientist could ask for in a mentor.

IISER Pune provided me a great opportunity to interact with many people. Though initially I detested the "common- lab arrangement", due to frequent disappearance of chemicals and instruments, gradually I started appreciating its benefits. Throughout my Ph. D., I received help

from numerous people right from brain- storming over my ideas, designing of my experiments and troubleshooting them eventually. Being a first Ph. D. student in the lab, I was bit scared of working on the protocols from the scratch. However, I would like to thank all of seniors (from other labs), who took out their valuable time to teach me molecular, microscopy techniques without any reluctance. Dr. Aparna Sherlekar and Dr. Rashmi Kulkarni have been more than "just a senior". Aparna was a faithful ear for all the problems and frustrations I have gone through during Ph. D., while, Rashmi's accurate guidance had always been useful not only in experiments but also in real life.

When it comes to philosophical discussion, I will never forget my routine "coffee with Sampada Mutalik". I guess Sampada has been only person who could tolerate me so many years, as I have been following her right from M S University in Baroda. Over the eight years our friendship became matured and was strong enough to support each other when we were going through rough phases. I look forward for more and more fun- filled years with a great thinker and friend of mine.

Darshika Tomer and Neha Nirvan require special mention. Though initially we were just "shopping-buddies", over the years both of them became my close companions and provided me immense support as well as care. They have survived through my daily rants and also guided me correctly to keep myself on my toes.

Speaking of rants, this acknowledgement will be incomplete without the mention of 'Katta gang'. Over 'raat ka chai', we had intense philosophical discussions. The topic of the discussions spanned almost every field of knowledge. My heartfelt thanks to the members of the gang- Saurabh Kadam, Dnyanesh Dubal, Swapnil Bodkhe, Vyankatesh Rajmane, Yashwant Chaugule, Vaibhav Wagh, Swapnil Warhade and Abhijit Pendse. It was fun to have night- long discussions with you guys.

I was honoured to have guidance and timely suggestions from the members of my Research Advisory Committee: Dr. Gayathri Pananghat, Dr. Mugdha Gadgil and Dr. Aswin Seshasayee. I remember, I

could anytime barge in Gayathri's office with my problems and she used to have instantaneous and exact solutions. Another person who was of tremendous help in resolving scientific problems during last two years of my Ph. D. was Dr. Nishad Matange.

Last, but certainly not least, I owe my deep gratitude to my thesis advisor, Dr. Chaitanya Athale. I would like to acknowledge his efforts and encouragement to improve my presentation skills, my ability to review the literature and pose the question. His constant fervour of exploring new fields in science, has led me to try my hands on exciting projects. Though we were not on good terms, especially in last few years of my Ph. D., I thoroughly appreciate his ability to keep his calm composure and let me follow my own decisions while working in the lab. I leave his lab with a knowledge which is far more than just a "science".

In closing, I must admit that I was privileged for being able to wake up every morning excited and happy about my job.

Manasi Gangan

November 2017

Pune, India

Abstract

Cell dimensions are defined for each species. Several cellular pathways guarantee faithful duplication of cell during binary fission. However, phenotypic heterogeneity has been observed in genetically homogenous populations. Advanced scientific studies have proved such fluctuations crucial in the development, for they have been associated with phenomena like division of labour and rise of survivors under harsh environmental situations.

Thesis provides an insight into possible mechanisms for the origin of variations in cell lengths of clonal *E. coli* populations. Series of experiments helped assess an elongation in *Escherichia coli* cells and its effect on the population length distributions.

In our studies with *E. coli* batch cultures, we observed a linear increase in cell length variation above growth rate 1 generation per hour. On the other hand, below this threshold phenotypic heterogeneity was observed to be reduced but it stayed almost constant for all the growth rates. 'The point of inflection' in the cell length variation was identified and correspond to the growth rate at which multi- fork replication is triggered in a bacterial cell. Molecular analysis of RecA dynamics showed increased occurrence of replication fork stalling events at high growth rates. We inferred that stochastic arrest in the replication increases multiplicatively at higher growth rates because of multi- fork replication in *E. coli* cell which in turn, increases the probable halts in cell division through SOS response, producing non- genetic cell length variation in an isogenic population.

Deviation from defined cell dimensions has previously been assigned to the fluctuations in gene expression. Ergo, next step was to probe the effect of replication stochsticity on genetic expression and connect it with phenotypic noise in isogenic populations of *E. coli* MG22. By examining hydroxyurea treated populations at different growth rates, we detected that the total gene expression noise can be modulated by the population growth rate. Growth rate dependent

dissection of the noise showed that intrinsic noise dominates in genetic circuitry, at higher growth rates. In recovered cells, intrinsic noise was observed to be linear with increasing stalls in DNA replication till HU concentration hits 13 mM but then it surprisingly drops monotonously for higher HU concentrations, breaching the connection. Extrinsic gene expression noise takes over at slower growth rates. We predict that at sub- lethal HU dosages, replication fluctuations, in highly proliferating populations can cause phenotypic variability through elevated intrinsic gene expression noise.

On a slightly different note, analysis of *E. coli* elongation, revealed the presence of inherent asymmetry in growth of an individual cell. We found that, an asymmetry in growth introduces difference in the cell division time of two sisters born at same time. Inspection of MreB dynamics showed bias in the distribution of MreB molecules along cell length, suggesting possible role of MreB cytoskeleton in growth bias. Though, its connection with phenotypic variation is not clear, we report a new finding as opposed to the classical description of *E. coli* growth.

The work combined with published data, sets a new paradigm for growth in an *Escherichia coli* cell.

Key words: Cell length, multi- fork replication, phenotypic variability, RecA, stochastic gene expression, asymmetric growth, MreB.

Contents

Thesis Synopsis

Title: Cell growth in *Escherichia coli*: study of fluctuations and asymmetry in cell extension

Name of the Student: **Manasi S. Gangan**

Roll number: 20123150

Name of the thesis advisor: Dr. Chaitanya Athale

Date of Registration: 2nd January 2012

Indian Institute of Science Education and Research (IISER), Pune, India.

Identical appearances of bacterial cells in an isogenic population, belies the fate of an individual cell. Environment dependent fluctuations in the phenotype have been observed for many populations. Studies show that though genetically identical, response of an each bacterial cell to the various signals from "unpredictable" niche is found to be variable.

In recent years phenotypic variations in the clonal populations have received tremendous attention. Rise of persisters, for instance, under selective growth conditions, has been attributed to inherent phenotypic heterogeneity in clonal populations. Phenotypic switch from growing state to dormant state takes place only for certain proportion of the population, while others either resume their growth or terminate their life cycle as the consequence of environmental changes[1,2]. Clear presence of sub- populations with identical genomic DNA has been hypothesized to serve two purposes. It saves the genetic material which can later on rejuvenate in optimal growth conditions[3]. Programmed death of the population serves as a nutrient source to the growing cells and thus, ensure their propagation[4]. Survivors of drastic environmental conditions often pose the problem of antibiotic resistance[5]. Cellular differentiation can also be studied through phenotypic variations. It is been proposed that phenotypical fluctuations

generate division of labour in the clonal population of pathogenic bacteria. That confers the advantage of evading the host as well as producing more bacteria at the same time[6].

Random fluctuations in the genetic circuit are considered as a prominent reason for phenotypic variations. Variable activity of the promoter or the regulatory proteins, may result into the loss of coordinated gene expression within the genetically identical populations[6]. Noise in gene expression facilitates physiological difference within identical cells, which in turn, exhibit differential response to growth conditions.

Segregation of various sub- cellular components, found in limited copies, can be error- prone. Hence, apparent symmetric division can create two cells with different molecular composition[7,8]. Moreover, bacterial cells are known to corner most of the damage in a cell bearing the oldest pole. The process is termed as aging[9]. The "oldest mother" cell tends to show different phenotypic characters, like elongation[10]. Thus, a bias in segregation contributes to phenotypic noise among identical sister cells.

This thesis explores the phenotypic variability in wild-type populations of *Escherichia coli*, from three different perspectives. In particular, cell size was the phenotype of choice and we studied variation in the cell lengths with reference to the modulation in growth rates, noise in gene expression and growth asymmetries. In an attempt to understand an effect of phenotypic heterogeneity on the spatial arrangement of the community, we extended our studies to compete two different strains of *E. coli* with difference in the extent of population cell length variation. Objectives divide thesis in following four parts:

(a) Threshold effect of growth rate on cell length variability.

(b) Correlation between stochastic DNA replication process and gene expression noise in growth rate dependent manner.

(c) Asymmetric growth of single *E. coli* cell.

(d) Colony competition and spatial patterns in isogenic populations growing from a central "homeland".

Chapter 3: Threshold effect of growth rate on cell length variability

Population growth rate greatly influences cell size as well as cell physiology in *E. coli*[11]. This section revolves around the connection between growth rate and phenotypic noise. In batch cultures of *E. coli*, we altered population generation time either by allowing growth under different temperatures or by changing growth media to modify the nutrient supply. Cell sizes were measured for mid- log phase populations and cell length variability is estimated in terms of coefficient of variation. Plot of cell length variability against population growth rate showed two distinct regimes, separated by a threshold corresponding to growth rate of 1 generation per hour. Cell length variability increases monotonously with growth rates greater than 1 generation per hour, while it is reduced and remains constant for cultures with growth rates slower than the threshold.

We identify the inflection point as trigger for multi- fork replication in *E. coli*[12]. We hypothesize that increased replication stall events at higher growth rates hold cell division on through SOS response pathway in an individual cell, which in turn, causes size variation in the population.

We supported our hypothesis by elevating per cell frequency of replication fork stalling with hydroxyurea treatment. We observed increase in cell length variation as a function of hydroxyurea dosage only in populations proliferating with high growth rate. We used colocalization of RecA loci with the bacterial DNA as a proxy for replication stall and correlated it with population cell length variability. As expected, we obtained linear relationship between two, in growth rate dependent manner. Our results were further buttressed

by an assessment of synchronous live cells grown in microfluidic device and on agar surface, which also enable us to study the effect at single cell level.

Thus, we propose stochastic replication arrest as a possible non- genetic mechanism for phenotypic fluctuation in an isogenic population of *Escherichia coli*.

Chapter 4: Correlation between stochastic DNA replication process and gene expression noise

In continuation with our findings, we delve into the link between DNA replication stall events and gene expression noise. We analysed isogenic populations of *E. coli* MG22 strain for this study. *E. coli* MG22 has been constructed in Elowitz lab (2001) and possesses single copy of CFP and YFP gene each on either arm of *E. coli* circular DNA, such that each gene copy is equidistant from OriC and is expressed under P_{lac} [13]. We measured the failure of correlation in the expression levels of two genes when replication arrest is frequented. We quantified average intensities per cell for CFP and YFP as an equivalent to the extent of an expression of respective genes, while variation in population cell lengths was considered as an output for replication halt.

We observed growth rate dependence in gene expression noise. At higher growth rate, fluctuations in the genetic circuit was increased, while at slower rated they were reduced in recovered cells. Interestingly, we found that growth rate also controls sub- component of the noise that contributes the most to total gene expression. Intrinsic noise influences total gene expression noise at high growth rates. Extrinsic noise dominates in the gene circuitry at slower growth. More importantly, an intrinsic noise responds linearly to the replication stall events only at sub- lethal hydroxyurea concentrations (< 10 mM) and hence correlates positively with cell length variation at higher proliferations, while extrinsic noise remains unaffected.

We conclude the chapter by postulating that random fluctuations in chromosomal replication, at sub- lethal HU concentrations, increase noise in gene expression within an individual cell, which in turn, promotes non- genetic phenotypic noise in an individual organism.

Chapter 5: Asymmetric growth of single *E. coli* cell

Chromosomal segregation associated with cell growth decides the placement of cytokinetic ring in *E. coli*[14]. In other words, expansion of *E. coli* cell not only affects the fidelity of the segregation of sub- cellular entities but also, can be the determinant in the generation of physiologically different daughters. This section dissects *E. coli* single cell to study its growth pattern. In order to increase the reliability of the results, we observe growth of *E. coli* surface and its membrane. Experiments were reinforced by analysing the segregation of nucleoids and correlating it with cell growth.

Our results showed a distinct difference between the growth in two cellular halves of a single cell. We further correlated it with MreB localization in the cell. We found that higher MreB content in cell half correlates with high speed of growth at the nearest pole. Growth asymmetry was found to be inheritable from one generation to other. Fast growing end continues to grow with higher speed in daughter generation, while daughter that inherits slow growing pole from mother develops a new pole which leads the growth. Interestingly, it gives rise to the time lag between the divisions of two sisters. Daughter cell with fast growing end from the earlier generation divides earlier than its sister which receives slow growing end.

We sum up this section with a novel finding that *E. coli* cell grows asymmetrically, as opposed to the classical description of its growth. However, we could not investigate the link between growth asymmetry in a single cell and population cell length variability.

Chapter 6: Colony competition and spatial patterns in isogenic populations growing from a central “homeland”

Cell size and shape are unique to every specie of bacteria and they determine the structure of the colony by arranging the cells in a particular pattern[15]. In this section we tried to explore the effect of size fluctuations on the colony make- up.

Experimental set up includes the growth of a culture consisting of two sub- populations, each expressing different molecular reporter, on agar surface. Result of the competition among two sub- populations was measured in terms of the proportion of area occupied by each population in colony. We compete isogenic populations of *E. coli* MG1655 transformed either with peGFP or with pmCherry. We also studied *E. coli* DH5α population under identical experimental set-up. However, due to the lack of an appropriate control and because of inconclusive results, experiment was suspended. In future, non- motile strains can be used for the competition, as a control experiment. Also, to arrange an even competition, cells can be made to express protein other than mCherry, as it hampers the growth in mutants.

In conclusion, work reviews the growth in an individual cell of an *E. coli* and quantitatively connects the fluctuations in it to the phenotypic heterogeneity found in the population.

References:

1. Smits, W. K., Veening, J.-W. & Kuipers, O. P. Phenotypic Variation and Bistable Switching. *Bact. Physiol. A Mol. Approach* 339–365 (2008).

2. Sousa, A. M., Machado, I. & Pereira, M. O. Phenotypic switching: an opportunity to bacteria thrive. *Sci. against Microb. Pathog. Commun. Curr. Res. Technol. Adv.* 252–

262 (2011).

3. Hilbert, D., Laboratories, M. D., Ã, P. J. P. & Hilbert, D. W. Sporulation of Bacillus subtilis. *Curr. Opin. Genet. Dev.* **7,** 579–586 (2004).

4. Engelberg-Kulka, H., Amitai, S., Kolodkin-Gal, I. & Hazan, R. Bacterial programmed cell death and multicellular behavior in bacteria. *PLoS Genetics* **2,** 1518–1526 (2006).

5. Shah, D. *et al.* Persisters: a distinct physiological state of E. coli. *BMC Microbiol.* **6,** 53 (2006).

6. Freed, N. E. *et al.* A simple screen to identify promoters conferring high levels of phenotypic noise. *PLoS Genet.* **4,** 2–7 (2008).

7. Huh, D. & Paulsson, J. Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Genet.* **43,** 95–100 (2011).

8. Huh, D. & Paulsson, J. Random partitioning of molecules at cell division. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 15004–9 (2011).

9. Lloyd-Price, J. *et al.* Asymmetric disposal of individual protein aggregates in Escherichia coli, one aggregate at a time. *J. Bacteriol.* **194,** 1747–1752 (2012).

10. Wang, P. *et al.* Robust growth of escherichia coli. *Curr. Biol.* **20,** 1099–1103 (2010).

11. Hill, N. S., Kadoya, R., Chattoraj, D. K. & Levin, P. A. Cell size and the initiation of DNA replication in bacteria. *PLoS Genet.* **8,** e1002549 (2012).

12. Cooper, S. & Helmstetter, C. E. Chromosome replication and the division cycle of Escherichia coli. *J. Mol. Biol.* **31,** 519–540 (1968).

13. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297,** 1183–6 (2002).

14. Norris, V., Woldringh, C. & Mileykovskaya, E. A hypothesis to explain division site selection in Escherichia coli by combining nucleoid occlusion and Min. *FEBS Lett.* **561,** 3–10 (2004).

15. Kysela, D. T., Brown, P. J., Casey Huang, K. & Brun, Y. V. Biological Consequences and Advantages of Asymmetric Bacterial Growth. *Annu. Rev. Microbiol* **67,** 417–35 (2013).

16. Taheri-Araghi, S. *et al.* Cell-size control and homeostasis in bacteria. *Curr. Biol.* **25,** 385–391 (2015).

17. Cho, H., McManus, H. R., Dove, S. L. & Bernhardt, T. G. Nucleoid occlusion factor SlmA is a DNA-activated FtsZ polymerization antagonist. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 3773–8 (2011).

18. Huisman, O., D'Ari, R. & Gottesman, S. Cell-division control in Escherichia coli: specific induction of the SOS function SfiA protein is sufficient to block septation. *Proc. Natl. Acad. Sci. U. S. A.* **81,** 4490–4 (1984).

19. M, C. J. and V. Cell growth and length distribution in Escherichia Cell Growth and Length Distribution in Escherichia coli. *J. Bacteriol.* **134,** 330–337 (1978).

20. Athale, C. A. & Chaudhari, H. Population Length Variability and Nucleoid Numbers in E . coli. *Bioinformatics* (2011).

21. Renzette, N. *et al.* Localization of RecA in Escherichia coli K-12 using RecA-GFP. *Mol. Microbiol.* **57,** 1074–85 (2005).

22. Wery, M., Woldringh, C. L. & Rouviere-Yaniv, J. HU-GFP and DAPI co-localize on the Escherichia coli nucleoid. *Biochimie* **83,** 193–200 (2001).

23. Schneider, C. a, Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of

image analysis. *Nat. Methods* **9,** 671–675 (2012).

24. Lutkenhaus, J. Assembly Dynamics of the Bacterial MinCDE System and Spatial Regulation of the Z Ring. *Annu. Rev. Biochem.* **76,** 539–562 (2007).

25. Chen, Y., Milam, S. L. & Erickson, H. P. SulA inhibits assembly of FtsZ by a simple sequestration mechanism. *Biochemistry* **51,** 3100–9 (2012).

26. n Sánchez-Gorostiaga A, Palacios P, Martínez-Arteaga R, Sánchez M, Casanova M, V. M. Life without Division : Physiology of Escherichia coli FtsZ-Deprived Filaments. *MBio* **7,** 1–10 (2016).

27. Den Blaauwen, T., Lindqvist, A., Löwe, J. & Nanninga, N. Distribution of the Escherichia coli structural maintenance of chromosomes (SMC)-like protein MukB in the cell. *Mol. Microbiol.* **42,** 1179–1188 (2001).

28. Simmons, L. a, Grossman, A. D. & Walker, G. C. Replication is required for the RecA localization response to DNA damage in Bacillus subtilis. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 1360–5 (2007).

29. William, M. FTSZ AND THE DIVISION OF PROKARYOTIC CELLS AND ORGANELLES. *Nat Rev Mol Cell Biol* **6,** 862–871 (2005).

30. Venkatesh, R. *et al.* RecX protein abrogates ATP hydrolysis and strand exchange promoted by RecA: Insights into negative regulation of homologous recombination. *Proc Natl Acad Sci U S A* **99,** 12091–12096 (2002).

31. Alice, A. MINIATURE escherichia coli CELLS DEFICIENT IN DNA. *Proc. Natl. Acad. Sci. U. S. A.* **6,** 321–26 (1966).

32. Zaritsky, A., Wang, P. & Vischer, N. O. E. Instructive simulation of the bacterial cell division cycle. *Microbiology* **157,** 1876–1885 (2011).

33. Errington, J. Bacterial morphogenesis and the enigmatic MreB helix. *Nat Rev Micro* **13,** 241–248 (2015).

34. Ackermann, M. *et al.* Self-destructive cooperation mediated by phenotypic noise. *Nature* **454,** 987–90 (2008).

35. Gangan, M. S. & Athale, C. A. Threshold effect of growth rate on population variability of Escherichia coli cell lengths. *Roy. Soc. open Sci.* **4,** 1–17 (2017).

36. Skerker, J. M. & Laub, M. T. Cell-cycle progression and the generation of asymmetry in Caulobacter crescentus. *Nat. Rev. Microbiol.* **2,** 325–37 (2004).

37. Aldridge, B. B. *et al.* Asymmetry and Aging of Mycobacterial Cells Lead to Variable Growth and Antibiotic Susceptibility. *Science* **335,** 100–104 (2011).

38. Brown, P. J. B. *et al.* Polar growth in the Alphaproteobacterial order Rhizobiales. *Proc. Natl. Acad. Sci.* **109,** 3190–3190 (2012).

39. Joyce, G. *et al.* Cell division site placement and asymmetric growth in Mycobacteria. *PLoS One* **7,** 1–8 (2012).

40. Guberman, J. M., Fay, A., Dworkin, J., Wingreen, N. S. & Gitai, Z. PSICIC: noise and asymmetry in bacterial division revealed by computational image analysis at sub-pixel resolution. *PLoS Comput. Biol.* **4,** e1000233 (2008).

41. Taniguchi, Y. *et al.* Quntifying E.coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **329,** 533–539 (2011).

42. Strahl, H., Bürmann, F. & Hamoen, L. W. The actin homologue MreB organizes the bacterial cell membrane. *Nat. Commun.* **5,** 1–11 (2014).

43. Hallatschek, O., Hersen, P., Ramanathan, S. & Nelson, D. R. Genetic drift at expanding

frontiers promotes gene segregation. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 19926–30 (2007).

44. Rainey, P. B. *et al.* The evolutionary emergence of stochastic phenotype switching in bacteria. *Microb. Cell Fact.* **10,** S14 (2011).

List of Publications:

Published articles:

1. **Gangan MS** and Athale CA. 2017 Threshold effect of growth rate on population variability of *Escherichia coli* cell lengths. ***Royal Society Open science***. 4: 160417

2. Chaphalkar AR, Jain K, **Gangan MS**, Athale CA (2016) Automated Multi-Peak Tracking Kymography (AMTraK): A Tool to Quantify SubCellular Dynamics with Sub-Pixel Accuracy. ***PLoSONE***. 11(12)

3. Beal J, Haddock-Angelli T, Gershater M, de Mora K, Lizarazo M, Hollenhorst J, et al. (2016) Reproducibility of Fluorescent Expression from Engineered Biological Constructs in *E. coli*. ***PLoS ONE*** 11(3) [the iGEM2015 team from IISER Pune is part of the consortium]

Manuscripts in preparations:

Gangan MS, Mishra P, Athale CA. Asymmetric growth of *E. coli* cell.

List of figures:

List of Tables:

List of Equation:

Abbreviations

DIC	Differential Interference Contrast
PDMS	Polydimethylsiloxane
LB	Luria- Bertani broth
YEB	Yeast Extract Broth
TB	Tryptone broth
PBS	Phosphate buffer saline
PFA	*para*- Formaldehyde
NA	Numerical aperture
BSA	Bovine serum albumin
IPTG	Isopropyl- β- D- 1- thiogalactopyranoside
DAPI	4', 6- diamino- 2- phenylindole
FM4- 64	N-(3-triethylammoniumpropyl)-4-(6-(4-(diethylamino) phenyl) hexatrienyl) pyridinium dibromide
SDS	Sodium dodecyl sulphate
DTT	1, 4- dithiothreitol
PVDF	Polyvinylidene fluoride
HRP	Horse radish peroxidase
HU	Hydroxyurea
TM	Trimethoprim
Ceph	Cephalexin
Chl	Chloramphenicol
Rif	Rifampicin
ROI	Region of interest
LOI	Line of interest

List of strains/ plasmids used:

Sr. No.	Strain Name	Source
1	*E. coli* MG1655	Described in Guyer et al., 1981[1]
2	*E. coli* DH5α	Described in Hanhan, 1989[2]
3	*E. coli ΔrecA*	CGSC JW26691
4	*E. coli ΔsulA*	CGSC JW09411
5	*E. coli ΔslmA*	CGSC JW56411
6	*E. coli* RP437	CGSC 12122[3]
7	*E. coli recA- gfp*	Described in Renzette et al.,2005[4]
8	*E. coli* MG22	Described in Elowitz et al.,2002[5]
9	*E. coli mreB- yfp*	CGSC 13021[6]

Sr. No.	Plasmid Name	Source
1	pBAD24	CGSC[7]
2	pGFP	Clontech
3	peGFP	This study
4	pmCherry	This study
5	pBAD24- hupA- GFP	Described in Wery et al.,2001[8]
6	pRecA- mCherry	This study
7	pBAD24- recA	This study

Chapter 1

Introduction

1.1 Growth and division in *Escherichia coli*

Bacteria like *Escherichia coli* offer an excellent model to study the cellular physiology. This gut commensal has rod shaped cell which is gram negative in nature. In other words, cell is surrounded by a lipopolysaccharide (LPS) layer, called as outer membrane. Periplasmic space separates outer membrane from cell membrane. A single layer of peptidoglycan (PG) resides in the peripalsmic space (Fig. 1.1)[9]. Peptidoglycan layer, made up of glycan strands connected by pentapeptides, is elastic in nature and remodelling of its layer is essential in the maintenance of the rod shape as well as the growth and division of an organism[9,10].

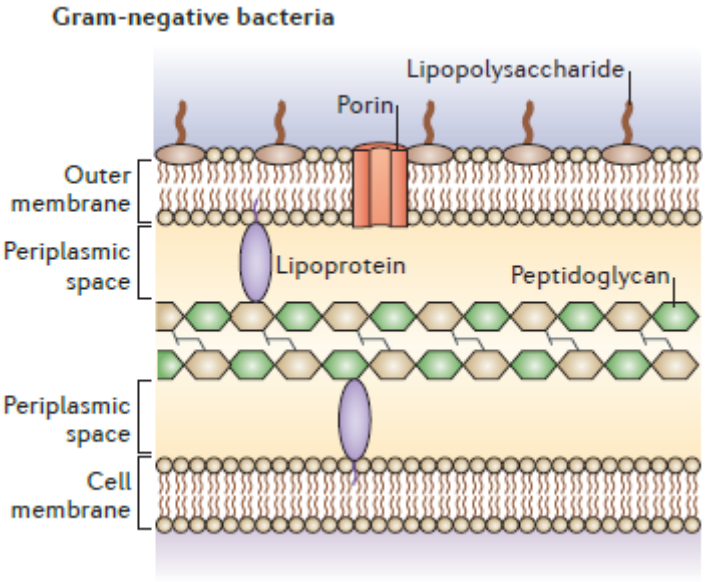


Fig. 1.1: Cell wall structure of gram negative microorganisms (Reprinted form Brown et al., 2015 with permission[11]. Copyright obtained from Springer/ Nature Publishing Group (Appendix E, C1)).

Growth and division in *E. coli* cell is a sequential interplay between two cytoskeletal proteins. A new- born *E. coli* cell elongates laterally till it reaches the division length which is approximately twice as that of new- born length. Elongation is driven by an actin homologue, MreB. Helical cables of MreB polymers run throughout the length of an *E. coli* cell, providing a template as well as the direction for the cell wall synthesis[12,13]. Growth control is switched to FtsZ, a tubulin equivalent, during septum formation. This cytoskeletal protein polymerizes into the ring at mid- cell and thus forms a scaffold for the recruitment of cell wall synthesizing machinery[14]. Gradual constriction of the Z- ring guides the partition of the two daughters by septum (Fig. 1.2).

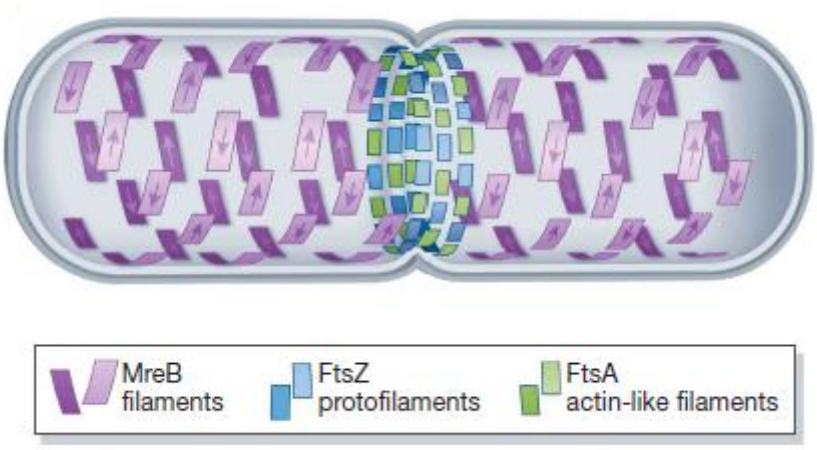


Fig. 1.2: Cytoskeleton machinery in *Escherichia coli*. Purple boxes represent short directed patches of MreB filaments arranged helically and are involved in the growth of the cell at side walls. FtsZ and FtsA polymers are indicated by blue and green boxes respectively. Z- ring formed at the mid- cell, is tethered to the membrane and thus stabilized by an actin homologue, FtsA (Reprinted from Juarez and Margolin, 2012 with permission[13]. Copyright obtained from John Wiley and Sons (Appendix E, C3)).

These two cytoskeletal components instruct the synthesis of murein sacculus. However, the mode of synthesis differs depending on the enzyme involved in the process. MreB interacts with PBP 2 through RodA in order to insert a new strand in existing PG layer, while, FtsZ employs PBP 3 to trigger the addition of new building blocks, required for constriction[9]. Recently, 'three- for- one' mechanism has been proposed to explain the addition of PG blocks during the elongation as well as the division of the cell (Fig. 1.3). Model postulates that three new strands of peptidoglycan are synthesized per one old strand called as 'docking strand' in murein sacculus. These strands are connected covalently to the strands (stress bearing strands) which are adjacent to the docking strand through transpeptidation. Cleavage of docking strand results in its replacement with newly synthesized peptidoglycan and subsequent expansion of the murein layer[15,16]. Model, thus, also considers the 'make- before- break' strategy proposed by Arthur Koch, in 1985 for the growth of rod shaped organisms[17].

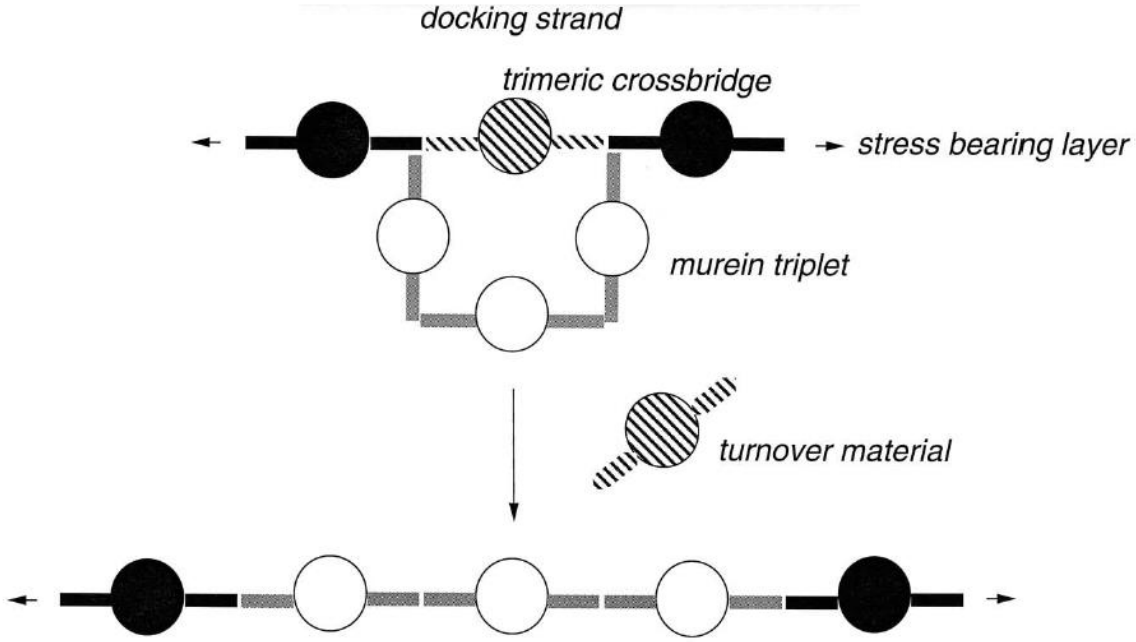


Fig. 1.3: Three- for- one mechanism of cell wall synthesis. Hatched circle indicates docking strand which is replaced by three new strands (white circles). Black circles stand for stress bearing strands (Reprinted from Sceffers and Pinho, 2005 with permission[16]. Copyright obtained from American Society of Microbiology (Appendix E, C4)).

Mode of construction slightly deviates during cytokinesis[18]. PG strand above the constriction site is considered as a docking strand to add three new strands. The process then propagates, each time generating newer and newer docking strands. The segregation happens by an action of transamidases that cleave docking strands (Fig. 1.4).

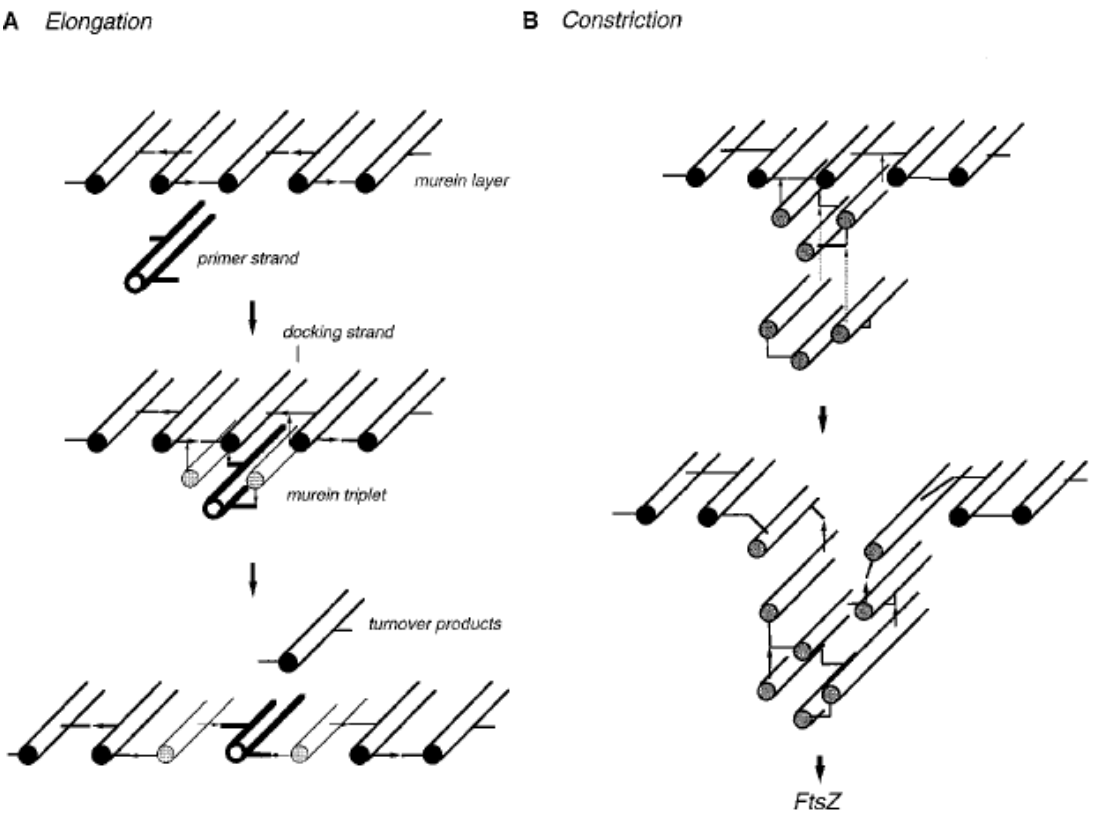


Fig. 1.4: Three- for- one mechanism in elongation (A) and constriction (B) events (Reprinted from Höltje, 1998 with permission[18]. Copyright obtained from American Society of Microbiology (Appendix E, C4)).

Support for the model comes from discovery of multienzyme complex (Fig. 1.5). It consists of (a) bifunctional transpeptidase- transglycosylases: PBP 1A, PBP 1B and PBP 1C, (b) monofunctional transpeptidases: PBP2 or PBP 3 (based on the mode of synthesis), (c) lytic transglycosylases: Slt 70, MltA and MltB, (d) D, D- endopeptidases: PBP 4 and PBP 7. Since the complex includes enzymes of two opposite classes, it is named as 'yin- yang' complex[15].

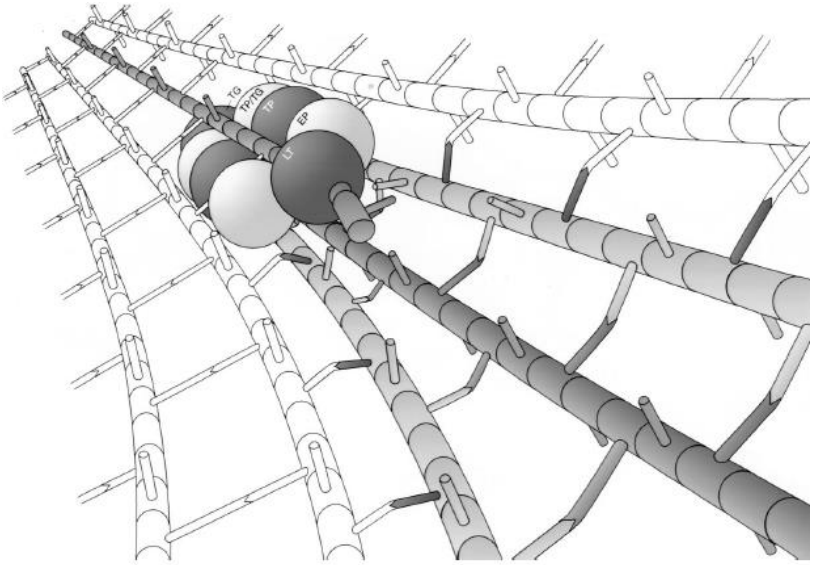


Fig. 1.5: Holoenzyme synthesizing the cell wall (Reprinted from Höltje, 1998 with permission[18]. Copyright obtained from American Society of Microbiology (Appendix E, C4)).

1.2 Cell length homeostasis *E. coli* population

Various models have been proposed to explain the maintenance of cell size in an organism. Two major paradigms were accounted for the observed homeostasis in population cell sizes. 'Sizer' mechanism predicts cell division of an organism after a threshold cell length is attended. In other words, 'Sizer' mechanism states that the size of a cell at division is independent of its birth size. Mechanism was shown to work using *Schizosccharomyces pombe* as model system by Peter Fentes, in 1977[19,20]. Later on in 2014, model was shown to fit the cell length homeostasis in *E. coli* populations[21]. Other mechanism named as 'Timer' explains that cell division occurs after a constant elapsed time and assumes that the generation time of a cell remains constant irrespective of its birth length[22].

However, observation of an individual *E. coli* cell growth refutes above mechanisms. Experimental evidences show that in an *Escherichia coli* cell doubling time is negatively correlated with new- born cell size, thereby ruling out 'Timer' model. Though average new-

born cell volume of population is exponentially dependent on population growth rate, single cell data systematically deviate from the growth law. The phenomenon cannot be explained with 'Sizer' mechanism[23]. In 2015, a new model was devised for cell size conservation. It can be explained in the context of *E. coli* cell cycle. The life cycle of *Escherichia coli* progresses thorough three consecutive phases: B or birth period, C or chromosomal duplication period and D or division period (Fig. 1.6). B period varies based on the growth rate of the cell, while C and D periods last for 40 and 20 minutes respectively. A complete *E. coli* cell cycle takes place within 60 mins[24]. Per division cycle the cell synthesizes a constant volume, irrespective of its initial new- born volume or size (shown as Δ in Fig 1.6).

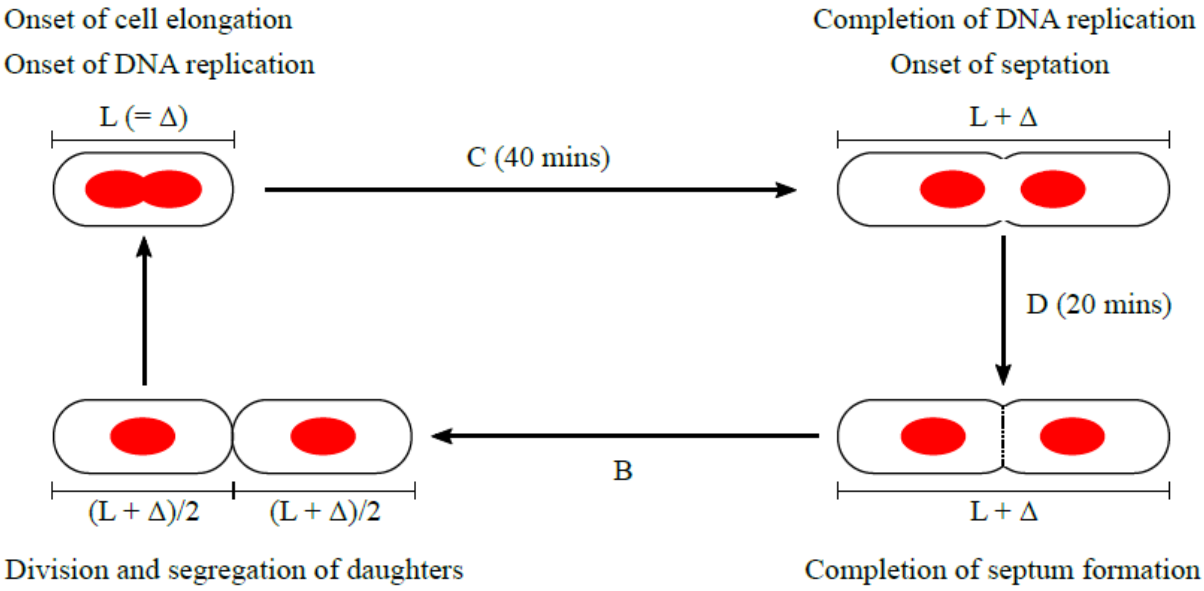


Fig. 1.6: BCD cell cycle of *E. coli* (Derived from Cooper and Helmstetter, 1968 and Taheri-Arighi et al., 2015[24,23]).

Strategy, known either as constant Δ model or as an 'adder principle' eventually helps converge population cell sizes within two- fold range of 2 µm and thus, maintains the homeostasis (Fig.

1.7)[23]. It highlights that a single new- born cell adds a constant volume (Δ) per cell cycle and then divides in the mid- plane, regardless of its birth size as well as generation time. And hence the respective birth cell length gradually tends to the modal value of the population size distribution. Thus, though model does not addresses the random occurrence of elongated cells in the clonal populations, it explains the strategy employed by an individual cell that exhibit the size deviation to attain size homeostasis.

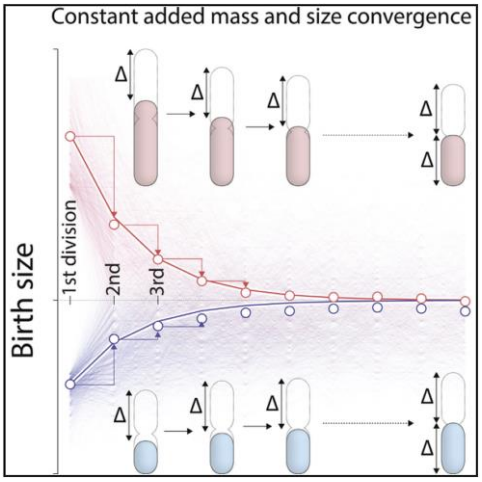


Fig. 1.7: Graphical representation of adder model (Reprinted from Taheri- Arighi et al., 2015 with permission[23]. Copyright obtained from Elsevier (Appendix E, C5)).

Rise of filamentous cells can be associated with cellular tendency to maintain the viability of the population through generations, under variable environmental conditions. Under optimal laboratory conditions, however, time required to complete one cell cycle is reduced to less than 60 mins and for the integration of replication and division within one life cycle, OriC is fired in earlier generations. Thus, as a result, an individual cell, at higher growth rate, encases 6- 8 actively progressing replication forks[24,25]. Situation demands tight regulation to maintain the temporal fidelity of the replication as well as cell growth. *E. coli* cell employs array of proteins/

pathways to ensure the temporal coupling between the DNA replication and segregation as well as cell growth and division.

1.3 Coupling between DNA replication and cell division

Major pathways that link DNA replication and cell division cycle in *E. coli* cell comprise nucleoid occlusion and SOS response[26,27]. Both of them target the polymerization of FtsZ, an initiator of cell division, over chromosomal DNA.

Nucleoid occlusion directly links chromosomal DNA with the placement of the septum within the cell. In *E. coli*, process of nucleoid occlusion employs a TetR family protein 'SlmA'. Two third region near OriC of *E. coli* genomic DNA, harbours approximately 21 binding sites for SlmA protein which accelerates the GTPase activity of FtsZ destabilizing Z- ring. Formation of septum across unsegregated chromosome is, thus, prevented by nucleoid occlusion (Fig. 1.8)[28–30].

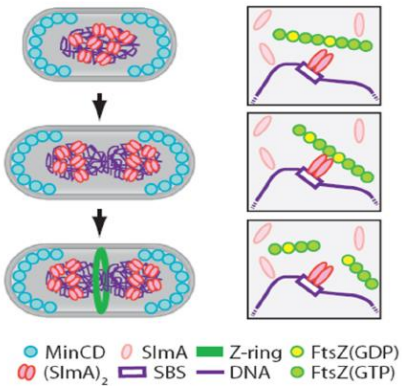


Fig. 1.8: Coupling between DNA replication and cell division in *E. coli* through nucleoid occlusion (Reprinted from Cho et al., 2011 with permission[29]. Copyright obtained from PNAS (Appendix E, C6)).

On the other hand, SOS response pathway delays division until the genetic material is being repaired. Replication fork is stalled at breached sites, an event which facilitates binding of DNA repair protein, RecA at nicked site[31]. Later event initiates SOS response not only to repair chromosomal DNA, but also to withhold cell division. RecA upon binding to processed ssDNA[32] undergoes conformational changes and activates the auto- cleavage activity of LexA, a repressor of SOS box proteins. Degradation of the repressor leads to the expression of SOS response proteins. SulA, one of the SOS proteins sequesters the monomer of FtsZ in its dimeric form and stops its polymerization[33–35]. Entire pathway has been summarized in Fig. 1.9.

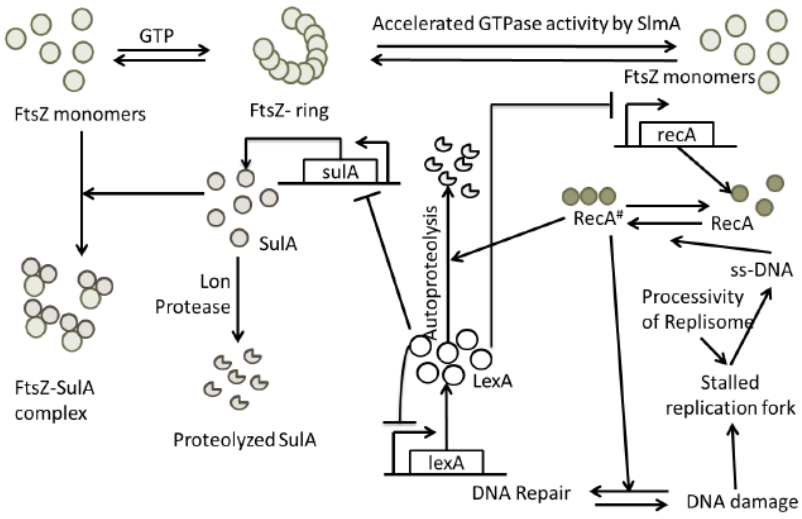


Fig. 1.9: Coupling between DNA replication and cell division in *E. coli* through SOS response pathway (Derived from D'Ari and Huisman, 1981,1984 and Chen et al., 2012 [33–35]).

1.4 Cell size variations in genetically identical populations

Though the cell cycle is tightly regulated in order to produce viable cells of average 2 µm size, length distribution of new- born wild- type cells extends up- to 10 µm (Fig. 1.10)[36] .

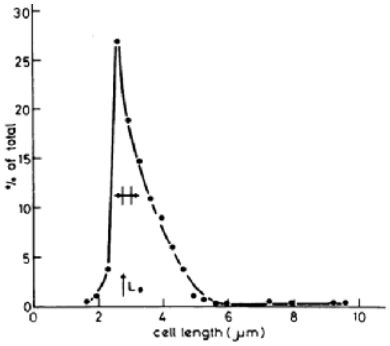


Fig. 1.10: Cell length distribution in the new- born populations of *E. coli* MG1655 (Reprinted from Cullum and Vincente, 1978 with permission[36]. Copyright obtained from American Society of Microbiology (Appendix E, C4)).

Sources that introduce phenotypic noise in the clonal populations of *E. coli* have been studied in the thesis. Growth conditions play decisive role in the cell size determination of an organism. Growth rate is known to shift the average of the population cell size. Fast growing cells are observed to be longer than the slower growing bacteria[37]. In addition, cell size also shows strong dependence on carbon availability in the niche[38]. For instance, in *Bacillus subtilis*, UgtP enzyme moonlights to keep check on the growth of an organism in nutrient dependent manner. Enzyme participates in well conserved pathway of glucolipid synthesis and can function as a sensor for carbon availability, which can directly be transduced to the divisome in order to regulate the division. Process thus, modulates the size of an individual cell at a given growth rate[37]. In *E. coli*, glucosyltransferase (OpgH) have been shown to connect the cell size with metabolic pathways in a growth rate dependent manner[39]. Apart from this, in batch cultures, growth phase dictates the physiology of bacterial cells. Difference in metabolisms specific to the growth phases, in turn, influence cell size of bacteria. Upon entry of the cells into stationary phase, an event called “reductive cell division” in triggered in bacteria so that already initiated

rounds of DNA replication and cell division are completed without any further growth. It results in the shift of average cell size towards left[40].

Phenotype of an organism is the culmination of complex interactions within genetic wiring. Temporal or spatial variations in the coordination of two or more genes can lead to the changes in phenotype. Noise in gene expression has been proposed to be advantageous under stress conditions by creating non- heterogeneity in genetically homogeneous population[41,5,42–44].

Huh and Paulsson, in 2001 has proposed that the fluctuations in the segregation of sub- cellular molecules result in the production of two daughters that quantitatively differ in their molecular composition, though superficially they look identical and hence give rise to the variations in the phenotypes[45,46].

Thesis addresses subcellular mechanisms that give rise to the cell length heterogeneity in the clonal populations of *E. coli*. We have focused our review majorly on the role of the population growth rates, fluctuations in the gene expression of a bacterium and the asymmetry observed during the division of the mother. Our approach uses the quantitative and molecular analysis of *E. coli* single cell and comparing it across the population.

1.5 Experimental approaches

Microscopic examination of an *E. coli* cell has a major share in this study. Fixed *E. coli* cells were imaged for the quantification of its length, on the other hand, live *E. coli* cell was captured for temporal studies of different cellular molecules. We generated fixed images of *E. coli* cells using DIC so that the cell size can automatically be measured by counting the pixels covered by the shadow of the cell. Confocal optical settings were used for recording temporal changes in an individual *E. coli* cell.

1. Differential Interference Contrast (DIC) microscopy:

DIC microscopy proves to be useful in capturing images of unstained samples. Optical set- up for DIC, enhances the contrast at the interfaces of the sample and the medium (Fig. 1.11). Resultant monochromatic virtual image has three dimensional effect. Minimal set up consists of a polarizer to produce plane- polarized light, which then enters the condenser. Light then passes through Wollaston prism, where it is split into two orthogonal wavefronts that penetrate into the sample. Optical density of the sample decides an extent of the retardation of these two wavefronts, which when recombined by second Wollaston prism, construct an interference pattern that highlights the boundaries of the sample creating an illusion of 3D image[47,48].

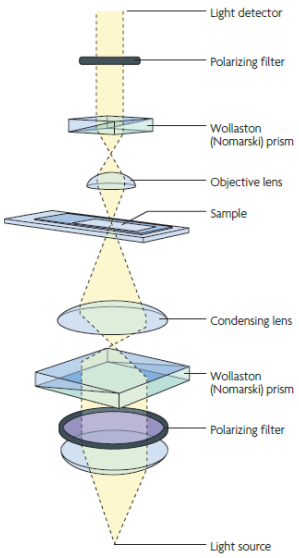


Fig. 1.11: Optical set- up of DIC microscopy (Reprinted from Rosenthal, 2009 with permission[48]. GNU free documentation license (Appendix E, C2)).

2. Confocal Laser Scanning microscopy:

Confocal laser scanning microscopy provides an improved image resolution. Classical fluorescence microscopy floods the sample with the light. Detector also collects the emission

that comes from unfocused part of the sample. Basic confocal configuration introduces pinhole between the light source and the sample, controlling an illumination of sample to the point (Fig. 1.12). Second pinhole, located between the sample and the detector, excludes out- of- focus flare before emitted light enters the detector. In order to amplify the signal, confocal set up uses photomultiplier tube as detector. Thus, confocal microscopy provides lateral and axial resolution[49,50].

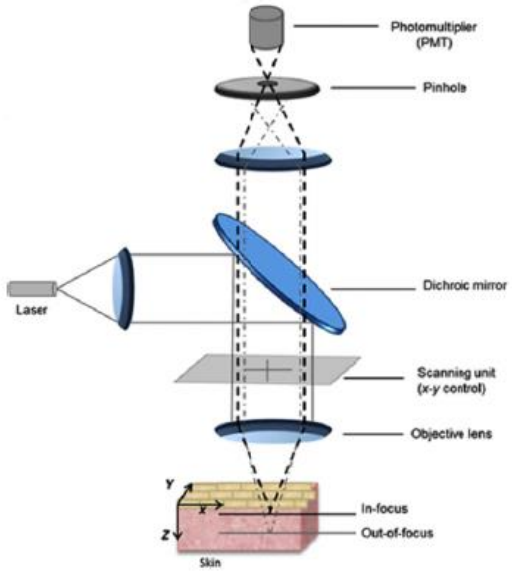


Fig. 1.12: Optical configuration of confocal laser scanning microscope (Reprinted from Rossetti, 2013 with permission[51]. Copyright obtained from INTECH open (Appendix E, C7)).

For microscopic time lapse observation, we grew *E. coli* cells either on agar pads or in continuous culture using 'mother machine'.

3. Lab- on- Chip:

Soft lithography technique has recently gained popularity in the field of quantitative microbiology. Technique is utilized to create microstructures that match the dimensions of bacteria and can confine them spatially during the time lapse imaging.

Standard procedure for creating microstructure involves spin coating of a photoresist on a silicon wafer. After the mask of desired micropattern is placed on the uniform coat of the photoresist, assembly is illuminated with UV light. Exposed part of the photoresist is cross-linked, while photoresist protected by the mask is washed away with the help of an organic solvent. Silicon wafer can then be used as a template to bake PDMS membranes, which when stuck to the glass surface make a microfluidic device (Fig. 1.13)[52].

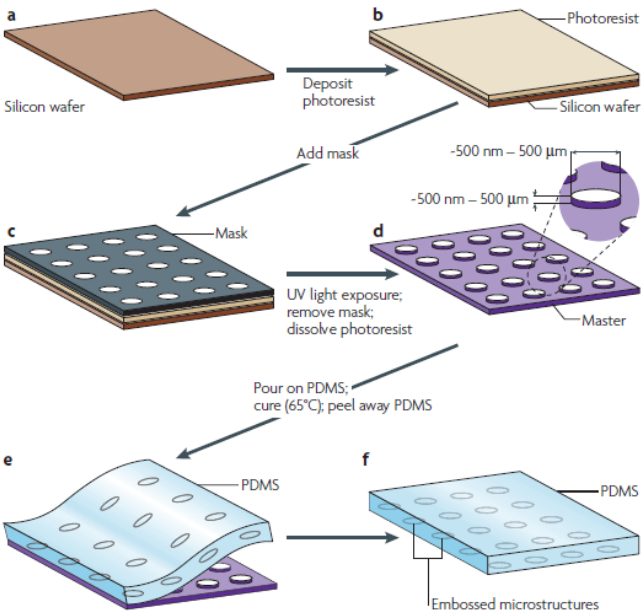


Fig. 1.13: Schematic of microfabrication (Reprinted from Weibel et al., 2007 with permission[52]. Copyright obtained from Springer/ Nature Publishing Group (Appendix E, C1)).

We imprinted agar pad with micro- pattern using an epoxy wafer. It created grooves, 2 µm deep and 1- 5 µm in the width. *E. coli* cells spread on such surface were confined into the

indentations. It helped in the restricted alignment of the cell and also reduced their spatial movement during the growth.

Growth of *E. coli* cells in microfluidic device was helpful to assess the population phenotype in continuous culture. Design has been described in Wang et al.,2010[53] and schematic has been shown in Fig. 1.14. However, we customized it by withdrawing nutrients from the source through the device (Chapter 2; Section 2.4). Modification was important to subject the same population to various environmental changes and analyse the changes in the cell size.

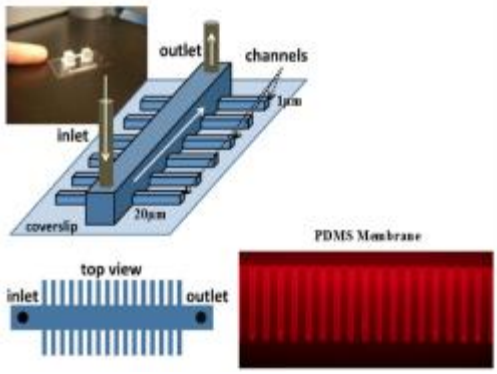


Fig. 1.14: Design of 'Mother machine'(Reprinted from Wang et al., 2010 with permission[53]. Copyright obtained from Elsevier (Appendix E, C5)).

4. Image Analysis

Analysis of an image was automated by using in- house MATLAB programme, which takes an advantage of shadow of the cell produced because of the DIC optics. Algorithm maps the image for the 'gradient' that exists because of the brighter object and its adjacent shadow[54]. Programme was useful in analysing over thousands of *E. coli* cells in the population which increased the reliability of our interpretations (Fig. 1.15).

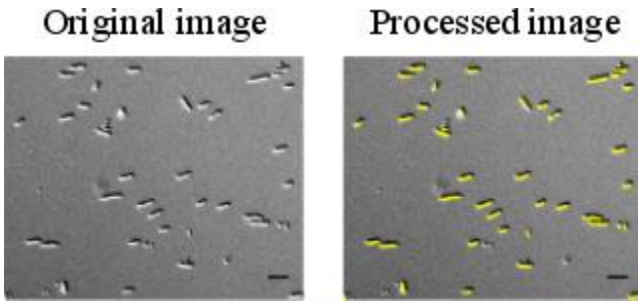


Fig. 1.15: Analysis of *E. coli* MG1655 populations for cell length using 'Gradient detection' code. Left image is raw DIC image of *E. coli* cell, while right image shows the corresponding cell contours (yellow lines) extracted and overlapped on the cell by an algorithm. Scale bar- 2 µm. (Adapted from Athale and Chaudhari, 2011[54]).

With our experiments we correlate growth rate with cell size fluctuations in isogenic populations of *Escherichia coli*. We further succeeded to link gene expression noise with cell length variation in growth rate dependent manner. Our single cell studies showed the presence of an asymmetry during the growth of an *E. coli*.

Chapter 2

Materials and methods

2.1 Bacterial strains and plasmids

Study has used various strains derived from *E. coli* K12. Although MG1655 (CGSC 6300) cells have been studied in major portion of this work, phenotypic variation in the populations of *ΔrecA* (JW26691), *ΔsulA* (JW09411), *ΔslmA* (JW56411) and MG1655 with genomic copy of *recA* replaced with *recA- gfp*[4] (Gift from Dr. G. P Manjunath. Construction of the strain is based on the work done by Steven Sandler group) have also been considered. For correlation studies between gene expression noise and cell length variation *E. coli* MG22[5] was used (Gift from Dr. Michael Elowitz). We also used *E. coli mreB- yfp*[6] (CGSC SX1466) and *E. coli* RP437[3] (CGSC 12122) to assess growth asymmetry and motility respectively in *E. coli* cell.

Plasmids were constructed for ectopic expression of RecA. These were derived from pGFP (Clontech, USA) and pBAD24 backbones respectively (Appendix B). Both the plasmids possess ampicillin resistance as marker. Expression of pBAD24- recA was induced with arabinose (0.2%. SRL, Mumbai, India). Nucleoid movement was tracked with extra-chromosomal expression of pBAD24- hupA- gfp[8] induced with 0.2 % Arabinose (Gift from Dr. Josette Rouviere-Yaniv). peGFP and pmCherry plasmids were backbone modifications of pGFP plasmid (Appendix B).

2.2 Growth conditions

Except for growth rate modulation experiments, secondary cultures of all *E. coli* strains, were inoculated at 1% concentration and grown at 37°C, 180 rpm in LB (HiMedia, Mumbai, India), till the population reaches mid- log phase.

Population growth rates were manipulated by changing either temperatures or nutrient source. In temperature modulation experiment 100 ml of bacterial cultures, with 1% inoculum in LB were grown either at 22°C or at 30°C or at 37°C or at 42°C, under constant sharing at 180 rpm. We chose six nutrient sources: LB, Yeast extract broth (YEB = 0.5% w/v yeast extract + 1% w/v NaCl), Tryptone broth (TB = 1% w/v tryptone + 1% w/v NaCl), M9 + Glucose (0.4% w/v), M9 + Succinate (0.9% w/v) and M9 + Acetate (0.5% w/v)[55]. All the minimal media were supplemented with thymidine (4 µg/ ml). Temperature of the experiment was maintained at 37°C with constant shaking at 37°C. All the media were made in deionized water and pH was ensured to be at 7. Growth was monitored and cells were harvested at every 0.5 hr, by measuring the optical density of the culture at 600 nm, till population entered the stationary phase.

2.3 Fixed cell imaging

Cells were washed with PBS and processed with 4% (w/v) PFA (Sigma- Aldrich, USA) in order to fix them. Except for *E. coli* MG22, all the cells were stained with 0.1 µg/ml DAPI (Sigma- Aldrich, USA) to observe nucleoids[54]. We used Axio- vision plan apochromat upright epifluorescence microscope (Carl Zeiss, Germany) or LSM 780, inverted confocal microscope (Carl Zeiss, Germany) to visualize the cells. Following microscope settings were used to image the different strains:

Strain	**Objective**	**Channels**
MG1655	40X/ NA 1.5 (Epifluorescence)	DIC and DAPI
ΔrecA	40X/ NA 1.5 (Epifluorescence)	DIC and DAPI
ΔsulA	40X/ NA 1.5 (Epifluorescence)	DIC and DAPI
ΔslmA	40X/ NA 1.5 (Epifluorescence)	DIC and DAPI
MG1655 + pRecA-mCherry	40X/ NA 1.5 (Epifluorescence)	DIC, DAPI and dsRed
***ΔrecA* + pRecA- mCherry**	40X/ NA 1.5 (Epifluorescence)	DIC, DAPI and dsRed
MG1655 + pBAD24- recA	40X/ NA 1.5 (Epifluorescence)	DIC and DAPI
***ΔrecA* + pBAD24- recA**	40X/ NA 1.5 (Epifluorescence)	DIC and DAPI
MG1655- *recA- gfp*	100X/ NA 1.5, oil (Epifluorescence)	DIC, DAPI and GFP
MG22	63X/ NA 1.4, oil (Confocal)	1. CFP laser: excitation: 430 nm, emission: 454- 516 nm 2. YFP Laser: excitation: 514 nm, emission: 514- 621 nm 3. DIC is coupled with CFP laser

		4. Detector used: GASP
MG1655- *mreB- yfp*	63X/ NA 1.4, oil (Confocal)	1. YFP laser: excitation: 514 nm, emission: 514- 621 nm 2. Detector used: GASP

Table 2.1: Microscope settings used for fixed cell imaging.

2.4 Microfabrication of mother machine

Micropattern for mother machine device was adapted from the published work from Suckjoon Jun's group[53], and was designed using CleWin software (WieWin Web, Netherlands). It was fabricated by using a ~100 nm layer of gold (for aligning the second layer) followed by spin-coating a 2 μm layer SU8-2 negative photoresist (Microchem, USA) onto a SiO_2 wafer using a spin coater model WS- 400B- 6NPP LITE (Laurell Tech. Corp., USA). Photoresist was cured by UV exposure with a mask (EVG, Austria) corresponding to trench and dead-end channels. Un-exposed photoresist was washed and a 20 μm layer of SU8-20 negative photoresist (Microchem, USA) spun and exposed to UV corresponding only to the trench, for curing.

2.5 Development of epoxy replica for microfluidics

To prepare PDMS membrane silicone elastomer and curing agent were mixed in 10 : 1 (w/w) proportion (Sylgard 184, Dow-Corning, USA). Mixture was then degassed for 1 hr, poured on silicon wafer and baked at 80°C for another 2 hrs. Cured PDMS membrane was used as a template to make epoxy replica of silicon wafer.

Mixture of resin and hardener (BondTite, Resinova, India) taken in the proportion of 8 : 10 (w/w) was degassed for 3 hrs. It was spread uniformly on the surface of PDMS membrane and incubated at 60°C, till the resin solidified. PDMS membrane was then separated from epoxy

surface. After confirming the conservation of the dimensions of the micropattern, the epoxy replica was used for microfluidic experiments.

2.6 Continuous cultures using microfluidic device

Cleaned PDMS membrane was bonded to the coverslip (22 X 22 mm^2), by plasma cleaning. Plasma was created from air at the vacuum pressure at 1 mbar (Emitech K050X, Quorum Technologies, U.K.). Device was passivated using 10 mg/ml of BSA (Sigma-Aldrich, Mumbai, India) for 1 hr, at 37°C, followed by an introduction of log- phase cells into the device by diffusion at 37°C for 1 hr. Fresh nutrients were flown in and the cells trapped in the channels were incubated with it for next 1 hr, before the device was attached with tubing (C- flex, ID = 0.020 cm, OD = 0.083 cm; Cole- Parmer, USA) and assembled on inverted confocal microscope. Nutrients were withdrawn from appropriate nutrient source, heating at 37°C, using syringe pump (PHD Ultra, Harvard Apparatus, USA) at the flow rate of 0.3 ml/hr. Following microscopic settings were used for continuous cultures of different strains. Images were taken at the interval of 2 mins.

Strain	**Objective**	**Laser used**	**Growth Medium**	**Time of the treatment (hr)**
MG1655	63X/ 1.4NA Oil	488 nm laser with DIC coupled to it.	LB	3 hr
MG1655	63X/ 1.4NA Oil	488 nm laser with DIC coupled to it.	M9 + succinate	3 hr
	63X/ 1.4NA		LB	1 hr

MG1655 + peGFP	Oil	488 nm laser with DIC coupled to it.	LB + HU	1 hr
			LB	2.5 hr
***ΔrecA* + peGFP**	63X/ 1.4NA Oil	488 nm laser with DIC coupled to it.	LB	1 hr
			LB + HU	1 hr
			LB	2.5 hr

Table 2.2: Microscope settings used for continuous cultures.

2.7 Agar pad imprinting

The agar was embossed with micropattern using an epoxy replica. The molten agar (LB + 2% bacteriological agar + inducer + selection) was poured on epoxy surface and allowed to solidify. Log phase cells were spread on the agar surface, and incubated a 37°C for 1 hr before the imaging. Following microscopy arrangements were used for different strains. Cells were scanned at the end of every 2 mins using inverted confocal microscope.

Strain	**Objective**	**Laser used**
MG1655 + pBAD24-hupA-GFP	63X/ 1.4 NA, Oil	488 nm laser with DIC coupled to it
MG1655 + pBAD24-hupA-GFP, stained with FM4- 64	63X/ 1.4 NA, Oil	1. 488 nm laser to image the nucleoid 2. 633 nm laser to image the cell membrane

Table 2.3: Microscope settings used for live cell imaging on agar pad surface.

2.8 Immunoblotting and densitometry to quantify cellular RecA levels

To detect the cellular levels of RecA in MG1655, before and after hydroxyurea (HU) treatment, cultures harvested prior drug treatment and after the recovery in LB, YEB and TB media. After equating optical densities of the cultures by appropriate dilutions, cultures were pelleted down and resuspended in 40 µl of 1X PBS and 10 µl of 5X SDS loading dye (250 mM Tris- Cl, pH 6.8, 10% SDS, 50% glycerol, 0.5% Bromophenol Blue and 500 mM DTT). Cells were lysed by heating the samples at 95°C for 10 mins with constant shaking at 700 rpm. Processed samples were centrifuged and supernatant was ran on 10% SDS gel at 120 V (BIO- RAD, USA). Samples were transferred to PVDF membrane (Immobilon-P transfer membrane, EMD Millipore Cooperation, USA). 5% milk powder in TBST buffer (Tris-Cl buffered saline, pH 7.4 and 0.1% Tween-20) was used to avoid non- specific reaction between primary antibodies and membrane surface. Membrane was incubated with anti-RecA antisera[56] (1:12000 dilution) at 4°C overnight. To increase the specificity of the reaction, anti- sera diluted to 12000 in blocking agent (5% milk powder in TBST buffer) was incubated with 100 ml of cell lysate of *ΔrecA* culture for 12 hrs at 4ºC. Next step was to obtain the cell lysate. *ΔrecA* population was grown to OD_{600nm} ~2.0, centrifuged and resuspended in PBST. Cells were burst opened by subjecting them to pulse sonication (ON: 1 sec, OFF: 3 secs, 30 cycles) for 2 mins. Anti-rabbit antibodies conjugated with horseradish peroxidase (Jackson ImmunoResearch, USA) were used at the concentration of 1:10000 to detect primary antibodies on the surface. Antibody was detected using luminol (Luminata™ Classico Western HRP substrate, EMD Millipore Cooperation, USA) by incubating it with the membrane and the luminescence measured (LAS 4000, GE healthcare, USA). Loading control was set by running the same samples on SDS gel and comparing their intensities. To analyse the blot in ImageJ, horizontal ROI was selected approximately at 38 kDa (Chapter 3, Fig. 3.8) and band density was obtained, in terms of area under the peaks, by using '*plot profile*' function. Percent area of each band (I_{band}) estimated by

using '*label peaks*' function, was normalized with respect to the maximum value I_{max}, as shown in equation 1. The resultant densities (I_{norm}) were plotted and compared for different growth conditions.

$$I_{norm} = \frac{I_{band}}{I_{max}} \qquad \text{Equation 1}$$

2.9 Bulk fluorimetry for RecA- GFP expression

5 ml cultures of MG1655- *recA- gfp* were grown in LB, YEB and TB for 1, 1.5 and 2 hrs respectively. After withdrawing 1 ml culture, rest of the cells were subjected to 30 mM hydroxyurea treatment for the time equivalent to 3 generations, followed by recovery in the same medium for another 3 generations. Sample populations were collected from all the media after recovery. Cultures thus, obtained were subjected to fluorimetric analysis in order to assess the cellular levels of RecA- GFP as response to hydroxyurea treatment. Culture density was measured at OD_{600nm}. Recovered cultures were diluted in order to match their density with cultures taken prior to the treatment. 50 µl of each culture was used to measure the fluorescence in 96- well, half- area, round bottom, black- plate (Corning USA). Samples were excited at 480 nm and fluorescence was measured at 520 nm using Varioskan Flash multifunctional plate reader (Thermo Scientific, USA). To correct for autofluorescence, fluorescence values recorded for MG1655 culture of similar density (negative control) in LB, YEB and TB were deducted from measured fluorescence of MG1655- *recA- gfp* samples in respective media. Per cell fluorescence was quantified by dividing corrected fluorescence with total number of cells in 50 µl.

2.10 Marking the cell centre by photobleaching

Repeated photobleaching was used to mark the cell midpoint of the cells membrane tagged with FM4- 64 (2 µg/ml, Invitrogen, USA), by increasing the laser (514 nm) intensity to 100% for 250 iterations (corresponding to 90 seconds) in an ROI approximately 0.3 µm wide (along the long-axis) at the mid-plane of the cell and perpendicular to the longitudinal axis of the cell, to reduce the fluorophore intensity to 0%. Images were acquired approximately for 10- 30 mins every 2- 4 seconds, and bleaching procedure repeated before the fluorescence recovery. Though 63X oil immersion was used for recording images, magnification was further increased by using optical zoom to varying scales for different fields of interests.

2.11 FRAP measurements of MreB- YFP

E. coli cells expressing MreB- YFP were bleached with 514 nm laser (100% intensity) at either of the cellular poles with 50 iterations for 1 sec, till the fluorophore intensity in the ROI drops to 0%. Images were recorded for next 10 mins at the interval of 1 sec in YFP as well as DIC channels. As a control for observational photobleaching of YFP, identical conditions were used to image same strain without intentional bleaching. Intensity values thus, acquired were used as reference values to correct the recovered intensities in bleached regions.

2.12 Drug treatment in batch cultures

E. coli cultures were grown either in LB or in M9 supplemented with succinate for 3 generations (1 hr for LB and 3 hrs for M9 + succinate) at 37°C, 180 rpm. Cells were then treated with an appropriate drug for next 3 generations. For MG22 cultures 2 mM of IPTG was added along with the drug, in order to induce the culture. Also, only for MG22 cultures, half of the culture volume was withdrawn right after the treatment and other half was subjected to

the similar growth conditions as that of the other *E. coli* strains. Cells were washed with PBS and were subjected to recovery under identical incubation conditions, after adding fresh growth medium (with or without required inducer) for 3 generations. Cells were harvested, fixed and imaged. Following are the details about the various drug treatments carried out for different *E. coli* strains.

Strain	**Drug**	**Concentration(s)**
MG1655	Hydroxyurea	0 mM
		10 mM
		30 mM
		70 mM
		100 mM
ΔrecA	Hydroxyurea	0 mM
		10 mM
		30 mM
		70 mM
		100 mM
ΔsulA	Hydroxyurea	0 mM
		10 mM
		30 mM
		70 mM
		100 mM
ΔslmA	Hydroxyurea	0 mM
		10 mM

		30 mM
		70 mM
		100 mM
***ΔrecA* + pRecA- mCherry**	Hydroxyurea	0 mM
		10 mM
		30 mM
		70 mM
		100 mM
MG1655- *recA- gfp*	Hydroxyurea	30 mM
***ΔrecA* + pBAD24 recA**	Hydroxyurea	30 mM
MG1655	Trimethoprim	1 µg/ml
MG22	A22	1 µg/ml
MG22	Cephalexin	10 µg/ml
MG22	Chloramphenicol	10 µg/ml
MG22	Rifampicin	1 µg/ml
MG22	Hydroxyurea	1 mM
		6.6 mM
		13.2 mM
		20 mM
		26 mM

Table 2.4: Drug concentrations used to treat the batch cultures of *E. coli* (Chapter 3 and 4).

2.13 Motility analysis

5 µl drop of polystyrene beads and log- phase cultures of RP437, DH5α, MG1655, *ΔrecA*, *ΔsulA*, *ΔslmA* was placed in a motility chamber which was devised by sticking double- back tape cut in the centre on a glass slide. After the suspension was put inside, chamber was sealed using glass coverslip. Assembly was observed and recorded using epifluorescence microscope (Carl Zeiss, Germany), under 40X/ 1.5 NA in DIC channel at the interval of 500 msecs for 2 mins.

2.14 Colony competition assay

E. coli cultures transformed either with peGFP or with pmCherry were mixed together at the cell density of approximately 10^6, in the proportion of 50 : 50. 2 µl of the mixed culture was placed on the surface of LB agar (2.5% LB + 2% bacteriological agar) to form a "homeland region" of approximately 1 mm of size. Growth of the colony was then observed under stereoscope (Olympus, Japan) under 0.63X magnification in GFP, RFP channels and in the Bright field at the end of every 12 hrs. To maintain the continuity in the orientation of colony while imaging, base of the petri plate was marked with a cross, which was then matched with the cross wires in the ocular piece of the microscope, as explained in Fig. 2.1.

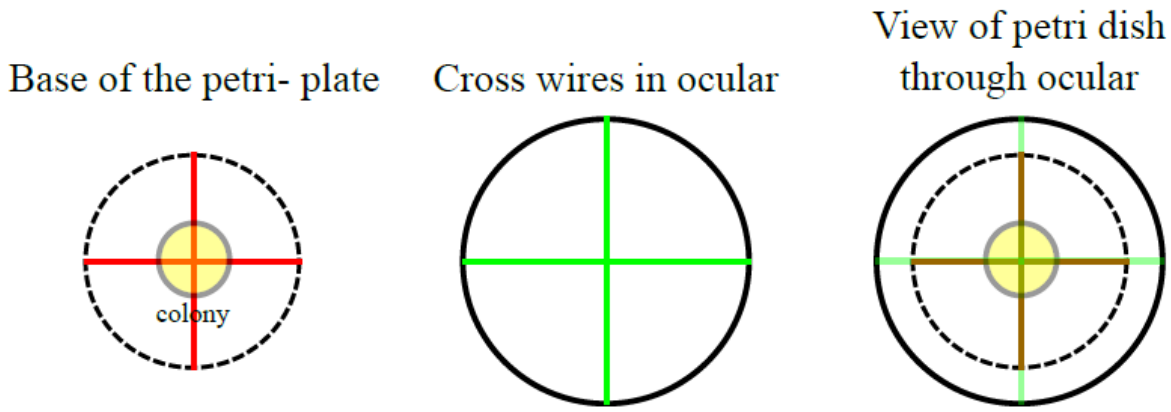


Fig. 2.1: Orientation of an *E. coli* colony in the field of view.

2.15 Data analysis and statistics

1. Bacterial growth measurements

Growth rates as well as doubling time of the batch cultures of *E. coli* MG1655 under different growth conditions were retrieved by fitting the population density measurements to the logistic function.

2. Cell length analysis and estimation of heterogeneity in the cell lengths

DIC images of the cells were analysed for cell length estimation, using in- house developed MATLAB '*gradient detection program*'[54]. Cell length distribution, obtained, was then fit to log- normal function to extract mean, variance of the distribution in MATLAB (Mathworks Inc. MA, USA), with '*lognpdf*' and '*lognstat*' functions (Appendix D, D.1). Validity of the test was tested statistically by Kolmogorov- Smirnov (KS) test for number of bins, n = 44 and significance level, $\alpha = 0.01$. The test statistics $D_{(\alpha, n)}$ is given as[57],

$$D_{(\alpha,n)} = \sqrt{\frac{-\ln(\alpha/2)}{2n}} \qquad \text{Equation 2}$$

Difference d_{max}, between cumulative distribution function of the fit and data was compared. Goodness of the fit holds true only if $d_{max} > D_{(\alpha, n)}$[57]. Cell length variability of the population

measured as the ratio of the mean cell length to the standard deviation in the length, which is also termed as the coefficient of variation.

3. ImageJ based analysis of the cells

Lengths of the cells grown in microfluidic device as well as MG1655- *recA*- *gfp* were analysed interactively in ImageJ (v1.5f)[58] from their respective DIC images, by drawing LOI along cell contour. For MG1655- *recA*- *gfp* strain, same contour was imposed on corresponding images in DAPI and GFP channels. Only those intensity peaks that showed overlap, upon merging the cell intensity profile in both channels were counted as a confirmation of colocalization between genomic DNA and RecA foci.

Kymographs of peGFP transformed cells in 'mother machine' were made in ImageJ using '*multiplekymograph plugin*'. LOI was drawn along selective channel lengths, in order to represent cellular dynamics over time.

4. Estimation of noise in the gene expression in MG22 populations

Images were processed and analysed in ImageJ (v1.50f) to determine the noise in gene expression as well as in the cell lengths of MG22 population. Images acquired in YFP channels were translated linearly by 1 pixel in X and Y directions to correct the offset between CFP and YFP channels. Cell contour was manually extracted from DIC images and then overlaid on corresponding CFP and YFP images. Grey value of each pixel covered by cell contour was measured and average intensity of CFP and YFP molecules within the cell was calculated. Length of the contour was assumed to be equal to the cell length. Intensity as well as cell length values, thus, obtained were used to calculate Variation in the gene expression and cell lengths in the population.

5. Analysis of growth asymmetry in *E. coli* MG1655 cells using MATLAB

Modified '*gradient detection programme*' developed in MATLAB (by Prangya Mishra, unpublished data) was used to detect the contours of DIC images of *E. coli* cells and track down their ends through the time. XY coordinates of the ends were used to calculate the total as well as cumulative growth of each end, using '*Euclidean distance formula*'. Cumulative distance thus obtained against time was fit to exponential function to find the growth rate of each end.

6. Assessment of distribution of MreB loci along the cell length

A macro was developed in ImageJ to automate the measurement of area occupied by MreB foci in each cell half. Code processes DIC image to extract an outline of the cell, which is divided into half and superimposed on corresponding image of MreB puncta in YFP channel. Puncta are then detected and their area is estimated by 'Analyse Particle' module. Macro outputs the ratio of sum area of MreB foci to the area of a cellular half in which foci have been detected (Appendix C, Appendix D, D.2).

For the analysis of FRAP, intensities corresponding to first 50 time points (50 secs) were considered. To correct for extensive fluorophore photobleaching observed during image acquisition, an exponential function was fit to experimental intensity values of reference ROI to obtain a decay rate of -0.005 AU/sec, which was then used to back- calculate the corrected intensity values for FRAP ROI, by dividing the acquired intensities with exponential decay rate. Corrected intensity values were normalized and fit to single exponential function to calculate mobile fraction and half recovery time of MreB molecules in bleached ROI. Corresponding DIC images of the cell were used to assess the growth near the poles using '*MultipleKymograph*' plugin and '*tsp050706*' macro in ImageJ.

7. Statistical analysis

GraphPad Prism 5 (GraphPad software) was used to apply an 'unpaired t test' to affirm the difference between the two distributions.

8. Motility analysis in various strains of *E. coli*

For motility analysis of *E. coli* cells, DIC time series was processed using ImageJ (v1.5f) 'Particle Tracker' plugin. Difference between the X and Y coordinates at two consecutive time points, was considered as the distance covered in a single step. Distribution of such displacements was fit to Gaussian distribution (Appendix D, D.3) to extract the average displacement (μ) and the standard deviation (σ). Later was then used to compute the coefficient of diffusion.

Results

Genesis of phenotypic diversity in *Escherichia coli* populations that occurs independently of variations in genetic make- up has been the focus of this thesis. To outline the possible reasons leading to the phenotypic heterogeneity, we made an attempt to assess the fluctuations in the cell lengths of a model organism *Escherichia coli*.

Each section in the 'result' part represents a different point of view used to investigate into the problem. We considered the effect of growth environment, genetic fluctuations and random segregation of sub- cellular molecules during cytokinesis. Though they seem to be three independent studies, we further set out to bridge them together. These chapters are as follows:

Chapter 3: Threshold effect of growth rate on the cell size distribution of an isogenic population of *Escherichia coli* MG1655

Chapter 4: Stochasticity in DNA replication process modulates the intrinsic noise in the genetic circuit of *Escherichia coli* MG22

Chapter 5: Asymmetric growth of an *Escherichia coli* cell

Last section is our venture into understanding the consequences of deviation in the cell size of an individual cell on the population. Preliminary studies that review the changes in the colony patterning due to size variations have been explained in:

Chapter 6: Colony competition and spatial patterns in isogenic populations growing from a central "homeland"

Chapter 3

Threshold effect of growth rate on the cell size distribution of an isogenic population of *Escherichia* coli MG1655

3.1 Motivation

Though, diverse shapes and sizes have been observed in prokaryotic kingdom, each micro – cell is programmed to maintain and replicate its unique, shape as well as size through the generations. A typical rod shaped *E. coli* cell, for instance, elongates during its growth phase while approximately maintaining a constant cell diameter, before it divides symmetrically around its mid- plane into two rod shaped daughter cells[59]. Size of a new- born *Escherichia* cell, in abundance of nutrients, averages around 2 μm[60]. Cell recruits various checkpoints to preserve its cell size. When *E. coli* cell extends by a constant mass in each cell cycle[23], these checkpoints control the spatio- temporal configuration of the septum within mother cell and generate two daughters of identical lengths.

MinCDE is the prominent system that protects the cell poles from FtsZ polymerization, thus minimizing the risk of formation of mini- cell or ghost cells in the population[61–63]. Nucleoid occlusion impedes untimely Z- ring polymerization in mid- cell before the genomic DNA is segregated, thus ensuring an unbiased distribution of genetic material between two sisters[29,64]. Temporal coupling between DNA replication and cell division is achieved with the help of SOS response. RecA, an initiator of SOS pathway, plays a pivotal role not only in repairing the DNA anomalies, but also it puts the cell division on hold till the DNA is repaired. After RecA molecules polymerize at break site on DNA, conformational changes in the protein enables it to interact with LexA and evoke its auto- protease activity[33,34]. LexA functions as a repressor for the expression of SOS box proteins. Drop in its cellular levels, triggers the manufacture of

proteins that are involved in DNA repair as well as blocking cell division. SulA, is a well-known inhibitor of cell division, sequesters of FtsZ monomers, thus shifting the equilibrium towards the depolymerisation of Z- ring [35].

Despite of several checks on cell partition, isogenic *E. coli* population shows occasional occurrences of elongated cells. Cell length distribution of wild- type *E. coli* cells growing under standard laboratory conditions, is always right- skewed log- normal function and extends to 10 µm[36], suggesting the episodes of random events in an individual cell, that makes cell escape the cytokinesis.

Filamentation occurs when cell growth goes uninterrupted without cell division. Several untimely events can cause hindrance to the septation. Observations of most of the elongated cells have shown the presence of unsegregated nucleoids that pauses the division and results in filamentation[65]. Deletion of nucleoid segregation protein, MukB leads to the similar effect[66]. Percent filamentous cells increase in the population when DNA repair system if frustrated. For instance, our earlier work with clonal population of *E. coli* DH5α and HB101 revealed that the cell size of an organism can reach upto 20- 25 µm (Fig. 3.1)[54]. Both organisms are known laboratory strains, popularly used for cloning. To avoid any genetic modifications in foreign DNA, both strains harbour genetic copy of mutated *recA* gene (*recA1*) that expresses disabled RecA protein, incompetent to repair damaged DNA. However, the protein can hinder the cell division through SulA mediated inhibition of FtsZ polymerization and eventual cell division. Division cycle skip, thus, ends up in the generation of an elongated cell.

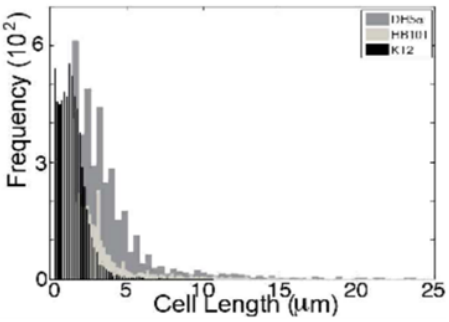


Fig. 3.1: Cell length distribution in *E. coli* populations with mutant copy of *recA* (Reprinted from Athale and Chaudhari, 2011 with permission[54]. Copyright obtained from Oxford University Press (Appendix E, C8)).

Apart from this, disturbed stoichiometry of numerous proteins involved in the cell division can postpone the septation. Studies have shown that modulation of the ratio of FtaZ : FtsA molecules within a cell, can suspend cytokinesis resulting in filamentation of an organism[67,68].

In classical microbiology, filamentation of bacterial cells has been seen as organismal response to environmental insults[69]. Recent focus on host- pathogen interactions have discovered that filamentation can be provide a great getaway to uropathogenic *Escherichia coli* cells from innate immunity[70]. Development of elongated phenotype has been suspected to promote the survival. In *Burkholderia pseudomallei*, for example, higher rate of survival is found to be coupled with higher incidents of filamentation in the population[71]. In *Proteus mirabilis* elongated cells could swarm easily to evade host defence system[72]. Though torrent of data have pinned environmental cues as a reason for switch of microorganisms to filamentous form, underlying molecular mechanisms are still uncovered. The chapter aims at unveiling the molecular dynamics that tune environmental modulations with occurrence of filamentation together in genetically identical population.

In 1958, Schaechter et al; device a growth law that implicated growth rate of an organism as a controller of the new- born cell volume[73]. Recent experiments have affirmed that changes in the growth rate of an organism, which in turn, has a great impact in its cell size[23].

This section probes into possible relationship between growth rate and cell length variations in the clonal populations.

3.2 Cell size distributions in *E. coli* batch cultures

For pilot experiments the clonal population of wild- type *E. coli* was grown in batches, under various growth conditions. We reviewed the fluctuations in the cell phenotype, when growth rate is changed through the modulation of physical factors like temperature and nutrient sources. Analysis stands as a starting point for further experiments in which we have attempted to introduce more complex growth conditions and test their effects on cell length distribution in bacterial community.

3.2.1 Population growth rate modulation

We proposed to change the population growth rate by tweaking the physical factors. Wild- type populations growing in LB were subjected to four different temperature conditions, for increasing temperature accelerates the rate of the biochemical reactions and hence the growth of the cell. In another set of experiments, nutrient source (carbon source) was changed in order to modulate the growth rate (Fig. 3.2 A and B). Microbial growth was followed over time till it hits the maximum carrying capacity (K) of the media and data was fit to the logistic growth equation (Equation 2) to extract the growth rate[74,75]:

$$N(t) = \frac{N(0)*K}{N(0)+(K-N(0))*Exp(-r*t)}$$ Equation 3

Where, N(0) is an initial population density and N(t) is the population density at time t. Growth rate is indicated by r, while K stands for the carrying capacity of the medium. Doubling time (t_d) was then computed simply as the inverse of growth rate, r[10].

$$t_d = \frac{1}{r}$$ Equation 4

Proliferation of the cells were higher at 42°C. While LB, being the rich source of nutrients accelerated the growth of bacteria to the maximum (Table 3.1 and 3.2).

Temperature (ºC)	**Carrying capacity**	**Growth rate (generation/ hr)**	**Doubling time (min)**
22	3.194	0.8986	66.77
30	3.011	1.604	37.41
37	2.918	2.067	29.03
42	2.48	2.01	29.85

Table 3.1: Growth rate of MG1655 populations at different temperatures.

Medium	**Carrying capacity**	**Growth rate (generation/ hr)**	**Doubling time (mins)**
LB	2.934	1.814	33.11
Yeast Extract	2.593	1.505	39.87
Tryptone	2.515	1.049	57.2
M9 + Glucose	3.028	0.9504	63.13
M9 + Succinate	2.464	0.4546	131.98
M9 + Acetate	1.43	0.2195	273.35

Table 3.2: Growth rate of MG1655 populations in different nutrient media.

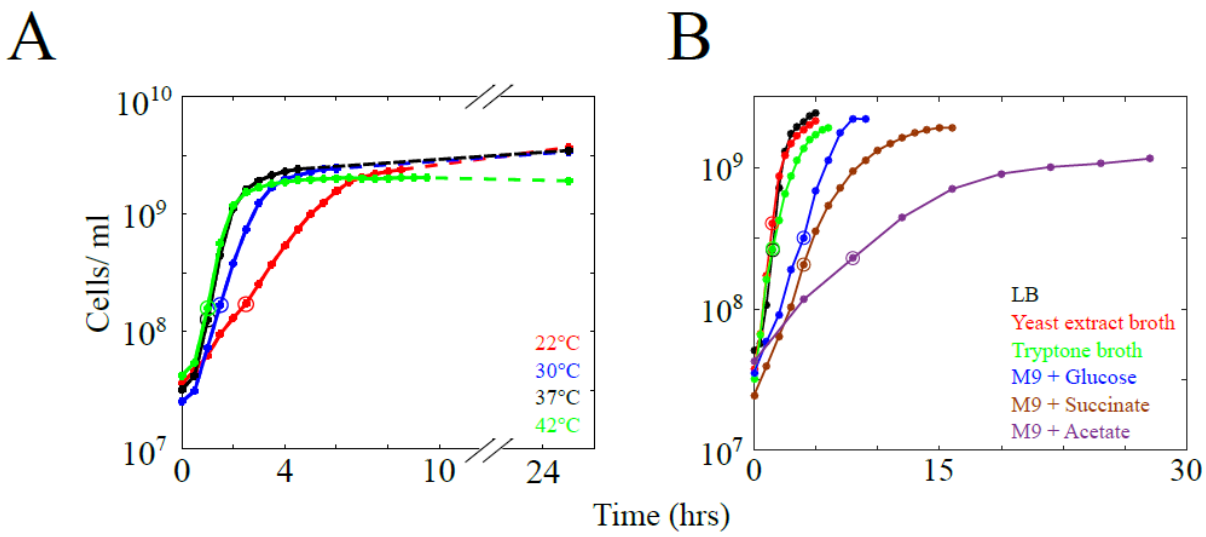


Fig. 3.2: Growth rate modulation of *E. coli*. The logarithmic growth of MG1655 population (A) grown in LB at different temperatures and (B) in different media has been plotted as a function of time on linear abscissa. Encircled time points on each growth curve represent mid-log phase from which cells were sampled for further analysis.

3.2.2 Effect of temperature changes on cell length variability

Cell lengths of isogenic populations were compared across the mid- log phase cultures to decouple results from the effect of growth phase on cell size. DIC snapshots of these cells were analysed for cell lengths by in- house '*Gradient Detection Programme*' developed in MATLAB[54]. Cursory observations revealed that the cell length distribution of the population grown in LB medium extends up to 10 µm, which is consistent with earlier reports[36]. Interestingly, the trend was found to be similar across all the temperatures, in spite of change in the population growth rates. We further, characterized the distributions by fitting the data to log- normal function (Appendix D, D.1). We retrieved similar values for average cell lengths as well as its variance at four temperatures (Fig. 3.3).

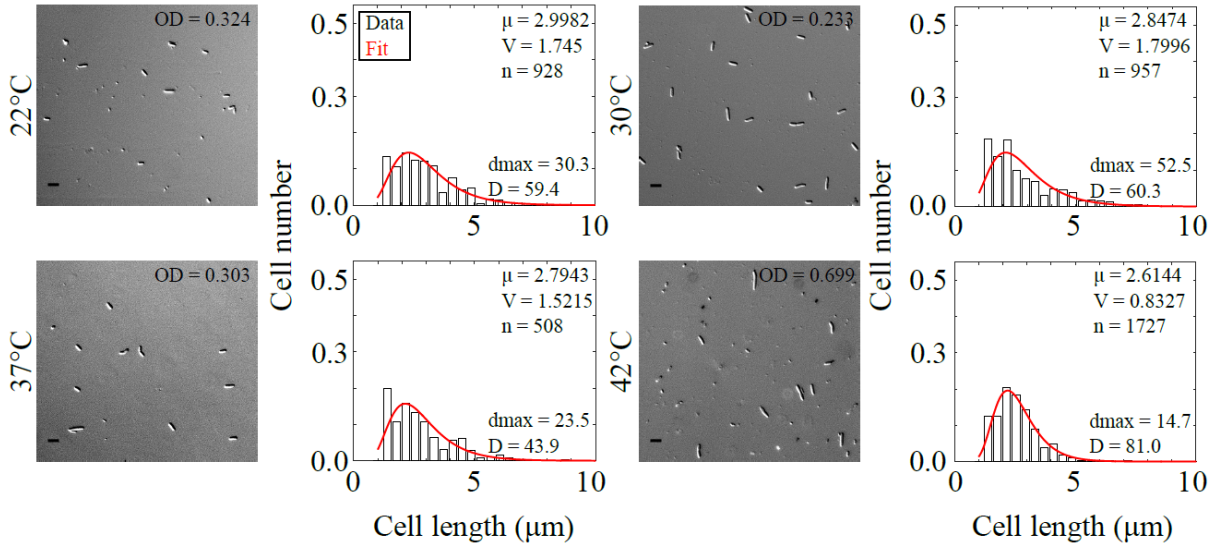


Fig. 3.3: Population cell size distribution at different temperatures. Cells grown at four different temperatures were harvested in the mid- log phase, fixed and imaged in DIC. Scale bar- 5 μm. The respective population cell length distribution fit to the log- normal function (solid red line) have been shown next to each image.

Phenotypic fluctuations were measured in terms of coefficient of variation for an individual population, which is a ratio of the standard deviation (σ) in population cell lengths to the average cell length (μ).

$$Cell\ length\ variability = \frac{\sigma}{\mu} \qquad \text{Equation 5}$$

Cell length variability in wild- type population, as validated by statistics, was similar across all the temperatures, which was reflected in the plot of cell length variability against population growth rate at four different temperatures (Fig. 3.4 A and B).

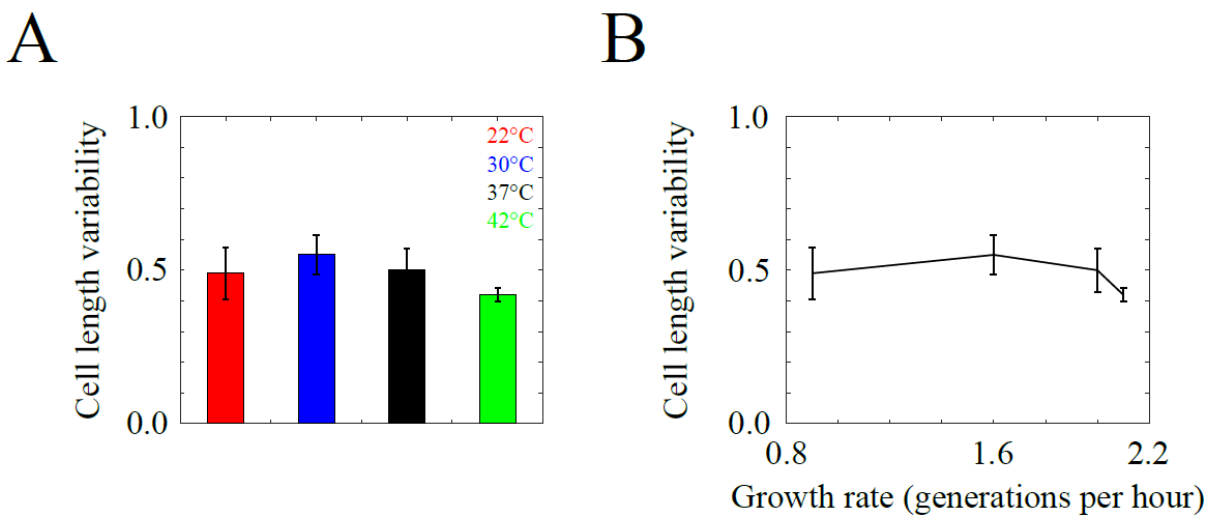


Fig. 3.4: Cell length variability across the temperature. (A) Coefficient of variation of population cell sizes is calculated for the populations grown at 22°C (red), 30°C (blue), 37°C (black), 42°C (green). Error bars- standard deviation. (B) Cell length variability in wild- type populations has been plotted with respective population growth rates. Error bars- standard deviation.

3.2.3 Effect of changes in nutrient availability on the cell size variation

When nutrient source was changed we found apparent decrease in average cell size with decrease in the growth rate. Observation was consistent to the earlier studies[76]. More analysis of the cell lengths by comparing them against log- normal equation (Appendix D, D.1) showed a gradual decrease not only in the average but also in the variance of the distribution with decrease in the growth rate (Fig. 3.5). For further detailed analysis, coefficient of variation was determined for each population propagated in different media and was correlated with respective growth rates.

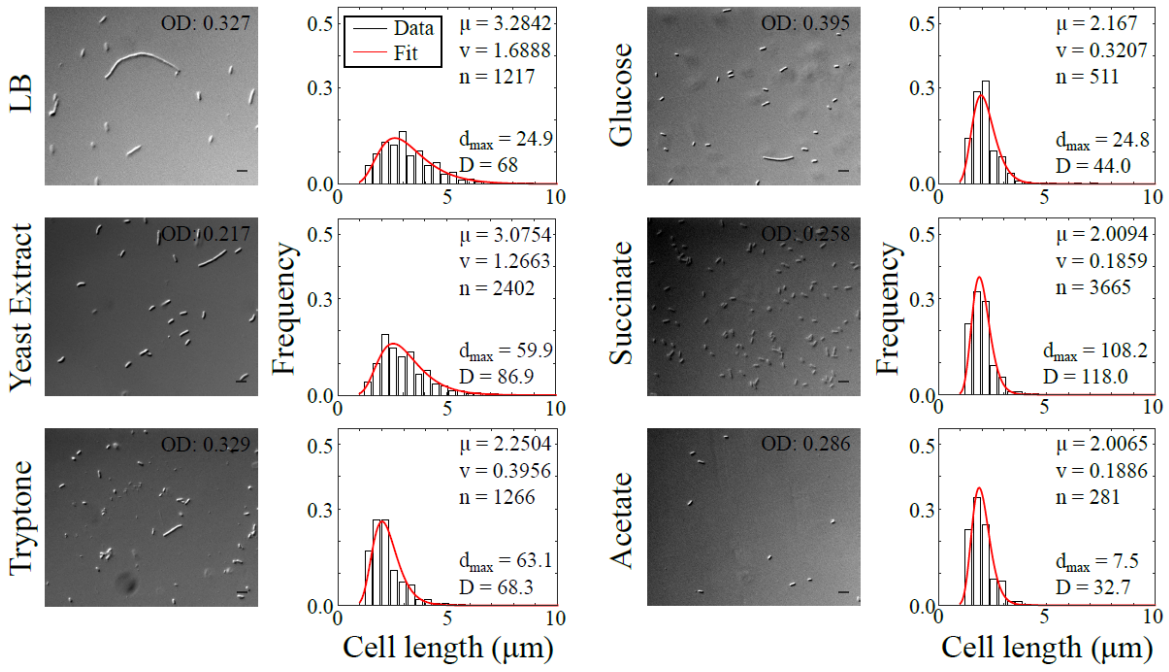


Fig. 3.5: Population cell size distribution in different growth media. Cells collected from mid-log phase were fixed and imaged in DIC for each nutrient condition. Scale bar- 5μm. Their respective cell length distribution was fit to log- normal distribution (solid red line), which has been validated statistically by KS- test.

Cell length analysis of the populations grown in different nutrient media showed two distinct regimes separated by a threshold corresponding to the growth rate of 1 generation per hour (red solid line). Cell length variability showed monotonous increase in the region corresponding to the growth rate greater than 1 generation per hour. However, cell length variability evaluated for the populations growing slower than the threshold growth rate was found to be similar (Fig. 3.6). Results clearly suggested the nutritionally imposed growth rate dependence of cell length variability. Assessment of steady- state batch cultures however, face major set- back in that cells lack synchronous growth. Asynchrony in an individual cell cycle introduces the cell cycle dependent noise in the lengths. Hence, we decided to examine variations in the cell lengths in more controlled environment.

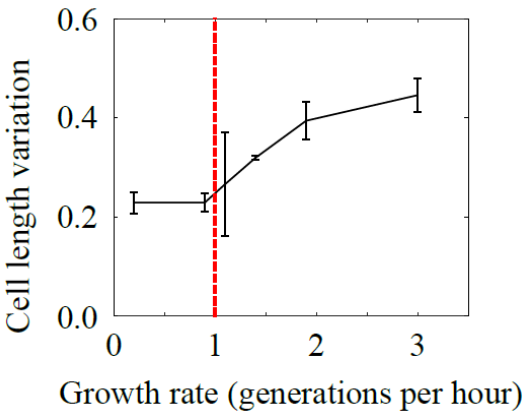


Fig. 3.6: Growth rate dependence of cell length variability. Population cell length variability measured in terms of coefficient of variation has been plotted as a function of respective growth rates in different nutrient media (C). Error bars- standard deviation.

3.3 Noise in the cells of the similar age

To study phenotypic heterogeneity in synchronous cultures, we assessed (a) micro- colonies originated from single *E. coli* mother cell and (b) individual cells at the time of division or birth. Analysis served to estimate the variations in *E. coli* cell lengths independent of the cell cycle stage.

3.3.1 Analysis of micro- colonies

Micro- colony spawned from single mother, consists of sisters with identical genomic DNA as well as cells in similar period of life cycle. Wild- type *E. coli* cells were grown on LB agar pad and their proliferation was captured at regular time intervals using confocal set- up. Though isogenic in nature, cells proliferated with different pace (Fig. 3.7 A and B). Distribution of cell lengths pooled from the analysis of 13 micro-colonies was similar to that of batch- cultures in LB and hence, was consistent to our previous results. Cell lengths of the population range between 1.5 µm to 8 µm. Modal value resided around 2.5 µm, while mean cell length was found to be 3.25 µm, with variance of 0.7688 (Fig. 3.7 C). Individual micro- colony was subjected to an interactive analysis for their cell lengths. Coefficient of variation estimated from the data was then compared with reference to the colonial growth rate (Fig. 3.7 D). Consistent with our previous data, growth rate of 1 generation per hour coincides with an inflection point, which implies that micro- colonies with high growth rate show more phenotypic fluctuations, while micro- colonies with slower proliferation had cells of similar sizes. Our results underlined that cell size variations show strong correlation with growth rate and also is independent of cell cycle phase.

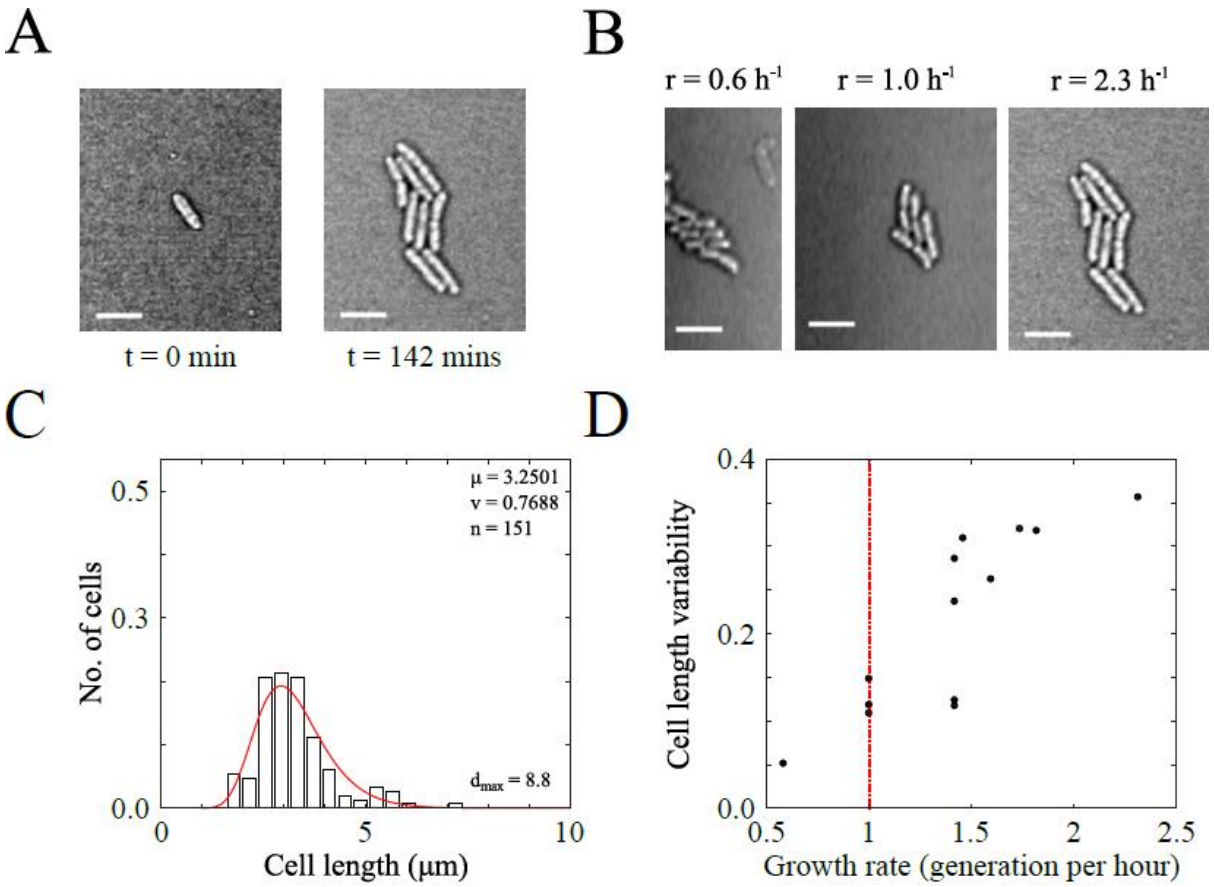


Fig. 3.7: Analysis of micro- colony. (A) Representative images of micro- colony proliferated from a single mother has been shown in DIC. Scale bar- 5μm. (B) End points of three colonies growing at different growth rates have been shown. Scale bar- 5 μm. (C) Frequency distribution of cell lengths pooled from over 13 micro- colonies was fit to log- normal function (solid red line) (D) Noise in cell lengths estimated for an individual micro- colony has been plotted as a function of its respective growth rate.

3.3.2 Analysis of continuous cell cultures

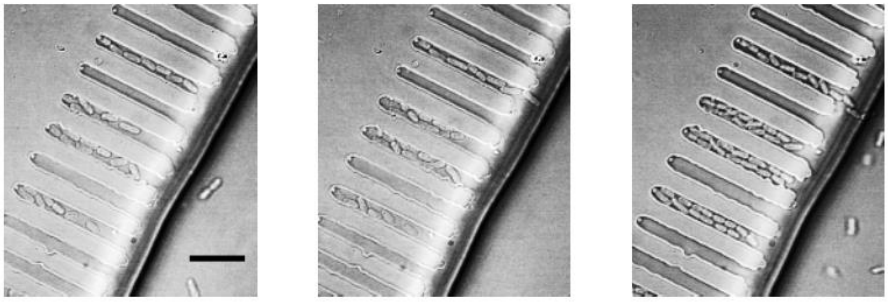
Further evidence came from continuous cultures of MG1655 cells. Measurement of the cell sizes either at birth or before cytokinesis, helped calculate noise in populations of identical age (Fig. 3.8 A and B). In two different experiments, we supplemented cells with LB and M9 + succinate. Fast growing cells show wider distribution of the birth lengths that extended from 2 µm to 5 µm. While, the distribution of the division lengths spread from 3 µm to 10 µm. Cell length distributions of cells grown in minimal media, were more compact at both the life events with considerable decrease in the average values as compared to the cells in LB (Fig. 3.8 C and D). Moreover, coefficient of variations, evaluated for each conditions, were in agreement with the values reported from batch cultures (Table 3.3).

Media	**Event**	**Mean (µm)**	**Coefficient of variation**
LB	Birth	2.729	0.3178
LB	Division	5.2694	0.3115
M9 + Succinate	Birth	1.4913	0.1567
M9 + Succinate	Division	3.2352	0.1393

Table 3.3: Cell length variability of MG1655 in continuous culture.

Thus, having established the correlation between cell size variation and population growth rate we moved on to search for an underlying mechanism.

A



t = 0 mins t = 40 mins t = 100 mins

B

Mother cell at t = 0 mins Mother cell at t = 30 mins New born cells at t = 32 mins

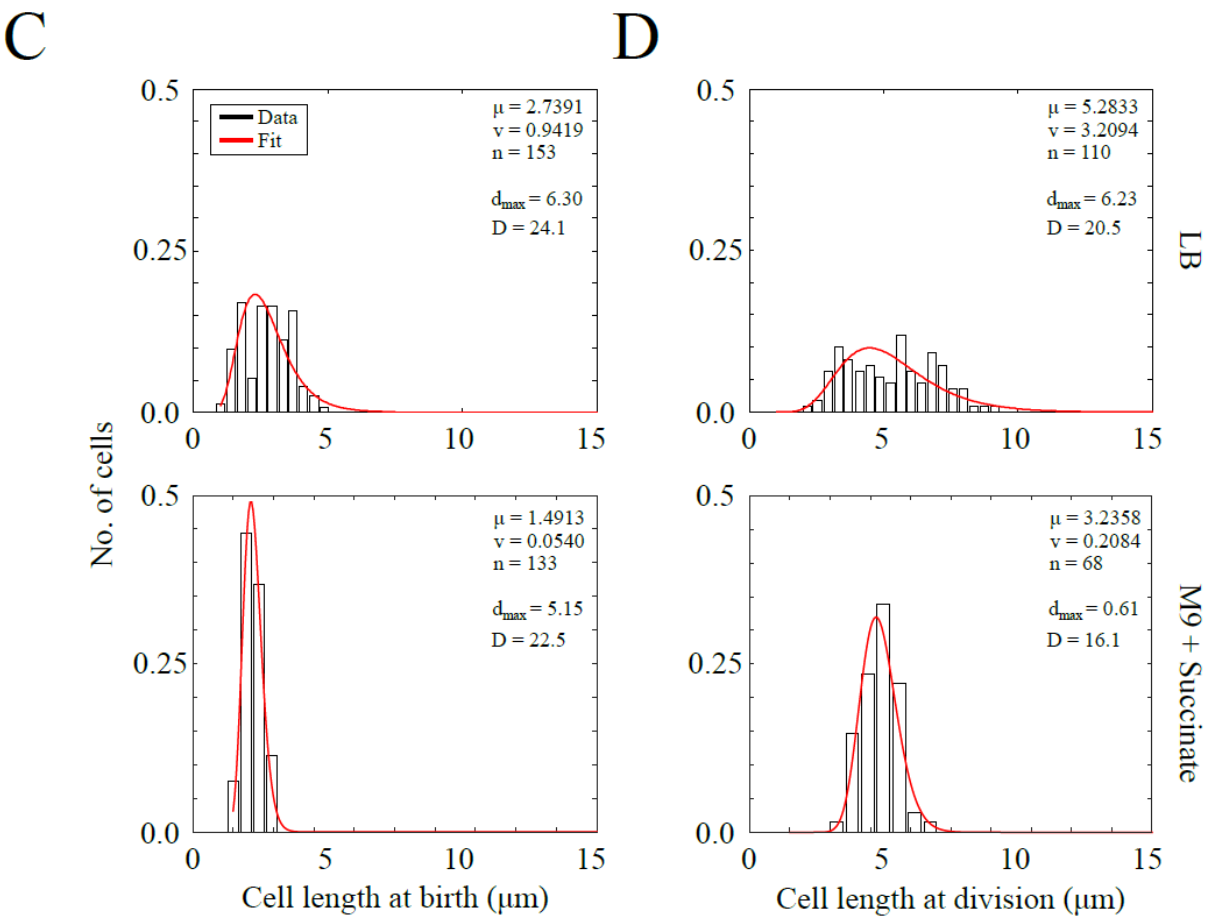


Fig. 3.8: Continuous cultures of *E. coli* cells. (A) Cells of MG1655 were grown in a microfluidic device and imaged in DIC (scale bar- 10 μm) (B) Cells growing in the channels are measured interactively for their lengths at the event of cytokinesis (dotted blue line) and at the birth of two daughter cells (dotted green and red lines) in ImageJ. (Scale bar- 5 μm). Population length distributions of MG1655 cells grown continuously either in LB or in M9 supplemented with succinate have been shown for (C) birth and (D) cell division and were fit to the log- normal distribution (solid red line).

3.4 Correlation between cell length variability and DNA replication

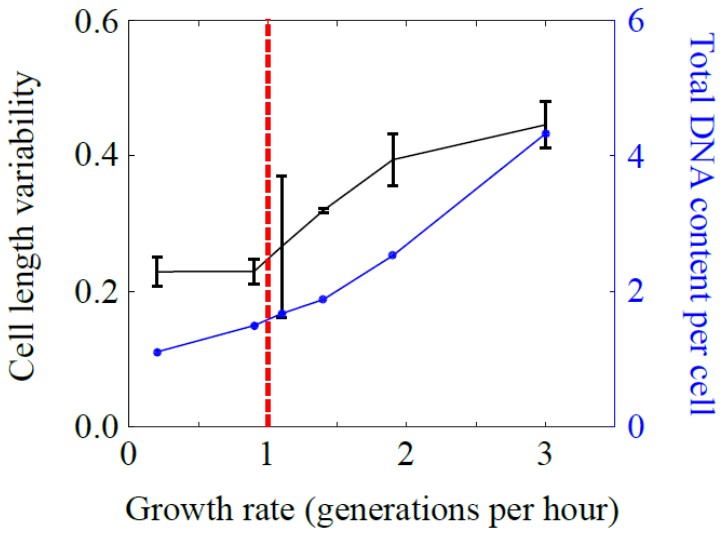
Growth rate determines the relative rate of DNA synthesis in an individual cell. Single cell proliferating at higher growth rate fires replication round more than once in order to compensate for reduced time for completing the cell cycle. Thus, cell with a doubling time less than 60 mins, harbours 4- 6 replication forks at a given time[24]. In other words, with increase in the growth rates, there is an increase in the relative rate of DNA synthesis as well as the genomic content of per cell. We quantified the genomic content of the cell (G) using a formula proposed by Cooper and Helmstetter based on multi- fork replication model of an *E. coli* cell[24].

$$G = \frac{\tau*(2^{(C+D)/\tau}-2^{D/\tau})}{C*ln2}$$ Equation 6

Values of C and D are taken to be constant and are 40 mins and 20 mins respectively. Doubling time of the culture retrieved from different experiments was substituted in place of τ, in order to compute the genomic content (G).

As predicted, plot of genomic content against population growth rate showed a trend similar to that of cell length variation (Fig. 3.9 A). Moreover, correlation between cell length variability and estimated genomic content of the cell for respective population growth rates, was revealed to be linear (Fig. 3.9 B), implying the role of replication process in the existence of phenotypic noise in isogenic populations.

A



B

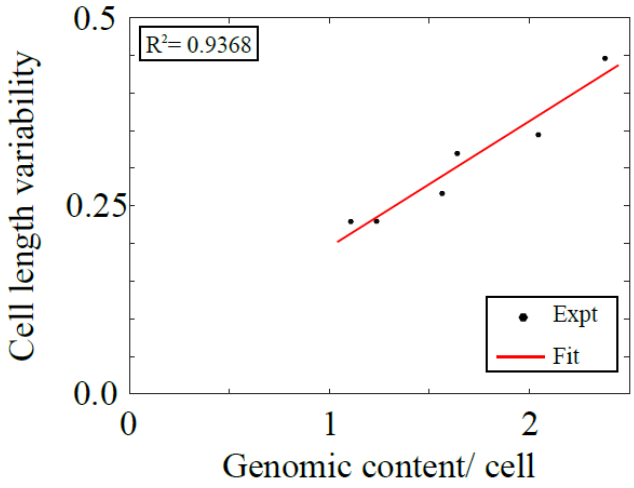


Fig. 3.9: Correlation between cell length variability and per cell genomic content (A) Cell length variability (black solid line) and genomic content of an individual cell (blue solid line) have been plotted as a function of growth rate. Red line corresponds to the growth rate of 1 generation per hour. Error bars- Standard deviation. (B) Cell length variation (black filled circles) plotted against genomic content per cell was fit to the linear function (red solid line).

In bacteria, the process of DNA duplication is tightly coupled with cell division. Hence, random fluctuations in earlier process can put later on the halt. Multifork replication increases the probability of stochastic replication fork stalling, which in turn, adds up in the events that lead to the cell division delay. These fluctuations at single cell level, can cause cell length variation in population.

We then proceeded to test the hypothesis, with the help of genomic mutations and environmental perturbations.

3.5 Estimation of cell size variation in genetically mutated populations

In *E. coli*, cell division is tightly coordinated with DNA replication. Mechanisms that couple these two events comprise SOS response pathway and nucleoid occlusion[27]. Stalled replication fork is repaired by RecA which also promotes expression of SulA, by inducing auto- cleavage of LexA, repressor protein for SOS box. SulA, an effector of SOS pathway, delays cell division by sequestering FtsZ molecules[34,35]. On the other hand, nucleoid occlusion pathway involves SlmA that evokes disassembly of Z- ring over DNA molecules, thereby preventing septation across genomic DNA[29]. Absence of any of these proteins can disrupt the coordination between replication and division. We characterized the populations of genetically disabled cells in order to check the effect of replication stochsticity on occasional filamentation in clonal population.

3.5.1 Cell length variability in deletion mutants of *E. coli*

Since, size variation was reckoned with presence of defective copy of RecA in our previous studies[54], perusing of SOS mutants became the prime focus of our studies. Cell length distributions of mid- log phase cultures *ΔrecA* and *ΔsulA* strains were compared with that of the wild- type. Deletion of *recA* gene was observed to cause increase in the frequency of elongated cells in culture. Also, cell length distribution extended to 15 µm, though there is no

significant shift in the mean cell length, as tested by log- normal function. Cell length distribution in *ΔSulA* populations, however, was found to be more compact than MG1655 (Fig. 3.10 A). Coefficient of variation scored for the population cell lengths showed that cell length variability in *recA* deletion mutants is higher than the wild- type, while it was minimized in the populations with zero SulA molecules to inhibit cell division (Fig. 3.10 B).

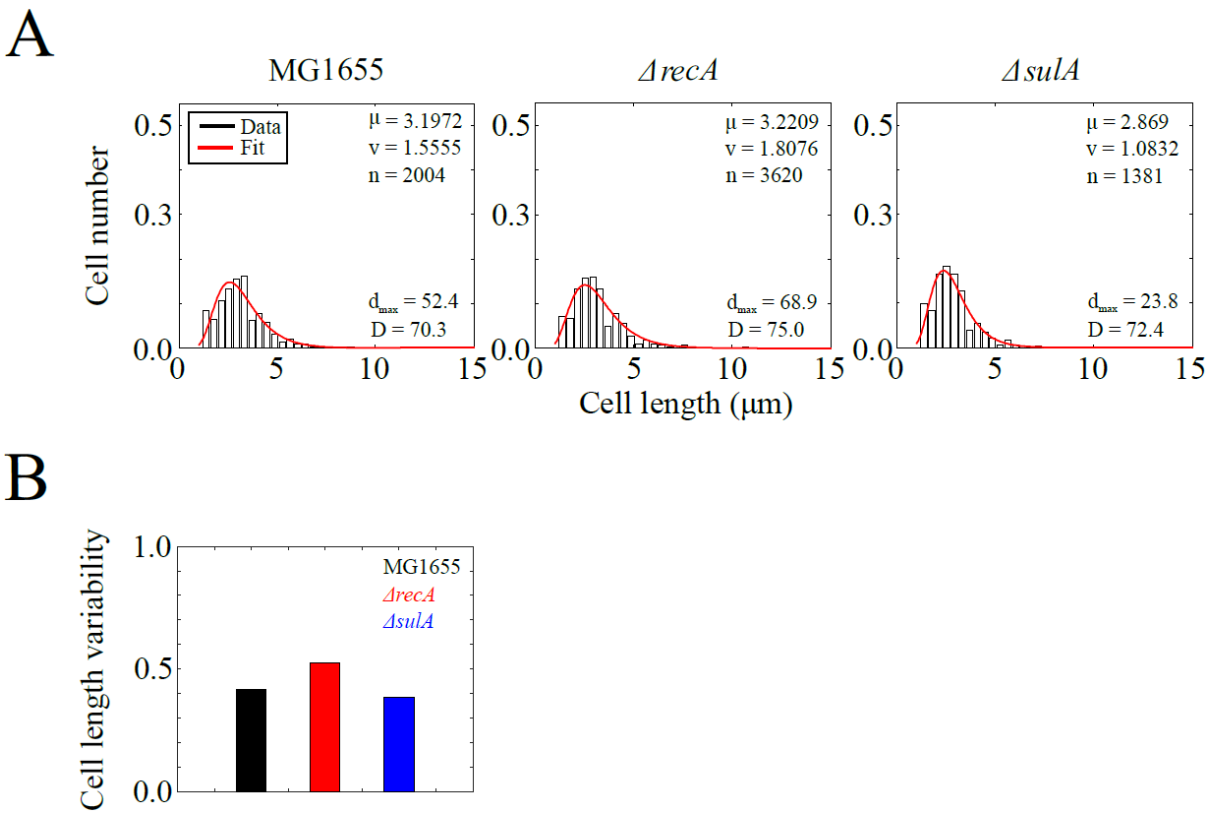


Fig. 3.10: Cell length variation in mutant populations. (A) Cell length distributions for isogenic populations of MG1655, *ΔrecA* and *ΔsulA* harvested in the logarithmic phase of the growth, were fit to the log- normal distribution (red solid line). (B) Cell size variation in each population has been estimated in terms of coefficient of variation.

3.5.2 Rescue of *ΔrecA* phenotype using ectopic expression of RecA

RecA is an initiator of SOS response pathway. Hence, its absence mitigates the ability of the cell to recuperate the stalled replication fork. Incompletely replicated as well as segregated, nucleoids, in that case, put the cell division on hold and cell continues to elongate into filamentous morphology. Hence, as a next step, we took a glance at the population length distribution of *ΔrecA* strain overexpressing RecA either as RecA- mCherry or as native RecA protein. Our objective was to observe for reshaping of size distribution, when RecA is provided externally to the system with deactivated RecA expression. In the populations harvested from mid- log phase, distribution of the cell lengths extended till 10 µm, similar to the populations of MG1655 expressing either RecA- mCherry or RecA ectopically (Fig. 3.11 A). Note that RecA- mCherry expression, though controlled by P_{lac}, did not require an addition of an inducer to the medium, owing to the leaky nature of the promoter (Appendix B). On the contrary, expression of RecA under the control of P_{BAD}, required addition of arabinose to the medium. Coefficient of variations calculated for cell sizes of four individual populations were found to be comparable (Fig. 3.11 B). Results were suggestive of involvement of RecA mediated SOS repair pathway of DNA replication in cell size variation.

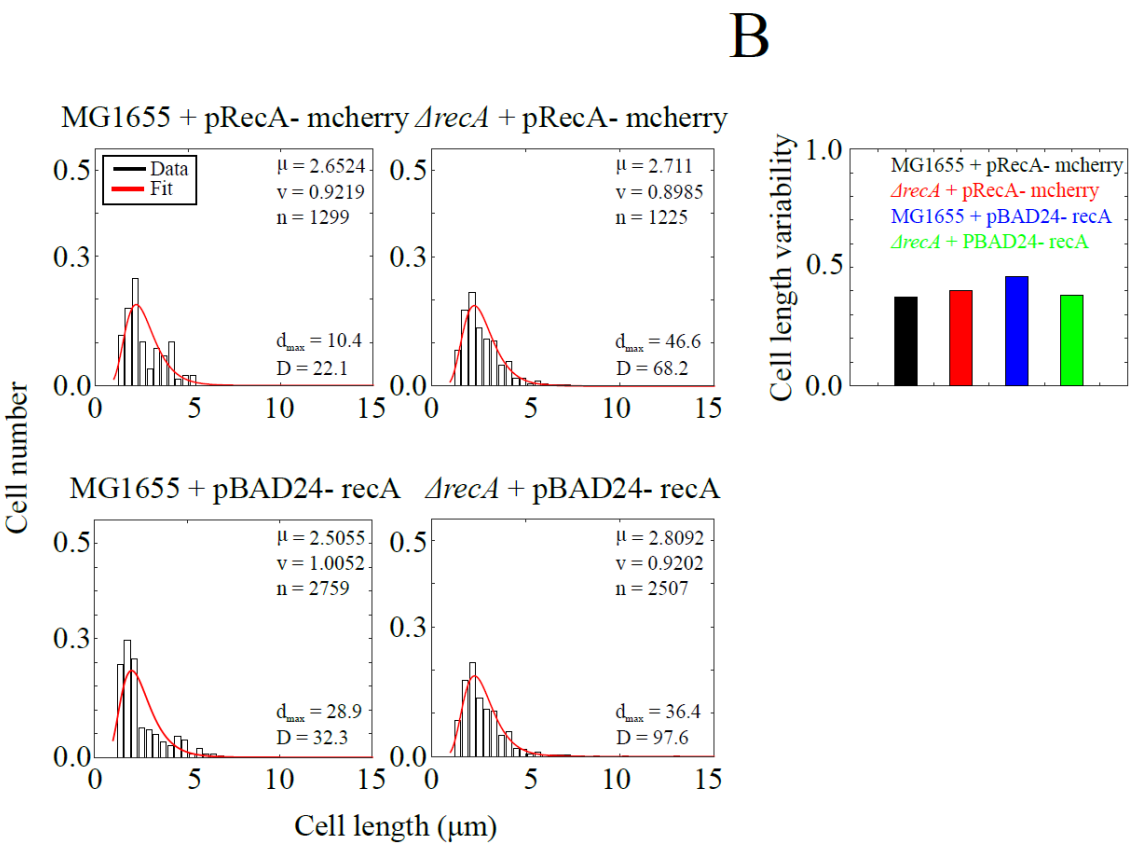


Fig. 3.11: Rescue of phenotype in *ΔrecA* populations by ectopic expression (A) Cell length distribution of log phase populations of MG1655 and *ΔrecA* transformed either with pRecA-mCherry or with pBAD24- recA plasmid were fit to the log- normal equations (red solid line) (B) Coefficient of variations measured for the same have been shown.

Further validation for this hypothesis was received by cataloguing the changes in size distribution in various *E. coli* populations, upon ameliorated frequency of replication halts. According to our plan, accelerated replication cease per cell, might highlight the effect of random replication stalls on cell size. Observation, when coupled with the analysis of deletion mutants would help confirm the molecular players in non- genetic size heterogeneity.

3.6 Effect of replication stalling on cell lengths

Hydroxyurea (HU) inhibits the penultimate step of de novo synthesis of dNTPs in bacterial cell, by blocking ribonucleotide reductase enzyme[77], though recent report has suggested that HU directly damages DNA rather than to interfere with RNR function[78]. Nevertheless, the strategy was used to increase the events of replication fork stalling to study the occurrence of filamentation in batch as well as continuous cultures. Results enabled us to connect random fluctuations in the replication process with phenotypic variation in the clonal populations.

3.6.1 Effect of hydroxyurea on batch culture population

Hydroxyurea treatment was carried out at two different growth rates achieved by growing cells in two different media, viz., LB and M9 + succinate. Size variation was studied with respect to the increasing concentrations of hydroxyurea. Four different strains, *viz.*, *ΔrecA*, *ΔslmA*, *ΔsulA* and *ΔrecA* bearing pRecA- mCherry plasmid were chosen for these studies and were compared against wild- type strain for cell length variability. Except for *ΔsulA* cells, all other strains when grown in LB, respond to the treatment, with linear increase in the size variation till hydroxyurea concentration reached to 30 mM. For slow growing cells, cell length variation was minimal and did not change with an increase in hydroxyurea concentration. Fast growing *ΔrecA* populations showed the highest fluctuations in the size, while *ΔSlmA* culture exhibited the trend exactly similar to that of wild- type. Trend of size fluctuations in *ΔrecA* mutants was restored to that of MG1655, upon overexpression of pRecA- mcherry protein in the cell (Fig. 3.12 A).

Moreover, treatment of MG1655 populations with trimethoprim increases the cell length variations at high growth rates (Fig. 3.12 B). Considering the inhibition of Dihydrofolate reductase and subsequent reduction in dNTP pool by trimethoprim[79], the results upheld our proposition that the random fluctuations in the replication as a possible mechanism for non-genetic size variation in clonal populations. We buttressed the observation further by observing similar effect in continuous cultures.

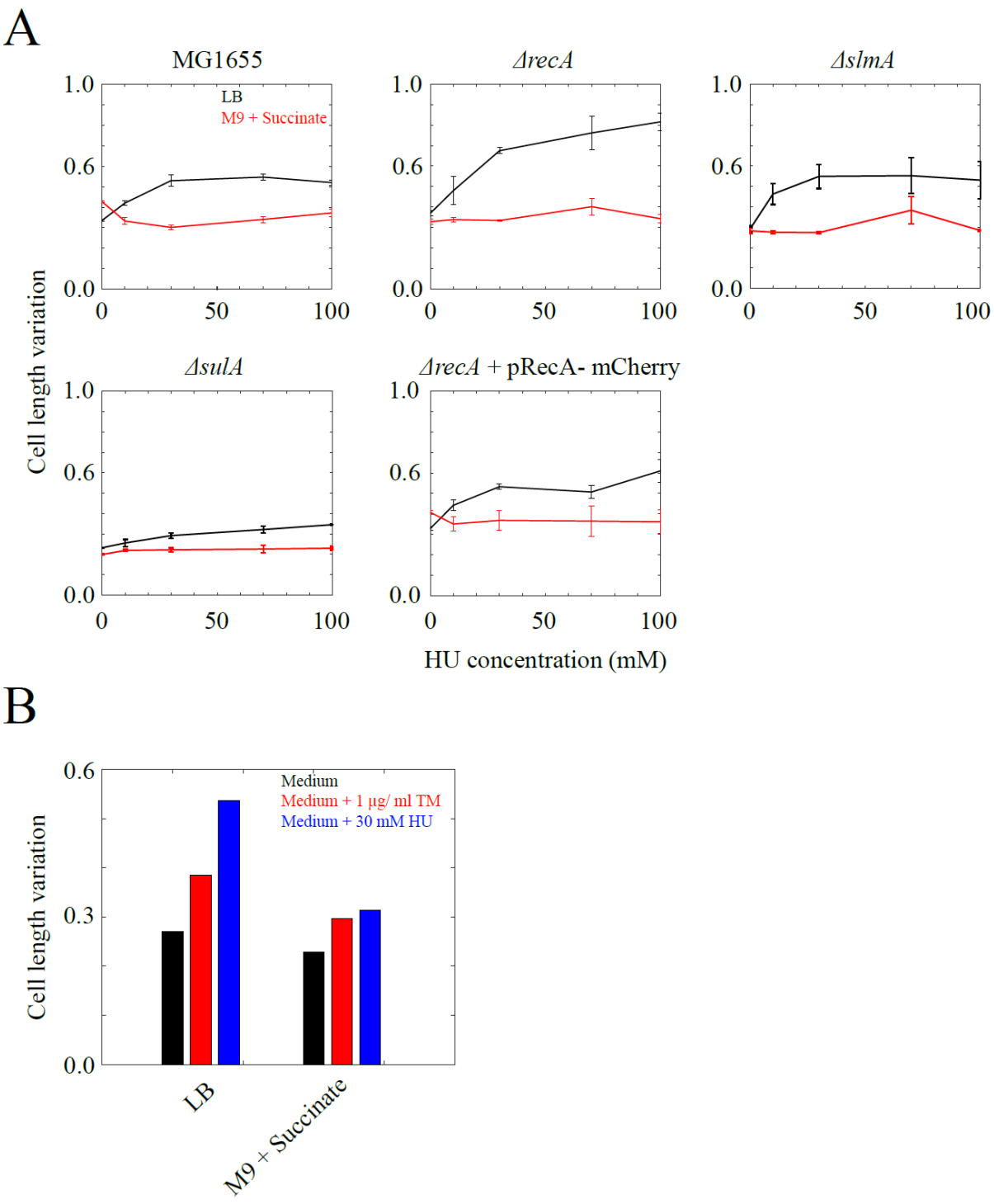


Fig. 3.12: Effect of replication stalling on population variation of cell length (A) Variations in the cell lengths in the populations of MG1655, *ΔrecA*, *ΔslmA*, *ΔsulA* and *ΔrecA* transformed with pRecA- mCherry grown either in LB (black solid lines) or in M9 + succinate (red solid lines) are plotted as the function of hydroxyurea concentration. Error bars- standard deviation.

(B) Cell length variations of wild- type populations are compared for the trimethoprim (red) and hydroxyurea (blue) treatments.

3.6.2 Effect of hydroxyurea on continuously growing populations

Continuous cultures provide an advantage of examining the phenotypic variations in live cells under experimental conditions. MG1655 and *ΔrecA* strains were chosen for this set- up. Analysis of MG1655 cells help conclude the effect of replication stops on cell size in wild type populations, while use of *ΔrecA* populations not only help authenticate our observations in wild- type populations but also supply an evidence for the possible role for SOS response in cell length variability. Both the populations (in different microscopy arrangements) were allowed to grow continuously in LB for 1 hour. Growth was followed sequentially by hydroxyurea treatment for 1 hr and recovery for 2.5 hr. Differential response to the hydroxyurea in isogenic populations was observed. In some of the microchannels, cells continued the growth to generate healthy cells of around 2 µm size (Fig. 3.13 A and C), while, other bacterial cells start elongating during recovery phase. However, not all the cells in channels elongate, contributing to the size variation (Fig. 3.13 B and D). In many other micro-channels, we observed that filamentous cells divided asymmetrically giving rise to two daughters, during their recovery. One of daughter cells continued growing normally, while in other daughter cell, division remained arrested, leading to the population size variation. Experiment, thus, supports our hypothesis which predicts the role of stochastic replication arrests in length heterogeneity. Moreover, it set the precedence for investigating into the molecular mechanism.

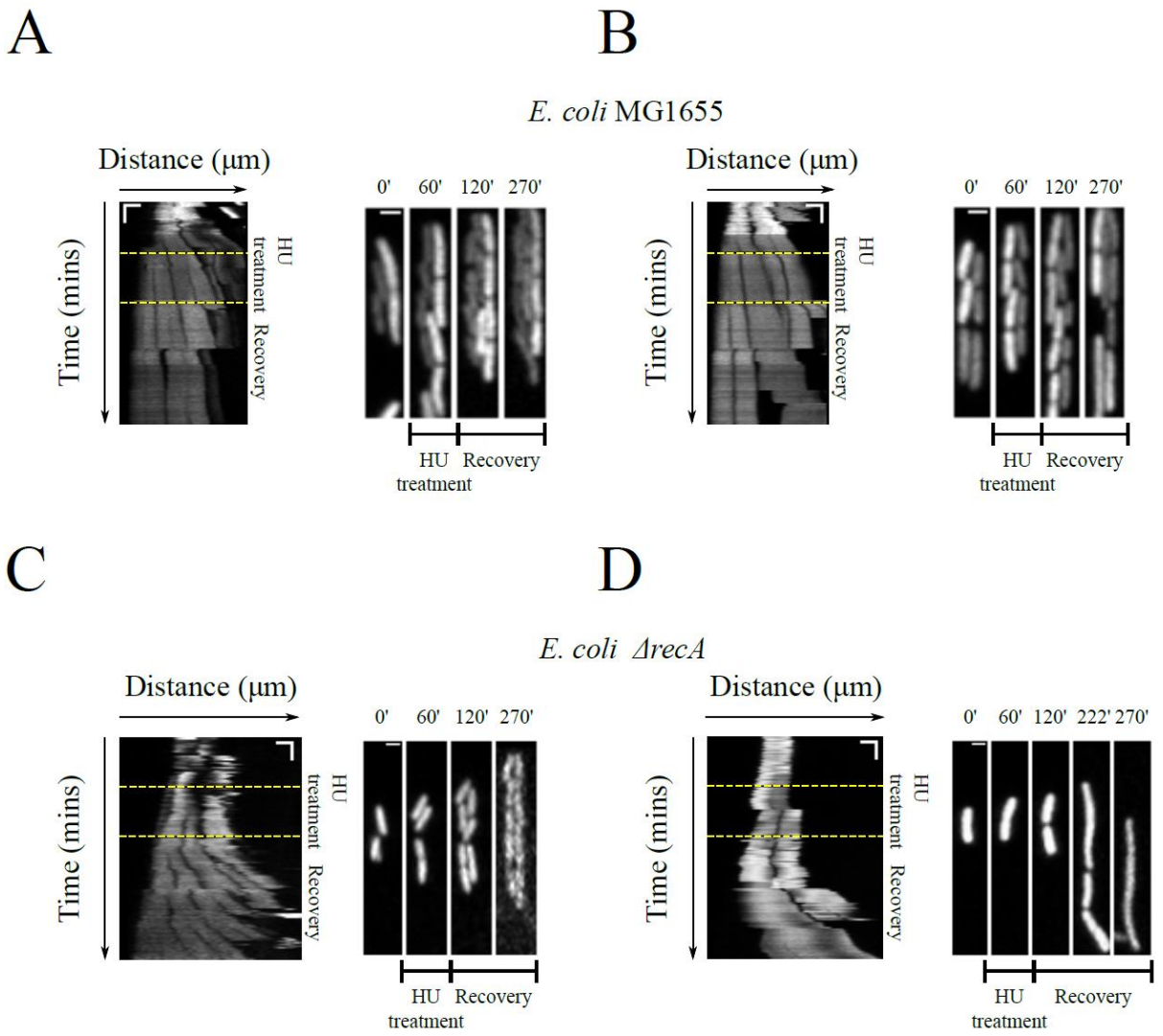


Fig. 3.13: Effect of hydroxyurea treatment on continuous cultures (A) and (B) represent two separate microchannels in mother machine device used to grow MG1655 temporally. While, (C) and (D) represent the growth of *ΔrecA* cells in microfluidic device through the time. Kymograph scale bar: Horizontal scale bar- 2 µm, Vertical scale bar- 100 secs. Montage scale bar: 2 µm.

3.7 Correlation between cell length variability and the frequency of replication fork stalling

RecA is known to get recruited at stalled replication forks to launch DNA repair mechanism termed as SOS response[31]. It makes RecA, an excellent marker to locate stalled replication fork. MG1655 cells expressing RecA- GFP endogenously, were grown in three different nutrient media (Fig. 3.14). We scored all the three populations for number of cells with RecA-GFP puncta coinciding with DAPI stained genomic DNA. Estimated proportion of population with RecA and nucleoid colocalization was found to be linear with growth rate. Treatment with hydroxyurea, indeed, increased the occurrence of such cells in growth rate dependent manner (Fig. 3.15 A and B). We further tested the proportion of the cells with reference to the cell length variation in respective populations. A linear correlation between the cell length variability and the replication fork stalling events (Fig. 3.15 C) was detected. It thus, establishes the molecular link between replication stochsticity and cell size variation in *E. coli* populations.

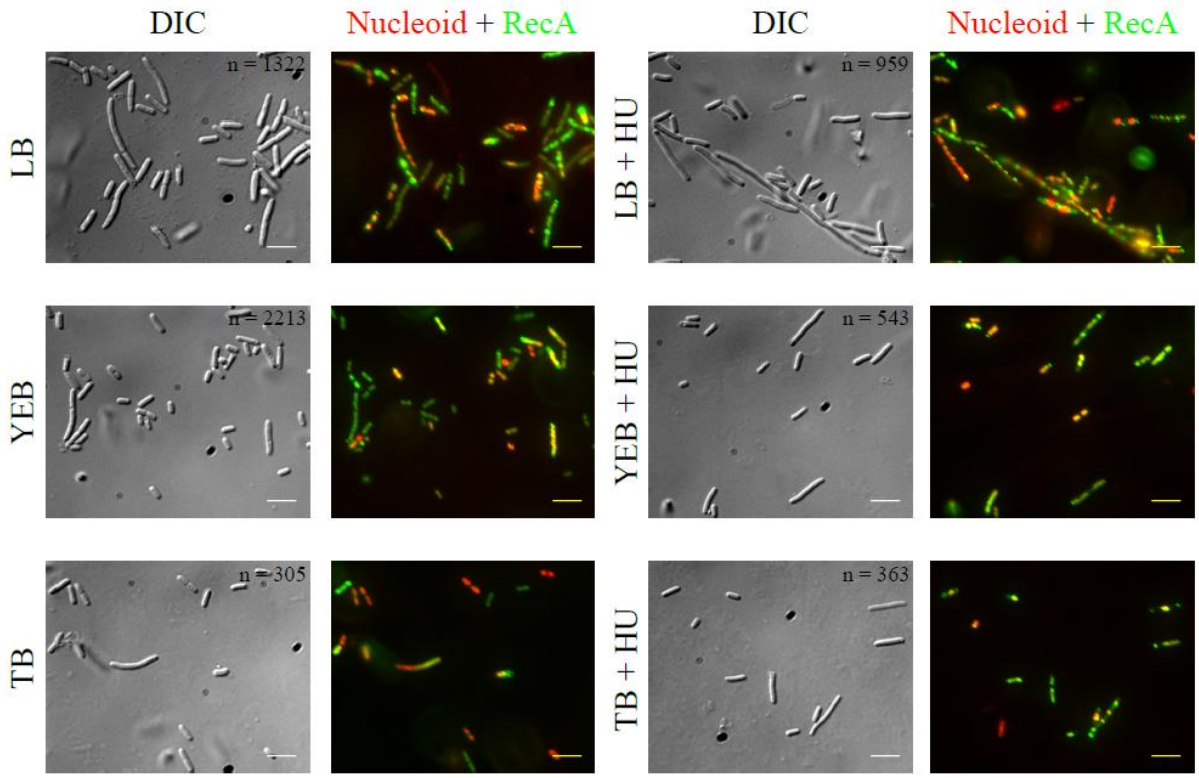


Fig. 3.14: Colocalization of RecA with the nucleoid. MG1655 expressing RecA- GFP endogenously was grown either in LB or in YEB or in TB. Mid- log phase culture was visualized in DIC, GFP and DAPI channels. First two columns show the control populations, while the two rightmost columns indicate the effects of hydroxyurea on bacterial populations. Scale bar- 5 µm.

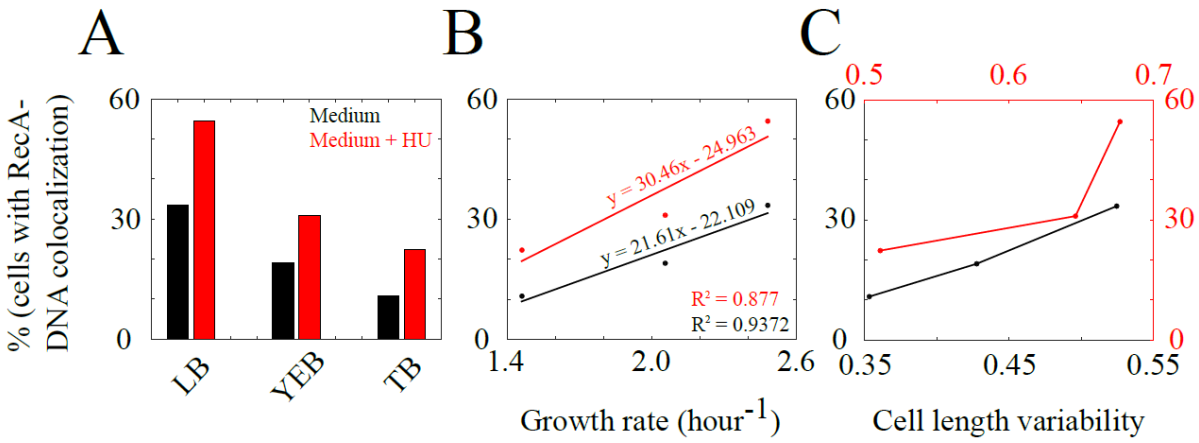


Fig. 3.15: Correlation of replication stalling and growth rate. Proportion of population with RecA colocalized with nucleoid was quantified for the control populations (black) as well as the hydroxyurea treated populations (red) and has been plotted as the function of respective growth media (A), population growth rate (B) and population cell length variation (C).

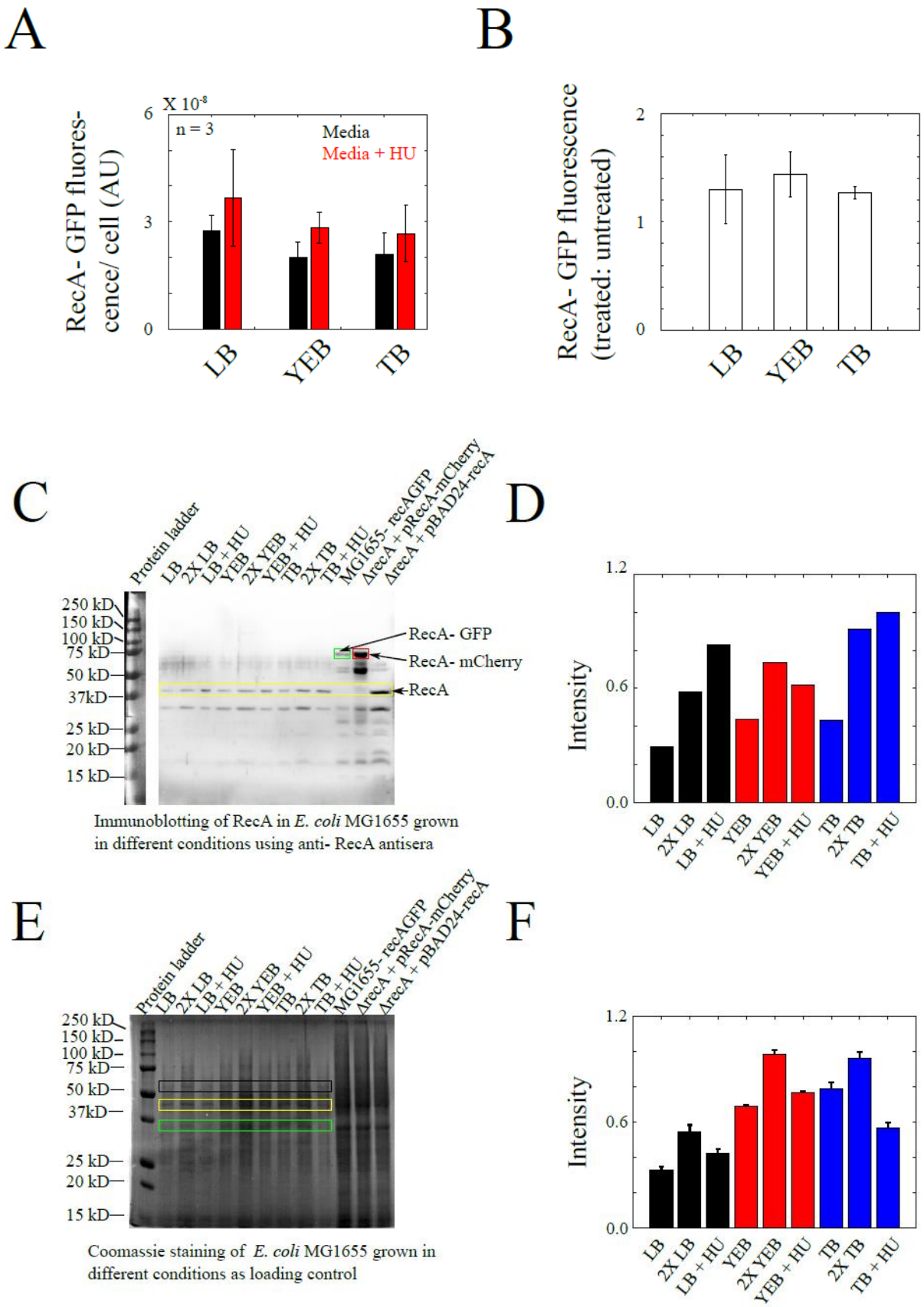
3.8 Expression of RecA as a response to increased DNA breaks

Relationship was further confirmed by measuring intracellular RecA levels after the introduction of hydroxyurea into the system and comparing them with RecA expression prior to the treatment. Fluorimetric analysis of MG1655 expressing endogenous RecA- GFP showed marginal increase in GFP signal after HU treatment at all the growth rates (Fig. 3.16 A). Fluorescence ratio of treated samples to that of untreated averaged around 1 (Fig. 3.16 B).

However, GFP is not a native protein to *E. coli*. We considered a possibility of retention of the tag in the cell, after the destruction of native part of the fusion by proteases. It might have interfered with our readings and hence cellular RecA levels in MG1655 cells subjected to identical growth conditions (Fig. 3.16 C) were probed, using rabbit anti- sera developed against *E.coli* RecA[56]. We observed that the RecA expression almost doubled up (Fig. 3.16 D), after HU treatment, while the levels of other cellular proteins, as analysed by Coomassie blue staining, were exactly comparable under both the conditions (Fig. 3.16 E and F).

In addition to this, endogenous RecA- GFP levels in MG1655- *recA*- *gfp* strain grown in LB were also analysed. (Fig. 3.16 G). Result obtained were comparable to RecA levels in wild-type strain. Both the organisms showed 2.5 times increase in RecA expression, upon HU treatment (Fig 3.16 H). Consistent with our earlier observations, expression of other proteins remain unaltered after changes in growth conditions (Fig. 3.16 I and J).

Experiment attests RecA as one of important molecular players in cell size heterogeneity in an isogenic populations, mediated by inherent stochasticity present in DNA replication process.



Immunoblotting of RecA in *E. coli* MG1655 grown in different conditions using anti- RecA antisera

Coomassie staining of *E. coli* MG1655 grown in different conditions as loading control

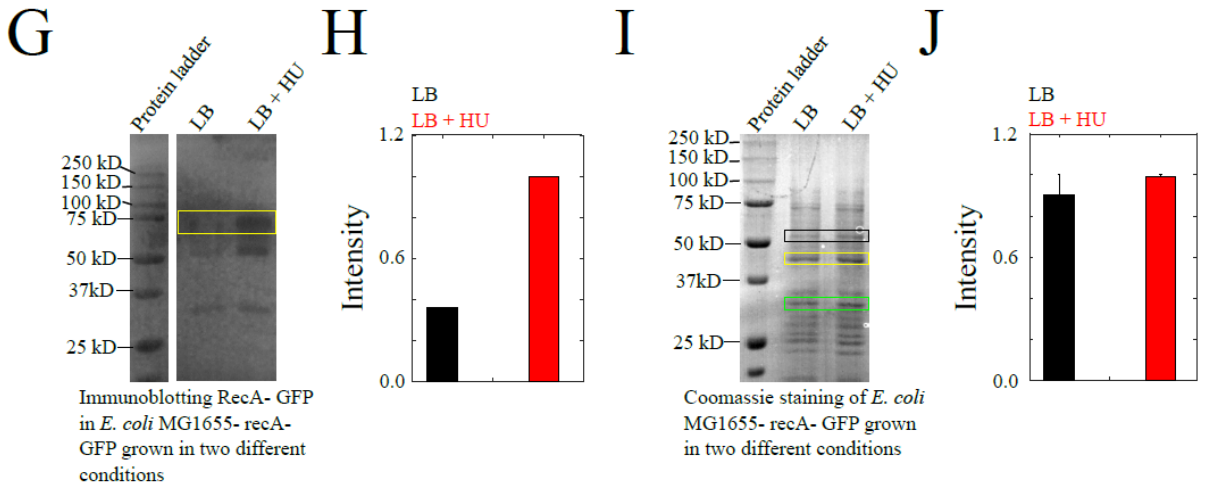


Fig. 3.16: Changes in the expression of *recA* gene after HU treatment (A) Fluorimetric analysis of RecA molecules by counting the GFP signal. The cells with endogenous *recA-gfp* gene under the native control were grown in LB, YEB and TB. Black bars represent untreated populations, while red bar stand for the population treated with hydroxyurea (B) The ratio of GFP signal in untreated cells to that in treated cells have been shown for LB, YEB and TB media (C) Immunoblot to probe RecA protein in HU treated and untreated populations grown either in LB or in YEB or in TB. Yellow box indicates the approximate positions for 38 KDa, while green and red represent the positions for 70 KDa proteins on the blot. (D) Densitometric quantification of an Immunoblot has been shown (E) Identical samples were analysed for protein load on SDS gel stained with Coomassie Blue. Random horizontal lanes (black, yellow and green) were selected, quantified and their averages have been shown in (F). Error bars- standard deviation (G) Western blot analysis of RecA- GFP levels in MG1655- recA- gfp strain prior HU treatment and after HU treatment. Yellow box represents the approximate position of 70 KDa protein on the blot (H) Quantification of the blot has been shown (I) The samples were analysed on SDS gel for loading control (J) Three lanes selected (black, yellow and green) were estimated for their intensity. Error bar- standard deviation.

3.9 Discussion

Diverse mechanisms bring robustness into the cell size of *E. coli*. Pathways regulate the cell division spatio- temporally, thereby maintaining the cell size of a new- born cell. For instance, MinCDE system oscillates between the two poles of an *E. coli* cell and prevent the polymerization FtsZ at the polar region. Process ensures the localization of septum away from the pole and forbids the formation of the mini- cells or the cells with different lengths[61–63,80–82]. SlmA mediated nucleoid occlusion inhibits the aggregation of divisome machinery in nucleoid occupied area[29]. Mechanism controls the cytokinesis spatio- temporally and connects it with the growth of the cell.

An average cell size of *E. coli* is determined by nutrient- imposed growth rate. At higher growth rate the cells have been found to be longer as compared to the cell growing slowly[76]. However, effect of growth rate on the population length distribution has not been discussed before. Here, we showed that the extent of deviation from average cell length, in wild- type populations of *E. coli* depends on the growth rate determined by the carbon source in an immediate environment. Cell length variability is biphasic with reference to the population growth rate. Variation in population lengths is proportionate with growth rates that demand the multi- fork replication in an individual cell. On the other hand, its value reduces but remains more or less constant when the cell proliferation time is long enough to easily accommodate one round of DNA replication per generation. Results obtained with batch cultures were further buttressed by studying cell length variation in a micro- colony as well as in the population growing continuously in microfluidic device. Fact that, micro- colony is originated from a single mother, gave us an advantage of studying isogenic cells with similar cellular age. Similarly, evaluation of cell lengths in new- born as well as dividing mother cells, help eliminate the noise that might be contributed by age dependent inconsistencies across the population.

Cellular filamentation has its close correspondence with DNA related metabolic process. Growth rate directs the extent of DNA synthesis in *E. coli* cell. Inspection of the connection between rate of DNA synthesis and size heterogeneity was obvious next step. We made two important observations: (a) the genomic content quantified per cell increased with the population growth rate, which, in fact, was in accordance with established reports[24,25] and (b) trend of increase in genomic DNA with population growth rate was similar to that of cell length variability, which was then construed as probable contribution to the size variation from DNA metabolism.

In *E. coli*, DNA replication is coupled to the cell division through two major pathways. One of them is nucleoid occlusion. The other is RecA mediated SOS response pathway. RecA is a well- known recombinase in prokaryotes, which repairs the DNA upon replication stalling[31]. During the repair process, it delays the cell division by activating SulA mediated FtsZ sequestering[35]. In the absence of cell division, cell resumes its growth governed by MreB[83] and becomes filamentous.

Our hypothesis proposes the probable multiplicative increase in the frequency of stochastic stalling of replication fork due to multifork replication at higher growth rates, postpones cell division event and gives rise to the elongated cell. Stochasticity is innate to the replication. Progression of replication fork can be stalled for various reasons. Replication process is slowed down at certain chromosomal loci due to the protein barriers which may include operator proteins or proteins that are associated with terminal sequences on the DNA. In certain situations, replisome may experience head- on collisions with RNA polymerase complex, which can cause pause in the replication. Another reason for replication fork stalling is an unavailability of building blocks or enzyme components. Considering basal level noise exists in all metabolic processes in a cell, there can be occasional stoichiometric imbalance in the numbers of enzymes that are part of replisome complex or biomolecules like dNTPs that are

substrate for DNAP. It can transiently stop DNA replication. DNAP processivity causes episodic dissociation of DNAP from DNA. These events may not contribute to the cell length variability, as recovery of replication pace in this cases, is immediate and hence does not require cell division to be postponed. However, RecA mediated recovery of replication fork, which has been arrested due to the occurrence of DNA nicks, plays significant role in the generation of cell length variability in isogenic populations, for SOS response thwarts cell division, while DNA repair takes place in a cell. Though the process happens in an individual cell, at population level it begets the heterogeneity in the lengths.

Hypothesis was tested by (a) disabling the DNA repair system and (b) Increasing the stop frequency of the replication fork. Idea was to observe an upsurge in the instances of halt in cell division in major proportion of the population, when the replication is paused frequently in a single cell.

We found increase in the percentage filamentation in the populations of *ΔrecA* mutant, which is reflected in its cell length variability. Coefficient of variation of this population was higher as compared to the wild type population. We noticed exactly the opposite effect when *sulA* gene was deleted from the genome, as cell length variability showed a slight decrease. According to the recent report HU treatment transiently halts DNA replication, recovery of which does not require RecA[78]. However, the stock of HU used in this experiment was previously sterilized and stored at 4ºC. It might have led to the degradation of HU into intermediates that damage DNA. Under such circumstances the recovery of the stalled replication fork needs RecA[78]. Loss of RecA protein, attenuates the recovery of the stalled replication in a cell. Replication arrest for prolonged time compels the cell elongate into the filamentous form. We predict the presence of mechanism(s) other than SOS response and nucleoid occlusion that fetter(s) the cell division back. Live cell imaging of *ΔrecA* cell, with nucleoid tagged by HupA- GFP on LB agar, clearly showed the presence of unorganized

nucleoids occupying most of the cell volume in an elongated cell, as opposed to a healthy cell that possessed compact DNA foci. Our result is in compliance with the discovery in 2013 by James Walker group. It states that incompletely replicated and segregated nucleoid at division site can inhibit the polymerization of FtsZ, without the help from Min system, SOS response and SlmA[84]. When SulA is knocked out from the system, cytokinesis is resumed even when DNA is being repaired by RecA. Since, nucleoid occlusion protects only region around the chromosomal DNA, rest of the cell except poles are prone to Z- ring formation in the absence of SulA. Thus, occurrence of filamentation are reduced in *ΔsulA* populations and its cell length distribution is more compact than MG1655. Moreover, phenotype of *recA* deletion mutant was rescued when RecA was overexpressed in the cell. Cell length distribution in these populations was comparable to that of wild- type. Note that we did not examine cell length variability in *lexA* mutants. However, *ΔlexA* mutation has been reported to be lethal[85]. In addition to this, *lexA3 (Ind-)* mutations (non- cleavable variant)[86] could have been considered for these studies. Though LexA molecules, being unresponsive to RecA mediated auto- cleavage, can show similar trend as that of *ΔsulA* mutants, under optimal conditions. But, upon HU treatment, LexA molecules will shut down SOS response, resulting in a trend similar to *ΔrecA*. Hence, we are not sure studies of lexA3 populations will support our results. Nevertheless, use of genetic mutants thus, (a) highlighted that replication fork stalling, indeed, engenders the cell size variation and (b) paved the path for further investigate into the molecular mechanism that links the replication stochasticity with population cell length variability.

Increased stall events per cell linearly increased the cell length variability in wild- type populations at high growth rates, validating our hypothesis. More importantly, phenotypic variation was further elevated in genomic knock- out of *recA*, which is subsequently reduced to the wild- type level by extra- chromosomal expression of RecA. While, *sulA* deletion makes the population unresponsive to the treatment. Importantly, cells responded to the drug treatment

at higher growth rate, while at slower growth rates cell length variability did not change with HU concentration. Multifork replication at high growth rate also multiplies the events of probable replication arrest per cell, upon DNA damage. While, at slow growth rates there are only two replication forks per cell, which does not alter the probability of replication stall even after the drug treatment. Here, we would like to mention that we did not check the functional efficacy of extra- chromosomally expressed RecA in *E. coli* cells. Scoring the survival rate of transformed *ΔrecA* populations and comparing it with that of its parent population after HU treatment or UV exposure could have enabled us to understand the functional integrity of the protein. Also, the extent of LexA cleavage after similar treatments could have been a nice way to probe the RecA function. These experiments not only could have increased the reliability of our data, but also could have backed up our assumption that SOS pathway connects stochastic DNA replication with size heterogeneity, with substantial evidences. Nevertheless, we filled this gap by observing the cell elongation pattern in continuous cultures under the influence of HU treatment as well as alignment of RecA foci with reference to genomic DNA at different growth rates.

Analysis of live cells growing in a microfluidic device, perhaps, was the crucial experiment in that, it provided us with a sense of the events that lead to the length variations in clonal population. In both the strains of *E. coli*, few cells getting elongated after the introduction of hydroxyurea into medium. Rest of the population, nonetheless, kept dividing. It causes significant increase in the standard deviation, with slight increase in average cell length. As a result, an increase in phenotypic fluctuations estimated in terms of coefficient of variation was seen for an isogenic populations. Moreover, experiment also spotlights the randomness of the process. Though all the cells were equally exposed to the drug, only selective cells chose to respond for the reasons that are not clear, currently. Propensity of the population to exhibit

differential response to the identical environment can be due to innate stochasticity in different biochemical pathways including DNA replication that are influential in scheduling the division.

Results can further be bolstered by the introduction artificial replication impediments on DNA through FROS. It will involve cloning of tetO sequences (array of 100- 500 sequences) at desired places on DNA. Tight association of tetR with these operator sequences will result into cease in the fork progression. This event, then can easily be studied with reference to the elongation of the cell into filamentous form.

Effect of stochastic replication on cell size variation was further accentuated with molecular study that evidently associated replication stalls with phenotypic fluctuations. Events of colocalization of RecA with chromosomal DNA were counted as the incidents of replication fork arrest. We observed increased occurrence of cells with RecA- DNA colocalization with the growth rate. Increase in such proportion in *E. coli* population also increased cell length variability linearly. The befitting experiment to understand the involvement of SOS response pathway in phenotypic noise would have been molecular probing of SulA and its correlation with cell length variation, as the expression of SulA is completely dependent on the activation of SOS response[4,87]. Nonetheless, our molecular quantification of cellular levels of RecA turned out to be useful. A distinct boost in RecA expression upon induced replication breach was noticed. It in turn, certified that the inhibition of cell division due to replication arrest happens through SOS response pathway.

Study, thus, facilitates the conclusion that the stochastic fails in DNA replication process can introduce non- genetic phenotypic fluctuations in the clonal populations of *Escherichia coli*.

Ramification of the cell size variations in MG1655 populations are not clear from currently available data. However, we assume that cell length variability in isogenic *E. coli* populations is a manifestation of variation in the response to the stalled replication fork across the cells. It

was best illustrated when we treated continuously growing *E. coli* cells with HU. A distinct difference in the responses exhibited by two sisters was observed. Such behaviour brings about scores of queries pertaining the disparity that lies underneath of seemingly identical cell bodies. In addition to that, it also encourages to gauge the tuning between gene expression and phenotypic heterogeneity after arrested replication fork disrupts the routine.

Chapter 4

Stochasticity in DNA replication process modulates the intrinsic noise in the genetic circuit of *Escherichia coli* MG22

4.1 Motivation

Fluctuations in the gene expressions have been postulated as the mechanism that actuates the phenotypic variability in an isogenic population[41,5,42–44,88]. Series of biochemical pathways define gene expression. Though these biochemical reactions are hard- wired and are expected to be deterministic in nature, they are prone to subtle and random changes owing to *in- vivo* crossovers between two different pathways and/ or involvement of more than one enzymes at a time. Stochastic fluctuations in the circuit, vital for protein synthesis, can end up in the variable protein levels over the time or across the cells in the populations. These unavoidable fluctuations, commonly referred as '*noise in the gene expression*' result into the quantitative difference in protein function and hence causes difference in the physical appearance as well as the behaviour of a single cell in the population, in spite of their identical genomic DNA and niche. With recent developments in science, noise in gene expression has increasingly been appreciated for its crucial role in the complex multi- cellular organizations in the community. Gene expression noise has been suspected to have evolutionary significance, as it generates division of labour in a clonal population. Numerous experiments provide evidence in the support of this proposition. In *Salmonella typhimurium*, gene encoding for flagellar protein was replaced with GFP expressing gene and then allowed to express under native promoter. However, fluorescence intensity differed from cell to cell in a micro- colony generated from a single mother, which in turn, was parsed as the presence of fluctuations at the promoter of the gene (Fig. 4.1). Variations in the promoter activity, is thought to introduce variations in the expression of flagellum on the cell surface. It culminates into the formation of two sub-

populations in *Salmonella* community. Section of the population that expresses flagella is employed in the host tissue invasion, while the population with suppressed flagellar expression has been suspected to be involved in the propagation of the pathogen in the host[41]. Noise present in the functioning of *E. coli* chemosensory system, through stochastic expression of the chemotactic proteins can be another example. Though the cell has evolved feedback loops for the filtration of the noise in the motility signals, theoretical model suggests that complete removal of the noise reduces the efficiency of the cellular motility. Stochasticity in the process has been believed to create phenotypic diversity, thereby increasing the chances of survival in the natural environment of an organism[88].

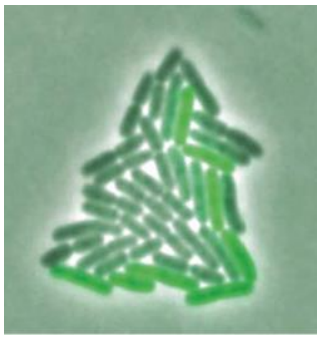


Fig. 4.1: Fluctuations in the expression of flagellar gene (measured in terms of GFP expression) in *Salmonella* micro- colony (Reprinted from Freed et al., 2008 with permission[41]. Copyright obtained from PLOS (Appendix E, C9)).

In 2002, stochasticity in the gene expression in an *E. coli* cell was analysed and compared across the population as well as through time. For these studies, YFP and CFP encoding genes on either arms of circular DNA of *E. coli* under the control of lactose and/ or IPTG. Both the genes were equidistant from OriC (Fig. 4.2)[5].

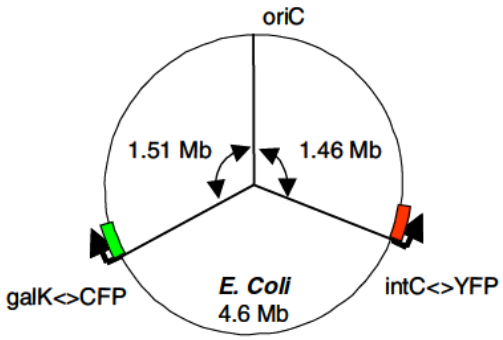


Fig. 4.2: Cartoon depicting the genomic DNA of *E. coli* MG22 (Reprinted from Elowitz et al., 2002 with permission[5]. Copyright obtained from the American Association for the Advancement of Science (Appendix E, C10)).

Considering the bidirectional replication in *E. coli*, equal number of copies for both the genes were expected in the cell. Intensity output for total CFP and YFP molecules in the cell was counted as an indirect measure for the expression of respective genes. It was observed that in the absence of an inducer, expression of CFP and YFP genes varied from cell to cell in the cultures of MG22 as well as RP22. The noise was reduced significantly when IPTG was added to the medium (Fig. 4.3). Interestingly, in *recA* deletion mutants the loss of correlation in the expression of two genes was further amplified even in the presence of an inducer (Fig. 4.3, subsection C, highlighted with red box). Authors predicted the absence of RecA, which is an essential protein in the rescue of the stalled replication fork, varies the copy number of genes in different parts of DNA, owing to the discontinued DNA replication and that might mirrored as the variation in the intensities[5].

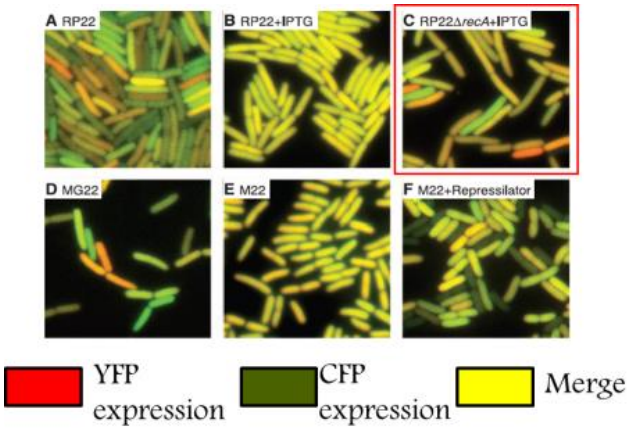


Fig. 4.3: Gene expression fluctuates in an isogenic populations of *E. coli* RP22. Red box highlights the effect of deletion of *recA* gene from the genome on the gene expression in the presence of an inducer (Reprinted from Elowitz et al.,2002 with permission[5]. Copyright obtained from The American Association for the Advancement of Science (Appendix E, C10)).

Two colour reporter assay decomposed gene expression noise into two distinct components. An *intrinsic noise* in genetic circuit is defined as an extent to which two copies of the gene, under identical control and environment fail to correlate their expression. For instance, in a cell the expression levels of CFP and YFP may vary with time. It introduces noise in the genetic circuit of the population (Fig. 4.4 A). On contrary, an *extrinsic gene expression noise* refers to the fluctuations in the expression of a gene from one cell to other. In other words, the expression of YFP and CFP can be correlated temporally within a cell. However, the correlation pattern may differ from one cell to another in population (Fig. 4.4 B)[5].

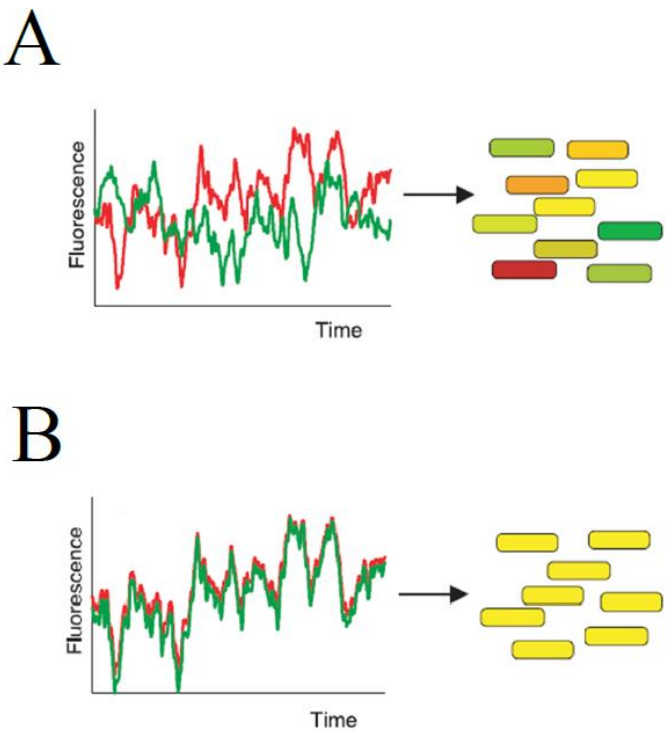


Fig. 4.4: Illustration of (A) an intrinsic noise and (B) an extrinsic noise in gene expression (Reprinted from Elowitz et al., 2002 with permission[5]. Copyright obtained from The American Association for the Advancement of Science (Appendix E, C10)).

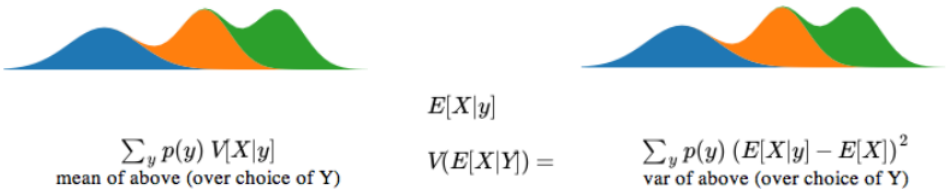
Experiment implements a statistical idea known as 'the law of total variance'[89], which decomposes the variance in a random variable 'X', as average component variance (unexplained variance) and variance of component averages (explained variance), as illustrated in Fig. 4.5. In probability theory, law of total variance[89] states that if X and Y are random variables on the same probability space and the variance of X is finite, then

$$Var(X) = E[Var(X|Y)] + Var(E[X|Y]) \qquad \text{Equation 7}$$

Term '$Var(E[X|Y])$' on right hand side of equation 7 is described as 'explained variance', for it measures the proportion to which a mathematical model accounts for variation in given data set. Whereas, '$E[Var(X|Y)]$', denotes the fraction of the variance in a dependent variable 'X'

that cannot correctly be predicted by explanatory variable 'Y' and hence is termed as 'unexplained variance'[89].

$$\text{Var}[X] = E\,[Var(X|Y)] + Var\,(E[X|Y])$$



$$V[X|y]$$

$$E(V[X|Y]) = \sum_y p(y)\, V[X|y]$$

mean of above (over choice of Y)

$$E[X|y]$$

$$V(E[X|Y]) = \sum_y p(y)\,(E[X|y] - E[X])^2$$

var of above (over choice of Y)

Fig. 4.5: Illustration of law of total variance. It depicts the marginal density plot and each colour represents constituent Gaussian density. The expectation and the variance of 'X' has been broken down by conditioning 'Y'.

In biological context, 'X' represents gene expression or output and 'Y' stands for extracellular inputs/ signals. Variation in 'X' can be generated due to the fluctuations in 'X' (gene expression) or variations in 'Y' (extracellular growth conditions). In a single cell, we can study the distribution of 'X' by keeping 'Y' fixed. Thus, an expectation of $Var(X|Y)$ gives us the stochasticity present in the process of sensing the environmental signals and is termed as 'intrinsic noise'. In population, extracellular signal 'Y' depends upon the micro- environment as well as the growth stage of the cell and hence varies from cell to cell. Since, $E[X|Y]$ is an average output for given input, variance of $E[X|Y]$ impersonates the extent to which each cell in the population responds to the signal. Thus, the last term in equation 7, denotes 'extrinsic noise'[90–92].

It, thus, makes an attractive model for vetting cell length variations in microbial populations, sharing common genomic DNA. We showed that stochastic DNA replication can cause non-genetic variations in *E. coli* phenotype (Chapter 3)[93]. Our work combined with present facts motivated us to study the nexus between gene expression noise and stochasticity in DNA replication. We used cell length variation in *E. coli* MG22 populations, as an indirect measure for the extent of replication arrest per cell, generated by treating the cells with increasing dosage of hydroxyurea. Cell length variation was then compared with gene expression noise and its sub- components to find the link.

4.2 Correlation between the gene expression of CFP and YFP genes in *E. coli* MG22 cells under optimal growth conditions

As a preliminary test, degree of correlation between CFP and YFP expression in the population propagated in LB at optimal laboratory conditions was examined. Cells were grown at 37º C, with constant shaking at 180 rpm. Expression of both fluorescence genes were induced by adding 2 mM IPTG to the medium. We analysed fixed images of MG22 cells in DIC and CFP, YFP channels to extract cell lengths and intensity values respectively (Fig.4.6). Experiment was, particularly, of importance, as it assisted us to understand (a) the trend of cell size distribution and cellular intensities of CFP and YFP in MG22 populations at control growth conditions and (b) whether the trend exhibited by MG22 was comparable with that of MG1655 (wild- type).

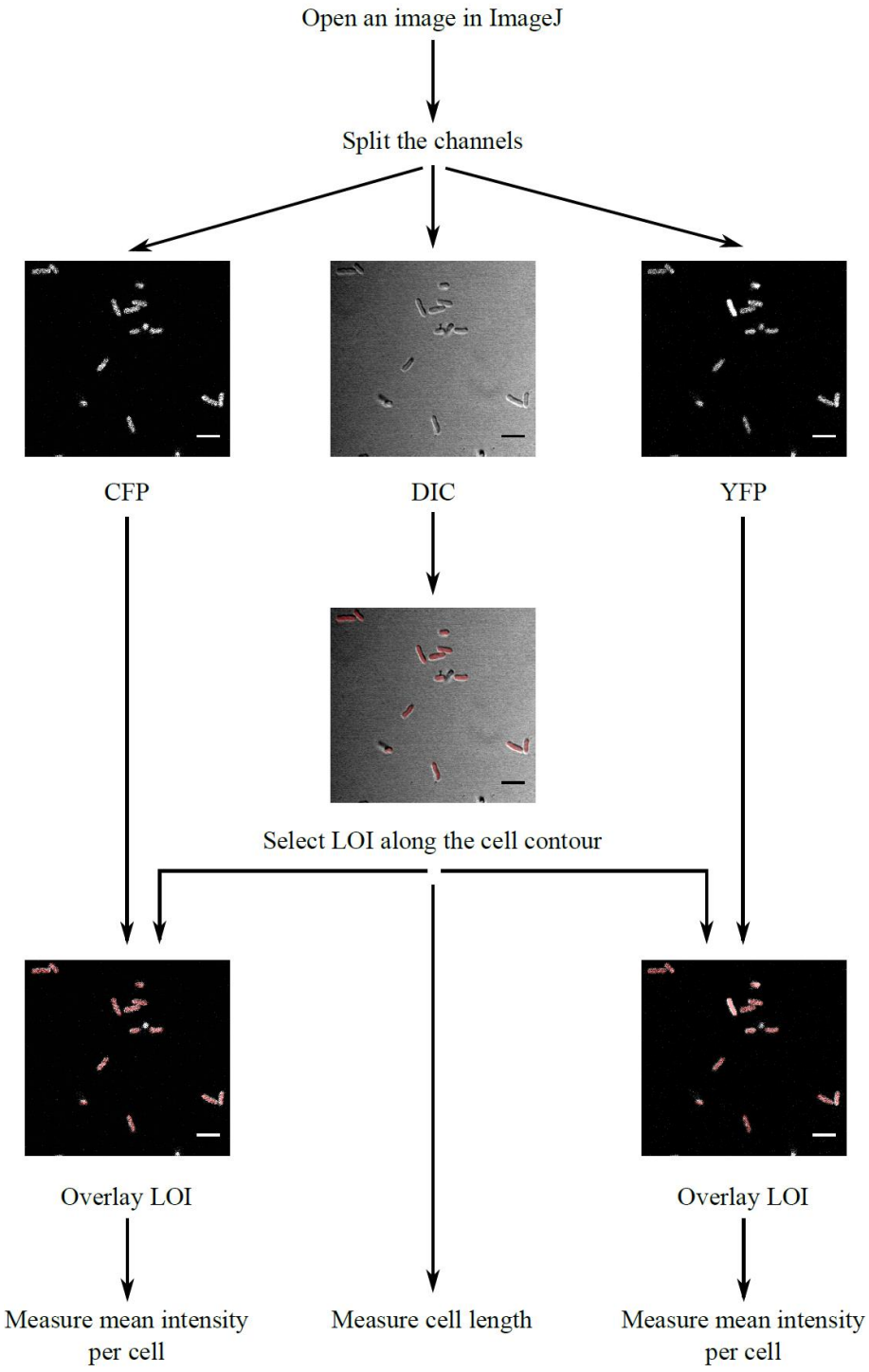


Fig. 4.6: Analysis of fixed images of *E. coli* MG22 using ImageJ. DIC image is used as a template to draw the cell contour (red line = 5 pixel) which is then superimposed on the respective images in YFP and CFP channel to extract raw mean intensity along the cell length. Scale bar- 10 µm.

Cell length distribution had an average of 2.188 µm and extended to 6 µm. Coefficient of variation for population cell lengths was calculated to be 0.302 and was comparable to our previous results (Fig. 4.7 A). YFP and CFP intensities in an individual cell were normalized with respect to the respective maximum intensities observed in the population. Though the population distributions of CFP and YFP intensities do not have same shape, their averages were similar (Fig. 4.7 B and C). Per cell average value for CFP (R_i) and YFP (G_i) is used to determine Mander's overlap coefficient, r[94] (equation 8). Though modal value of the distribution was 1, shape of the distribution was negatively skewed and reached to 0.8 (Fig. 4.7 D).

$$r = \frac{\sum_i R_i * G_i}{\sqrt{[\sum_i (R_i)^2 * \sum_i (G_i)^2]}}$$ Equation 8

Analysis provided a starting point for further assessment, as it was suggestive of variation in the expression of CFP and YFP genes (a) in a single cell and (b) across the population. With this primary knowledge, we proceeded to perturb the growth environment and its subsequent effects on the gene expression noise in MG22 populations.

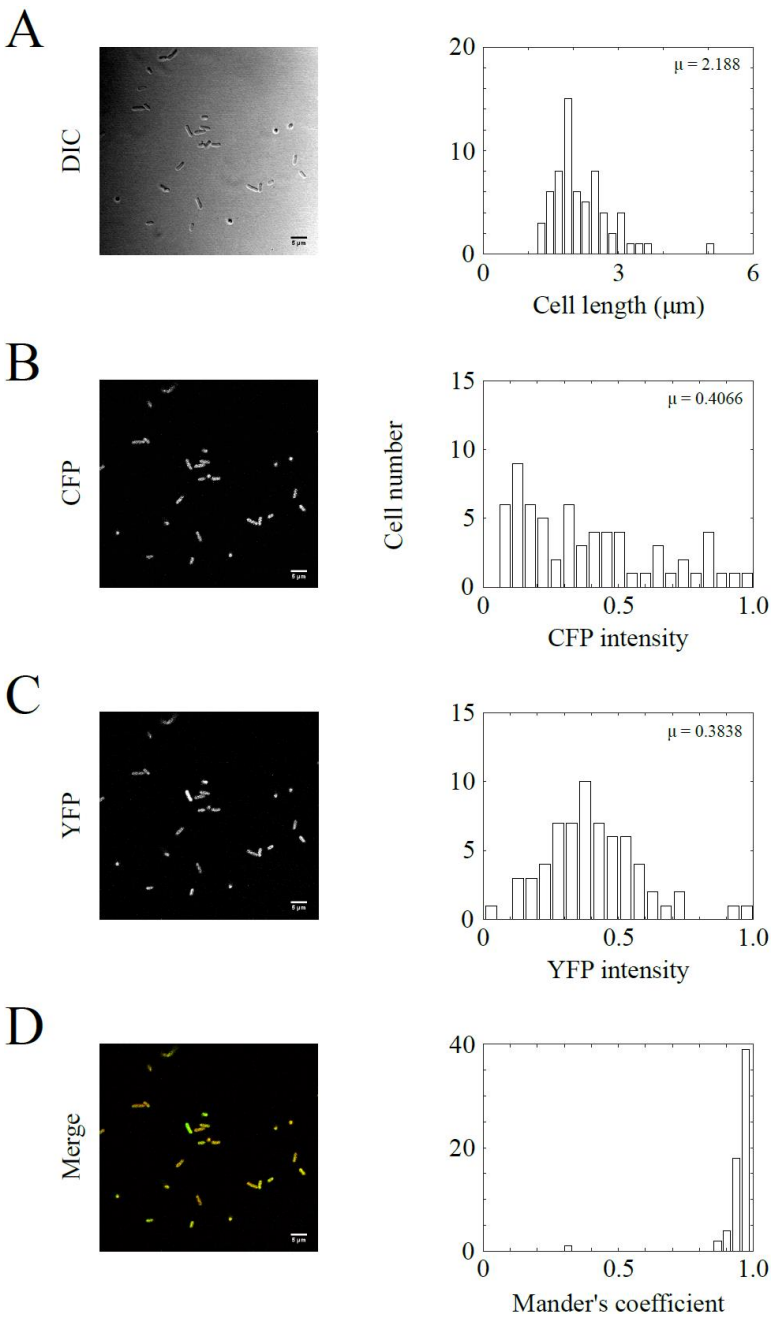


Fig. 4.7: Correlation of CFP and YFP expression in MG22 under ordinary growth conditions (A) Depicts the still image of MG22 in DIC (left panel) and the distribution of their cell lengths (right panel). (B) and (C) Corresponding fluorescence images in CFP and YFP are shown in left panels and their population intensity distributions have been represented in the right panels (D) Left panel indicates the merge of CFP (red) and YFP (green) channels. Mander's coefficient (Equation 8) between the two colours was evaluated and its distribution has been shown in right panel. Scale bar (snapshot)- 5 μm.

4.3 Introduction of noise in genetic network using drug treatment

Cells were grown either in LB or in M9 supplemented with succinate to attune the doubling time of the populations. Cells in LB, grew with the doubling time of 20 mins, while in minimal media, cells complete one round of life cycle in 70 mins. Following the drug treatment for the time equivalent to three generations in each medium, cells were fractionated into two aliquots. One of them was fixed to observe the changes in the cell after treatment, while other was allowed to recover for another three generations. Thrust of the experiment was to assess the quantitative relationship between stochastic replication fork stalling and gene expression noise in clonal population across the threshold growth rate of 1 generation per hour. Variation in cell lengths as well as gene expression in the population was calculated using following formulae:

$$Cell\ length\ variability, CV^2 = \frac{\sigma_L^2}{\mu_L^2} \qquad \text{Equation 9}$$

Where, σ_L is standard deviation in the population cell lengths and μ_L is average cell length. The gene expression noise is calculated using formulae dictated in Elowitz's work[5]:

$$Intrinsic\ noise, \eta_{int}^2 \equiv \frac{\langle (C_i - Y_i)^2 \rangle}{2\langle C_i \rangle \langle Y_i \rangle} \qquad \text{Equation 10}$$

$$Extrinsic\ noise, \eta_{ext}^2 \equiv \frac{\langle C_i Y_i \rangle - \langle C_i \rangle \langle Y_i \rangle}{\langle C_i \rangle \langle Y_i \rangle} \qquad \text{Equation 11}$$

$$Total\ noise, \eta_{tot}^2 \equiv \frac{\langle C_i^2 + Y_i^2 \rangle - 2\langle C_i \rangle \langle Y_i \rangle}{2\langle C_i \rangle \langle Y_i \rangle} \qquad \text{Equation 12}$$

C_i and Y_i represent the average intensity in an individual cell in population.

Hydroxyurea treatment was used to increase the episodes of replication fork stalling. We used five different concentrations of HU ranging from 1 mM to 26 mM, as cell length variability shows linear trend in that regime (Chapter 3; Fig. 3.12). We also, set several controls to compare the effects of hydroxyurea on cell length variability and gene expression. Control cells underwent drug treatments that impeded either protein expression or cell division, so that we could assess the effect of replication stall on gene expression noise more reliably. Rifampicin

and chloramphenicol are known to impede transcription and translation respectively[95,96]. Since, both the processes are vital in the synthesis of proteins, disturbance in them can perturb the gene expression. Hence, rifampicin and chloramphenicol were used at the sub- lethal concentrations to induce an increase in total gene expression noise. Defects in nucleoid segregation can also lead to the fluctuations in gene expression. It was achieved by dismantling MreB cytoskeleton using A22[97–99], as MreB skeleton has been shown to play a crucial role in the nucleoid segregation[66,100]. Cephalexin inhibits FtsI, a protein acting downstream to FtsZ during cytokinesis and its treatment serve a good control for the changes in the cell size[101,102]. Concentrations of these antibiotics used in experimental set- ups, have been specified in Chapter 2, table 2.10.

4.3.1 Study of 'noise' after the treatment

Preliminary inspection of the fixed images right after treatment gave an impression that the noise in treated populations is comparable to that in the control populations. ImageJ based quantification of the images reflected the observations. In addition, we also observed uniform reduction in cell density in case of cultures grown in M9 + succinate, though the optical densities of the cultures were recorded to be similar to that of LB grown populations. Reason for the difference in the cell densities is not clear to us (Fig. 4.8).

Cell length variability heterogeneity scored for all the experimental set- ups revealed that irrespective of the type of drug treatment, coefficient of variation for cell sizes was higher at high proliferation rates and is minimal when cells are growing at slower pace (Fig. 4.9 A). It was consistent with our earlier observations with MG1655 populations (Chapter 3, Fig. 3.12 A). However, we observed an abnormal hike in cell length variation estimated for control populations at higher growth rate. Because of it, length heterogeneity in control appeared to be comparable with that of cephalexin treated cells and higher than phenotypic noise obtained for populations treated either with A22 or with chloramphenicol or with rifampicin or with range

of hydroxyurea concentrations. Upon going through the data, we noticed a presence of a filamentous cell in control populations. Its length was measured to be 60 μm. Consequently, addition of this data point caused rightward shift in average cell length as well as it increased the standard deviation in population cell length in the data pool, inclusion of which in final quantification across three population inflated cell length variation. Cell length variability, in respective growth media was found to be the highest for cephalexin treated cells. While, cell length variations in the populations treated with A22 or chloramphenicol or rifampicin or different concentrations of HU were found to be comparable. Trend was observed to be same in both at both the growth rates (Fig. 4.9). Surprisingly, total gene expression noise was similar at both the growth rates and did not exhibit any peculiar pattern across the different experimental set- ups (Fig. 4.9 B). Dominance of the component noise in the genetic wiring, however, was found to be growth rate dependent. Intrinsic noise contributed significantly to the total noise, when the cells were propagating with higher rates in LB (Fig. 4.9 C). On the other hand, cells cultured in minimal medium had variation in the gene expression, mainly because of the introduction of an extrinsic noise (Fig. 4.9 D). Nevertheless, no particular trend was observed for intrinsic as well as extrinsic noise across the populations cultivated in different environmental conditions. Moreover, we noticed elevation in error bars (standard deviation) in population intrinsic as well as extrinsic noise in control growing in LB (Fig. 4.9 C and D). Though this anomaly appears to be similar to the one observed in case of cell length variability, its origin is not clear to us.

It was clear from the results that the cells did not exhibit the effect of drug after 3 generations of treatment. Our previous experience with HU treatment showed cell elongation during recovery phase (Chapter 3, Fig. 3.13) and hence, it was necessary to scrutinize recovered populations for phenotypic as well as genotypic noise.

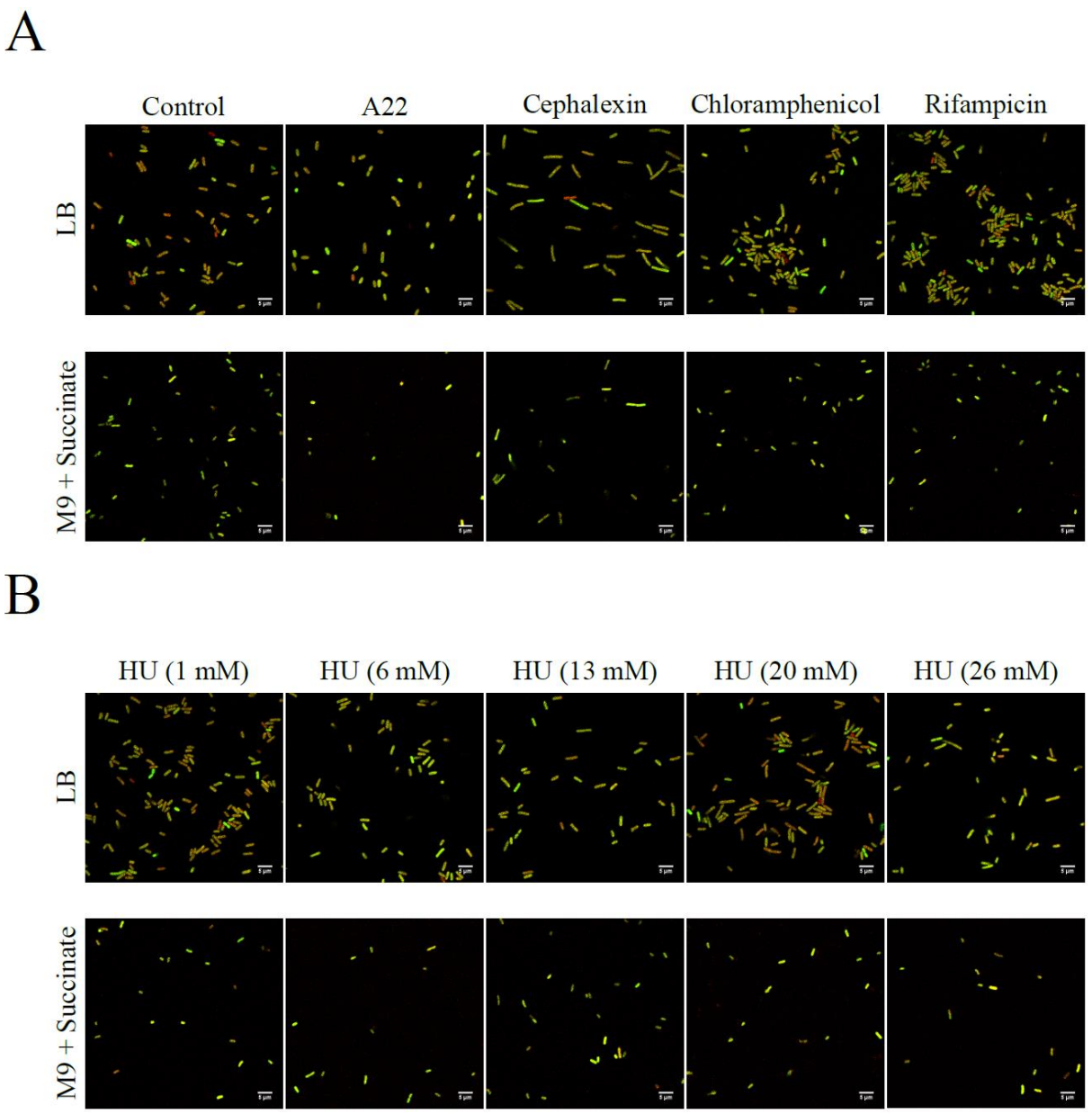


Fig. 4.8: Snapshots of *E. coli* MG22 grown either in LB (upper panel) or in M9 + succinate (lower panel) after drug treatment. Corresponding images in CFP (red colour) and YFP (green colour) were merged to represent the noise in cell lengths as well as gene expression in (A) control cells and (B) cells treated with increasing concentrations of hydroxyurea. Scale bar- 5 µm.

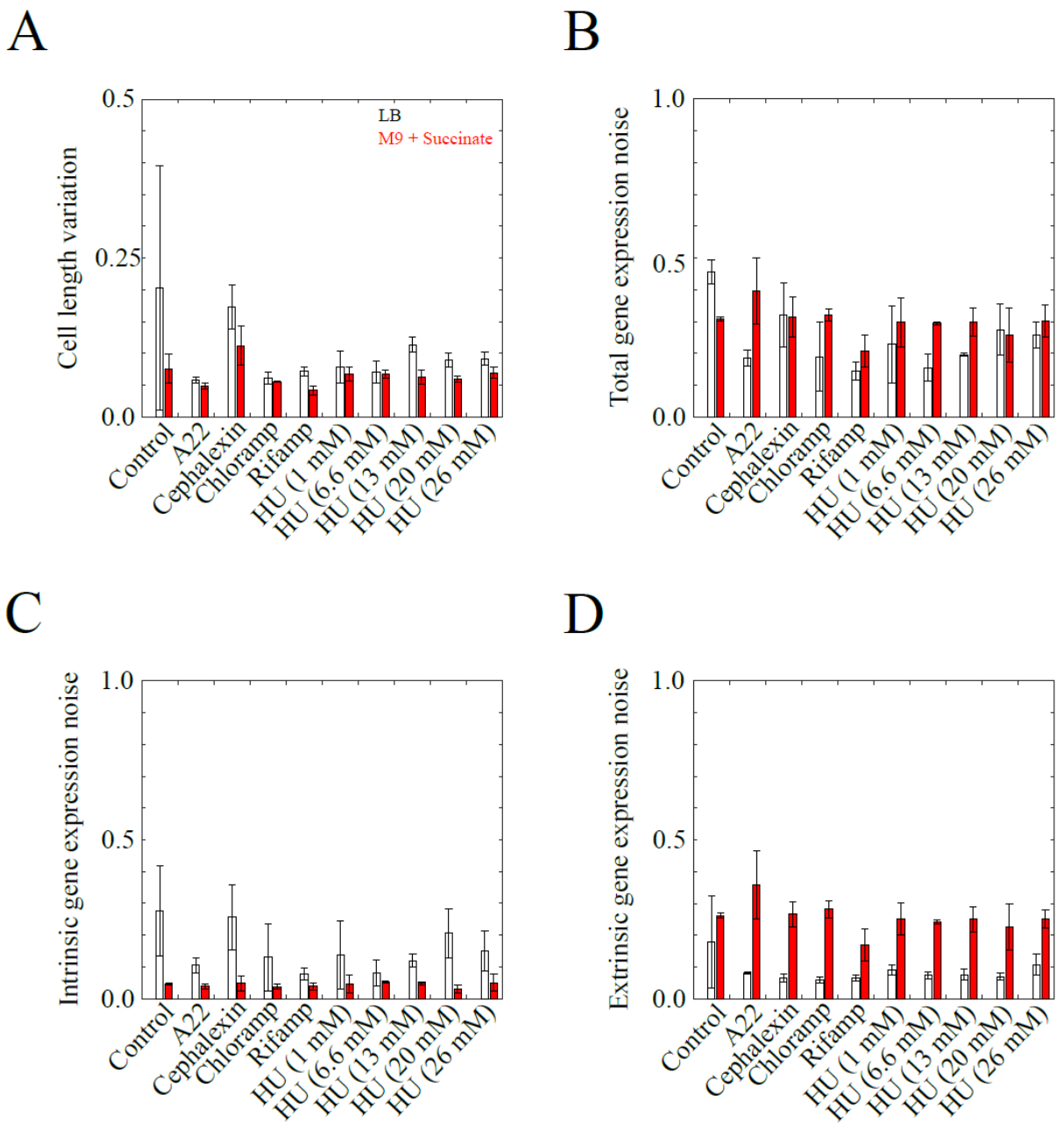


Fig. 4.9: Quantification of 'noise' in *E. coli* MG22 treated with drugs. MG22 populations grown in LB (white bars) or in M9 + succinate (red) were analysed to assess (A) population cell length variability (B) total noise in gene expression (C) intrinsic noise in gene expression and (D) extrinsic gene expression noise. Error bars- standard deviation.

4.3.2 Study of 'noise' after the recovery

Recovered MG22 cells (Fig. 4.10 A and B) were analysed interactively for (a) cell length variability (b) total gene expression noise (c) intrinsic gene expression noise and (d) extrinsic gene expression noise. Unlike the noise evaluated for the cells withdrawn right after the treatment, we observed significant increase in both phenotypic as well as total genotypic noise for cells growing with high growth rate than that for the cells grown in minimal medium (Fig. 4. 10). This section details the assessment of drug- recovered cells, harvested from control experiments.

Cell length variability was similar for control cultures growing at two different growth rates. It can be attributed to the design of an experiment. Cells harvested after recovery are in the late log phase. At higher densities, physiology of the cell differs from the physiology of the cell in the mid- log phase. An individual cell becomes smaller and more uniform in the size, owing to the switch into stationary phase physiology[40]. Hence, cell length variability decreases as the culture approaches the stationary phase. As expected, in cephalexin treated populations size heterogeneity was highest at both the growth rates. Variation in the cell lengths was same for the cultures treated either with A22 or with chloramphenicol or with rifampicin and was comparable to the size variation in untreated cells (Fig. 4.11 A).

Untreated population showed no significant difference in total gene expression noise, when grown in LB and minimal medium. Similar values were quantified for cephalexin treated populations at both the growth rates. However, fast growing cells showed higher disturbance in gene expression, upon A22 or chloramphenicol or rifampicin treatment. Cells with slower proliferation rates had gene expression noise similar to that of control (Fig. 4.11 B).

Interestingly, in recovered cells also, growth rate determines the component that contributes the most to total gene expression noise. At higher growth rates, intrinsic noise forms major

proportion of genotypic noise, and extrinsic noise dominates at lower growth rates. Result, thereby, confirmed our earlier observations.

In LB, intrinsic noise as well as total gene expression noise were the highest for chloramphenicol and rifampicin treated populations, followed by the populations treated with A22. Extrinsic noise, was however, similar across all the populations (Fig. 4.11 C and D). After we evaluated the controls, we examined the effect of replication stall events on gene expression noise.

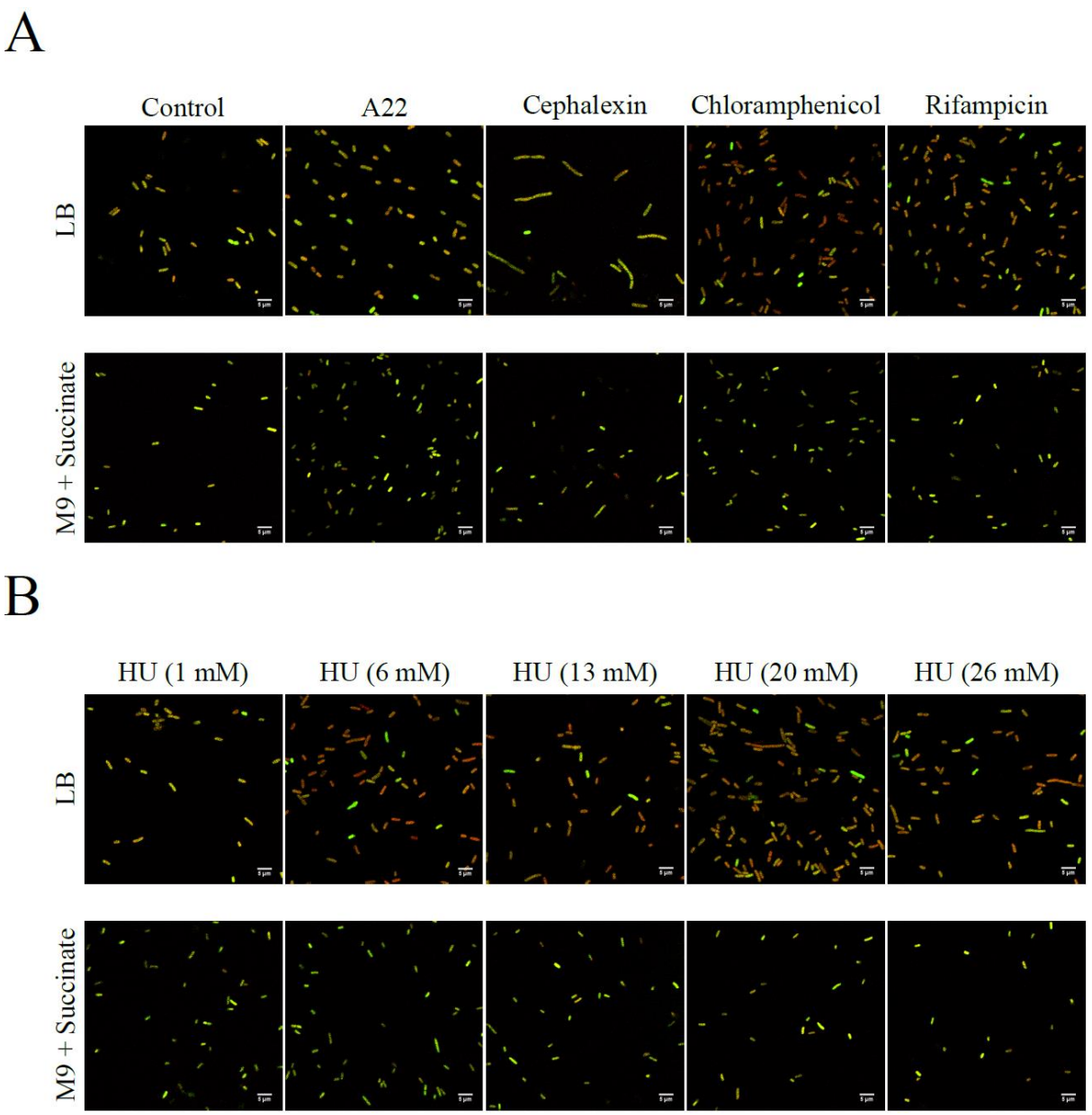


Fig. 4.10: Snapshots of recovered *E. coli* MG22 cultures grown either in LB (upper panel) or in M9 + succinate (lower panel) after drug treatment. Corresponding images in CFP (red colour) and YFP (green colour) were merged to represent the noise in cell lengths as well as gene expression in (A) control cells and (B) cells treated with increasing concentrations of hydroxyurea. Scale bar- 5 µm.

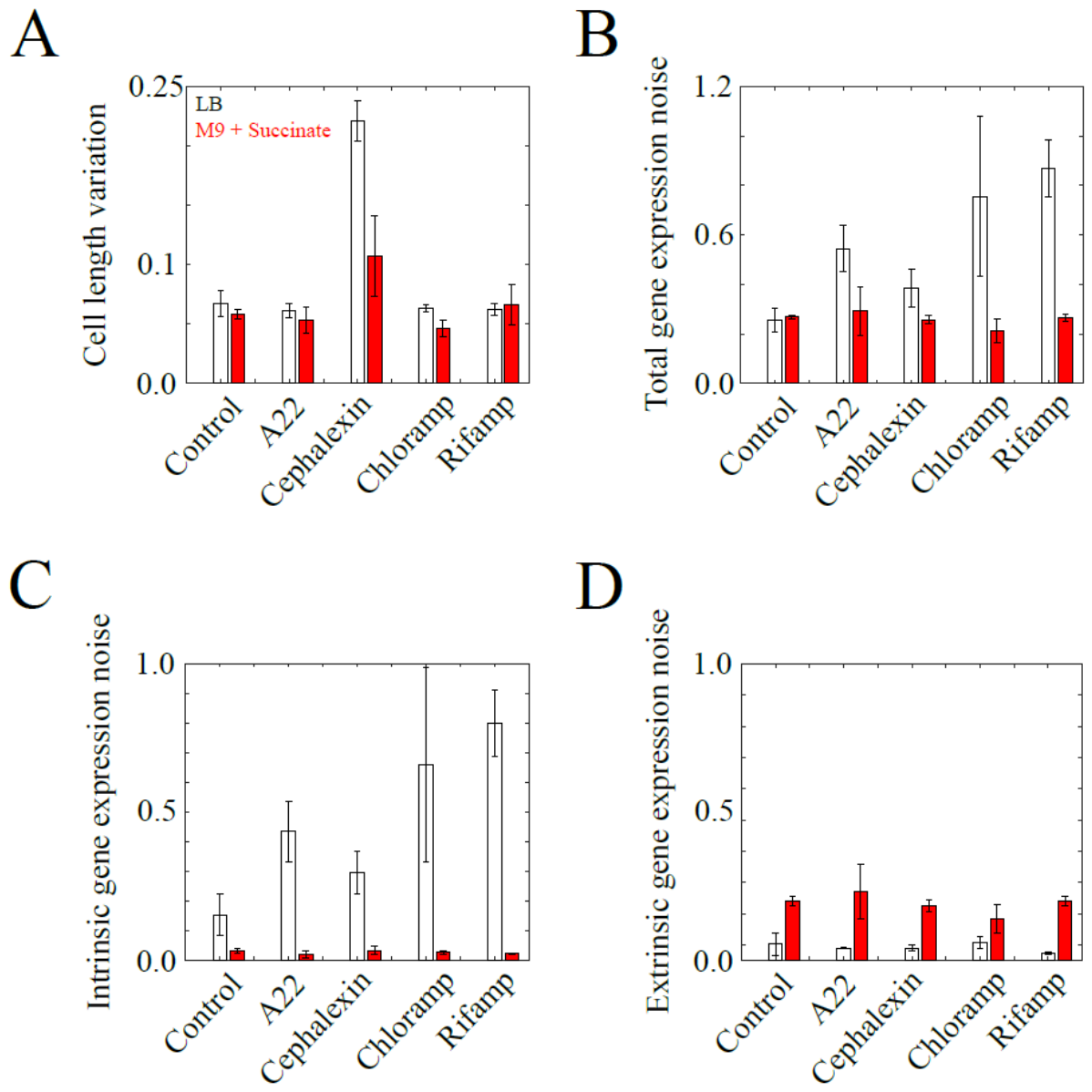


Fig. 4.11: Response of *E. coli* MG22 population to drug treatment. *E. coli* MG22 population was grown either in LB (white bars) or in M9 medium supplemented with succinic acid (red bars) and treated with different drugs. Effect of drug treatments on (A) Phenotypic noise, (B) Total gene expression noise, (C) Intrinsic gene expression noise and (D) Extrinsic gene expression noise has been shown. Error bars- standard deviation.

4.4 Effect of replication fork stalling on the gene expression

MG22 cells were treated with increasing concentration of hydroxyurea ranging from 1 mM to 26 mM for 1 hour, to frequent the replication fork stalling. Based on the results obtained from control cells (Section 4.3), we decided to analyse the cells which were allowed to recover for 3 generations after drug treatment (Section 2.12, table 2.10). Consistent to our previous results, cell length variability (black solid line) increased as a function of increasing hydroxyurea dosage at higher growth rates, while remained unaffected for slow growing cells. We observed that in LB grown cells, intrinsic noise (blue solid line) increased till hydroxyurea concentration hits 13 mM and then showed monotonous dip. At lower growth rates, intrinsic noise was found to be at the basal level. Extrinsic noise (red solid line), however, did not show the correlation with increasing instances of replication fork arrests, in LB as well as in minimal medium. Total genotypic fluctuation (green solid line) followed exactly the same pattern as that of its constituent noise that contributed the more (Fig. 4.12 A and B). Based on our results we propose one more possible mechanism in which random replication stalls influence the intrinsic gene expression noise and might result in non- genetic phenotypic fluctuations in an isogenic population of *E. coli*.

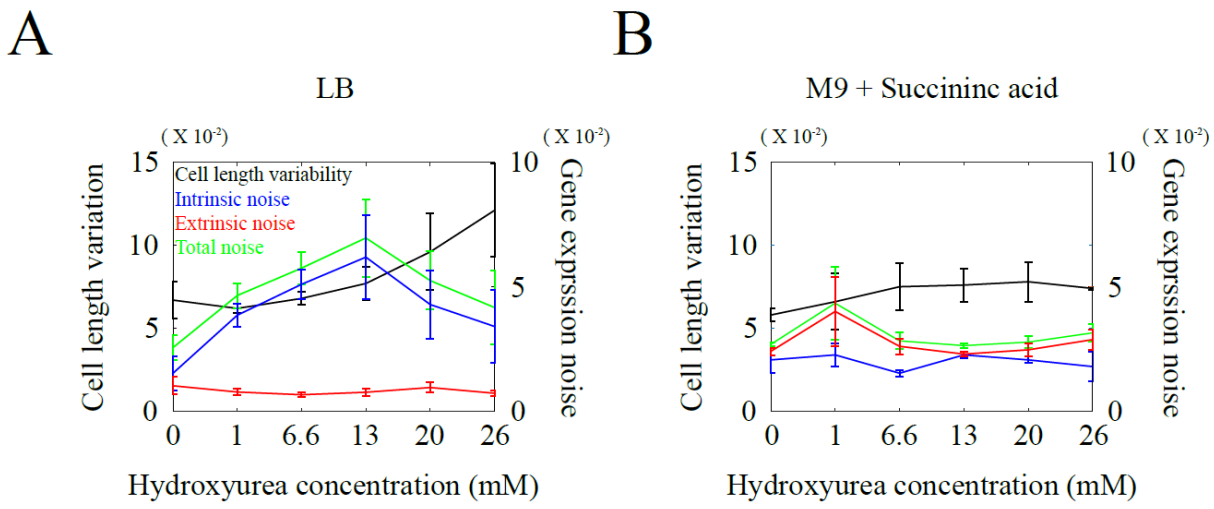


Figure 4.12: Effect of increased frequency of replication stall on genotypic and phenotypic noise. Variations in cell length (left axis) and gene expression (right axis) are plotted as a function of increasing concentration of hydroxyurea in MG22 populations grown either in (A) LB or in (B) M9 medium supplemented with succinic acid. Error bars- standard deviation.

4.5 Discussion

Genetic wiring and its fine tuning is of paramount importance for an existence of the life. A tight controller mechanism, hence, employed to ensure an efficiency of the gene expression. In recent years, however, it has been established that despite of the recruitment of barrage of molecules for the maintenance of the fidelity, gene expression can never be a deterministic process. Substantial noise is introduced in the genetic circuitry through ever fluctuating levels and activities of the biomolecules involved in the gene expression. Existence of noise in genetic circuit then subsequently can result in different physical appearances of genetically identical cells. These differences can also shepherd the cellular decision as well as their fate that eventually influence the structure of the community.

Initial efforts were mainly focused on the mathematical representation of noise as well as theoretical prediction of its sources. Nevertheless, in recent years attempts are made to assign

functional significance to the noise in gene expression. Here, we undertake the study in order to detail the role of the random replication stalls in genetic fluctuations and its resultant outputs.

We analysed MG22 populations[5], which under optimal growth conditions express CFP and YFP encoding gene. Though the strain has modification in its genomic DNA, its population cell length distribution is comparable with MG1655. Thus, we confirmed that the genetic modifications do not contribute to the cell size variation. Shape of intensity distribution for CFP and YFP molecules were different from each other, however, their averages were comparable. It implied the presence of noise in the genetic circuit under common environmental growth conditions. It was further confirmed by the left- skewed shape of distributions drawn for Mander's coefficient estimated for levels of CFP and YFP molecules in individual cell in the population.

Primary results, then drove us to our principle objective, which was to assess the genetic noise with changes in the growth rate as well as extent of replication stalls. Though growth rate appeared to have significant contribution in drug mediated effects on *E. coli* cultures, we also observed that the population response to the various drug treatments has a lag period. 'Noise' in the test populations was within the standard deviation of that of the control populations after treatment. Significant difference in the noise values appeared only when cells were allowed to recover for three generations. Results were in consent with our earlier experiments with continuous cultures of MG1655 which were subjected to hydroxyurea treatment. Filamentation was observed during recovery and not during the introduction of drug into the media[93]. We suspect that the time duration designed for the treatment and the subsequent recovery of the cells accounts for such behaviour. Since, the sub- lethal concentrations of antibiotics have been used, we predict that an individual cell take some time to accumulate the required concentration of the drug and exhibit the effect. From our experiment we realise that this time amounts to that of required to complete 6 generations and therefore the actual effect of the chemical is seen

after recovery. When treated cells are resuspended into fresh media, lag in the recovery denotes the time needed for the dilution of the drug in a cell because of cell divisions.

Noise in gene expression as well as cell sizes was comparable for untreated populations. We assigned this to the experimental design. Untreated cells are in early stationary phase. Increased population density reduces the cellular growth rate. Population growth rates are found to be similar for both the growth media and hence the noise levels are also found to be similar.

Our experiment decomposed the gene expression noise in growth rate dependent manner. In cells recovered from the drug treatment, the total gene expression noise was higher at higher growth rates, while it was reduced in a population with slower growth. In addition to this, in a fast growing population intrinsic noise was higher in the gene circuit. At the slower rate extrinsic noise takes over. Extrinsic noise is due to the cell to cell variability of the gene expression in the populations. It stems from the cell to cell differences in the enzyme levels or the cell volume or the cell age[103]. Lower levels of extrinsic noise in our experiment suggests that the total gene expression noise is independent of the physiological state of the cell at higher propagation rates. However, what conditions the growth rate dependent behaviour in gene expression noise is not clear from our results.

When replication halts were amplified in fast growing *E. coli* cells, an increase in the cell length variability after recovery was observed, indicating increased replication defects. Intrinsic noise showed proportionate increase with an increase in the probability of replication stalls. However, when plotted as a function of hydroxyurea dosage, intrinsic noise hits the maxima at HU concentration of 13.2 mM and then decreases linearly, as opposed to the cell length variation that increases monotonously.

Detailed analysis of population intensities and cell lengths pointed that at higher HU concentrations, average difference between CFP and YFP intensities decreases without a

significant change in the average per cell intensities of CFP and YFP molecules. Also, we observed comparable average cell lengths accompanied by gradual rise in the standard deviation with an increase in HU dosage. It generates linear decline in an intrinsic noise at greater HU concentrations, while phenotypic noise increases linearly. Hence, the correlation between an intrinsic noise and cell length variability is abolished for HU concentration above 13.2 mM.

Though the reason behind this strange trend in intensity values is not clear, we suspect that at lethal HU doses (>10 mM)[77], population starts eliminating the selective cells that harbour greater fluctuations in the genetic circuit, which in turn dampens the intrinsic gene expression noise in population by reducing the average difference between CFP and YFP intensities per cell. In addition, elimination of selective elongated cells adds on to the standard deviation without significant changes in the average of the length distribution. It results in the increase in the ratio of 'standard deviation to the average cell lengths', which has been used as a measure to determine the cell length variability in the population. Our interpretations, however, require further validation from single cell studies in continuous cultures as well as theoretical calculations.

In conclusion, we propose that the frequent arrest in the DNA replication process, may create the difference in the genomic copies of CFP and YFP gene. Difference is then propagated into the downstream processes like transcription and translation which is mirrored as a difference in the number of copies for CFP and YFP molecules within a single cell. It may end up with introducing variation in the abundance in the molecules required for the completion of cell cycle. At the same time, DNA repair system shuts down the cytokinesis[93]. Both the events lead to the non- genetic phenotypic diversity in the clonal population.

Work reiterates the role of fluctuations in the origin of phenotypic variation that are independent of genetic differences. Most importantly, it poses availability of essential biomolecules as one more aspect to study. Substantial variations in the levels of cellular molecules propagate noise not only across the population but also through the generations. Study of existence of asymmetries in growth and division in a single cell, thus, become necessary to understand the localization of sub- cellular entities in two compartments of mother cell. Asymmetric distribution of sub- cellular molecules in mother prior to its division gives birth to the twins that might differ in the molecular composition, thereby existing different physical appearance during the course of their respective cell cycle.

Chapter 5

Asymmetric Growth of *Escherichia coli* cell

5.1 Motivation

Asymmetric growth and division in eukaryotic cells are known to help create two cells which are physically as well as functionally different without undergoing genetic manipulations. Progenies that are generated from such division, survive different fates, which in turn, marks the beginning of complex developmental process[104–108]. Thus, asymmetry has been proved to be an integral part in the process of growth as well as division in eukaryotes.

Though micro- organisms were thought to assume simple life- style, recent discoveries have associated complex cellular organizations with microbial colonies[109]. Moreover, numerous bacteria have evolved their distinct asymmetric morphological forms which are then used for their own survival and propagation in the habitat they live.

For instance, *Caulobacter crescentus*, an alpha- proteobacterium gives rise to two different cells viz. stalk cell and swarmer cell. Stalk cell remains attached to the substratum, while swarmer cell explores the niche[110]. Asymmetric segregation, thus, helps not only in finding the places with abundance of nutrient and hence the survival, but also in sequestering the damage, as stalk cell retains the older cell components and newly synthesized biomaterial is directed into the swarmer cell[110]. The strategy is used to propagate healthy progenies.

Asymmetric growth is accountable for survival in seemingly symmetric cells, in some other cases. Asymmetric division for spore formation in *Bacillus* species plays a prominent role in the conservation of species in drastic environments[111]. Sometimes, asymmetry is not apparent during growth. However, the molecular composition of the two cellular halves differs significantly, as one of them is relatively old and confers asymmetry to the cytokinesis, giving

rise to daughter cells that may exhibit differential response to the environmental cues. *Mycobacterium tuberculosis* has been studied for unipolar growth. Though the organism undergoes symmetric cell division, the two daughters are known to differ in the susceptibility to the antibiotics[112]. Another rod shaped microorganism, *Rhizobium*, is known to extend at one of the poles[113]. Unipolar growth in *Rhizobium* and *Mycobacterium* has been attributed to the lack of MreB mediated cell wall construction, because of which poles remain active regions of growth resulting into asymmetric growth of an organism[114].

In many microorganisms, though, asymmetric growth has not been reported, random segregation of subcellular biomolecules can chance produce different daughters sharing identical genomic DNA as well as superficial appearance. Inherent stochasticity in essential processes that determine the extent of gene expression evidently, affect the abundance of protein molecules among two progenies. Alterations in the availability of such molecules, further magnifies noise in the downstream biochemical reactions culminating into the generation of two or more sub- populations with different fates.

In *E. coli*, too, deviation in the placement of the septum from mid- plane has been observed many cells in the population[115]. Localization of nucleoids in the cellular sub- compartments is believed to dictate the site of septation in an *E. coli* cell[116–119]. Here, we analysed the growth of an individual *E. coli* cell in order to understand if there exists growth fluctuations that affect the compartmentalization of biochemical molecules and nucleoids which, in turn, alters the site of division giving rise to two rod shaped cells with different dimensions.

5.2 Growth measurements of DIC time lapse images

We began our studies with processing time- lapse series of DIC images of an *Escherichia coli* cell. Studies help understand the surface expansion pattern over time.

5.2.1 Spatial restriction of *E. coli* cells on agar pad

Major setback of agar pad analysis is the change in the spatial coordinates of the cell, due to cellular movement. To cancel possible translation of the cell on agar surface and correct our growth measurements, cells were confided in space, by imprinting agar pad with 1 µm width (Fig. 5.1 A and B). *E. coli* cells with nucleoids labelled ectopically with histone- like HupA, were grown in the indentations (Fig. 5.1 C). Absence of cellular movement was confirmed by measuring the angle between cell contour and abscissa of an image frame for three different time points separated by an interval of 22 mins. GraphPad Prism 5 (GraphPad software) used to verify the significance returned the p value as 0.1304 when distribution at t = 0 mins was compared against t = 22 mins. Similarly, p values quantified for the comparison between t = 0 mins and t = 44 mins as well as t = 22 mins and t = 44 mins were 0.5029 and 0.1982 respectively. Averages of corresponding distribution of angles, thus, were comparable as tested by an 'unpaired t test', validating minimal or no movement of the cell on the agar surface, during its growth (Fig. 5.1 D).

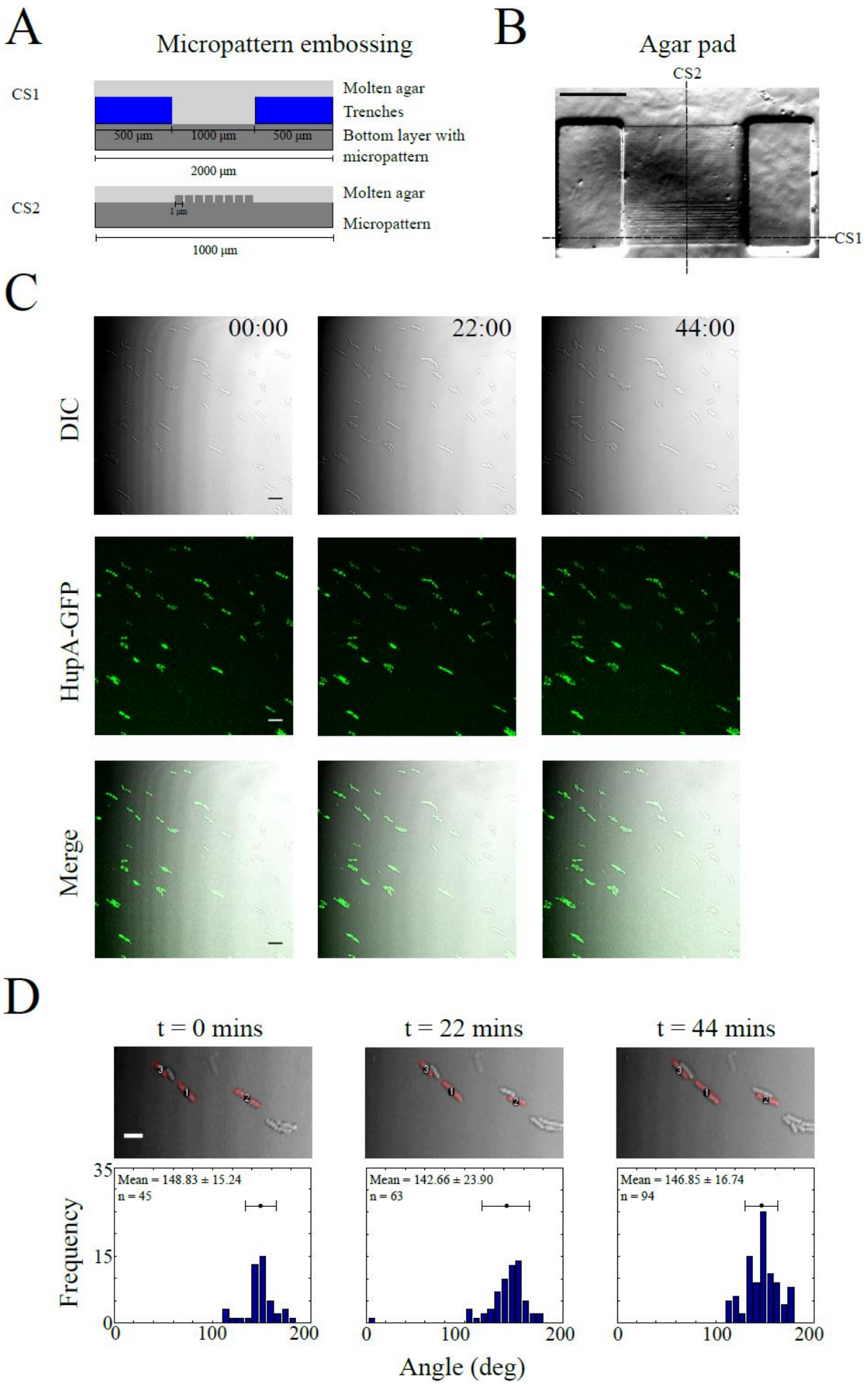

A
Micropattern embossing
CS1
Molten agar
Trenches
500 µm
1000 µm
500 µm
Bottom layer with micropattern
2000 µm
CS2
Molten agar
1 µm
Micropattern
1000 µm
B
Agar pad
CS2
CS1
C
00:00
22:00
44:00
DIC
HupA-GFP
Merge
D
t = 0 mins
t = 22 mins
t = 44 mins
Mean = 148.83 ± 15.24
n = 45
Mean = 142.66 ± 23.90
n = 63
Mean = 146.85 ± 16.74
n = 94
Frequency
35
15
0
0
100
200
Angle (deg)

Fig. 5.1: Agar pad imprinting. (A) Front (CS1) and lateral (CS2) cross- section of epoxy resin used for embossing agar pad (B) Micro- patterned agar pad has been shown. Scale bar- 500 µm (C) Growth of *E. coli* MG1655 transformed with pBAD24- hupA- GFP at three different time points at the interval of 22 mins in DIC (first row) and GFP (second row) channels have been shown. Scale bar- 5 µm (D) Distribution of cell angles at corresponding time points have been shown. Error bars- standard deviation.

5.2.2 Use of in-house detection code to estimate the growth at each cellular end

Growth was followed by in- house MATLAB algorithm (developed by Prangya Mishra, unpublished), which tracks down XY coordinates of two ends of the cell contour through time, as a proxy for growth in each half of the cell (Fig. 5.2 A). Validity of the software was tested by using it to estimate the movement of beads on agar surface and imaged under identical experimental conditions and found to be negligible which is around 0.6 µm for 22 mins (Fig. 5.3 A and B). Growth of an each pole at the end of a time interval was determined with reference to the initial time point (t = 0 mins), in terms of Euclidean distance and named as cumulative growth. Cumulative growth of the two cellular ends was determined till the cytokinesis take place and then fit to exponential function to extract growth rates at two ends. (Fig. 5.2 B). We found significant difference, as supported by 'unpaired t test', between growth rates of two ends of the bacterium, with an average of 0.0179 µm/ min (Fig. 5.2 C and D).

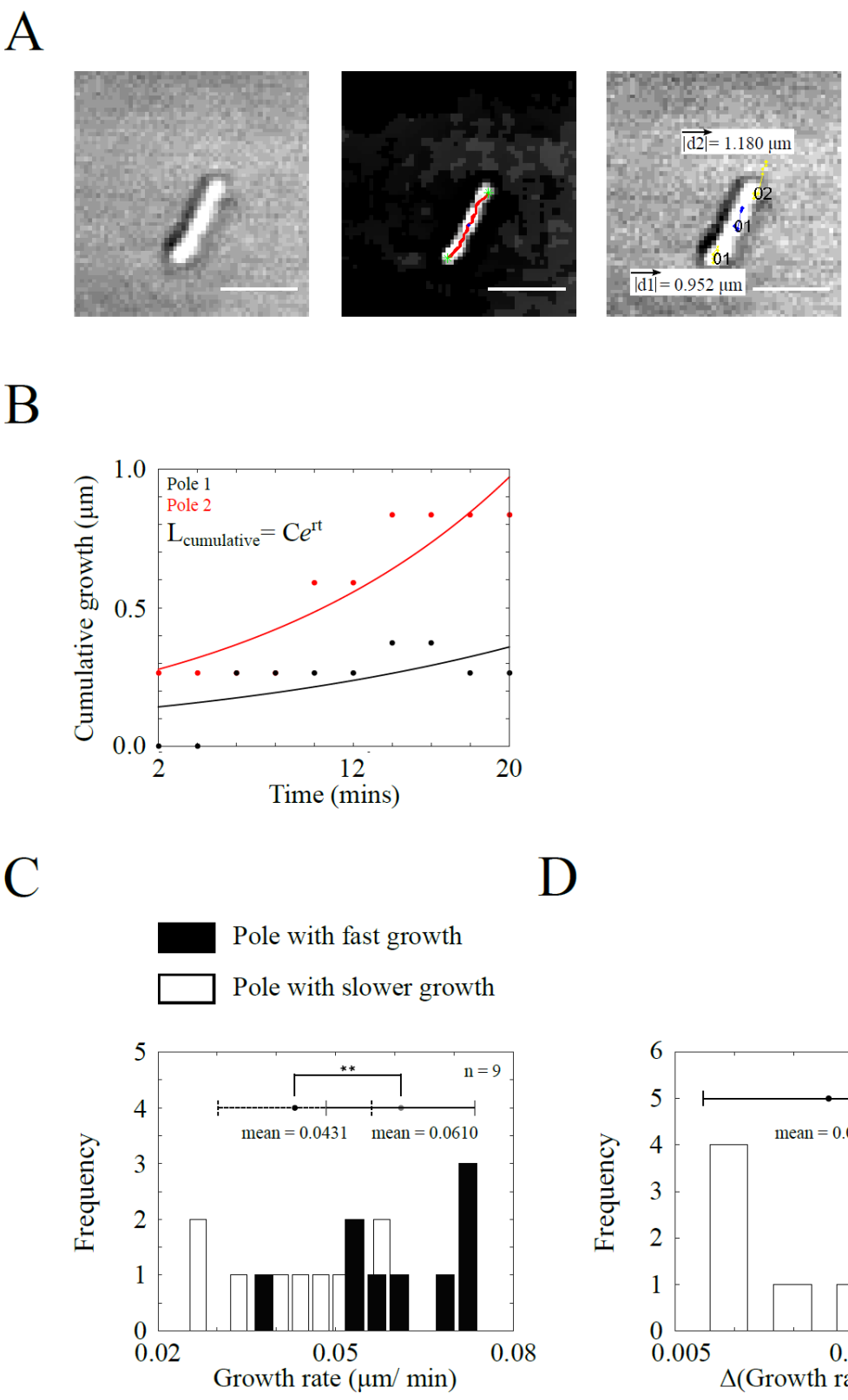


Fig. 5.2: Analysis of *E. coli* DIC image using MATLAB program (A) DIC time-series of a single *E. coli* cell (*left*) was tracked over time. Results were overlaid on first time- frame with

(*middle*) cell contour (red line) and poles (green asterisk) marked for first time- point. Tracks of pole movement (yellow) and the cell centroid movement (blue) over time (*right*). Scale bar- 5 µm (B) Growth of each pole at every interval from initial time point was measured (filled circles) and fit to an exponential function (solid lines), which in essence, estimated C value as 0.2419 and 0.1288 as well as growth rate, r as 0.0695 and 0.05117 µm/ min for fast growing end (red filled circles) and slow growing end (black red circles) (C) Distributions of pole growth rate are compared and the significant difference between their respective averages have been validated using unpaired t test ($n = 9$; $df = 16$; $p = 0.0081$). Error bars- standard deviation. (D) Distribution of the difference between the growth rates of the two cellular ends with average value of 0.0179 µm/ min has been depicted. Error bars- standard deviation.

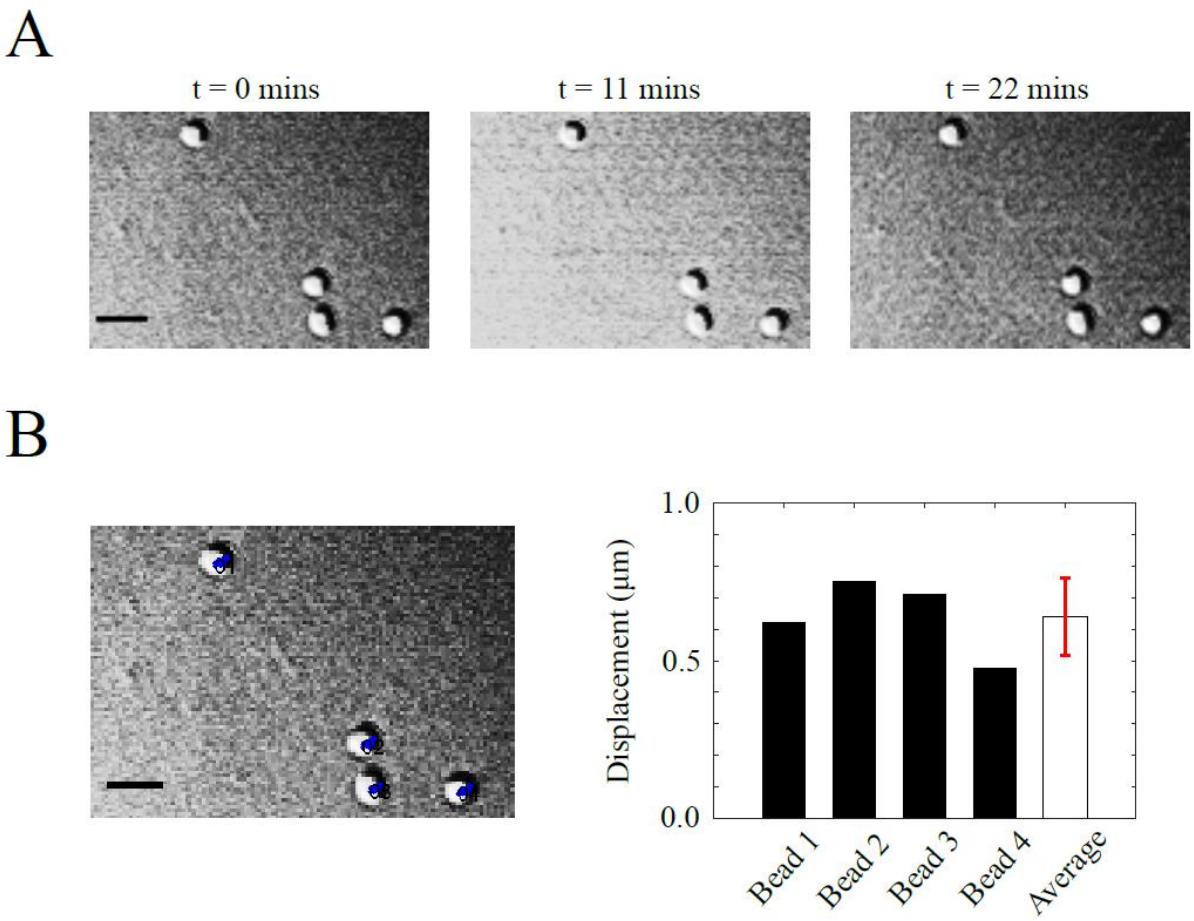


Fig. 5.3: Quantification of bead displacements using MATLAB programme (A) Time lapse images of 1 µm beads at the end of every 11 mins, Scale bar- 2 µm. (B) Trajectories of beads (blue) have been superimposed on the first time lapse image of the beads (left). Their respective displacements (black bars) and the average displacement (white bar) have been shown. Error bars- standard deviation.

5.3 Evidence from membrane analysis to support the presence of bias in the of growth an *E. coli* cell

Escherichia coli cell does not have natural visible surface marker, which virtually makes it symmetric in nature. We created a marker at mid- plane of the cell, by repeatedly bleaching cell membrane labelled with FM4- 64, before fluorescence intensity in dark region equilibrates

with that of rest of the membrane (Fig. 5.4 A and B). Temporal representation of such membrane clearly showed asymmetry in the growth of an *E. coli* cell. More interestingly, an apparent dislocation of bleached mark towards opposite direction of the growing end could be seen, indicating generation of newer geometric centre of the cell with time due to growth asymmetry (Fig. 5.4 C). Estimated slopes for two edges of the kymograph between two consecutive bleach pulses were distinct from each other. Quantification confirmed our observations (Fig. 5.4 D). Difference in two distributions of growth rates of the ends was statistically significant. And its average was found to be 0.0628 μm/ min. Result not only helps provide an evidence of presence of growth bias in *E. coli* cell, but also eliminates the possibility of apparent growth asymmetry in the cell due to the interactions with neighbouring cells, as we observed similar trend in cells physically associated with their neighbouring cells (Fig. 5.5).

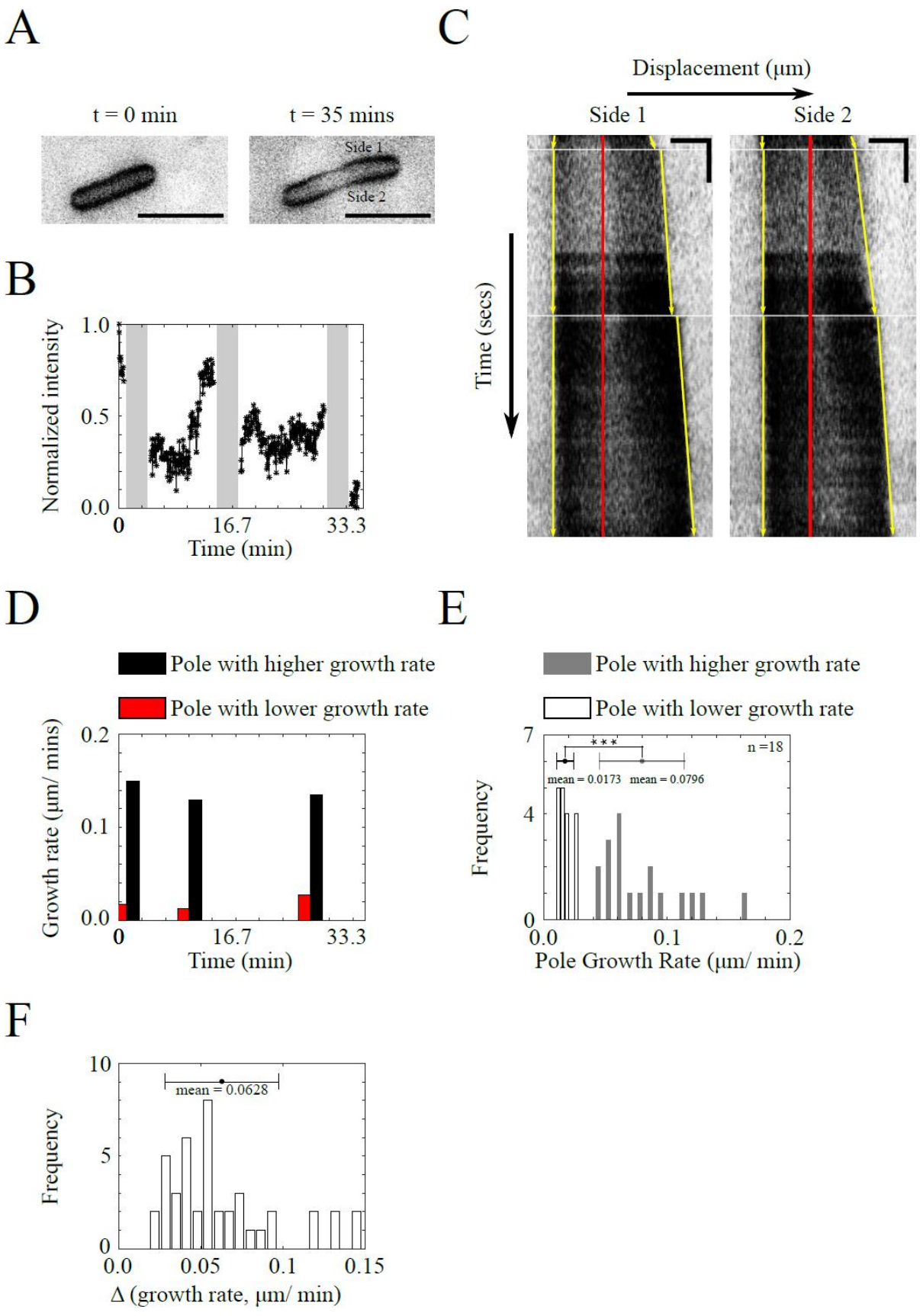

A
t = 0 min
t = 35 mins
Side 1
Side 2
B
Normalized intensity
1.0
0.5
0.0
0
16.7
33.3
Time (min)
C
Displacement (μm)
Side 1
Side 2
Time (secs)
D
Pole with higher growth rate
Pole with lower growth rate
Growth rate (μm/ mins)
0.2
0.1
0.0
0
16.7
33.3
Time (min)
E
Pole with higher growth rate
Pole with lower growth rate
Frequency
7
4
0
***
n =18
mean = 0.0173
mean = 0.0796
0.0
0.1
0.2
Pole Growth Rate (μm/ min)
F
Frequency
10
5
0
mean = 0.0628
0.0
0.05
0.1
0.15
Δ (growth rate, μm/ min)

Fig. 5.4: *E. coli* surface expansion is asymmetric across its mid- plane during growth (A) Cell membrane of *E. coli* cell were labelled using FM4-64 (inverted image) and a line of interest was used to generate kymographs for each side across the major axis of the cell. Scale bar- 5 µm. (B) FM4- 64 intensity in the bleached region is plotted as a function of time. Grey regions represent the time elapsed during bleach pulses (C) In the kymograph (inverted) dark region marks the labelled poles and the bright gap in the middle the bleach-mark. Bleach duration has been indicated with horizontal white line. Solid yellow arrows specify the LOI to evaluate the pole growth between two consecutive bleaching cycles and solid red line denotes the progression of bleached region over time. Horizontal scale bar- 2 µm; vertical scale bar: 150 secs (D) Growth rates of each end of the cell between two consecutive bleach pulses have been shown. Red and black bars represent poles with slow and high growth rate respectively (E) Pair- wise growth rates obtained from the kymographs are sorted based on the pole with more (grey) and less growth (white). The mean ± S. D. of pole growth rates were found to be significantly different by applying an unpaired t-test ($n = 18$; $df = 34$; $p < 0.001$) (F) Distribution represents the difference between the instantaneous growth rates of a pair of cellular ends in a time period between cumulative bleaching of the mid- cell ($n = 18$). Error bars- Standard deviation

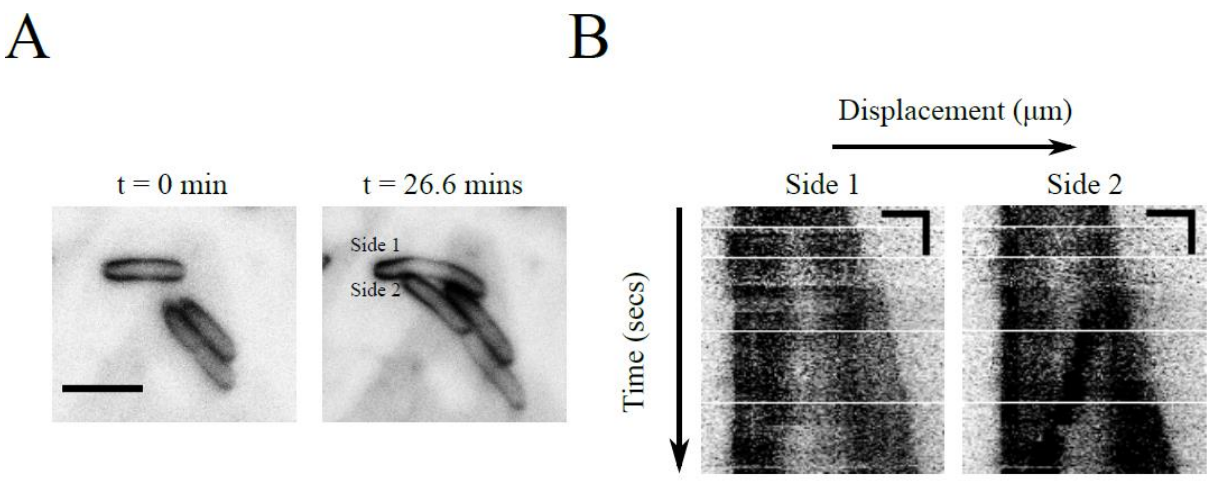
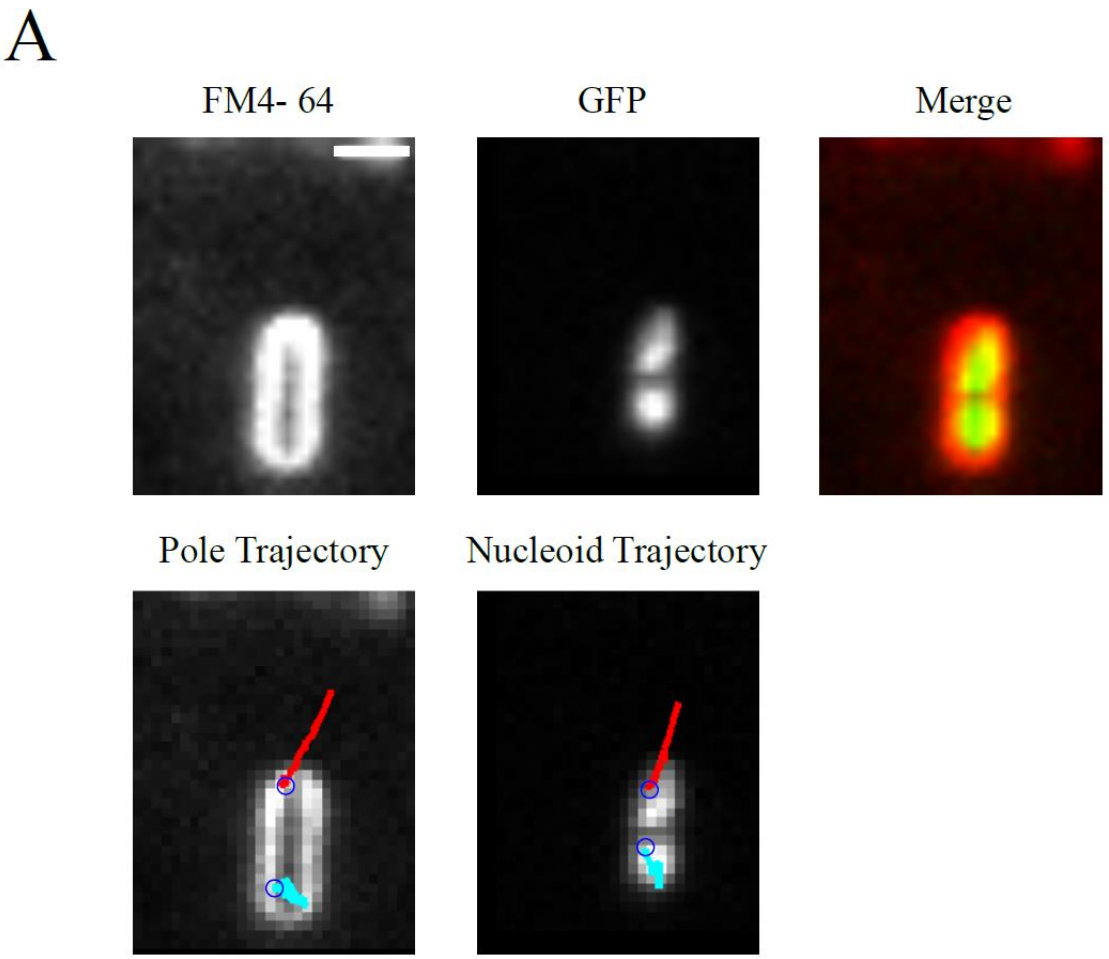
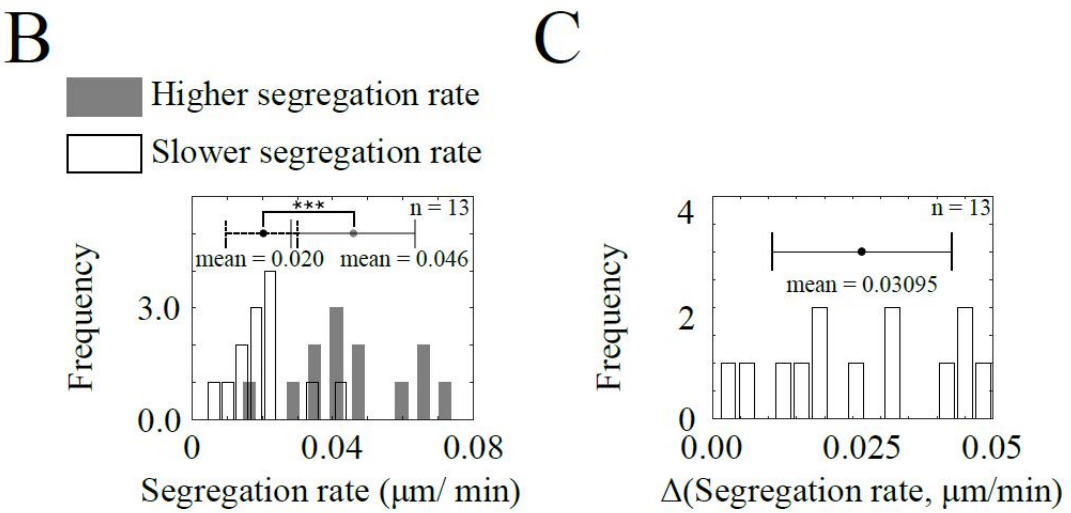


Fig. 5.5: Observed growth asymmetry in the cell is not because of force applied by the other cell in the vicinity (A) Representative images of *E. coli* cell stained with FM4- 64 (inverted) have been shown for t = 0 mins and t = 35 mins. Scale bar- 5 µm (B) Kymographs (inverted) generated for both the sides of the cell around its major axis have been depicted. (Horizontal scale bar- 2 µm; vertical scale bar- 100 secs).

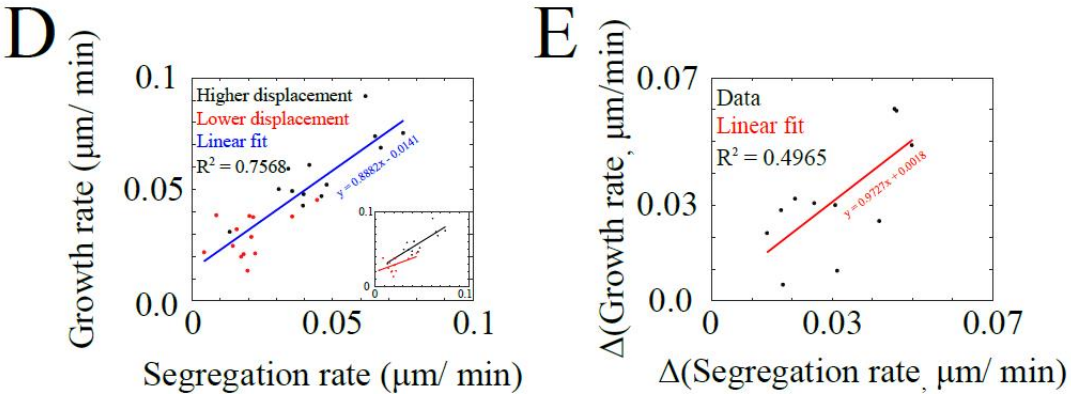
5.4 Effects on the movement of sub- cellular molecules

Asymmetry in growth introduces bias in the movement of two separating intracellular objects during cytokinesis. Displacements of two sister nucleoids were found to be significantly different (Fig. 5.6 A, B and C). Image analysis for pole growth revealed that the nucleoid near to the fast growing pole, displaces more than its duplicate (Fig. 5.6 A and D). Difference between segregation rates of two nucleoids forms a linear relationship with the difference between the growth rates of the two poles. Slope of the linear function is near 1 (Fig. 5.6 E), implying that the extra distance covered by the nucleoid is contributed by the biased growth of the cell towards its nearest pole. Thus, after confirming the asymmetric growth in *E. coli* cell, we proceeded to find the molecular mechanism underneath.


A
FM4- 64
GFP
Merge
Pole Trajectory
Nucleoid Trajectory


B
Higher segregation rate
Slower segregation rate
***
n = 13
mean = 0.020
mean = 0.046
Frequency
3.0
0.0
0
0.04
0.08
Segregation rate (µm/ min)

C
n = 13
mean = 0.03095
Frequency
4
2
0
0.00
0.025
0.05
Δ(Segregation rate, µm/min)


D
Higher displacement
Lower displacement
Linear fit
R² = 0.7568
y = 0.8882x - 0.0141
Growth rate (µm/ min)
0.1
0.05
0.0
0
0.05
0.1
Segregation rate (µm/ min)

E
Data
Linear fit
R² = 0.4965
y = 0.9727x + 0.0018
Δ(Growth rate, µm/min)
0.07
0.03
0.0
0
0.03
0.07
Δ(Segregation rate, µm/ min)

Fig. 5.6: Sister nucleoids segregate to different distance prior to cell division (A) Images of *E. coli* cells membrane- stained with FM4- 64 and with GFP labelled HupA (*first row*) are overlaid with their respective trajectories with open blue circles as their start point (*second row*). Scale bar- 2 µm (B) Distributions of the segregation rates of two sister nucleoids vary significantly in their average values, as tested by unpaired t test (n = 13; df = 24; p = 0.001). Error bars- standard deviation (C) Figure depicts distribution of difference in the displacement rates of two sister nucleoids. Error bars- standard deviation (D) Pole growth rates are plotted as a function of segregation rates of their nearest nucleoids. Black and red filled circles categorized cell poles based on their growth rates in a pair. Data was fit with linear function (blue solid line). Inset depicts two linear functions (solid black and red lines) fit to individual set of pole growths (black and red filled circles) (E) Difference in growth rates of poles in a pair was plotted against the difference in segregation rates of sister chromosomes and fitted to a linear function (red solid line).

5.5 Molecular basis for asymmetric growth

MreB cytoskeleton plays central role in governing the synthesis of new cell wall as well as the maintenance of the cell shape[83,120,121]. We proposed to study MreB localization within the cell as a possible mechanism for asymmetry in *E. coli* cell. We analysed distribution of MreB molecules in fixed as well as live *E. coli* cells, expressing MreB- YFP under the control of native promoter[6].

We developed an ImageJ macro that can divide DIC image of the cell into two halves and use those ROI to detect and determine the sum area of MreB loci in each cell half from corresponding YFP image (Fig. 5.7 A, Appendix C, Appendix D, D.2). Analysis of 15 cells

showed the significant difference in MreB area distribution in two halves of the cell (Fig. 5.7 B and C).

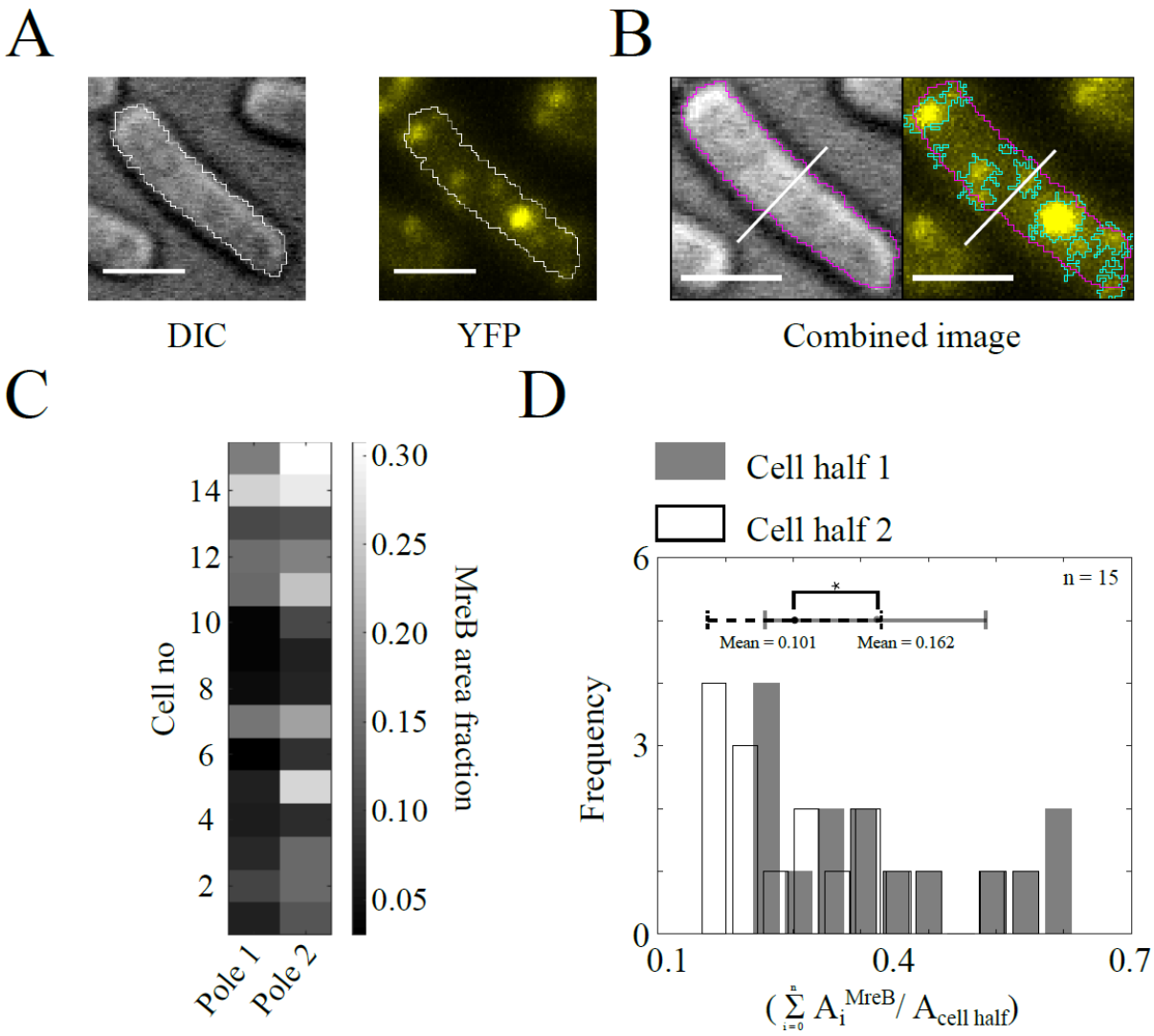


Fig. 5.7: Spatial distribution of MreB loci shows bias toward one of the cellular halves (A) Still image of *E. coli* cell bearing YFP tagged MreB in DIC (*left panel*) and YFP (*middle panel*) channel. Scale bar- 2 µm (B) DIC image was processed to obtain the cell outline (pink solid line) and geometric centre (white solid line) which were overlaid on YFP image to detect and quantify the total area occupied by MreB puncta in each cell half (*right panel*). Scale bar- 2 µm (C) Colourmap used to illustrate the difference in the area occupied by MreB in either halves of the cell (n = 9), the frequency distributions of which have been shown in (D). Difference in two distributions is tested by unpaired t test (n = 15, df = 28, p = 0.0306). Error bars- standard deviation.

Data was further bolstered by bleaching an area near one of cellular ends and observe recovery of MreB- YFP molecules for next 50 seconds (Fig. 5.8 A and Fig. 5.9 A). Our outputs faced an error because of an unintentional bleaching of fluorophore during the image capture. In order to cancel the errors in the measurements, we estimated intensity decay rate, λ, for experimental time, t, by fitting intensity data (I) collected from the reference cell to an exponential equation.

$$I_{ref}^{t} = I_{ref}^{0} * e^{(-\lambda t)} \quad \text{Equation 13}$$

Decay rate was then used to back- calculate the corrected intensities ($I_{correct}$) from the raw intensities (I_{FRAP}) at time 't', in the bleached area (Fig. 5.9 B, C and D and Fig. 5.10).

$$I_{correct}^{t} = {I_{FRAP}^{t}}\Big/{e^{(-\lambda t)}} \quad \text{Equation 14}$$

Corrected intensities were subjected to the full scale normalization[122].

$$I_{normalize}^{t} = \frac{(I_{correct}^{t} - I_{correct}^{bleach})}{(I_{correct}^{prebleach} - I_{correct}^{bleach})} \quad \text{Equation 15}$$

Where, $I_{correct}^{bleach}$ represents the intensity of ROI after bleaching, while, $I_{correct}^{prebleach}$ stands for ROI intensity before bleaching.

Normalized intensities were fit to the single exponential function (Fig. 5.8 C and Fig. 5.10) to retrieve the values of mobile fraction (a) as well as half recovery time (t_{half}) of MreB molecules in the bleached area.

$$f(x) = a * (1 - exp(-b * x)) \quad \text{Equation 16}$$

Half recovery time was calculated as follows:

$$t_{half} = -\frac{\ln(0.5)}{(-b)} \quad \text{Equation 17}$$

Distribution of mobile fractions showed existence of two populations, with 43% of the cells showing the value of mobile fraction greater than 0.5 in the bleached region. Remaining cells had mobile fraction of MreB equal or less than 0.5 (Fig. 5.8 D). Measurements of the growth

rates at the end nearest to the bleaching area showed that the mobile fraction of MreB is more at the ends that have higher growth rates end, while slow growing ends have less MreB molecules (Fig. 5.8 B and E, Table 5.1). Half recovery time of the intensities, ranged from 5 mins to 50 mins. It showed linear relationship with mobile fractions and was proportional with growth rate of the nearest pole (Fig. 5.11 A, B and C).

Though the results need further confirmations from more experiments, we predict the function of MreB in creating growth bias in *E. coli* cell.

Growth rate of the bleached end (µm/min)	Growth rate of the opposite end (µm/min)	Mobile fraction of MreB YFP	Half recovery time (secs)
0.063	0.005	0.8490	29.27
0.063	0.007	0.5846	19.23
0.021	0.056	0.3034	14.2
0.007	0.049	0.4779	20.64
0.021	0.059	0.4539	14.85
0.07	0.028	0.6082	18.67
0.063	0.007	0.5794	18.64
0.077	0.014	0.8844	46.21
0.021	0.047	0.3216	18.39
0.011	0.058	0.4025	9.2
0.016	0.053	0.4523	20.0
0.016	0.063	0.4623	15.19
0.0527	0.095	0.649	20.45
0.016	0.037	0.4962	16.03

Table 5.1: Correlation between pole growth rates and mobile fraction of MreB in *E. coli*.

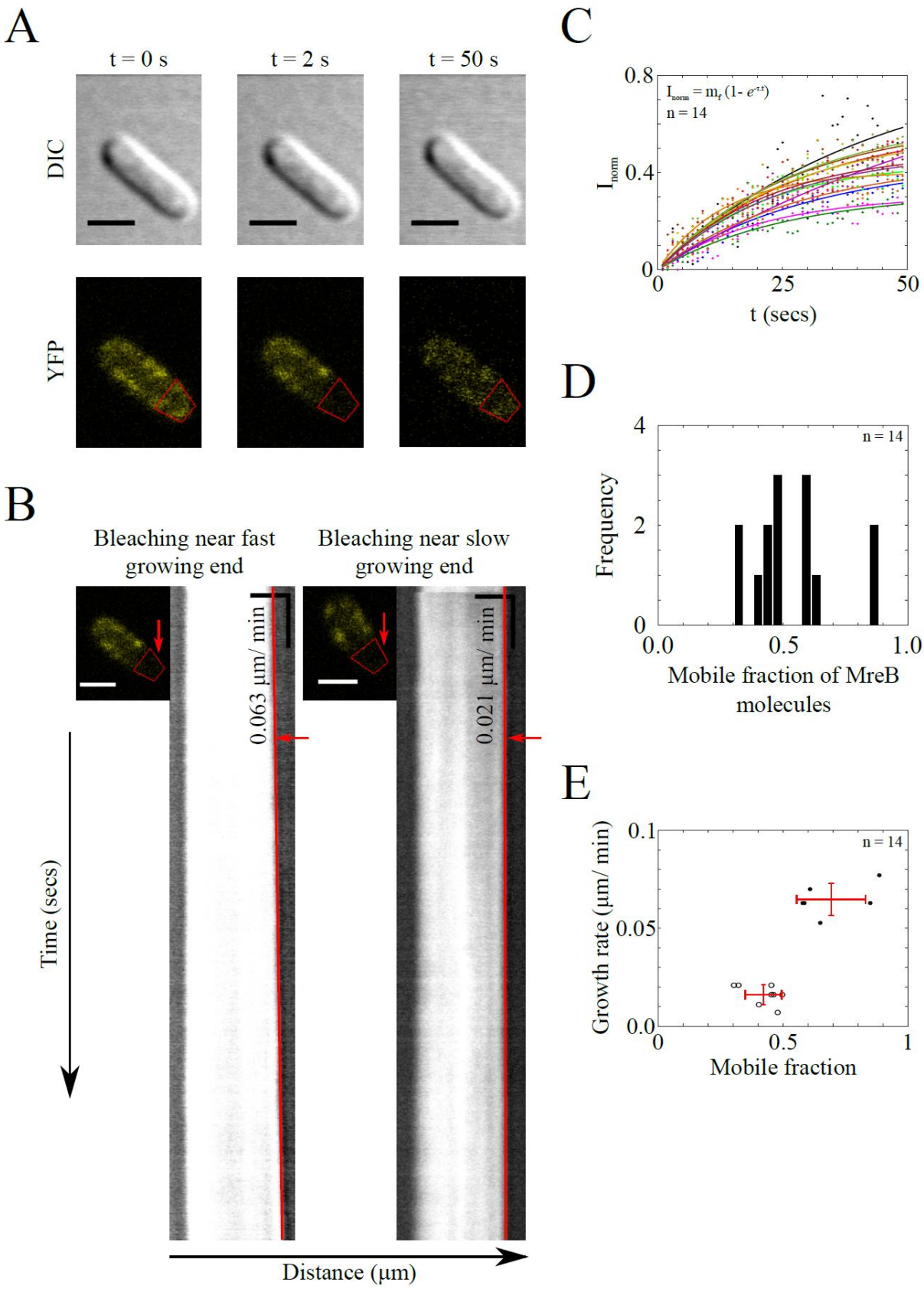

A
t = 0 s
t = 2 s
t = 50 s
DIC
YFP
B
Bleaching near fast growing end
Bleaching near slow growing end
0.063 µm/ min
0.021 µm/ min
Time (secs)
Distance (µm)
C
$I_{norm} = m_f (1- e^{-rt})$
n = 14
I_{norm}
t (secs)
D
n = 14
Frequency
Mobile fraction of MreB molecules
E
n = 14
Growth rate (µm/ min)
Mobile fraction

Fig. 5.8: MreB dynamics shows positive correlation with the growth rate with the nearest pole (A) *E. coli* cell with endogenous MreB YFP (*lower left panel*) was bleached near one of the poles (*lower middle panel*) and observed for fluorescence recovery (*lower right panel*) in YFP and in DIC (*upper panel*) channels. Scale bar- 2 µm (B) Kymographs were created from DIC images of *E. coli* cell. Slope of its edge (solid red line) was used to estimate the pole growth rate. Red solid arrows underline the fact that growth rate of the pole closest to the bleached ROI. Horizontal scale bar- 2 µm; Vertical scale bar- 50 secs (C) Normalized YFP intensities were fit to single exponential function to determine mobile fraction (m_f) and half recovery time of MreB molecules in bleached region (Equation 16). Each colour refers to an individual *E. coli* cell (n = 14) (D) Frequency distribution of MreB mobile fractions from regions near pole show two distinct groups. (E) Growth rates of bleached poles have been plotted against their respective MreB mobile fractions. Difference between growth rates (vertical error bar) and mobile fractions (horizontal error bar) of the resultant two populations was verified by an unpaired t test. (Growth rate: n = 14, df = 12, $p < 0.0001$; Mobile fraction: n = 14, df = 12, $p = 0.0004$). Error bars- standard deviation.

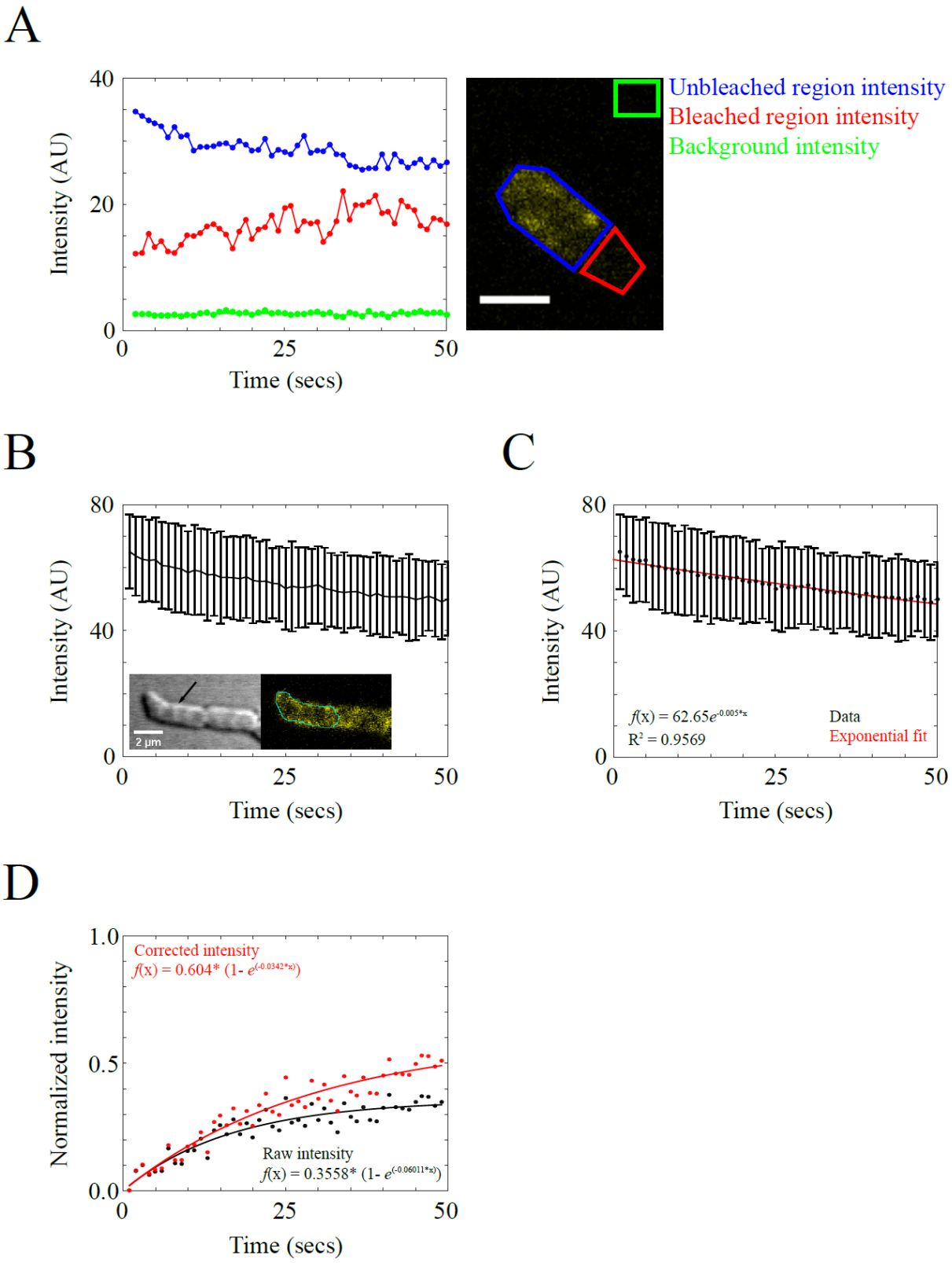


Fig. 5.9: Intensity values for post- bleaching period were corrected for YFP photo- bleaching (A) Intensities from bleached region (red), unbleached region (blue) of the cell and from background (black) have been plotted for first 50 secs of the experiment (B) Intensity values

from reference region were plotted as function of a time. Inset shows the DIC as well as YFP snapshots of the representative reference cell (indicated by arrowhead) at t = 0 secs. Reference region has been marked with cyan boundaries. (C) Average intensities obtained from 11 reference cells were fit to an exponential decay equation (Equation 13) to calculate the decay rate (0.005 s^{-1}) (D) Normalized intensities of bleached region before correction (black) and after correction (red) were plotted for first 50 secs and fit to single exponential function (Equation 16).

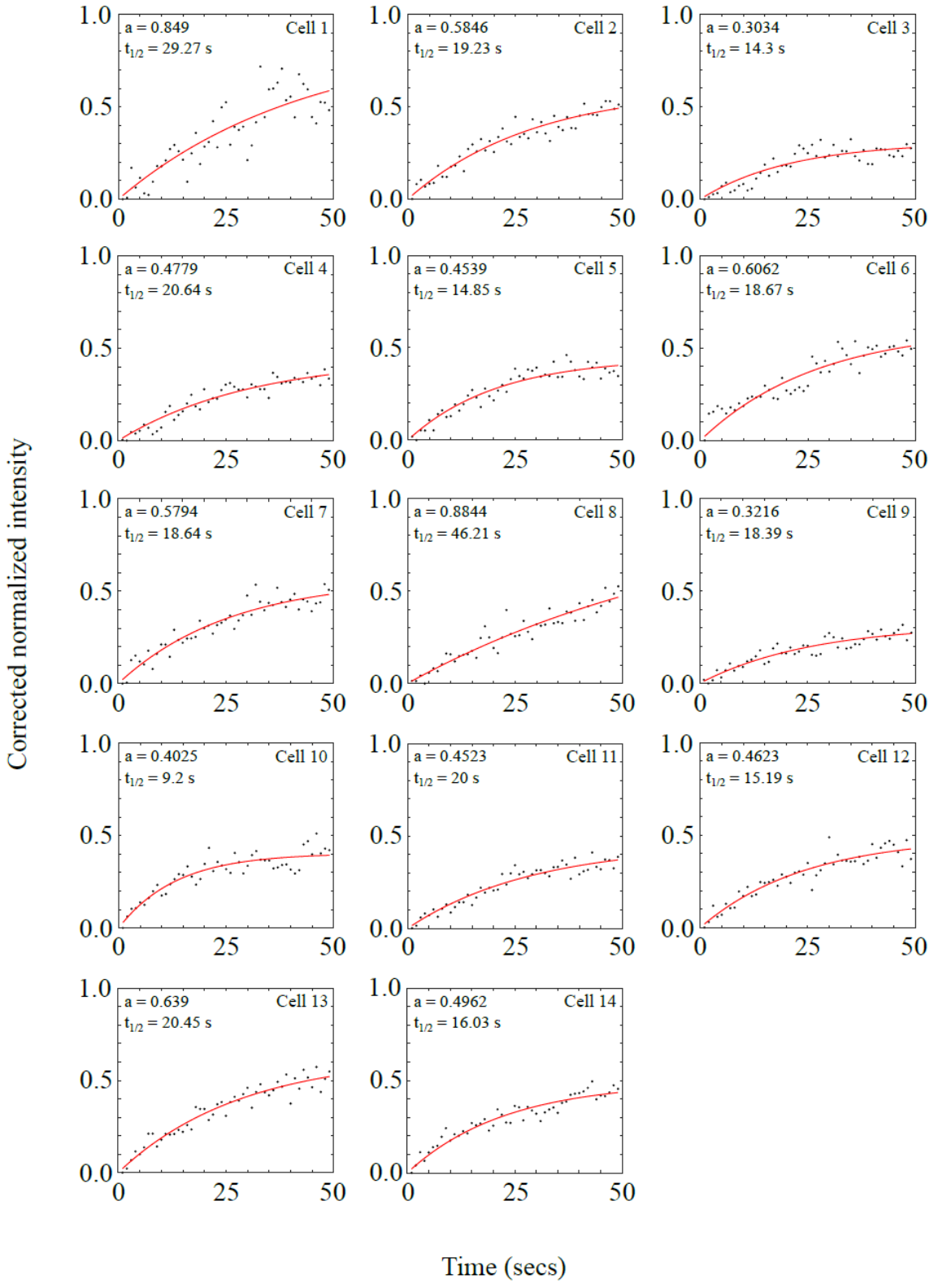


Fig. 5.10: Analysis of fluorescence recovery in bleached ROI. Corrected intensities of bleached regions were tested against single exponential function (Equation 16) to extract the values of mobile fraction of MreB molecules (a) and half recovery time (t_{half}).

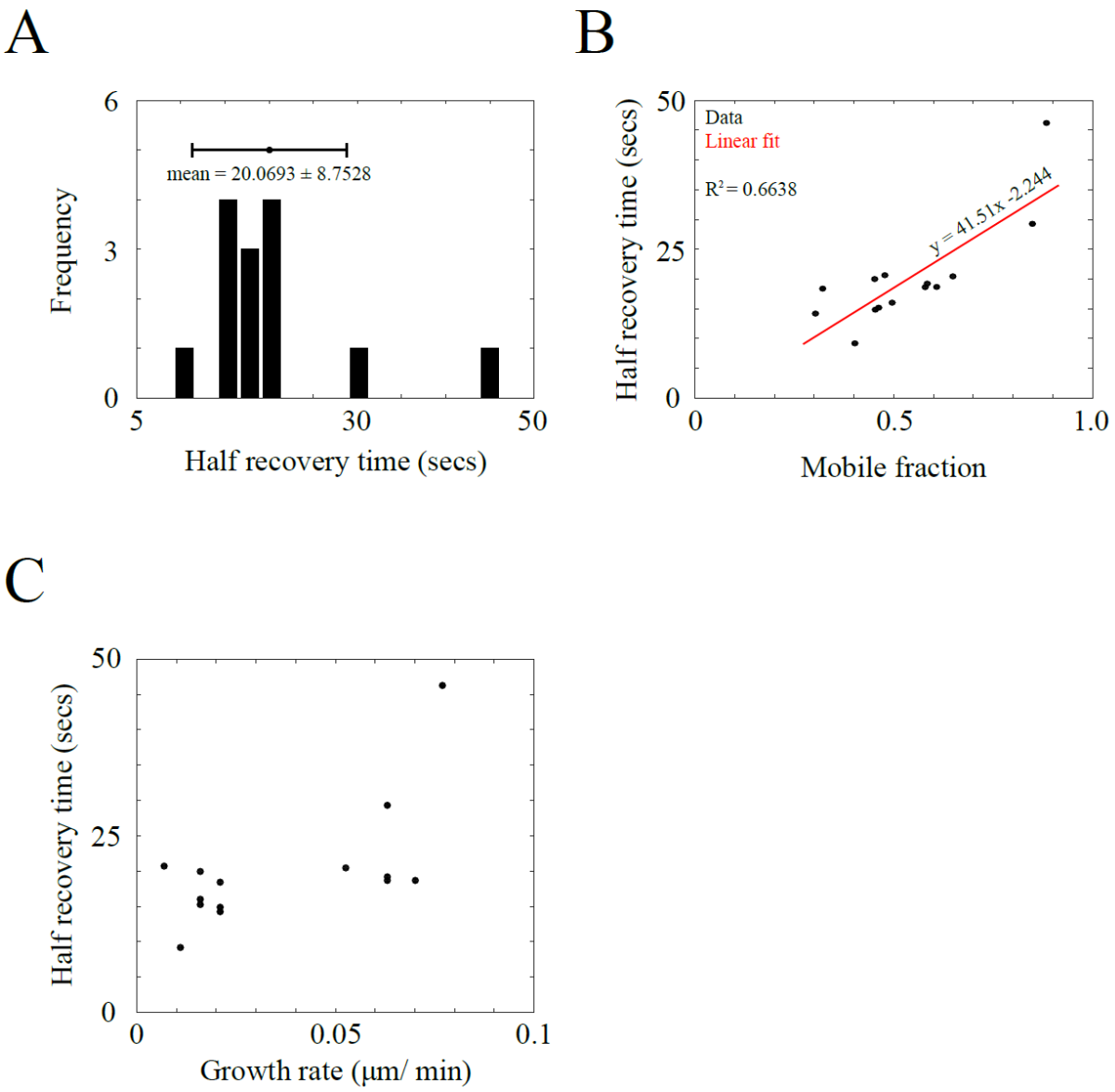


Fig. 5.11: Half recovery time of MreB YFP molecules is independent of their mobile fraction (A) Distribution of half recovery time of MreB YFP molecules obtained after bleaching *E. coli* cell (n = 14) has been depicted (B) Half recovery time has been plotted as the function of corresponding mobile fractions of MreB- YFP molecules in the bleached cellular end (n = 14) and are fit to the linear function (solid red line) (C) Half recovery time has been plotted against the growth rate of their nearest pole (n = 14).

5.6 Inheritance of growth asymmetry

Next, we studied the propagation of asymmetry through generations. Main aim of the analysis was to explore the connection of the cellular aging with growth asymmetry.

5.6.1 Analysis of two consecutive generations

Cells were labelled with FM4- 64 and the poles fluorescing with the highest intensity were tracked with an ImageJ '*particle tracker plugin*'. Since, FM4- 64 dilutes out with each cell cycle, only two consecutive generations could be analysed. Asymmetry was inherited from mother to its daughters (Fig. 5.12 A). However, we found a peculiar pattern. Fast growing end of the mother continues to be the fast growing pole in daughter also. On the other hand, daughter that inherits slow growing end from its mother, develops new pole that grows with higher rate (Fig. 5.12 B).

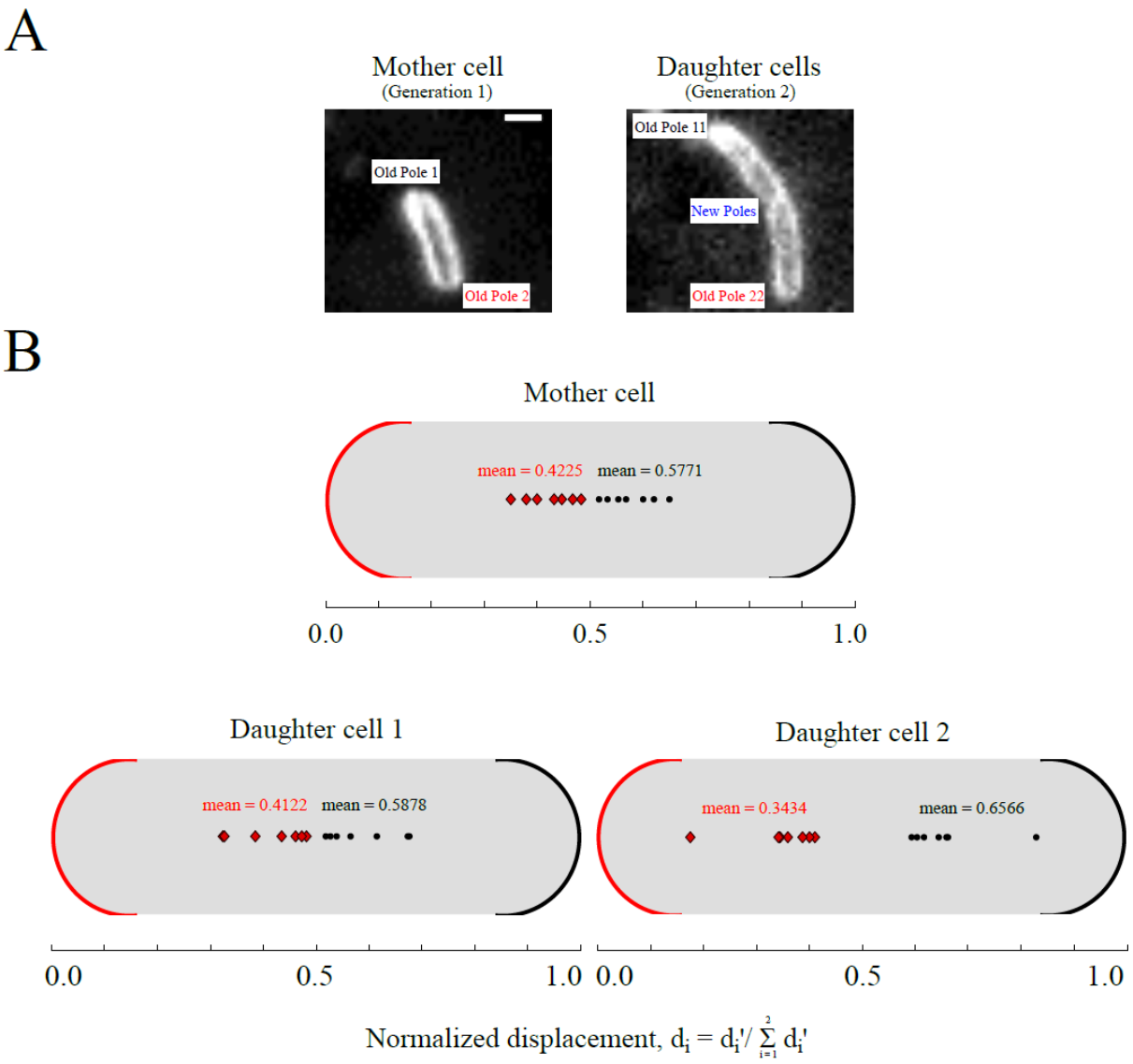


Fig. 5.12: Growth asymmetry is inherited from one generation to the next (A) Inheritance of the poles from mother *E. coli* cell labelled with FM4- 64 to its two daughters has been elucidated (B) Sum normalized growth at each end of the cell was estimated and has been illustrated in the figure along with the pole inheritance for two consecutive generations. Black filled circles signify high growth, while red filled diamonds stand for slower growths (n = 7).

5.6.2 Analysis of micro- colony generated from single mother

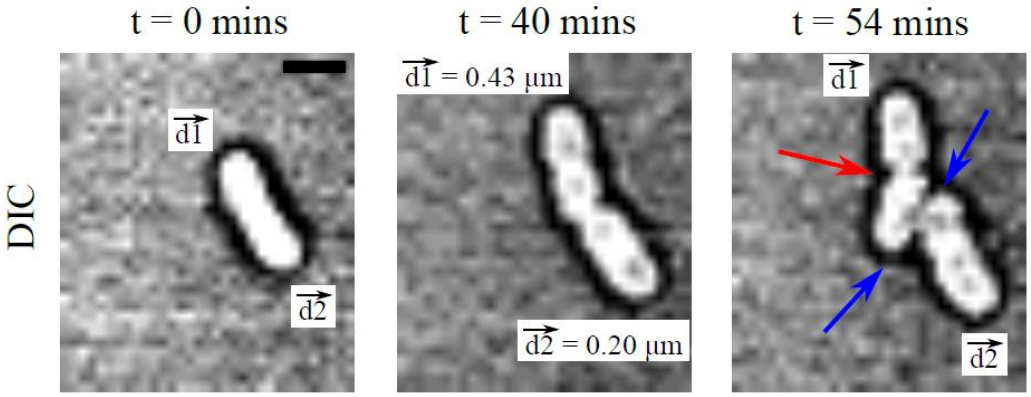
Bias in cellular growth was found to introduce a difference in cell division time of two identical daughters (Fig. 5.13 A and B). Cellular growths were assessed in six micro- colonies spawned from single mother cells for four generations. A schematic has been used to explain the inheritance of the poles from parental generations to the progenies. Number assigned to each end designates its age (Fig. 5.13 C). We observed the association of fast growing pole with less generation time in a pair of sisters. In other words, daughter that receives fast growing end from the mother divides earlier than its twin sister with slow growing end from its mother. Trend was found to be propagated in future generations also (Fig. 5.13 D and E), though it was found to be declined with time (Fig. 5.13 F) which was reflected in the asymmetry ratio which is nothing but the sum normalized difference in the doubling times in a pair of sisters.

$$Asymmetry\ ratio = \frac{(t_{d_{max}} - t_{d_{min}})}{(t_{d_{max}} + t_{d_{min}})} \qquad \text{Equation 18}$$

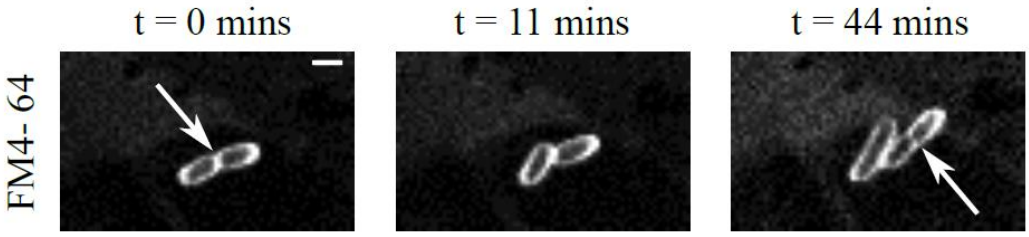
Where $t_{d_{max}}$ is the division time of the slow growing sister, while $t_{d_{min}}$ represents doubling time of fast growing sibling.

Average asymmetry ratio showed a linear decrease with generations. Moreover, the ratios obtained for various pairs of the sisters became more and more clustered around the average value with increase in time. (Fig. 5.13 G).

A



B

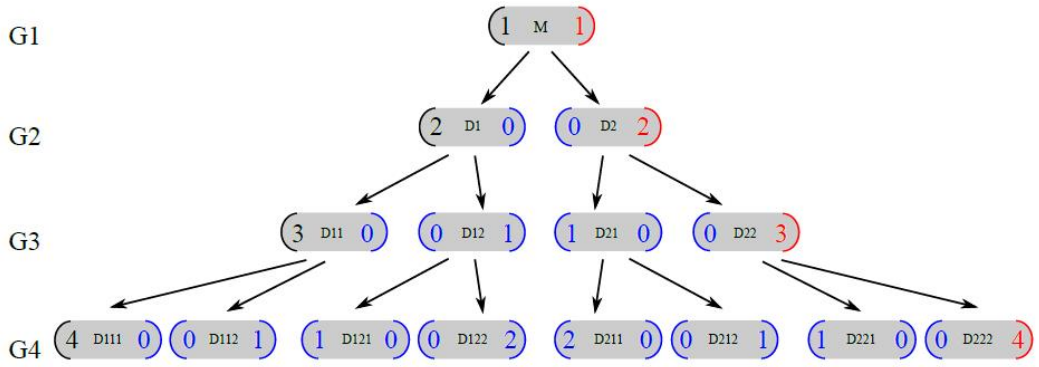


Septum formation in mother

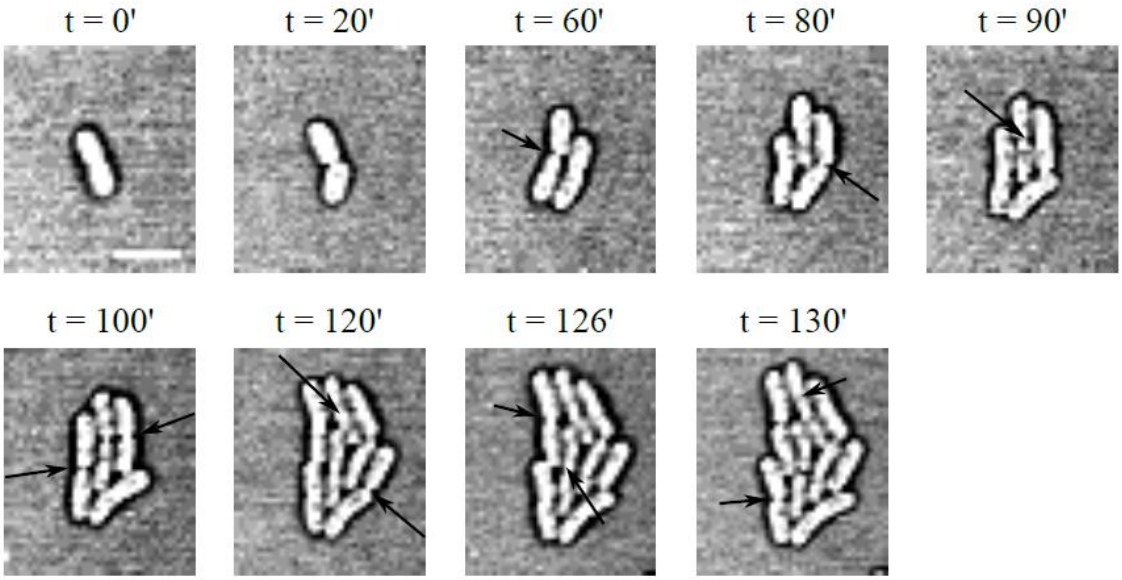
Birth of two daughters

Septum formation in one of the daughters

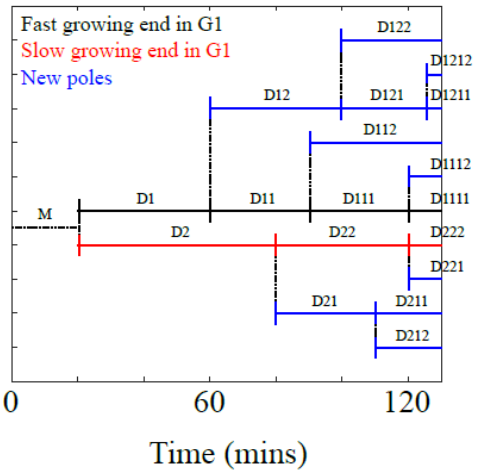
C



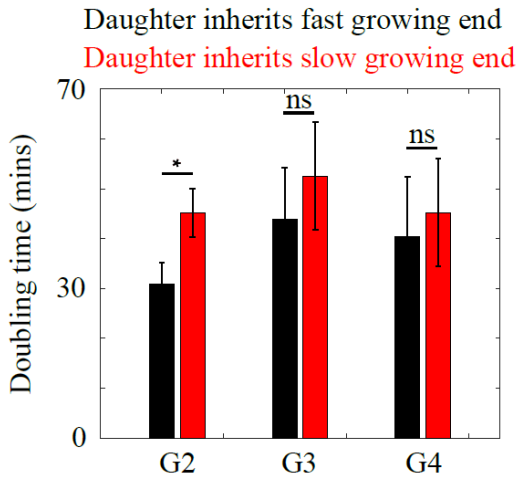
D


t = 0'
t = 20'
t = 60'
t = 80'
t = 90'
t = 100'
t = 120'
t = 126'
t = 130'

E


Fast growing end in G1
Slow growing end in G1
New poles
M
D1
D11
D111
D1111
D12
D121
D1211
D122
D1212
D112
D1112
D2
D22
D222
D221
D21
D211
D212
0
60
120
Time (mins)

F


Daughter inherits fast growing end
Daughter inherits slow growing end
Doubling time (mins)
70
30
0
*
ns
ns
G2
G3
G4

G

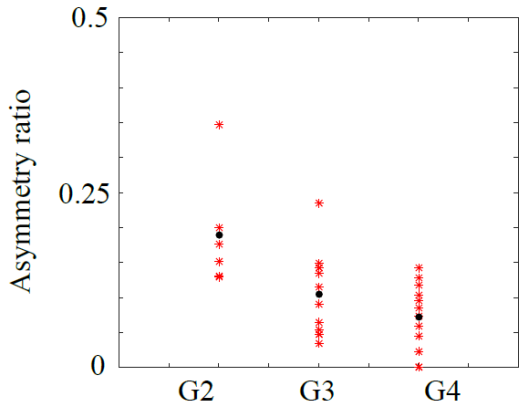

Asymmetry ratio
0.5
0.25
0
G2
G3
G4

Fig. 5.13: Single cell generation time for two sisters differs because of asymmetric growth of the mother cell (A) DIC time series of growth of *E. coli* cell shows that Daughter cell that inherits fast growing pole (d1) from mother divides earlier (marked by red arrow) than its sister that inherits slow growing pole (d2) from the mother cell. Scale bar- 2µm. Blue arrows indicate the generation of new poles after division event in mother cell (B) Time taken to complete one cell cycle is different for two sister *E. coli* cells membrane stained with FM4- 64 dye. Scale bar- 2 µm (C) Schematic of pole inheritance in an *E. coli* micro- colony has been shown for four generation. Black and red coloured poles denote an inheritance of mother cell (indicated as M) poles into daughter (indicated as D) generations, while new poles formed as a result of cell division are shown by blue colour. Numbers near the pole indicate the age of the corresponding poles in the micro- colony (D) Montage of growing colony at different time point depicts the variation in the generation times of the two sisters. Scale bar- 3 µm (E) Association of an inheritance of the growth asymmetry (horizontal lines) with division time asymmetry between two sister cells has been shown with the analysis of a microcolony sprawled from one mother. Black solid line represents inheritance of fast growing end. Slow growing end is shown with red solid line, while blue solid lines represent the inheritance of newly developed ends after division. Vertical solid lines denote the division events, while black vertical discontinuous lines joins two siblings generated from same mother (F) Average doubling time of two daughters in each generation has been shown. Black and red bars denote the doubling time of the cells inheriting fast growing and slow growing ends respectively. Significance in the difference between the doubling times of the two sisters was tested using unpaired t test. Error bars- standard deviation (G) Difference between the generation times of two sisters has been presented as asymmetry ratio (ratio of difference between the two doubling times to the addition of two doubling times) and plotted against its corresponding generation (no. of micro- colonies analysed- 6).

5.7 Discussion

Though *E. coli* exhibits perfectly symmetrical rod shape, molecular studies show asymmetric distribution of damage towards the older pole of the cell[123]. Division of this cell gives rise to the daughters with different fates. Stochasticity in the partition of sub- cellular components has been postulated as the source of heterogeneity in an isogenic population[124].

In this section we studied single cell growths in *E. coli* and observed the bias in the growth distribution in two halves of the cell. Since, our findings are contradictory to the classical microbiology description of an *E. coli* cell growth we increased the reliability of our observations through the experimental settings as well as by studying the growth of the cell with reference to the temporal changes in (a) cell surface (b) cell membrane and (c) nucleoid segregation.

Pole growth was mapped as an indicator for growth in the two cell arms. Pole growth is not to be misconstrued as the presence of active growth regions at poles. From different experiments, we found that the average growth rate of fast growing end or leading end in *E. coli* is between 0.05- 0.08 µm/ min, while growth rate of slow growing end or lagging end ranges from 0.01- 0.04 µm/ min. Asynchronous cells showed asymmetric growths, irrespective of their growth rates on agar surface. Hence, we inferred that the growth asymmetry is independent of the growth rate of an individual cell.

Preliminary analysis of the cell growth was done using in- house MATLAB programme. In order to remove spurious tracking the lateral coordinates of the cell were confined on the agar surface by etching it with micro- patterns. Bead movement was tracked to validate MATLAB programme. We expected bead displacement equal to or around 0 µm/ min, as exhibited by inanimate object under thermal noise. *E. coli* cell, in its native form, does not have surface markers which can be used as a guide to dissect its growth. Hence we created a marker by

bleaching the FM4- 64 tagged membrane at the geometric centre of the cell. Over the time, an apparent translation of the mark on *E. coli* surface was observed. We attribute this to the creation of newer mid- planes with the growth biased towards one of the halves. This experiment is of importance, in that we observe asymmetry in the growth irrespective of the presence of any structure or inconsistencies in agar surface in immediate vicinity of an organism. Growth bias was more evident when asymmetry was noticed in the displacement of the nucleoids during the segregation. Due to asymmetric growth, destined locations for segregated chromosomes are spaced unevenly from the site of chromosomal replication. And hence the segregation bias is introduced into the displacements of two sister nucleoids.

Rod shaped cells like *Mycobacteria* and *Rhizobia* which exhibit asymmetric growth lack MreB[113,114]. In these organism, poles retain active growth. On the contrary, MreB is an important protein in maintenance of rod shape and governance of cell wall synthesis in *E. coli*[83,121]. Our analysis of fixed as well as live cells showed polarization of MreB localization in the cell. In order to eliminate the risk of interpreting the data limited by resolution capacity of microscope, MreB dynamics was examined in live cells with the help of FRAP. Upon quantifying the recovery of MreB molecules in the bleached polar regions of various cells, we could clearly categorized them into two distinct populations. 6 out of 14 cells analysed, showed higher fluorescence recovery as compared to other cells. Higher mobile fraction in FRAP region implied higher concentration of MreB in nearby region. As more MreB molecules flow in, fluorescence in that region is almost recovered. It confirmed our conclusions drawn from fixed cell images. Moreover, a strong correlation between the growth rate of each end and the mobile fraction of MreB molecules could be seen. Leading end always had mobile fraction of MreB greater than 0.5, while lagging end of the cell possessed mobile fraction less than 0.5. Result suggested the plausible role of MreB in the generation of asymmetry. However, as mentioned earlier, our experiments faced the drawback of resolution limit as well as stability

of MreB- YFP molecules, as experiments suffered from quick loss of YFP fluorescence during imaging. Super- resolution imaging accompanied with stable construct of MreB will be useful. "Sandwich" construct described in earlier studies, in which mTomato encoding gene was introduced within the reading frame of genomic copy of *mreB* gene[125], can be considered as suitable for these studies, as it has been proven to be stable with time. Since, the construct is stable it can help to follow MreB dynamics and its possible contribution in growth asymmetry in the growth through more than two generations.

Unipolar growth in *Mycobacteria* generates two daughters which have different susceptibility towards antibiotic treatment[112]. Hence, we probed into the inheritance of growth asymmetry and its effect on *E. coli* population. Inheritance of asymmetry in *E. coli* was found to be different than that in *Mycobacterium*. Leading end continued to grow with higher rate in subsequent generations, while cell inheriting lagging end of the mother, developed a new leading end. We observed that it created time lag between the divisions of two sisters in initial generations on agar surface. But, as time progressed, time lag decreased. Similar phenomenon observed in *Mycobacterial* cells is known to give rise to the daughter cells that differ in their antibiotic susceptibility[112]. In *E. coli*, the cell cycle is governed by 'adder principle' of homeostasis[23], as opposed to the 'timer' model in *Mycobacterium*. According to the principle, cell does not proceed to cytokinesis before it adds up constant mass. Hence, we reasoned that the time lag in the division cycle of the twins does not necessarily lead to the cell length variability.

However, with current studies we are unable to comment on the significance of asymmetric growth on the cellular physiology. Follow up experiments are required to conclude its utility in the cell. Though study of the cells through several generation can give us a clear picture, current experimental set up is not suitable for such observations. Tracking cellular ends becomes tedious as well as increasingly unreliable in the colony speedily spreading on the agar

surface. A careful design of microfluidic device will extremely be helpful in such scenario. Microfluidic experiments can be combined with fluorescence tracking of the poles by tagging polar proteins in *E. coli*. It will help enhance accuracy of the results. Simultaneous studies of aging pattern in the cells will be a nice extension to these.

Results received from these suggested experiments in association with existing conclusions will help revise the concept of the growth for a single cell of *Escherichia coli*.

Growth dynamics undoubtedly turns out to be of utmost importance in the maintenance of cell dimensions. Any alterations in it, can affect the shape of an organism. Nonetheless, it is more important to understand its subsequent effects on the population.

Chapter 6

Colony competition and spatial patterns in isogenic populations growing from a central "homeland"

6.1 Motivation

Morphology of microorganisms range right from simple rounds to complex helical shapes[126,127]. These shapes have been known to be evolved as an adaptation to particular environments[128]. For instance, the crescent shape of *Caulobacter crescentus* gives organism an advantage of propagating younger populations more efficiently and successfully in the fresh water[129]. Modern bacteria can change their morphology as a response to the environmental changes in their habitat[69,130]. These changes can be nutritional, in which case, bacilli like bacteria metamorphosed into durable and long- lasting spores[111]. Environmental cues also decide the patterning of the populations. *Myxococcus xanthus* population, for example, has cells aligned parallel to each other along the long axis, in the process called 'rippling', important for predation of the prey colony[131].

In their natural niche, however, different species of microorganisms share the same space[132]. Such biofilms consists of different organisms with different shapes. How cell shape and size determine the structure of the colony is one of unresolved questions in microbiology. Understanding of the dynamics of these biofilms has become necessary as they play cardinal role in many chronic diseases, formation of human commensal microbiome, development of antibiotic resistance and bioflueling.

In this section we attempted to extend our previous work with cell length variability to shed light on the connection between cell sizes and colony morphology.

We investigated it using "colony competition assay" devised by Hallatschek et al in 2007[133]. In the original set up, two populations of *E. coli* DH5α expressing either YFP or RFP were mixed together to form a homeland region. Resultant colony was sectioned with every sector distinctly segregating the two "sub- populations". This genetic de- mixing in the colony was attributed to the random fluctuations that occur at the smaller scale. Changes in the cells that are present at the expanding front can leave its signature in terms of specific cellular arrangements within the colony as colony progresses away from the homeland region (Fig. 6.1).

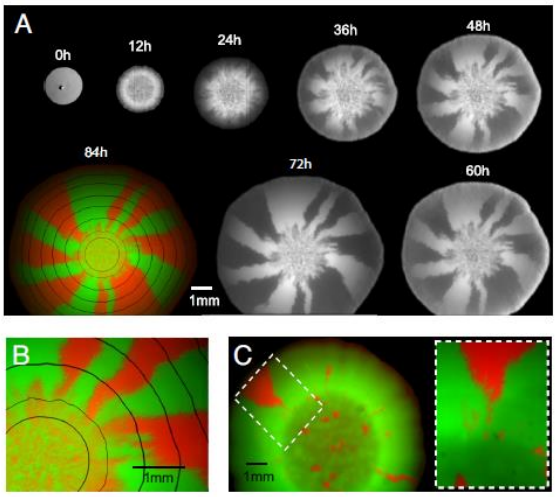


Fig. 6.1: Formation of defined sectors in a mixed culture of *E. coli* DH5α growing on agar surface (Reprinted from Hallatschek et al., 2007 with permission[133]. Copyright obtained from PNAS (Appendix E, C6)).

In other words, composition of the homeland region can influence the choreography of an entire community. It formed the basis of our study where we tried to introduce phenotypic noise in the homeland region and study its effect on the colony structure.

6.2 Experimental design

Experiment was broken down in two parts. First part involved determination of motility of *E. coli* strains under study, whereas in second part was planned to deal with the competition between two different *E. coli* strains.

1. Determination of the motility

Motility of the various strains of *E. coli* was recorded as explained in section 2.13 of chapter 2. Time lapse sequence thus obtained for each strain was then analysed using ImageJ (v1.50f). Background of these recorded DIC images was removed by subtracting 'a constant number' from every pixel of every frame of the film. Though the value of the constant varied from one movie to another, it was same for all the frames in one movie. Processing gave the images false appearance of fluorescence images in which only those pixels covered by cell body had highest intensity. Processed images when fed to '*particle tracker*' plugin in ImageJ, centroid of the cell was detected and traced over time to return the values of its X and Y coordinates in every frame. Euclidean distance covered by different cells at every time interval was then fit to Gaussian function, after the data was sum normalized. Use of '*pdf*' function in MATLAB helped to calculate a fit, when its average value was fixed at '0 μm' for non- motile strain. Instantaneous displacements were fit to the Gaussian using 'free mean', for motile cells (Appendix D, D.3). Standard deviations calculated accordingly, were then, in turn, used to determine the diffusion coefficient for different *E. coli* strains.

2. Colony competition assay

E. coli strains differed in the marker plasmids were competed with each other as described in section 2.14, chapter 2. Growth of the colony from 'homeland region' was captured in GFP, RFP and bright field channel of a stereoscope at the interval of 12 hours. Primary analysis included mapping of percent area occupied by each strain in a colony in order to understand

whether phenotypic heterogeneity confer any (dis)advantages to the strain. Image analysis was carried out in ImageJ (v1.50f). Bright field images were processed in order to extract area of the colony, while images in GFP and RFP channels gave area occupied by different strain in the colony. After applying '*Otsu thresholding*', the area was quantified using '*Analyze particle*' module in ImageJ. Assessment of the competition was based on the increase or decrease in the percentage area of any of the strain over time.

6.3 Determination of motility of *E. coli* strains

Motility is one of the dominant strategies that has been evolved to help organism explore the niche for food as well as for shelter. It grants cell survival advantage over other non- motile cells. Hence, in our experiments, it was necessary to determine the chemotactic ability of each *E. coli* strain under consideration, so that we could avoid the competition between motile and non- motile strain. Analysis, thus, helped us to eliminate the possibility of outgrowth of the motile strains over other non- motile bacteria because of their ability to surf the surface.

Movements of circular beads (Diameter: 1 μm), *E. coli* DH5α and *E. coli* RP437 were examined as experimental controls. Motion studied in bead solution reflected the random movement of inanimate objects due to thermal activities. DH5α served control for non- motile bacteria. RP437 is a bonafide strain that shows motility[3] and was analysed as a control for directed movement of bacteria (Fig. 6.2 A).

Single step taken by objects in 'X' and 'Y' directions were calculated (Section 6.2) and their frequency distributions were tested for Gaussian function (Appendix D, D.3) to find average displacements (μ) as well as standard deviation (σ).

$$f(x) = \frac{1}{\sigma * \sqrt{2 * \pi}} * e^{\frac{-(x-\mu)^2}{2*\sigma^2}}$$ Equation 19

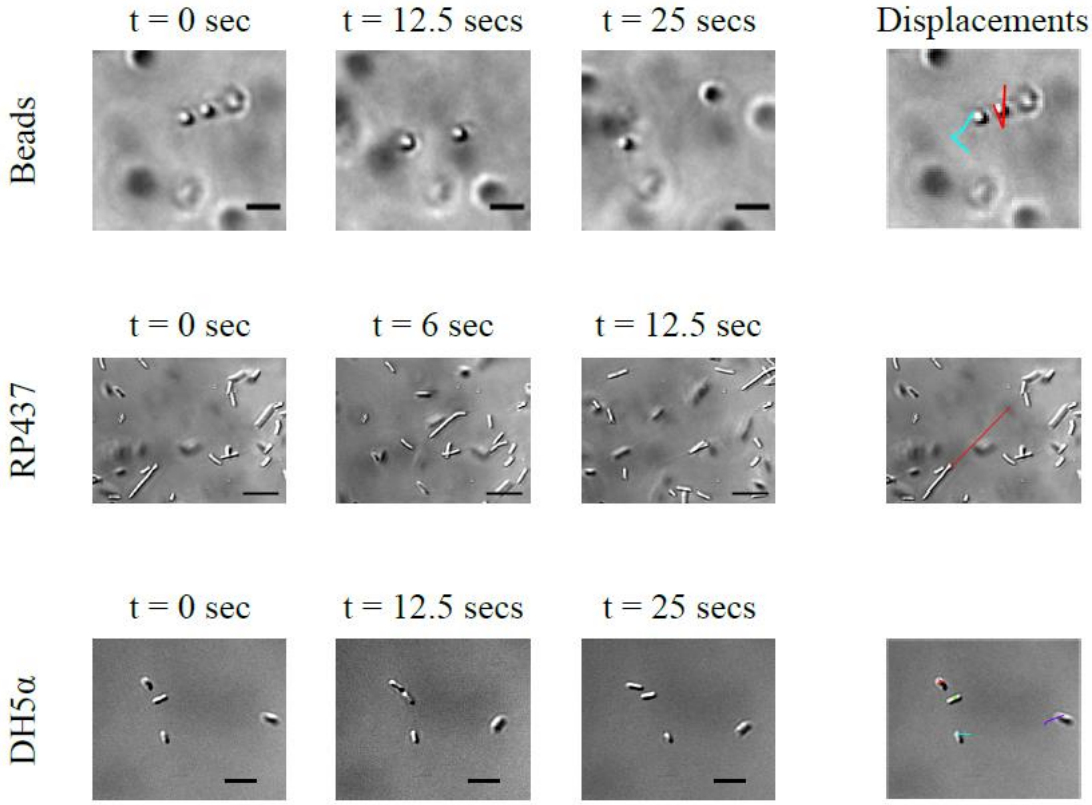
The later was substituted in the formula: $D = \frac{\sigma^2}{2t}$, Equation 20

to estimate the coefficient of diffusion for each strain for the duration of an experiment (t)[134]. Distance covered by beads and DH5α in a single step was averaged at 0 µm, while for RP437 showed average movement of 0.173 µm in both the directions. As expected, RP437 was reckoned with the highest coefficient of diffusion. On the other hand, beads and DH5α had diffusion coefficient of 0.1 and 0.03 μm^2/ sec, (Fig. 6.2 B). We verified the reliability of our results by comparing the diffusion coefficient of beads extracted from the experiment with the diffusion coefficient (D) computed for an object of 1 µm diameter (Fig. 6.2 C), using Stokes-Einstein equation[134]:

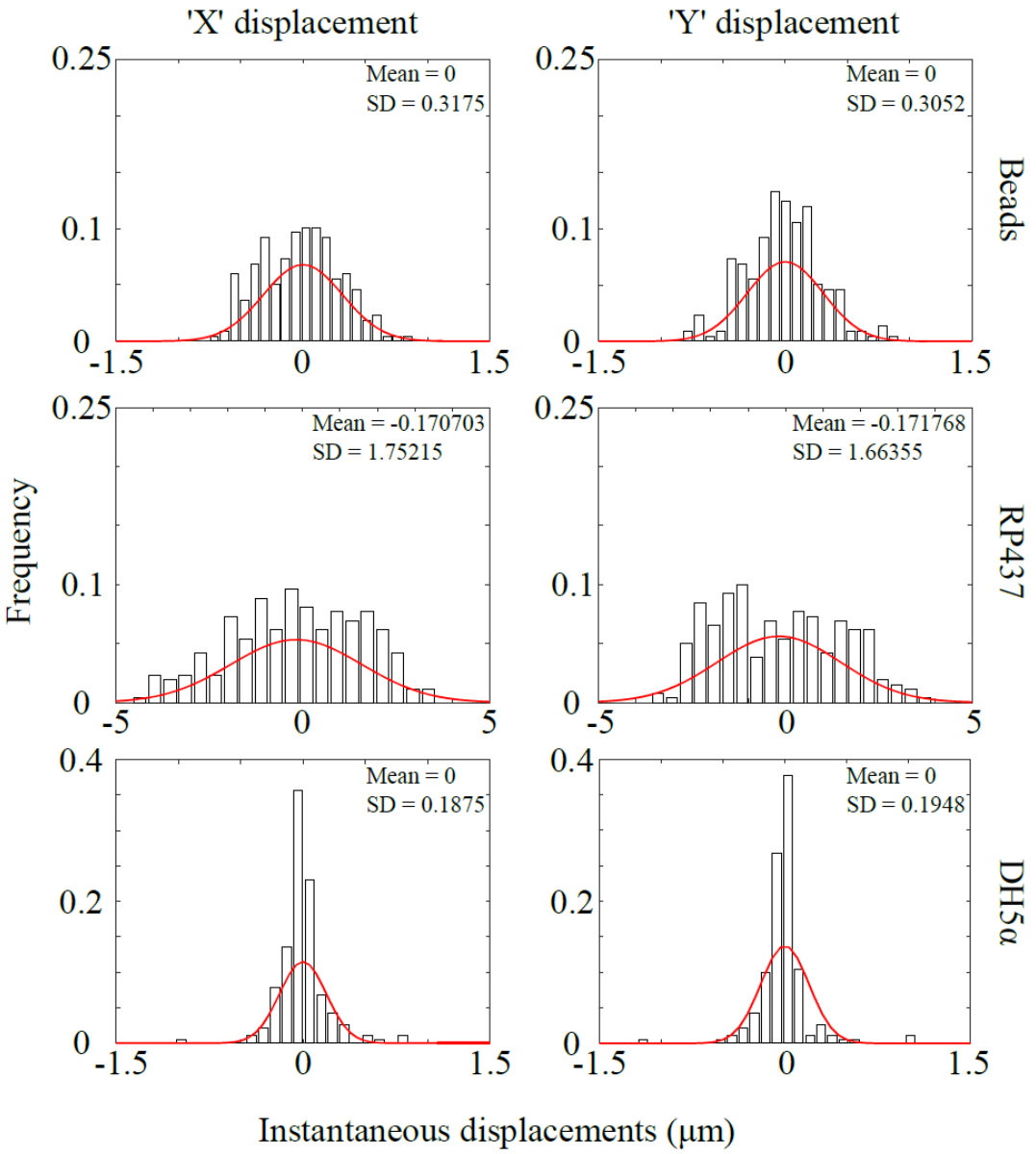
$$Diffusion\ coefficient, D = \frac{KT}{6 * \pi * r * \rho}$$ Equation 21

K denotes Boltzmann's constant, T is the temperature. The radius of the object and the viscosity of the medium has been presented by r and ρ respectively.

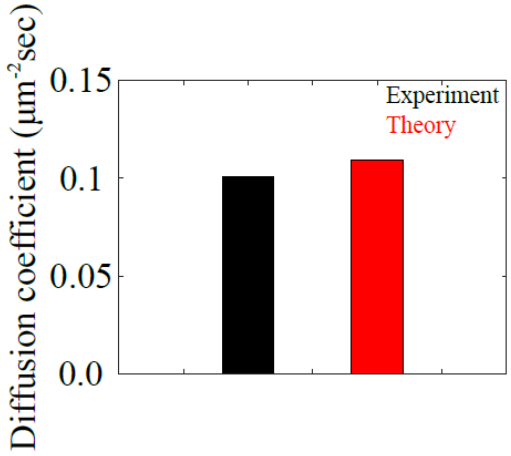
A


t = 0 sec
t = 12.5 secs
t = 25 secs
Displacements
Beads
t = 0 sec
t = 6 sec
t = 12.5 sec
RP437
t = 0 sec
t = 12.5 secs
t = 25 secs
DH5α

B


'X' displacement
'Y' displacement
Mean = 0
SD = 0.3175
Mean = 0
SD = 0.3052
Beads
Mean = -0.170703
SD = 1.75215
Mean = -0.171768
SD = 1.66355
RP437
Mean = 0
SD = 0.1875
Mean = 0
SD = 0.1948
DH5α
Frequency
Instantaneous displacements (μm)

C


Diffusion coefficient (μm-2sec)
Experiment
Theory

D

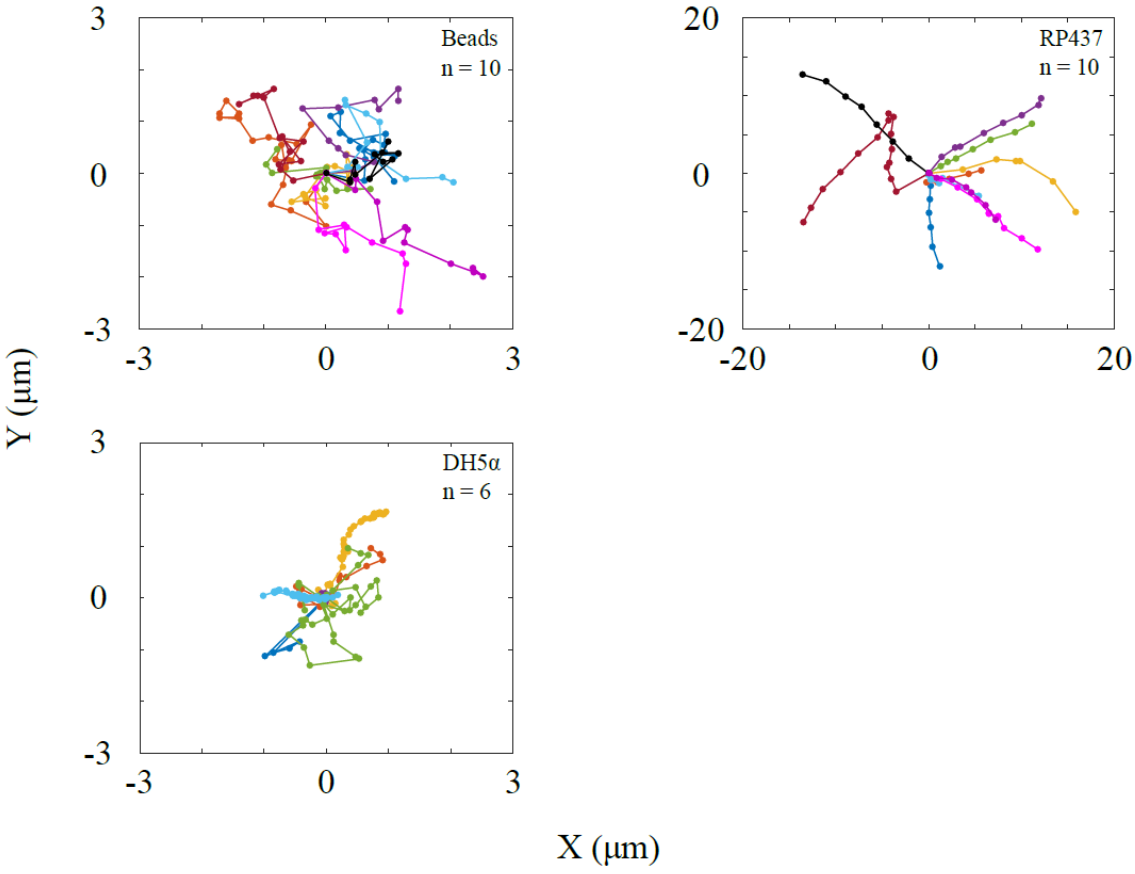
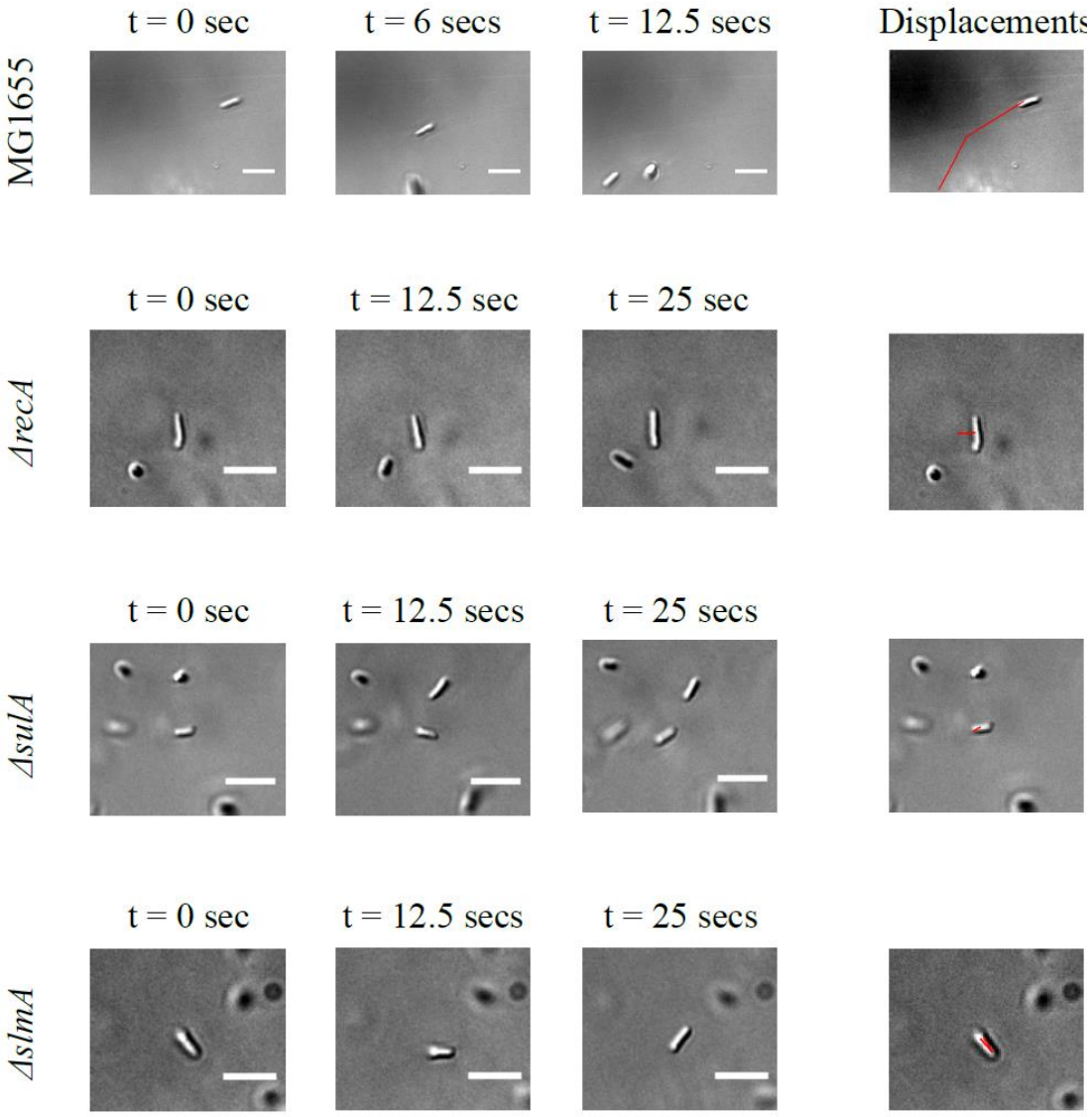


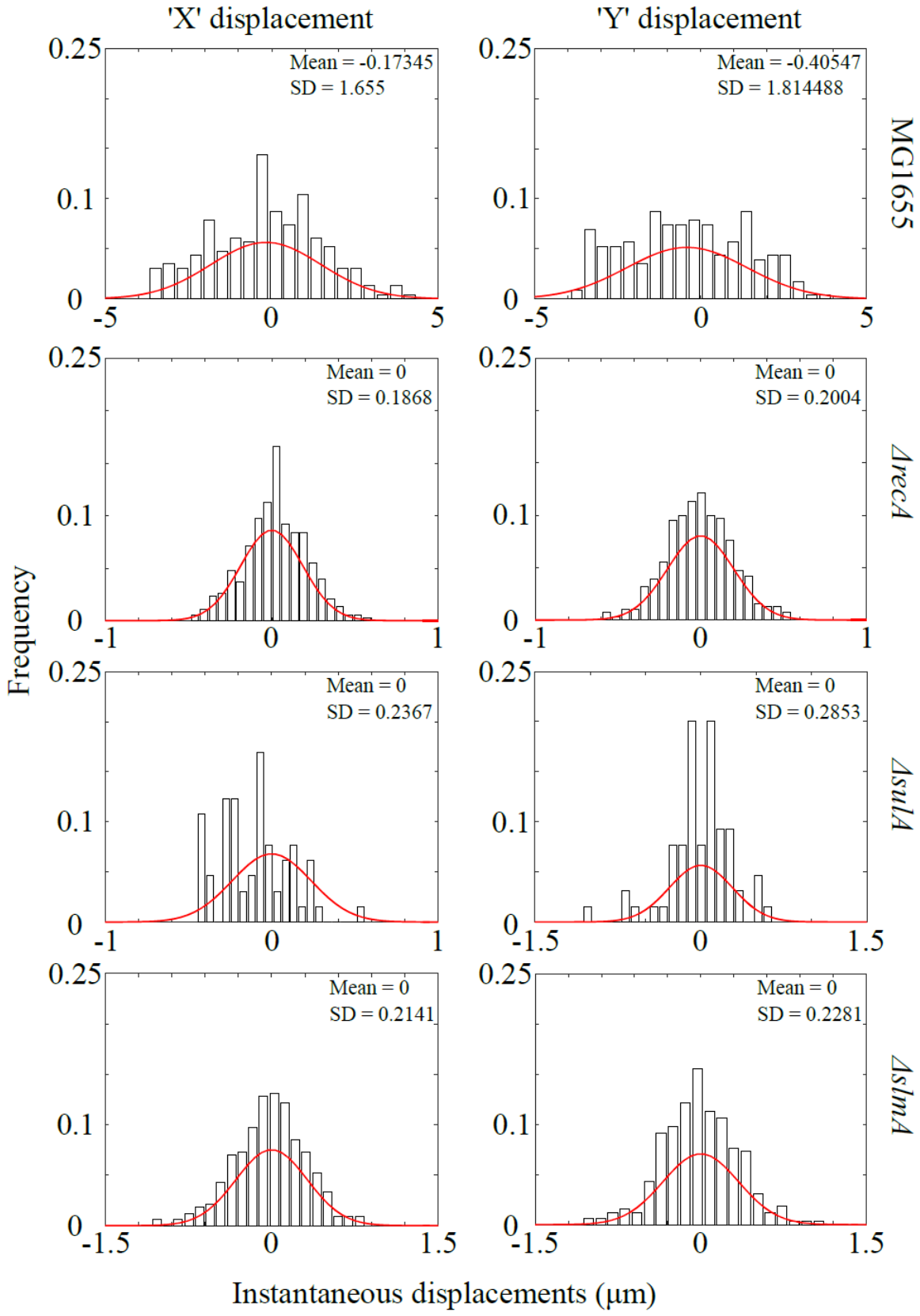
Fig. 6.2: Motility analysis of control *E. coli* strains (A) Time lapse images of beads (upper row), RP437 (middle row) and DH5α (lower row) have been shown. The rightmost column represents the trajectories of respective object overlaid on the first time frame. Scale bar- 5 µm (B) Sum normalized frequency distribution of a distance covered by bead (upper row), RP437 (middle row), DH5α (lower row) in a single step taken in X (left column) and Y (right column) direction have been indicated. Except for RP437, the Gaussian fit were calculated by fixing the average at '0 µm' for Beads and DH5α. For RP437 (motile strain) average of the fit was optimized by MATLAB algorithm (C) Diffusion coefficient of beads obtained from an experiment is compared with diffusion coefficient of an object of 1 µm size (D) X, Y plots of two dimensional movement of beads and *E. coli* RP437 and DH5α strains have been shown. Solid lines stand for no. of cells considered for the studies. Solid lines stand for no. of cells considered for the studies.

Once, we established our controls, we then, analysed motility in MG1655, *ΔrecA*, *ΔsulA* and *ΔslmA* strains. Time lapse images showed that MG1655 is a motile strain, while others were non- motile (Fig. 6.3 A). Analysis of the distributions of their displacements were in accordance with it (Fig. 6.3 B). In addition, 2 D plots of X and Y movements of MG1655, *ΔrecA*, *ΔsulA*, *ΔslmA* and their comparison with that of beads, RP437 and DH5α, also corroborated our observations. XY trajectories of non- motile strains tend to return to the same location before they were directed away from the origin, as opposed to the XY trajectories of motile strains which were directed and progressively moved away from the origin with time (Fig. 6.2 D and Fig. 6.3 C). Our observations when translated into the estimation of diffusion coefficient associated with each strain, revealed that the diffusion coefficient of MG1655 is comparable to that of RP437. All the non- motile strains have diffusion coefficient ranging between 0.3-0.5 (Fig. 6.4, Table 6.1).

A


t = 0 sec
t = 6 secs
t = 12.5 secs
Displacements
MG1655
t = 0 sec
t = 12.5 sec
t = 25 sec
ΔrecA
t = 0 sec
t = 12.5 secs
t = 25 secs
ΔsulA
t = 0 sec
t = 12.5 secs
t = 25 secs
ΔslmA

B


'X' displacement
'Y' displacement
Mean = -0.17345
SD = 1.655
Mean = -0.40547
SD = 1.814488
MG1655
Mean = 0
SD = 0.1868
Mean = 0
SD = 0.2004
ΔrecA
Mean = 0
SD = 0.2367
Mean = 0
SD = 0.2853
ΔsulA
Mean = 0
SD = 0.2141
Mean = 0
SD = 0.2281
ΔslmA
Frequency
Instantaneous displacements (μm)

C

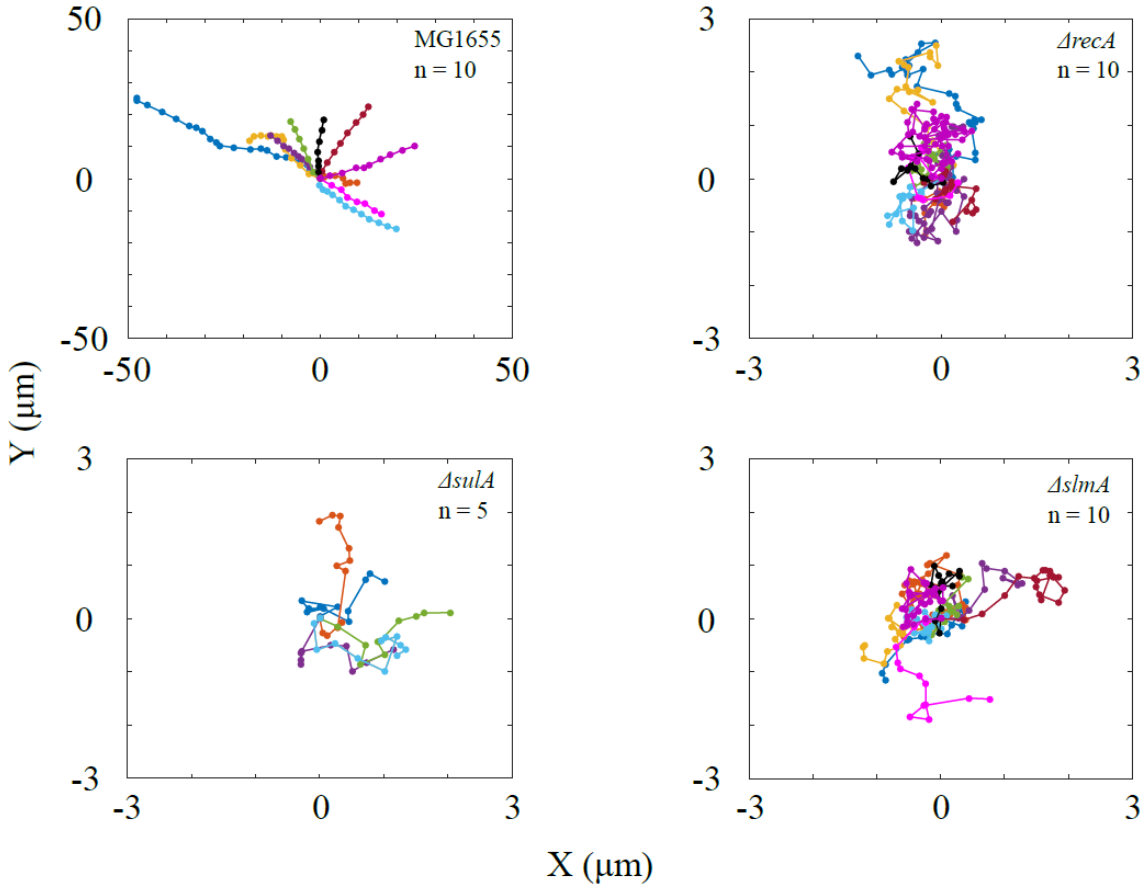


Fig. 6.3: Motility analysis of *E. coli* strains (A) Temporal presentation of movements exhibited by MG1655 (first row), *ΔrecA* (second row), *ΔsulA* (third row) and *ΔslmA* (fourth row) have been depicted. Their respective trajectories have been shown in fourth column. Scale bar- 5 µm (B) Sum normalized distribution of instantaneous displacements of MG1655, *ΔrecA*, *ΔsulA* and *ΔslmA* strains have been represented. For motile strain (i.e. MG1655), a default average value given by an algorithm was used to calculate the Gaussian fit. On the other hand, average value was fixed to 0 µm, for rest of the non- motile strains (C) X, Y plots of two dimensional movement of beads and different *E. coli* strains have been shown. Solid lines stand for no. of cells considered for the studies.

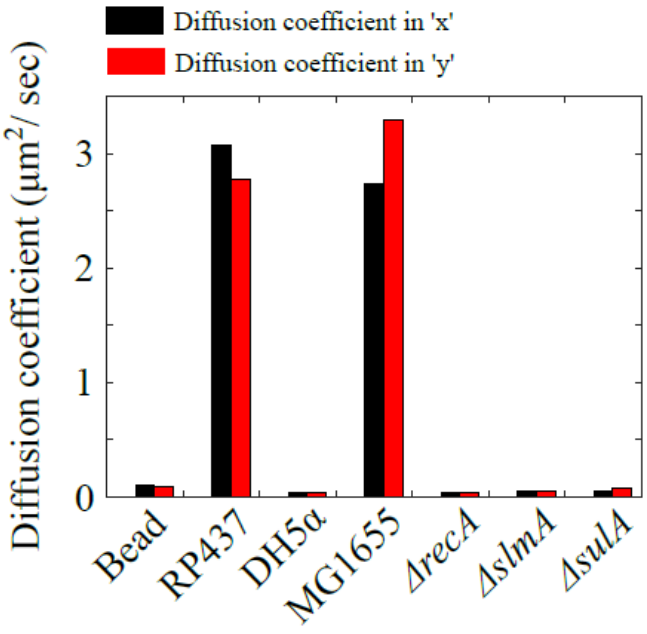


Fig. 6.4: Diffusion coefficient of beads and different *E. coli* strains.

Sr. no	***E. coli* strain**	**Diffusion coefficient in x (μm^2/ sec)**	**Diffusion coefficient in y (μm^2/ sec)**
1	Bead (Diameter = 1 µm)	0.1008	0.0931
2	RP437	3.0700	2.7674
3	DH5α	0.0351	0.0379
4	MG1655	2.7390	3.2923
5	*ΔrecA*	0.0349	0.0402
6	*ΔslmA*	0.0458	0.0520
7	*ΔsulA*	0.0560	0.0814

Table 6.1: Diffusion coefficient of different strains of *E. coli* in LB were obtained by tracking their displacements in ImageJ and testing the distribution of instantaneous displacements against Gaussian function to retrieve the statistical deviation for each strain.

Once, we determined the motility of the strains we decided the pairs of competitors as follows:

pmCherry	**peGFP**
MG1655	MG1655
DH5α	DH5α
ΔrecA	*ΔrecA*
ΔsulA	*ΔsulA*
ΔslmA	*ΔslmA*
ΔrecA	*ΔsulA*
ΔrecA	*ΔslmA*
ΔslmA	*ΔsulA*

Table 6.2: Pairs of competitors for colony competition assay.

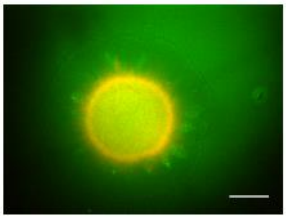
6.4 Competition between *E. coli* MG1655 (green) vs. *E. coli* MG1655 (red)

First competition was set up between MG1655 transformed with pGFP and MG1655 transformed with pmCherry. Though it is motile, analysis of *E. coli* MG1655 gave us preliminary idea of growth pattern in co- cultures of wild- type. Colony growth was followed at the end of every 12 hrs using bright field microscopy. Fluorescence was recorded in GFP and RFP channels (Fig. 6.5 A and B). Growth of the colony was linear till 36^{th} hr and then remained stationary (Fig. 6.5 C). A distinct spatial pattern of growth for two strains could be observed. Though, GFP fluorescence did not sustain for more than 24 hrs, growth of the cells expressing mCherry showed the segregation of two strains, as colony sprawled away from the homeland region (Fig. 6.5 A). Estimated red fluorescing area in the colony always resides near

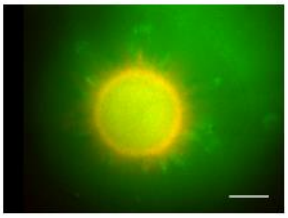
50% (Fig. 6.5 B). Our experiment showed (a) In wild- type motile bacteria, each sub- strain occupies half the colony area (b) The expression of pGFP plasmid reduces with time and hence needs modification similar to that of pmCherry.

A

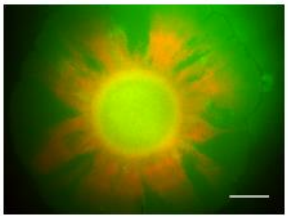
t = 12 hrs



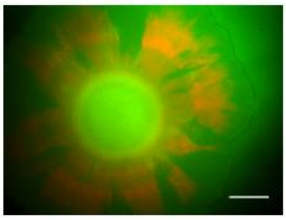
t = 24 hrs



t = 36 hrs

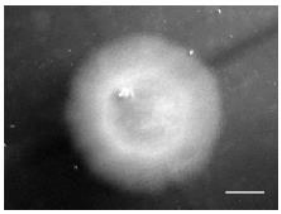


t = 48 hrs

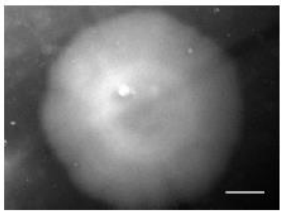


B

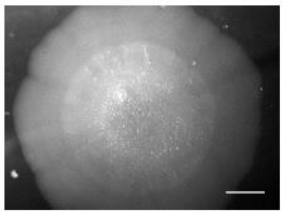
t = 12 hrs



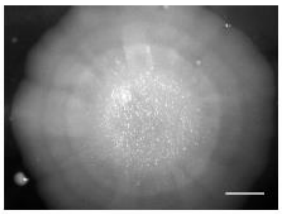
t = 24 hrs



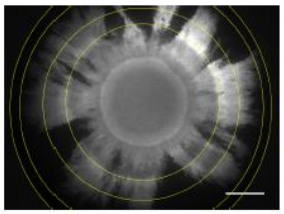
t = 36 hrs



t = 48 hrs



Colony growth



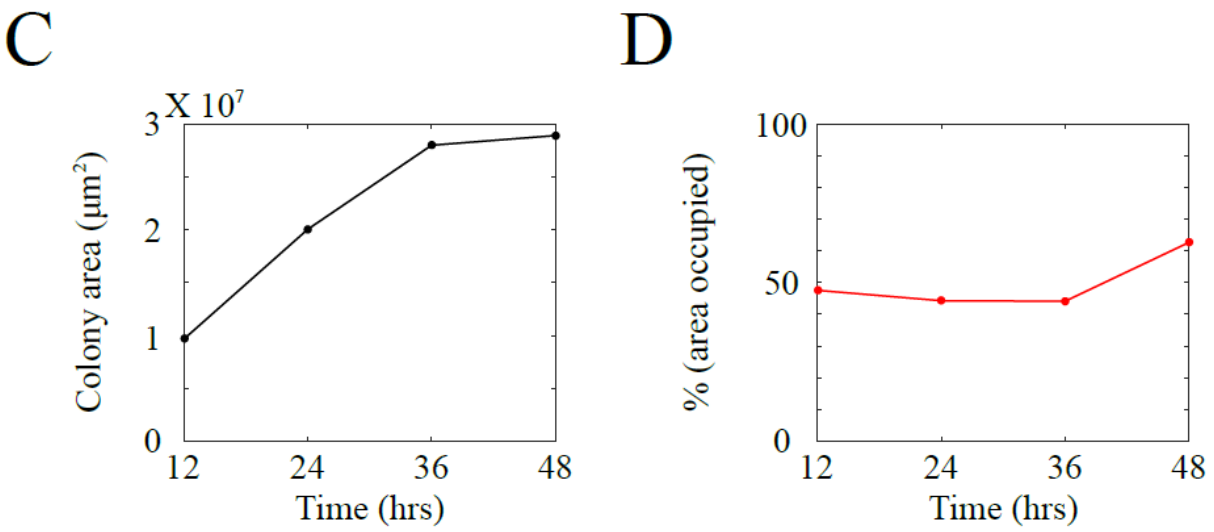


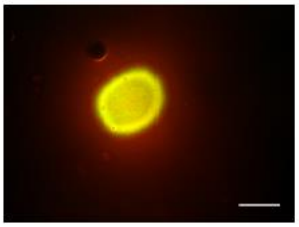
Fig. 6.5: MG1655 vs. MG1655 (A) Growth of the co- culture of *E. coli* MG1655 transformed either with pGFP or with pmCherry was followed at the interval of 12 hrs in GFP and RFP channels. Scale bar- 1 mm (B) An approximate periphery (yellow circles) of the colony at the end of each interval was extracted form corresponding bright field images at each time point (explained in section 6.2) and overlaid on the 48th hour image of mCherry expressing cells. Scale bar- 1 mm (C) Growth of the colony has been shown as a function of the time (D) Percent area occupied by cells bearing pmCherry has been plotted against time.

6.5 Competition between *E. coli* DH5α (green) vs. *E. coli* DH5α (red)

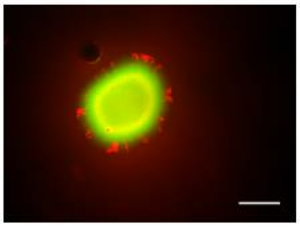
General problem using pGFP for expression of GFP in cell, was that one needs to incorporate IPTG in media, in order to activate the promoter. That puts temporal limitations on the expression as well as on experiment. On the contrary, modification in pGFP in order to convert it into pmCherry plasmid, made mCherry expression constitutive, as major portion of operator element was removed in the process (Appendix B, Section B.1). We used pmCherry backbone to ensure constitutive expression of GFP beyond 12 hrs (Fig. 6.6 A). mCherry gene was substituted with that of eGFP. New plasmid was named as peGFP (Appendix B, Section B.1). *E. coli* DH5α cells were transformed either with pmCherry or with peGFP plasmids and mixed together in 50 : 50 proportion in their log phase. Linear growth of the colony from homeland region was observed at every12 hrs (Fig. 6.6 A, B and C). At initial time points, we observed segregation in two colours. However, as time progressed, GFP signal took over almost all the area in colony, while growth of the cells bearing pmCherry was compromised (Fig. 6.6 A). In addition to this, we also noticed the spread of GFP expressing cells in Z- direction, instead of the growth restrained only in XY plane. Experimental outcomes were reflected in the percent area calculated for both the sub- strain. Percent area covered by green cells (green solid line) was near to 100%, while percent area for red cells (red solid line) decreased linearly with time (Fig. 6.6 D).

A

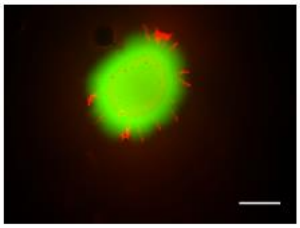
t = 12 hrs



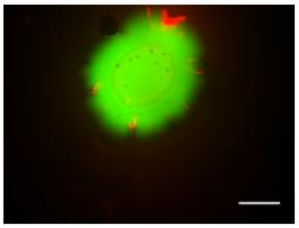
t = 24 hrs



t = 36 hrs

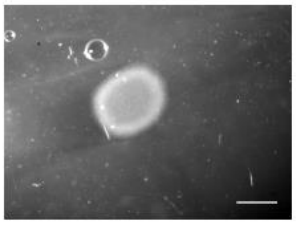


t = 48 hrs

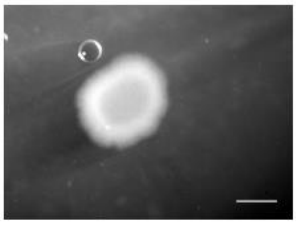


B

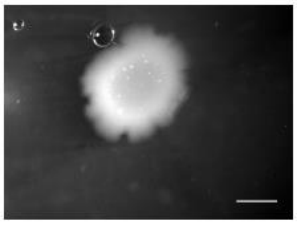
t = 12 hrs



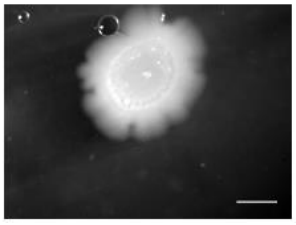
t = 24 hrs



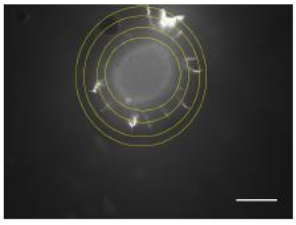
t = 36 hrs



t = 48 hrs



Colony growth



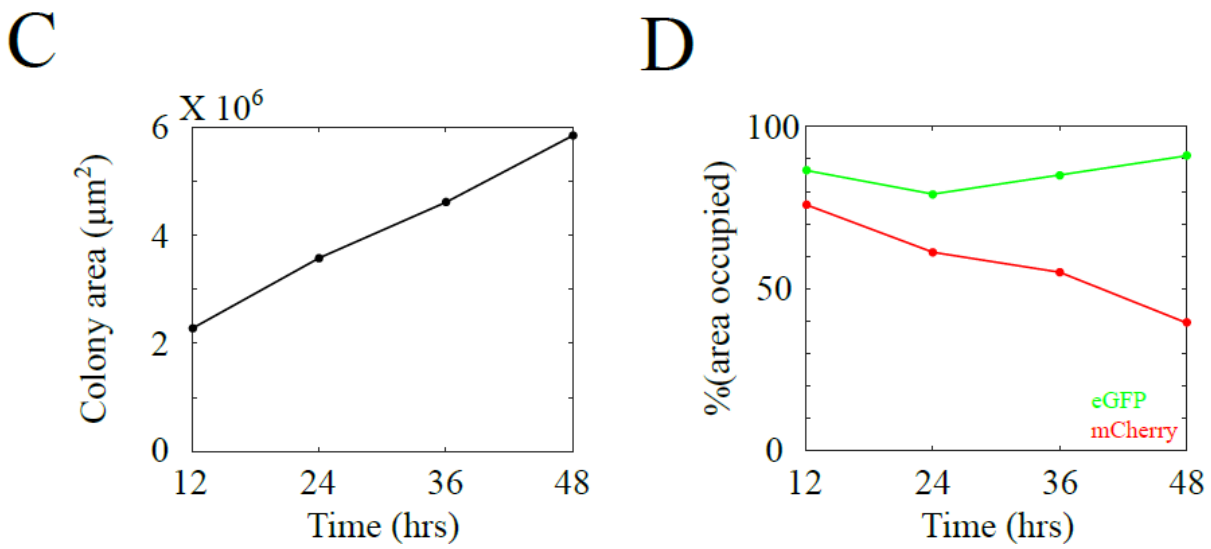


Fig. 6.6: DH5α vs.DH5α (A) Mixed culture of DH5α expressing either eGFP or mCherry was observed at the interval of 12 hrs. Scale bar- 1 mm (B) An image of *E. coli* cells transformed with pmCherry taken at the end of 48th hour was overlaid with outlines of the same colony extracted from bright field images (explained in section 6.2) at earlier time points. Scale bar- 1 mm (C) Colony growth has been plotted as a function of the time (D) Proportion of area covered either by eGFP signal (green solid line) or by mCherry signal (red solid line) has been plotted as a function of the time.

6.6 Discussion

Shape of an organism patterns the colony or the biofilm in its niche[135]. In an experiment carried out by Hallatschek et al, in 2007, it was revealed that the sectoring of two sub- strains of the populations bearing two different colours can be traced back to the arrangement of the two cells in the home- land region[133]. Crux of our experiment was to analyse an influence of population cell size variation on the arrangement of a single cell within homeland region, which can affect the patterning of the colony. We planned to change the cell length distribution of an

inoculum by mixing the two *E. coli* strains with different genetic background, such that the homeland region would be noisier as far as cell lengths are considered. Experiment was carried out in two successive steps. In the first part, we determined the motility of *E. coli* strains with an objective of deciding the pair for competition, as non- motile strain competing against motile strain involves a risk of motile strain outgrowing non- motile partner. Our analysis showed us that MG1655, also a wild- type strain, is motile, while other strains like *ΔrecA*, *ΔsulA* and *ΔslmA* are non- motile. Distributions of instantaneous displacements by *E. coli* DH5α have modal value near 0 μm. We attribute this to the no. of cells that have been tracked over time. Only six DH5α cells were analysed. In addition, it is evident from two- dimensional XY plots that DH5α cells wandered at the same place, before they showed a movement away from their original location. We could find some of the cells taking small steps during the motion, which is why instantaneous displacements reside near 0 μm. This behaviour was not observed for rest of the non- motile strains like *E. coli ΔsulA*. This observation can further be strengthen by analysing the movement of more *E. coli* DH5α cells.

Second step of the experiment, however, suffered from a drawback of not having non- motile, wild- type population. Spatial spread of motile cells is different from non- motile populations. That changes the construction of colonies of respective strains. Hence, motile MG1655 cells does not provide with valid control to analyse the colonies formed by other non- motile strains under identical experimental conditions. Also, the constitutive expression of mCherry protein in high copies in mutant cells like DH5α cost the cell its fitness. Growth of DH5α transformed with pmCherry in suspension was slower than DH5α cells bearing peGFP. We predict that the observed outgrowth of eGFP expressing cells over pmCherry transformed cells can be attributed to the slower growth rate of mCherry expressing population. We propose that growth of DH5α with pmCherry plasmid in batch cultures needs to be recorded systematically and

compare with growth in parent strain (DH5α) as well as DH5α with peGFP in order to confirm the growth defects.

In future experiments, mCherry can be replaced with other variant of fluorescence proteins viz. CFP or RFP so that balanced and healthy growth of both the competitors can be achieved in the proliferating colony. Use of microfluidic experiments, can allow to subject the community to different environments sequentially, thereby mimicking the natural habitat an organism.

In conclusion, though preliminary, these studies open opportunity to fathom the relationship between shape variations and the communal behaviour in lower organisms.

Chapter 7

Conclusion and Outlook

Classical microbiology describes *E. coli* cell growth as bidirectional and symmetric across the mid- plane[10]. Plane of division exactly coincides with the mid- cell, giving rise to the twins with identical genetic as well as phenotypic make- up. However, it fails to explain the variation in the external appearance or the occasional changes in the behaviour of a single cell in an isogenic population[136]. Nevertheless, Phenotypic heterogeneity in the clonal population has been shown to reckon with cell differentiation and formation of persister population[137,138]. Non-genetic fluctuations in phenotype creates sub- populations that co- exist together and exhibit a differential response to the surroundings[139].

In previous studies, stochastic fluctuations in the genetic circuits of an organism have been implicated in the genesis of phenotypic variations[43,140]. In addition to this, errors in the segregation of sub- cellular contents in twin sisters can serve as the source of phenotypic heterogeneity[45,46]. Thesis reviews the extension in an *E. coli* cell as well as the mechanisms that cause fluctuations in it. Our approach extends from a single cell to the clonal populations of *E. coli*.

Our work connected stochastic halts in replication process with population cell length variation in growth rate dependent manner. Analysis of micro- colonies generated from single mother as well as study of cell sizes in continuous cultures showed that cell length variability is independent of the cell age. We also provided with the molecular links. Our analysis of RecA dynamics in *E. coli* before and after HU treatment showed us that stochastic temporary suspensions in DNA replication pauses cell division through SOS response pathway. Cell that skips the division cycle, elongates into the filamentous form. We propose that the random occurrence of these events in an individual lead to the skewed cell length distribution extending

beyond the two- fold of average cell length of *E. coli*. We further, investigated the relationship between replication stochasticity with gene expression noise. Correlation between intrinsic noise and the extent of replication arrest could be observed at sub- lethal dosage of hydroxyurea. The relationship was breached after hydroxyurea concentration was increased beyond 13 mM. We also noticed the effect of the growth rate on the noise components of gene expression. Fast growing cell exhibit high levels of intrinsic noise, while in slow growing cultures extrinsic noise dominates the genetic network. We concluded that stochastic replication process can also generate phenotypic variations in an isogenic population by elevating intrinsic noise levels in genetic circuitary.

When we tested the growth of an individual cell in wild- type population, asymmetry in the growth distribution across the cell length was observed. We confirmed asymmetry with help of various assays. Asymmetry was found to be propagated through consecutive generations and follows a peculiar pattern. Asymmetry generates the time differences in the division cycle of the two identical sisters. More importantly, it was revealed from our FRAP studies that bacterial cytoskeleton (MreB) may be instrumental in generating a bias in the growth of an individual cell. However, we predict that it does not contribute to the population cell length variation.

As a future step, studies can be extended into quantifying an effect of phenotypic variability on a microbial community. Colony competition assay could serve the same purpose. Higher occurrence of cell filamentation in the ‘homeland region’ can affect the arrangement, which ultimately will change the boundaries of the sectors.

In addition to this, our studies can be extended to understand the fate of the filamentous cell. Division in naturally occurring elongated cells can be followed spatio- temporally. Molecular tags to proteins relevant in cell division or nucleoid segregation can help not only in

understanding the future of an elongated cell but also in finding the probability of filamentation in an offspring.

Filamentation can be investigated with reference to the aging of the cell. The studies can give us insight into whether there is any correlation between segregation of cellular damage and probability of filamentation, at a single cell level. Moreover, in our experiment, only a certain proportion of the cells in the population respond to the drug treatment, while others continue their division cycle and propagate. This observation can be supported with aging studies or by following the pattern that corners the cellular damages to understand the generation and the importance of phenotypic heterogeneity. In addition, the elongation pattern of a single cell can be extrapolated to study its interaction with the neighbouring cells and the arrangement of their successors either in the colony or in the biofilm.

Appendix

A. Selection of growth media

Propagation rate of *E. coli* cells were modulated by changing the nutrient supplement. Though minimal medium supported with different carbon source was the choice, the doubling time achieved was below 60 mins. In order to attain the growth rate that necessitates the cell to fire replication multiple times, following media compositions were considered:

(a) Decomposition of Luria- Bertani broth into two different broths

(b) Supplement M9 medium with variable glucose concentrations as a sole carbon source

A.1 Deriving media from Luria- Bertani medium

Luria- Bertani is a nutrient rich complex medium which has been designed for optimized laboratory growth of bacterial cultures as follows (Hi- media laboratories, M1245):

Ingredient	**Grams/ Litre**
Yeast extract	5
Tryptone	10
Sodium chloride	10
Final pH (at 25°C)	7.5 ± 0.2

Table A.1: Composition of Luria- Bertani broth (cited from Hi- Media laboratories, M1245).

Doubling time of *E. coli* MG1655 in LB at 37°C with constant aeration at 180 rpm has been shown to be around 30 mins (Chapter 3, Table 3.1 and 3.2). We proposed to segregate yeast extract and tryptone so that the nutrient value of the medium will be lowered affecting the

growth rate of the population. We postulated following two compositions and named them as yeast extract broth (Table A.2) and tryptone broth (Table A.3).

Ingredient	**Grams/ Litre**
Yeast extract	5
Sodium chloride	10
Final pH (at 25°C)	7.5 ± 0.2

Table A.2: Composition of yeast extract broth.

Ingredient	**Grams/ Litre**
Tryptone	10
Sodium chloride	10
Final pH (at 25°C)	7.5 ± 0.2

Table A.3: Composition of tryptone broth.

Doubling time of *E. coli* MG1655 in these two media was tested using plate reader assay. We used 96 well plate (UV transparent, flat bottom, Costar, Corning, USA) that can accommodate maximum volume of 400 µl per well. In order to avoid the media spill out of the well during the course of experiment, we fixed the volume of the system to 200 µl. The media was inoculated with *E. coli* MG1655 at 1% concentration. Evaporation of the medium was avoided by covering its surface with 20 µl of sterile mineral oil (Sigma- Aldrich, USA). Growth was measured for the culture at 37°C with constant shaking at 600 rpm, at the interval of 10 mins for 3 hrs till the saturation was observed. Growth was monitored at 600 nm, using plate reader (Varioskan, Thermo Fisher Scientific, USA). Growth in yeast extract broth (YEB) and tryptone broth (TB) was compared against the growth in LB. Temporal changes in the culture density

were fit to logistic equation (Fig. A.1) to extract the growth rate of the population in YEB as well as in TB (Table A.4). Cells grown in YEB and TB doubled with growth rate of 1.2324 and 0.8634 generations per hour respectively (Table A.4). Growth rate is surely less than that in LB and triggers the multi- fork replication in the cell.

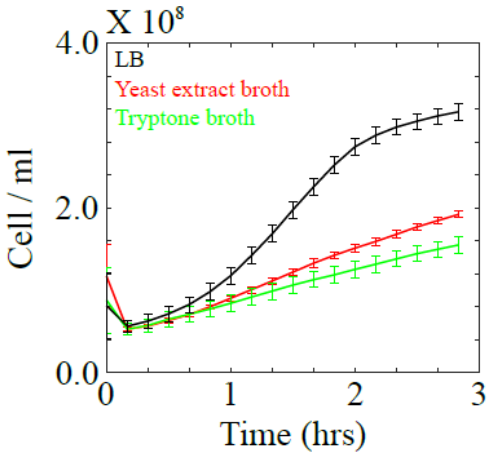


Fig. A.1: Growth measurement of *E. coli* MG1655 in LB, yeast extract broth and tryptone broth respectively.

Medium	**Carrying capacity**	**Growth rate (generation/ hr)**	**Doubling time (mins)**
LB	0.4347	1.719	34.9
Yeast Extract	0.2518	1.2324	48.7
Tryptone	0.2460	0.8634	69.5

Table A.4: Estimated *E. coli* population growth rates in yeast extract broth and tryptone broth.

In bulk experiments, when system was scaled up to 100 ml, growth rate was observed to increase further and doubling time reduced to 39 and 57 mins when the cells are cultured in yeast extract and tryptone broth respectively (Chapter 3, Table 3.2). This difference can be attributed to the small system as well as the presence of mineral oil layer that may cut off the

oxygen, thus limiting its availability. Small volume, moreover, also amounts for decreased carrying capacity. However, the experiment confirmed that the use of yeast extract or tryptone as a sole source of energy can support the microbial growth, with the growth rates in the regime of multi- fork replication. Hence, both the media were then employed in the growth modulation experiments.

A. 2 Supplementing minimal medium with different concentrations of glucose

In addition to our attempt of devising new media, we also tweaked the percent glucose supplement to minimal media. Concentration of glucose was ranged from 0.1- 1%. While, 4 µg/ ml of thymidine was added to compensate for the mutation in pyrimidine synthesising pathway of *E. coli* MG1655. Cultures were grown and monitored in exactly the same way, described in earlier section. Growth rate (Table A.5) was retrieved for each of the nutrient condition by fitting the growth with logistic function (Fig. A.2 A). We observed no significant change in the growth rates when cultures were supported with different concentrations of glucose (Fig. A.2 B). It was in the range of 0.4- 0.5 generations per hour.

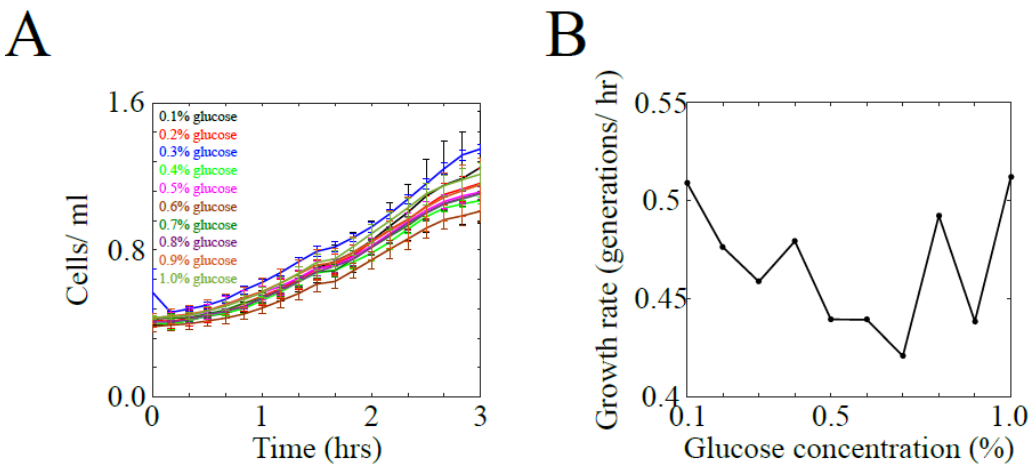


Fig. A.2: *E. coli* growth in minimal media containing different concentrations of glucose. (A) Growth measurements of *E. coli* MG1655 at different glucose concentrations have been plotted as a function of time (B) Population growth rates at different glucose concentrations have been shown.

Medium	**Carrying capacity**	**Growth rate (generation/ hr)**	**Doubling time (mins)**
M9 + 0.1% glucose	0.5974	0.50922	117.8
M9 + 0.2% glucose	0.5218	0.47622	126
M9 + 0.3% glucose	0.8266	0.45864	130.8
M9 + 0.4% glucose	0.3885	0.47928	125.2
M9 + 0.5% glucose	0.5011	0.43932	136.6
M9 + 0.6% glucose	0.5315	0.43922	136.6
M9 + 0.7% glucose	0.6977	0.4206	142.7
M9 + 0.8% glucose	0.3961	0.4923	121.9
M9 + 0.9% glucose	0.5515	0.4383	136.9
M9 + 1.0% glucose	0.4234	0.51204	117.2

Table A.5: Growth rates of *E. coli* MG1655 in minimal media supplemented with different glucose concentrations.

Since, the growth of *E. coli* in yeast extract broth or tryptone broth distinctly changes the doubling time as compared to M9 supplemented with various concentrations of glucose we used YEB and TB to alter the population growth rates and observe the effect on the cell length variability.

B. Construction of Plasmids

Following plasmids were constructed for the purpose of these studies:

1. pmCherry

2. peGFP

3. pRecA- mCherry

4. pBAD24-recA

B.1 Construction of pmCherry and peGFP plasmids:

Plasmids were constructed in order to express either GFP or mCherry proteins in *E. coli* cells and make the cell visible while imaging. Plasmids were derived from a commercial pGFP plasmid (Clontech, USA)[141–144], with PUC19 origin and ampicillin resistance gene (beta lactamase, bla) as a marker. GFP expression is controlled by P_{lac}, located 110 bp upstream to the ORF of *gfp*. The plasmid has two 'Multiple Cloning Sites' at N- as well as C- termini of the *gfp* gene. To repress the expression of the gene O1 and O2 sites have been introduced to downstream and upstream of the lac promoter, respectively. While, 5' MCS site acts as a putative O3 operator sequence.

pmCherry plasmid was constructed by replacing *gfp* gene with that of *mCherry*. The later was amplified from pmCherry- N1 plasmid (Clontech, USA). The following primers were used in Taq polymerase (Bangalore Genei, India) mediated PCR (Table B.1).

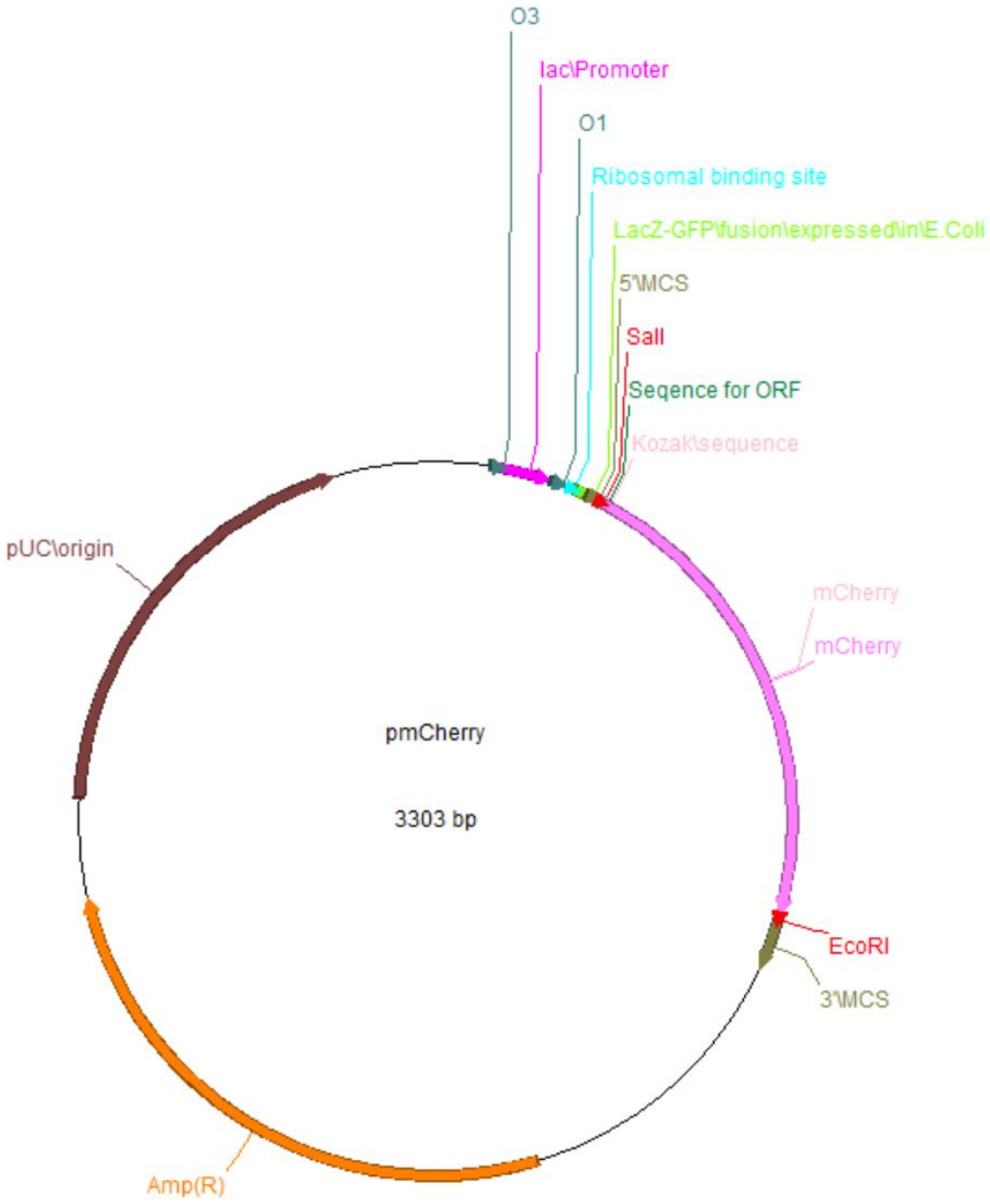
Primer	Orientation	RE site	Sequence (5'- 3')
mCherry_f	Forward	SalI	AAAGTCGACAAAAATGGTGAGCAAGGGC
mCherry_r	Reverse	EcoRI	CGGGAATTCCTACTTGTACAGCTCG

Table B.1: Sequence of the primers used to amplify pmCherrry and peGFP genes in 5'- 3' direction.

Purified amplicon and the pGFP plasmid were digested sequentially by EcoRI and SalI (NEB, USA) at 37°C for 5 hrs each. Digested fragments were gel extracted (QIAquick Gel Extraction Kit, Qiagen, USA) and ligated together in a system containing vector : insert ratio equal to 1 : 6, using T4 DNA ligase (Bangalore Genei, India). Ligation mixture was transformed into *E. coli* DH5α competent cells and scored for the ampicillin resistance on agar plates[145].

Initial screening of the colonies involved growth in LB for 2 hrs, followed by induction with 1 mM IPTG at 37ºC and 180 rpm, though later on it was noticed that mCherry expression in DH5α cells was constitutive or independent of IPTG induction. Cells were harvested, fixed and observed under upright epifluorescence microscope (Carl Zeiss, Germany) for fluorescence using dsRed filter (Fig. B.1 A and B).

A


O3
lac\Promoter
O1
Ribosomal binding site
LacZ-GFP\fusion\expressed\in\E.Coli
5'\MCS
SalI
Seqence for ORF
Kozak\sequence
pUC\origin
mCherry
mCherry
pmCherry
3303 bp
EcoRI
3'\MCS
Amp(R)

B

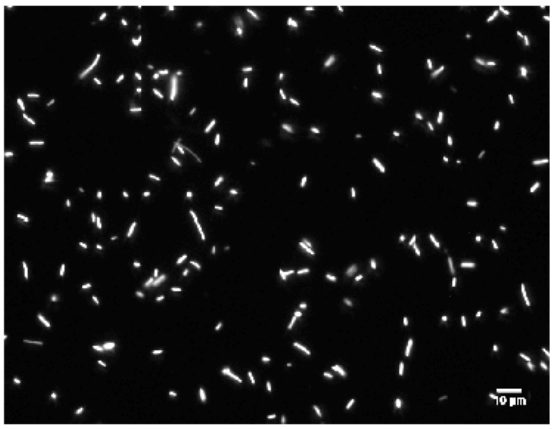
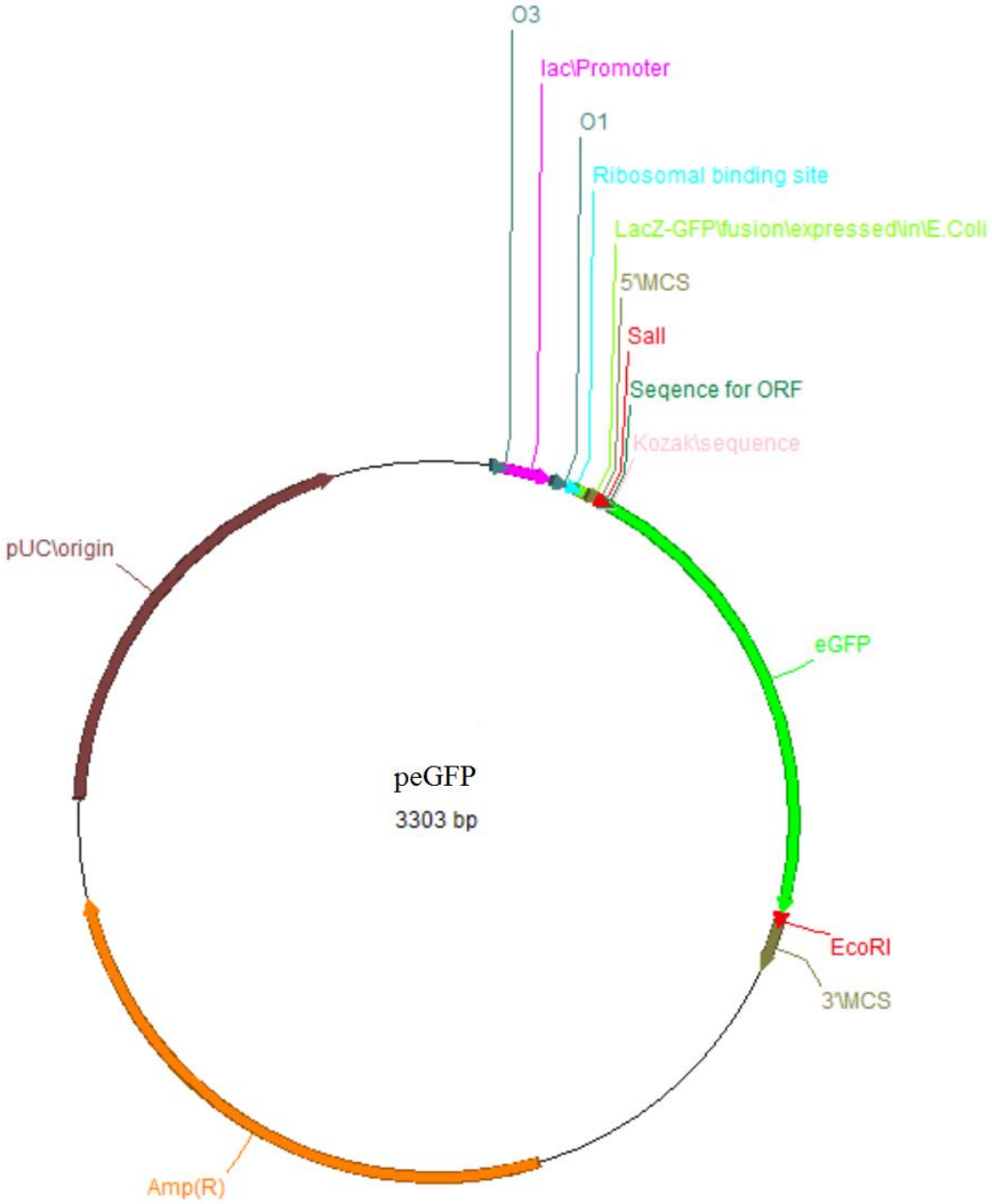


Fig. B.1: pmCherry construction (A) Plasmid map of pmCherry (B) *E. coli* DH5α cells, induced with 1 mM IPTG, express mCherry. Scale bar- 10 µm.

In the making of peGFP plasmid, the ORF of mcherry was deleted from pmCherry plasmid and was replaced with *eGFP* ORF. eGFP gene was amplified using mcherry primers (Table B.1) from a commercial plasmid peGFP- N1 plasmid (Clontech, USA). pmCherry plasmid and eGFP gene was digested sequentially with EcoRI and SalI at 37°C, for 5 hrs each, and then ligated in the ratio of 1 : 6. Cells transformed with ligation mixture were selected for ampicillin resistant. Plasmid was confirmed by observing the fluorescence in the transformed cells when excited at 488 nm under microscope (Fig. B.2 A and B).

A


O3
lac\Promoter
O1
Ribosomal binding site
LacZ-GFP\fusion\expressed\in\E.Coli
5'\MCS
SalI
Seqence for ORF
Kozak\sequence
pUC\origin
eGFP
peGFP
3303 bp
EcoRI
3'\MCS
Amp(R)

B

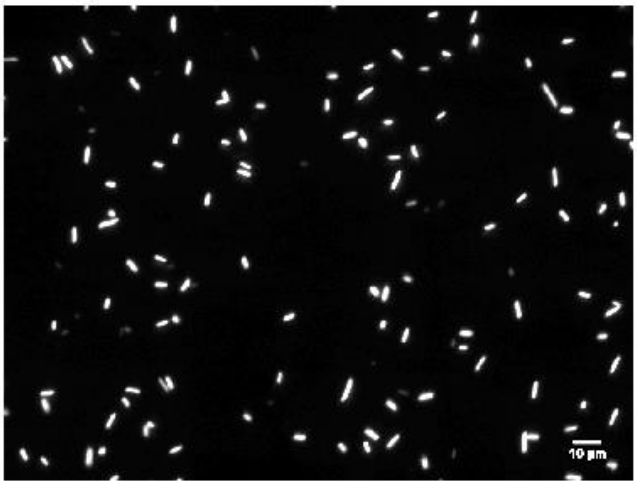


Fig. B.2: peGFP construction (A) plasmid map of peGFP (B) *E. coli* DH5α cells flooded with eGFP. Scale bar- 10 µm.

B.2: Construction of pRecA- mCherry plasmid:

Another modification in pmCherry plasmid involves an introduction of *recA* gene at the N-terminal of mCherry gene. *recA* gene was amplified from genomic DNA of *E. coli* MG1655[146] using these two primers, with Taq polymerase (Table B.2).

Primer	**Orientation**	**RE site**	**Sequence (5'- 3')**
RecA-mcherry_f	Forward	HindIII	GCTAAGCTTATGGCTATCGACGAAAA
RecA-mcherry_r	Reverse	SalI	GCAGTCGACATAAAATCTTCGTTAGTTTCT

Table B.2: Sequence of the primers used to amplify *recA* gene in 5' to 3' direction.

Amplicon and plasmid, both were digested sequentially with HindIII (NEB, USA) and SalI at 37°C, 5 hrs each. Digested vector and insert were ligated together in the ratio 1 : 6. Transformed colonies were scored for ampicillin survival and confirmed by visualizing induced cells under dsRed filter for RecA foci along the cell length (Fig. B.3).

A

B

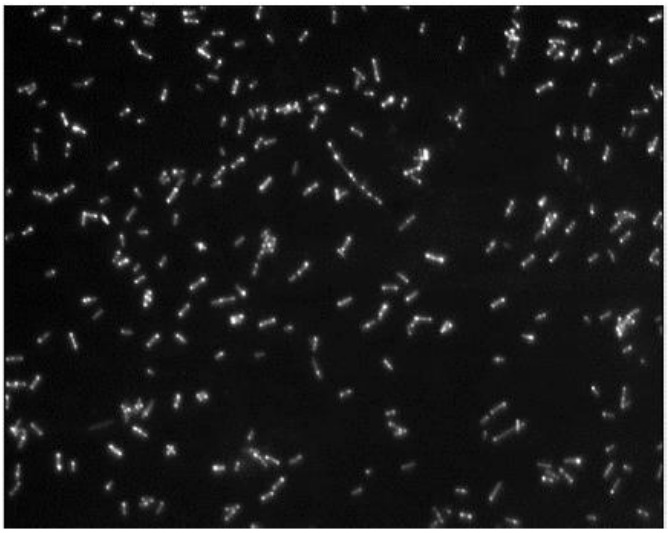


Fig. B.3: pRecA- mcherry construction (A) Plasmid map of pRecA- mcherry (B) *E. coli* DH5α cells expressing RecA- mcherry.

Expression of pmCherry plasmid or other plasmid derived from pmCherry was constitutive and independent of IPTG induction. We attribute this to (a) Destruction of putative O3 site due to digestion at SalI. This might result into reduced suppression by lac operator protein and hence increase in the promoter activity. (b) Reduction in the distance between the promoter seqΔuence and the ORF from 110 bp to 82 bp due to the restriction digestion.

B.3 Construction of pBAD24-recA plasmid:

recA amplification from *E. coli* genomic DNA was carried out using following two primers (Table B.3).

Primer	Orientation	RE site	Sequence (5'- 3')
pBAD24-recA_f	Forward	NheI	AAA GCT AGC ATG GCT ATC GAC GAA AA
pBAD24-recA_r	Reverse	XbaI	GCC TCT AGA TTA AAA ATC TTC GTT AGT TT

Table B.3: Sequence of the primers used to amplify *recA* gene in 5' to 3' direction.

pBAD24 plasmid was cut open by digesting it with NheI and XbaI at 37°C, 3 hrs. Similar treatment digested *recA* gene. The insert and the vector was mixed in the proportion of 6 parts to 1 part. Colonies were selected using LB agar containing ampicillin (Fig. B.4). Colony PCR served as the confirmation for selected transformed colonies. Following primers were used for colony PCR (Table B.4).

Primer	Orientation	Sequence (5'- 3')
pBAD24- f	Forward	GAC GCT TTT TAT CGC AAC TCT CTA CTG TTT CT
pBAD24-recA_r	Reverse	GCC TCT AGA TTA AAA ATC TTC GTT AGT TT

Table B.4: Sequence of the primers used for the colony PCR in 5' to 3'direction.

A

Fig. B.4: pBAD24-recA construction. Plasmid map of pBAD24- recA.

C. Development of ImageJ macros for analysing fluorescent puncta in cellular halves

Macro was originally programmed to analyse the sum area occupied by MreB foci in each half of an *E. coli* cell. Macro operates in three steps.

DIC image of an *E. coli* cell as well as the corresponding image of MreB puncta in YFP channels are opened in ImageJ

Part I: Divide cell into two halves

Macro selects the DIC image window, process it to obtain the coordinates of the two halves of the cell.

1. Image is enhanced contrast (0.4%, normalization)

2. Filter: '*Unsharp mask*'[147] is applied twice; Parameters: radius – 5 and mask – 0.9

3. Image is then thresholded using 'Otsu method'[148]. The step helps extract cell shape and area. The later parameter is estimated using '*Analyse particle*'[149]

4. Image undergoes post- threshold processing in following sequence: (i) Fill holes (ii) Erode (iii) Dilate

5. Image is then skeletonized

6. Line joining the end points of the skeleton is drawn

7. Perpendicular bisector for this line is calculated in order to divide the skeleton into two equal parts

8. Additional six coordinates are calculated on each sides of perpendicular bisector using its slope

9. Two different polygons covering each half of the cell were synthesized and their ROI were saved and overlapped on YFP image

Part II: Overlay of selected ROI on fluorescence image

Macro then selects YFP window in order to process it as well as overlay it with ROI obtained from DIC image.

1. Filter '*Unsharp mask*' is applied on YFP image. Parameters: radius – 5, mask – 0.9

2. Otsu thresholding

3. Overlay ROI separately and duplicate the area

4. Select each duplicated image window separately and apply '*Analyse particle*' to detect and quantify the area of MreB focus

Part III: Evaluation of area

1. Area of the particles detected in each half is added together and divided by the area of the cell

2. Numbers are displayed in the result window

D. Source codes

Study uses various algorithms in order to make image analysis as well as statistical interpretation of data, more objective and accurate. Following sub- sections provide source codes for the programs developed during the course of these studies, either in MATLAB or in ImageJ.

D.1 Fitting Log- normal function to the population cell length distribution

The code, designed in MATLAB, fits the cell length distributions of *E. coli* populations to log-normal function to return the values for average cell length and variance in the distribution (Chapter 3) and then statistically validates it by calculating goodness- of- fit. For the later part, an algorithm employs a non- parametric 'Kolmogorov- Smirnov statistical test' (Chapter 2, Section 2.15)[57]. Code accepts text files and displays the figure in which distribution fit to log-normal equation.

File name: LognormalFitMG.m

```
%Manasi Gangan, IISER Pune
%20140927: Modified Chaitanya Athale, IISER Pune

close all; clear all;
%read in the data
a=importdata('FileName.txt');
N = length(a)%no. of samples
%plot the experimental data based on a histogram
x=linspace(1,15,44); %x-vals
%experimental data freq. distr
[p,q]=hist(a,x);
p1=p./max(p); %max normed
p2=p./sum(p); %sum normed
%figure(1),hold on, subplot(1,2,1), bar(q,p1,'b'),title('max normed')
figure(1),bar(q,p2,'b')%,title('sum normed')


%== FIT
%create an x-axis range of lengths for the fit based on the min-max of
data
b=linspace(min(a),max(a),500);
%fit a lognormal distribution
pd=fitdist(a,'lognormal');
```

```
%generate the pdf of the fit
y=lognpdf(q,pd.mu,pd.sigma);%plot the pdf as a function of the x-vals
y2= y./sum(y);%sum normed (area under curve = 1)
%extract the logn parameters (mu, var)
[M,V]=lognstat(pd.mu,pd.sigma)


%figure(1), hold on, subplot(1,2,1),plot(q,y,'-
r'),legend('Expt','Fit'),
%xlabel('length'),ylabel('freq');

figure(1), hold on, plot(q,y2,'-r')%,legend('Expt','Fit'),
%xlabel('length'),ylabel('freq');
xlim([0,10])
ylim([0,0.55])

%cumulative sum used to calculate chi-square statistic
p3=cumsum(p2);
%figure(2), plot(q,p3,'-ob');% O: observed

%cumulative sum of the fit
y3=cumsum(y2);
%hold on, plot(q,y3,'-xr');% E: expected
%legend('Observed','Expected')
%text(4,1.2,sprintf('No. of observed classes: %i',length(q)));
%here n=no. of classes
%hence: df = n - 1

%chi-square statistic
SSD = sum((((p3-y3).*N).^2)./(N.*y3));
SSD;
df = length(q)-1; %degrees of freedom


%KS- test statistics
F1=N*p3;
F2=N*y3;

d=F1-F2;
max(d)

D=sqrt((-log(0.001/2))/2*N)
```

D.2 Finding mid- plane of the cell

Macro has been developed in ImageJ, in order to divide the cell image at the mid- point of the cell and estimate the total area occupied by the molecule of interest in either halves of the cell. Macro also extracts the values for cell length as well as cell area (Chapter 5 and Appendix C). Macro is applied to the copped images of the cell in DIC and YFP channels. It first processes DIC image to return the cellular dimensions and then relates the derived ROIs with corresponding fluorescence image of the cell (Appendix C).

File name: CellMIdPlaneDetection.ijm

```
//This macro divides the cell in two halves and counts area of the
particles in each half.
//Requires plugin "Measure ROI" to be installed.
//Requires v1.5a.
//Developed by Manasi Gangan, IISER Pune on 2016/01/09.

//Extracts cell shape.
selectWindow("DIC_ImageName.tif");
getDimensions(width, height, channels, slices, frames);
run("Enhance Contrast...", "saturated=0.4 normalize");
run("Unsharp Mask...", "radius=5 mask=0.90");
run("Unsharp Mask...", "radius=5 mask=0.90");
setAutoThreshold("Otsu dark");
//run("Threshold...");
//setThreshold(128, 255);
setOption("BlackBackground", false);
run("Convert to Mask");
run("Convert to Mask");
run("Fill Holes");
run("Erode");
run("Dilate");
//run("Erode");
//run("Dilate");
//run("Erode");
//run("Dilate");
run("Analyze Particles...", "size=750 pixel add");

//finds width, length and area of the cell and divides it into two
halves.
newImage("Output_1", "8-bit black", width, height, 1);
selectWindow("Output_1");
roiManager("Select", 0);
run("Invert");
run("Make Binary");
run("Median...", "radius=5");
run("Skeletonize");
run("Analyze Particles...", "size=0 add");
```

```
roiManager("Select", 1);
List.setMeasurements;
  x1 = List.getValue("FeretX");
  y1 = List.getValue("FeretY");
  length = List.getValue("Feret");
  degrees = List.getValue("FeretAngle");
//print(degrees);
  if (degrees>90)
     degrees -= 180;
  angle = degrees*PI/180;
  x2 = x1 + cos(angle)*length;
  y2 = y1 - sin(angle)*length;
m1= (y1- y2)/(x1- x2);
m2 = -1/ m1;
m3 = atan(m2);
m4 = atan(m1);

roiManager("Select",0);
run("Measure Roi");
w = getResult("Roi_Width", 0);
l = getResult("Roi_Length", 0);
//print(w,l);
selectWindow("Results");
run("Close");

List.setMeasurements;
  //print(List.getList); // list all measurements;
  xC = List.getValue("X");
  yC = List.getValue("Y");

x3 = xC + (w/2)*cos(m3);
y3 = yC + (w/2)*sin(m3);
x4 = xC - (w/2)*cos(m3);
y4 = yC - (w/2)*sin(m3);

x5 = x3 - (l/2)*cos(m4);
y5 = y3 - (l/2)*sin(m4);
x6 = x4 - (l/2)*cos(m4);
y6 = y4 - (l/2)*sin(m4);

x7 = x3 + (l/2)*cos(m4);
y7 = y3 + (l/2)*sin(m4);
x8 = x4 + (l/2)*cos(m4);
y8 = y4 + (l/2)*sin(m4);

makePolygon(x3, y3, x4, y4, x6, y6,x5,y5);
roiManager("Add");
makePolygon(x3, y3, x4, y4, x8, y8,x7,y7);
roiManager("Add");

//Quntifies the area of the fluorecesent puncta in each half of the
cell.
selectWindow("FluorescenceImageName.tif");
roiManager("Select",2);
run("Duplicate...", " ");
run("Subtract...", "value=50");
run("Unsharp Mask...", "radius=5 mask=0.90");
setAutoThreshold("Otsu dark");
```

```
//run("Threshold...");
//setThreshold(109, 255);
run("Convert to Mask");
run("Convert to Mask");
run("Analyze Particles...", "size=0.07 display clear");
N = nResults;
selectWindow("Results");
N=0;
total_area_Left=0;
for (a=0; a<nResults(); a++) {
    total_area_Left = total_area_Left + getResult("Area",a);
}
//print (total_area_Left);

selectWindow("Results");
run("Close");

selectWindow("Image2.tif");
roiManager("Select",3);
run("Duplicate...", " ");
run("Subtract...", "value=50");
run("Unsharp Mask...", "radius=5 mask=0.90");
setAutoThreshold("Otsu dark");
//run("Threshold...");
//setThreshold(109, 255);
run("Convert to Mask");
run("Convert to Mask");
run("Analyze Particles...", "size=0.07 display clear");
N = nResults;
selectWindow("Results");
N=0;
total_area_Right=0;
for (a=0; a<nResults(); a++) {
    total_area_Right = total_area_Right + getResult("Area",a);
}
//print (total_area_Right);

selectWindow("Results");
run("Close");


selectWindow("Image2.tif");
roiManager("Select",0);
run("Measure");
A = getResult("Area", 0);
//print(A);
selectWindow("Results");
run("Close");
run("Measure Roi");
w1 = getResult("Roi_Width", 0);
l1 = getResult("Roi_Length", 0);
//print(w1,l1);
selectWindow("Results");
run("Close");

LR1 = total_area_Left/A;
RR1 = total_area_Right/A;
```

```
LR2 = total_area_Left/l1;
RR2 = total_area_Right/l1;

setResult("Area", 0, A);
setResult("Width", 0, w1);
setResult("Length", 0, l1);
setResult("Left", 0, total_area_Left);
setResult("Right", 0, total_area_Right);
setResult("LRatio1", 0, LR1);
setResult("RRatio1", 0, RR1);
setResult("LRatio2", 0, LR2);
setResult("Rratio2", 0, RR2);

selectWindow("Image2.tif");
makeLine(x3, y3, x4, y4);
```

D.3 Fitting Gaussian function to the distribution of instantaneous displacements

An algorithm has been programmed in MATLAB, in order to fit Gaussian function to instantaneous displacements recorded for beads and different strains of *E. coli* in order to determine the average displacement and the deviation in its value exhibited by respective strain. It was used to calculate the diffusion coefficient for the object under experiment in a media with viscosity equivalent to that of water (Chapter 6, Fig. 6.2).

File name: GaussianFunction.m

```
% This code fits gaussian function to the sum- normalized distribution
of instantaneous displacements
% and returns the value for "standard deviation", based on chosen
average of the distribution
% Developed by Manasi Gangan, IISER pune on 8th July 2016

close all; clear all;

a = importdata('FileName.txt'); %call the data

[p,q] = hist(a,20);
p2 = p./sum(p); %sum normalization of the frequency

%Calculations of the parameters
mu1 = mean(a) %for motile strains
mu2 = 0; %for non- motile strains
sigma = sqrt(sum((a- mu(1/2)).^2)/length(a)) %mu1 or mu2 is chosen
based on the nature of the strain

%Defining fit
x = linspace(-2,2,300);
y = pdf('normal', x, mu(1/2), sigma);
y2 = y./sum(y); %sum normalization


%Figure
bar(q,p2,'w');
hold on
plot(x3, y2,'r','Linewidth',1.5);

xlabel('Instantaneous displacement (/um)');
ylabel('Frequency');
title('Displacement');
```

E. Copyrights

Sr. No.	Publisher	Journal	Comments/ Link
C1	Springer/ Nature Publishing Group	Nat. Rev. Microbiol.	https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=ba07b5b5-8f1b-422a-ba35-6f7fd5729560
		Nat. Rev. Microbiol.	https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=ba07b5b5-8f1b-422a-ba35-6f7fd5729560
C2	Nature Publishing Group	Nature cell boil.	Permission granted according to "GNU free documentation license"
C3	John Wiley and Sons	EMBO	https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=8624d99a-8482-4005-9b46-523e0acb1068
C4	American Society of Microbiology	Microbiol. and Mol. Biol. Rev.	"ASM authorizes an advanced degree candidate to republish the requested material in his/her doctoral thesis or dissertation. If your thesis, or dissertation, is to be published commercially, then you must reapply for permission"
		J. Bacteriol.	
C5	Elsevier	Curr. Biol.	https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=a1f6c865-625b-47ca-abd0-caf55f1f5a92
			https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=a2861d8a-edd2-481a-abb7-9e4bdcc61425
C6	Proceedings of the National Academy of Sciences of the United States of America	PNAS	Permission granted according to "Liberalization of PNAS copyright policy" for non- commercial use
C7	INTECH open	Confocal Laser Microscopy- Principles and Applications in Medicine, Biology and the Food Sciences (Book)	https://creativecommons.org/licenses/by/3.0/
C8	Oxford University Press	Bioinformatics	https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=53ab5869-be49-4114-b9d9-9cb687e7e62c
C9	PLOS	PLoS Genetics	"Copyright: © 2008 Freed et al. This is an open-access article

			distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.”
C10	The American Association for the Advancement of Science	Science	https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=5e0010ca-0f55-4e54-96a1-5df12bbde1b7

References

1. Guyer, M. S., Reed, R. R., Steitz, J. A. & Low, K. B. Identification of a sex-factor-affinity site in E. coli as gamma delta. *Cold Spr.Harb.Symp.Quant.Biol.* **45,** 135–140 (1981).

2. Hanahan, D. Studies on transformation of Escherichia coli with plasmids. *J. Mol. Biol.* **166,** 557–580 (1983).

3. Parkinson, J. S. Complementation analysis and deletion mapping of Escherichia coli mutants defective in chemotaxis . Complementation Analysis and Deletion Mapping of Escherichia coli Mutants Defective in Chemotaxis. *J. Bacteriol.* **135,** 45–53 (1978).

4. Renzette, N. *et al.* Localization of RecA in Escherichia coli K-12 using RecA-GFP. *Mol. Microbiol.* **57,** 1074–85 (2005).

5. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science (80-.).* **297,** 1183–1186 (2002).

6. Taniguchi, Y. *et al.* Quntifying E.coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **329,** 533–539 (2011).

7. Guzman, L. L. M. *et al.* Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose PBAD Promoter. *J. Bacteriol.* **177,** 4121–4130 (1995).

8. Wery, M., Woldringh, C. L. & Rouviere-Yaniv, J. HU-GFP and DAPI co-localize on the Escherichia coli nucleoid. *Biochimie* **83,** 193–200 (2001).

9. Pazos, M., Peters, K. & Vollmer, W. Robust peptidoglycan growth by dynamic and

variable multi-protein complexes. *Curr. Opin. Microbiol.* **36,** 55–61 (2017).

10. Pelczar, M. J. J., Chan, E. C. S. & Krieg, N. R. *Microbiology*. (McGraw Hill Education, 2001).

11. Brown, L., Wolf, J. M., Prados-Rosales, R. & Casadevall, A. Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi. *Nat. Rev. Microbiol.* **13,** 620–630 (2015).

12. Margolin, W. Sculpting the Bacterial Cell. *Curr. Biol.* **19,** (2009).

13. Juarez, J. R. & Margolin, W. A bacterial actin unites to divide bacterial cells. *EMBO J.* **31,** 2235–2236 (2012).

14. Busiek, K. K. & Margolin, W. Bacterial actin and tubulin homologs in cell growth and division. *Curr. Biol.* **25,** R243–R254 (2015).

15. Höltje, J. & Heidrich, C. Enzymology of elongation and constriction of the murein sacculus of Escherichia coli. *Biochimie* **83,** 103–108 (2001).

16. Scheffers, D. & Pinho, M. Bacterial cell wall synthesis: new insights from localization studies. *Microbiol. Mol. Biol. Rev.* **69,** 585–607 (2005).

17. Koch, A. L. & Doyle, R. J. Inside-to-outside Growth and Turnover of the Wall of Gram-positive Rods. *J. theor. Biol.* **117,** 137–157 (1985).

18. Höltje, J. Growth of the Stress-Bearing and Shape-Maintaining Murein Sacculus of Escherichia coli. *Microbiol. Mol. Biol. Rev.* **62,** 181–203 (1998).

19. Fantes, P. A. CONTROL OF CELL SIZE AND CYCLE TIME IN SCHIZOSACCHAROMYCES POMBE. *J. Cell Sci.* **24,** 51–67 (1977).

20. Sveiczer, A., Novak, B. & Mitchison, J. M. The size control of fission yeast revisited. *J. Cell Sci.* **109,** 2947–2957 (1996).

21. Robert, L. *et al.* Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism. *BMC Biol.* **12,** 17 (2014).

22. Wang, P., Hayden, S. & Masui, Y. Transition of the blastomere cell cycle from cell size-independent to size-dependent control at the midblastula stage in Xenopus laevis. *J. Exp. Zool.* **287,** 128–144 (2000).

23. Taheri-Araghi, S. *et al.* Cell-size control and homeostasis in bacteria. *Curr. Biol.* **25,** 385–391 (2015).

24. Cooper, S. & Helmstetter, C. E. Chromosome replication and the division cycle of Escherichia coli. *J. Mol. Biol.* **31,** 519–540 (1968).

25. Zaritsky, A., Wang, P. & Vischer, N. O. E. Instructive simulation of the bacterial cell division cycle. *Microbiology* **157,** 1876–1885 (2011).

26. Boer, P. A. J. De, Cook, W. R. & Rothfield, L. I. BACTERIAL CELL DIVISION. *Annu. Rev. Genet.* **24,** 249–247 (1990).

27. William, M. FTSZ AND THE DIVISION OF PROKARYOTIC CELLS AND ORGANELLES. *Nat Rev Mol Cell Biol* **6,** 862–871 (2005).

28. Wu, L. J. *et al.* Noc protein binds to specific DNA sequences to coordinate cell division with chromosome segregation. *EMBO J.* **28,** 1940–52 (2009).

29. Cho, H., McManus, H. R., Dove, S. L. & Bernhardt, T. G. Nucleoid occlusion factor SlmA is a DNA-activated FtsZ polymerization antagonist. *Proc. Natl. Acad. Sci. U. S.*

A. **108,** 3773–8 (2011).

30. Wu, L. J. & Errington, J. Nucleoid occlusion and bacterial cell division. *Nat. Rev. Microbiol.* **10,** 8–12 (2012).

31. Simmons, L. a, Grossman, A. D. & Walker, G. C. Replication is required for the RecA localization response to DNA damage in Bacillus subtilis. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 1360–5 (2007).

32. Cox, M. M. The RecA protein as a recombinational repair system. *Mol. Cell* **5,** 1295–1299 (1991).

33. O'Ari, O. H. and R. An Inducible DNA Replication- Cell Division Coupling Mechanism in E. coli. *Nature* **290,** 797–799 (1981).

34. Huisman, O., D'Ari, R. & Gottesman, S. Cell-division control in Escherichia coli: specific induction of the SOS function SfiA protein is sufficient to block septation. *Proc. Natl. Acad. Sci. U. S. A.* **81,** 4490–4 (1984).

35. Chen, Y., Milam, S. L. & Erickson, H. P. SulA inhibits assembly of FtsZ by a simple sequestration mechanism. *Biochemistry* **51,** 3100–9 (2012).

36. M, C. J. and V. Cell growth and length distribution in Escherichia Cell. *J. Bacteriol.* **134,** 330–337 (1978).

37. Weart, R. B. *et al.* A metabolic sensor governing cell size in bacteria. *Cell* **130,** 335–347 (2008).

38. Wang, J. D. & Levin, P. A. Metabolism, cell growth and the bacterial cell cycle. *Nat Rev Microbiol* **7,** 822–827 (2009).

39. Hill, N. S., Buske, P. J., Shi, Y. & Levin, P. A. A Moonlighting Enzyme Links Escherichia coli Cell Size with Central Metabolism. *PLoS Genet.* **9,** (2013).

40. Nyström, T. Stationary-Phase Physiology. *Annu. Rev. Microbiol.* **58,** 161–181 (2004).

41. Freed, N. E. *et al.* A simple screen to identify promoters conferring high levels of phenotypic noise. *PLoS Genet.* **4,** 2–7 (2008).

42. Paulsson, J. Models of stochastic gene expression. *Phys. Life Rev.* **2,** 157–175 (2005).

43. Kaern, M., Elston, T. C., Blake, W. J. & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* **6,** 451–464 (2005).

44. Raser, J. M. & O'Shea, E. K. Noise in gene expression: origins, consequences, and control. *Science (80-.).* **309,** 2010–2013 (2005).

45. Huh, D. & Paulsson, J. Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Genet.* **43,** 95–100 (2011).

46. Huh, D. & Paulsson, J. Random partitioning of molecules at cell division. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 15004–9 (2011).

47. Lasslett, A. Principles and Applications of Differential Interference Contrast Light Microscopy. **20,** 9–11 (2006).

48. Rosenthal, C. K. Milestone 8: Contrast by interference. *Nat. Cell Biol.* **11,** S11–S12 (2009).

49. Minsky, M. Memoir on Inventing the Confocal Scanning Microscope. *Scanning* **10,** 128–138 (1987).

50. Webb, R. H. Confocal optical microscopy. *Rep. Prog. Phys* **59,** 427–471 (1996).

51. Rossetti, F. C., Depieri, L. V. & Maria Vitória Lopes Badra Bentley. in *Confocal Laser Microscopy- Principles and Applications in Medicine, Biology and the Food Sciences* (INTECH Open, 2013).

52. Weibel, D. B., DiLuzio, W. R. & Whitesides, G. M. Microfabrication meets microbiology. *Nat Rev Microbiol* **5,** 209–218 (2007).

53. Wang, P. *et al.* Robust growth of Escherichia coli. *Curr. Biol.* **20,** 1099–103 (2010).

54. Athale, C. A. & Chaudhari, H. Population length variability and nucleoid numbers in Escherichia coli. *Bioinformatics* **27,** 2944–2948 (2011).

55. Lark, N.-C. C. and K. G. Cytological Studies of Shift-Up into Rich Medium Cytological Studies of Deoxyribonucleic Acid Replication in Escherichia coli 15T- : Replication at Slow Growth Rates and After a Shift-Up into Rich Medium. *J. Bacteriol.* **104,** (1970).

56. Venkatesh, R. *et al.* RecX protein abrogates ATP hydrolysis and strand exchange promoted by RecA: Insights into negative regulation of homologous recombination. *Proc Natl Acad Sci U S A* **99,** 12091–12096 (2002).

57. Zar, J. H. *Biostatistical analysis*. (Delhi, India: Pearson Education Inc., 2008).

58. Schneider, C. a, Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9,** 671–675 (2012).

59. Youngren, B., Nielsen, H. J., Jun, S. & Austin, S. The multifork Escherichia coli chromosome is a self-duplicating and self-segregating thermodynamic ring polymer. *Genes Dev.* **28,** 71–84 (2014).

60. Koppes, L. H., Woldringh, C. L. & Nanninga, N. Size variations and correlation of

different cell cycle events in slow-growing Escherichia coli. *J. Bacteriol.* **134,** 423–433 (1978).

61. de Boer, P. a, Crossley, R. E. & Rothfield, L. I. Roles of MinC and MinD in the site-specific septation block mediated by the MinCDE system of Escherichia coli. *J. Bacteriol.* **174,** 63–70 (1992).

62. Raskin, D. M. & de Boer, P. a. Rapid pole-to-pole oscillation of a protein required for directing division to the middle of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4971–6 (1999).

63. Hale, C. a, Meinhardt, H. & de Boer, P. a. Dynamic localization cycle of the cell division regulator MinE in Escherichia coli. *EMBO J.* **20,** 1563–72 (2001).

64. Männik, J. *et al.* Robustness and accuracy of cell division in Escherichia coli in diverse cell shapes. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 6957–62 (2012).

65. n Sánchez-Gorostiaga A, Palacios P, Martínez-Arteaga R, Sánchez M, Casanova M, V. M. Life without Division : Physiology of Escherichia coli FtsZ-Deprived Filaments. *MBio* **7,** 1–10 (2016).

66. Den Blaauwen, T., Lindqvist, A., Löwe, J. & Nanninga, N. Distribution of the Escherichia coli structural maintenance of chromosomes (SMC)-like protein MukB in the cell. *Mol. Microbiol.* **42,** 1179–1188 (2001).

67. Lutkenhaus, J. O. E. FtsA is localized to the septum in an These include : FtsA Is Localized to the Septum in an FtsZ-Dependent Manner. *J. Bacteriol.* **178,** 7167–7172 (1996).

68. Rico, A. I., Krupka, M. & Vicente, M. In the beginning, escherichia coli assembled the

proto-ring: An initial phase of division. *Journal of Biological Chemistry* **288,** 20830–20836 (2013).

69. Justice, S. S., Hunstad, D. A., Cegelski, L. & Hultgren, S. J. Morphological plasticity as a bacterial survival strategy. *Nat. Rev. Microbiol.* **6,** 162–168 (2008).

70. Justice, S. S., Hunstad, D. A., Seed, P. C. & Hultgren, S. J. Filamentation by Escherichia coli subverts innate defenses during urinary tract infection. *Proc. Natl. Acad. Sci.* **103,** 19884–19889 (2006).

71. Chen, K., Sun, G. W., Chua, K. L. & Gan, Y. Modified virulence of antibiotic-induced *Burkholderia pseudomallei* filaments. *Antimicrob. Agents Chemother.* **49,** 1002–9 (2005).

72. Belas, R. & Suvanasuthi, R. The Ability of Proteus mirabilis To Sense Surfaces and Regulate Virulence Gene Expression Involves FliL , a Flagellar Basal Body Protein The Ability of Proteus mirabilis To Sense Surfaces and Regulate Virulence Gene Expression Involves FliL , a Flagellar. *J. Bacteriol.* **187,** 6789–6803 (2005).

73. Schaechter, M., MaalOe, O. & Kjeldgaard, N. O. Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of Salmonella typhimurium. *J. Gen. Microbiol.* **19,** 592–606 (1958).

74. Kendrick, B. A. G.-M., Captain, I. M. S. & Pai, M. K. The Rate of Multiplication of Micro-organisms . *Proc. R. Soc. Edinburgh* . (1911).

75. Pearl, R. & Reed, L. J. On the Rate of Growth of the Population of the United States Since 1790 and its Mathematical Representation. *Proc Natl Acad Sci U S A* **6,** 275–288 (1920).

76. Hill, N. S., Kadoya, R., Chattoraj, D. K. & Levin, P. A. Cell size and the initiation of DNA replication in bacteria. *PLoS Genet.* **8,** e1002549 (2012).

77. Sinha, N. K. & Snustad, D. P. Mechanism of inhibition of deoxyribonucleic acid synthesis in Escherichia coli by hydroxyurea. *J. Bacteriol.* **112,** 1321–4 (1972).

78. Nazaretyan, S. A. *et al.* Replication rapidly recovers and continues in the presence of hydroxyurea in *Escherichia coli*. *J. Bacteriol.* JB.00713-17 (2017). doi:10.1128/JB.00713-17

79. Hitchings, G. H. Mechanism of Action of Trimethoprim-Sulfamethoxazole : I. *J. Infect. Dis.* **128,** S433–S436 (1973).

80. Alice, A. MINIATURE escherichia coli CELLS DEFICIENT IN DNA. *Proc. Natl. Acad. Sci. U. S. A.* **6,** 321–26 (1966).

81. Lutkenhaus, J. Assembly Dynamics of the Bacterial MinCDE System and Spatial Regulation of the Z Ring. *Annu. Rev. Biochem.* **76,** 539–562 (2007).

82. Arumugam, S., Petrašek, Z. & Schwille, P. MinCDE exploits the dynamic nature of FtsZ filaments for its spatial regulation. *Proc. Natl. Acad. Sci. U. S. A.* **111,** E1192-200 (2014).

83. Errington, J. Bacterial morphogenesis and the enigmatic MreB helix. *Nat Rev Micro* **13,** 241–248 (2015).

84. Cambridge, J., Blinkova, A., Magnan, D., Bates, D. & Walker, J. R. A Replication-inhibited unsegregated nucleoid at mid-cell blocks Z-ring formation and cell division independently of SOS and the SlmA nucleoid occlusion protein in Escherichia coli. *J. Bacteriol.* **196,** 36–49 (2014).

85. Ennis, D. G., Fisher, B., Edmiston, S. & Mount, D. W. Dual role for Escherichia coli RecA protein in SOS mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **82,** 3325–9 (1985).

86. Lin, L. L. & Little, J. W. Isolation and characterization of noncleavable (Ind-) mutants of the LexA repressor of Escherichia coli K-12. *J. Bacteriol.* **170,** 2163–2173 (1988).

87. Justice, S. S. *et al.* Differentiation and developmental pathways of uropathogenic *Escherichia coli* in urinary tract pathogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 1333–1338 (2004).

88. Patnaik, P. R. Noise in Bacterial Chemotaxis: Sources, Analysis, and Control. *Bioscience* **62,** 1030–1038 (2012).

89. Weiss, N. A. *Course In Probability*. (Pearson, 2005).

90. Hilfinger, A. & Paulsson, J. Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proc Natl Acad Sci U S A* **2011,** 12167–12172 (2011).

91. Bowsher, C. G. & Swain, P. S. Identifying sources of variation and the flow of information in biochemical networks. *Proc Natl Acad Sci U S A* **109,** E1320–E1328 (2012).

92. Fu, A. Q. & Pachter, L. Estimating intrinsic and extrinsic noise from single-cell gene expression measurements. *Stat. Appl. Genet. Mol. Biol.* **15,** 447–471 (2016).

93. Gangan, M. S. & Athale, C. A. Threshold effect of growth rate on population variability of Escherichia coli cell lengths. *Roy. Soc. open Sci.* **4,** 1–17 (2017).

94. MANDERS, E. M. M., VERBEEK, F. J. & ATEN, J. A. Measurement of co-localization of objects in dual-colour confocal images. *J. Microsc.* **169,** 375–382 (1993).

95. Jardetzky, O. Studies on the Mechanism of Action of Chloramphenicol. I. THE CONFORMATION OF CHLORAMPHENICOL IN SOLUTION. *J. Biol. Chem.* **238,** 2498–2508 (1963).

96. Wehrli, W. Rifampin: mechanisms of action and resistance. *Rev. Infect. Dis.* **5 Suppl 3,** S407–S411 (1983).

97. Bean, G. J. *et al.* A22 disrupts the bacterial actin cytoskeleton by directly binding and inducing a low-affinity state in MreB. *Biochemistry* **48,** 4852–4857 (2009).

98. Kruse, T. *et al.* Actin homolog MreB and RNA polymerase interact and are both required for chromosome segregation in Escherichia coli. *Genes Dev.* **20,** 113–124 (2006).

99. Carballido-López, R. The bacterial actin-like cytoskeleton. *Microbiol. Mol. Biol. Rev.* **70,** 888–909 (2006).

100. Kruse, T., Møeller-Jensen, J., Løbner-Olesen, A. & Gerdes, K. Dysfunctional MreB inhibits chromosome segregation in Escherichia coli. *EMBO J.* **22,** 5283–5292 (2003).

101. Pogliano, J., Pogliano, K., Weiss, D. S., Losick, R. & Beckwith, J. Inactivation of FtsI inhibits constriction of the FtsZ cytokinetic ring and delays the assembly of FtsZ rings at potential division sites. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 559–64 (1997).

102. Drawz, S. M. & Bonomo, R. A. Three decades of beta -lactamase inhibitors. *Clin. Microbiol. Rev.* **23,** 160–201 (2010).

103. Singh, A. & Soltani, M. Quantifying intrinsic and extrinsic variability in stochastic gene expression models. *PLoS One* **8,** (2013).

104. Knoblich, J. a. Asymmetric cell division during animal development. *Nat. Rev. Mol.*

Cell Biol. **2,** 11–20 (2001).

105. Zgurski, J. M., Sharma, R., Bolokoski, D. A. & Schultz, E. A. Asymmetric Auxin Response Precedes Asymmetric Growth and Differentiation of asymmetric leaf1 and asymmetric leaf2 Arabidopsis Leaves. *Plant Cell* **17,** 77–91 (2005).

106. Gonzalez, C. Spindle orientation, asymmetric division and tumour suppression in Drosophila stem cells. *Nat. Rev. Genet.* **8,** 462–72 (2007).

107. Petricka, J. J., Van Norman, J. M. & Benfey, P. N. Symmetry breaking in plants: molecular mechanisms regulating asymmetric cell divisions in Arabidopsis. *Cold Spring Harb. Perspect. Biol.* **1,** a000497 (2009).

108. Hohm, T., Preuten, T. & Fankhauser, C. Phototropism: Translating light into directional growth. *Am. J. Bot.* **100,** 47–59 (2013).

109. Zusman, D. R., Scott, A. E., Yang, Z. & Kirby, J. R. Chemosensory pathways, motility and development in Myxococcus xanthus. *Nat. Rev. Microbiol.* **5,** 862–872 (2007).

110. Skerker, J. M. & Laub, M. T. Cell-cycle progression and the generation of asymmetry in Caulobacter crescentus. *Nat. Rev. Microbiol.* **2,** 325–37 (2004).

111. Hilbert, D., Laboratories, M. D., Ã, P. J. P. & Hilbert, D. W. Sporulation of Bacillus subtilis. *Curr. Opin. Genet. Dev.* **7,** 579–586 (2004).

112. Aldridge, B. B. *et al.* Asymmetry and Aging of Mycobacterial Cells Lead to Variable Growth and Antibiotic Susceptibility. *Science* **335,** 100–104 (2011).

113. Brown, P. J. B. *et al.* Polar growth in the Alphaproteobacterial order Rhizobiales. *Proc. Natl. Acad. Sci.* **109,** 3190–3190 (2012).

114. Joyce, G. *et al.* Cell division site placement and asymmetric growth in Mycobacteria. *PLoS One* **7,** 1–8 (2012).

115. Guberman, J. M., Fay, A., Dworkin, J., Wingreen, N. S. & Gitai, Z. PSICIC: noise and asymmetry in bacterial division revealed by computational image analysis at sub-pixel resolution. *PLoS Comput. Biol.* **4,** e1000233 (2008).

116. Norris, V., Woldringh, C. & Mileykovskaya, E. A hypothesis to explain division site selection in Escherichia coli by combining nucleoid occlusion and Min. *FEBS Lett.* **561,** 3–10 (2004).

117. Binenbaum, Z., Parola, a H., Zaritsky, a & Fishov, I. Transcription- and translation-dependent changes in membrane dynamics in bacteria: testing the transertion model for domain formation. *Mol. Microbiol.* **32,** 1173–82 (1999).

118. Goehring, N. W. & Beckwith, J. Diverse paths to midcell: assembly of the bacterial cell division machinery. *Curr. Biol.* **15,** R514-26 (2005).

119. Bramkamp, M. & van Baarle, S. Division site selection in rod-shaped bacteria. *Curr. Opin. Microbiol.* **12,** 683–8 (2009).

120. Reimold, C., Defeu Soufo, H. J., Dempwolff, F. & Graumann, P. L. Motion of variable-length MreB filaments at the bacterial cell membrane influences cell morphology. *Mol. Biol. Cell* **24,** 2340–9 (2013).

121. Strahl, H., Bürmann, F. & Hamoen, L. W. The actin homologue MreB organizes the bacterial cell membrane. *Nat. Commun.* **5,** 1–11 (2014).

122. Ellenberg, J. *et al.* Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis. *J. Cell Biol.* **138,** 1193–

206 (1997).

123. Lloyd-Price, J. *et al.* Asymmetric disposal of individual protein aggregates in Escherichia coli, one aggregate at a time. *J. Bacteriol.* **194,** 1747–1752 (2012).

124. Tyagi, S. E. coli, What a Noisy Bug. *Science (80-.).* **329,** 518–519 (2010).

125. Ursell, T. S., Trepagnier, E. H., Huang, K. C. & Theriot, J. A. Analysis of Surface Protein Expression Reveals the Growth Pattern of the Gram-Negative Outer Membrane. *PLoS Comput. Biol.* **8,** (2012).

126. Nei, M. Selectionism and neutralism in molecular evolution. *Mol. Biol. Evol.* **22,** 2318–2342 (2005).

127. Cabeen, M. T. & Jacobs-Wagner, C. Bacterial cell shape. *Nat. Rev. Microbiol.* **3,** 601–610 (2005).

128. Kysela, D. T., Randich, A. M., Caccamo, P. D. & Brun, Y. V. Diversity Takes Shape: Understanding the Mechanistic and Adaptive Basis of Bacterial Morphology. *PLoS Biol.* **14,** 1–15 (2016).

129. Persat, A., Stone, H. A. & Gitai, Z. The curved shape of caulobacter crescentus enhances surface colonization in flow. *Nat. Commun.* **5,** 1–9 (2014).

130. Young, K. D. The Selective Value of Bacterial Shape. *Microbiol. Mol. Biol. Rev.* **70,** 660–703 (2006).

131. Berleman, J. E., Chumley, T., Cheung, P. & Kirby, J. R. Rippling is a predatory behavior in Myxococcus xanthus. *J. Bacteriol.* **188,** 5888–5895 (2006).

132. Smith, W. P. J. *et al.* Cell morphology drives spatial patterning in microbial

communities. *Proc. Natl. Acad. Sci.* **114,** E280–E286 (2017).

133. Hallatschek, O., Hersen, P., Ramanathan, S. & Nelson, D. R. Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 19926–30 (2007).

134. Berg, H. C. *Random walks in biology*. (Princeton University Press, 1993).

135. Kysela, D. T., Brown, P. J., Casey Huang, K. & Brun, Y. V. Biological Consequences and Advantages of Asymmetric Bacterial Growth. *Annu. Rev. Microbiol* **67,** 417–35 (2013).

136. Rainey, P. B. *et al.* The evolutionary emergence of stochastic phenotype switching in bacteria. *Microb. Cell Fact.* **10,** S14 (2011).

137. Sousa, A. M., Machado, I. & Pereira, M. O. Phenotypic switching: an opportunity to bacteria thrive. *Sci. against Microb. Pathog. Commun. Curr. Res. Technol. Adv.* 252–262 (2011).

138. Shah, D. *et al.* Persisters: a distinct physiological state of E. coli. *BMC Microbiol.* **6,** 53 (2006).

139. Ackermann, M. *et al.* Self-destructive cooperation mediated by phenotypic noise. *Nature* **454,** 987–90 (2008).

140. Raj, A. & van Oudenaarden, A. Stochastic gene expression and its consequences. *Cell* **135,** 216–226 (2008).

141. Fire, A., Harrison, S. W. & Dixon, D. A modular set of lacZ fusion vectors for studying gene expression in Caenorhabditis elegans. *Gene* **93,** 189–190 (1990).

142. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **111,** 229–233 (1992).

143. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prashert, D. C. Green Fluorescent Protein as a Marker for Gene Expression. *Science (80-.).* **263,** 802–805 (1994).

144. Inouye, S. & Tsuji, F. I. Aequorea green fluorescent protein. *FEBS Lett.* **341,** 277–280 (1994).

145. Sambrook, J. & Russell, D. W. *Molecular Cloning: A LABORATORY MANUAL*. (Cold Spring Harbor Laboratory Press, 2001).

146. Chen, W. & Kuo, T. T. A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. *Nucleic Acids Res.* **21,** 2260 (1993).

147. Mitra, S. K., Li, H., Lin, I. & Yu, T. A New Class of Nonlinear Filters for Image Enhancement. *coustics, Speech, Signal Process. 1991. ICASSP-91., 1991 Int. Conf. on. IEEE, 1991* 2525–2528 (1991).

148. Vala, H. J. & Baxi, P. A. A Review on Otsu Image Segmentation Algorithm. *Int. J. Adv. Res. Comput. Eng. Technol.* **2,** 387–389 (2013).

149. Hartig, S. M. Basic Image Analysis and Manipulation in ImageJ Basic Image Analysis and Manipulation in ImageJ. *Curr. Protoc. Mol. Biol.* (2013). doi:10.1002/0471142727.mb1415s102

ROYAL SOCIETY OPEN SCIENCE

rsos.royalsocietypublishing.org

Research

Cite this article: Gangan MS, Athale CA. 2017 Threshold effect of growth rate on population variability of *Escherichia coli* cell lengths. *R. Soc. open sci.* **4**: 160417.
http://dx.doi.org/10.1098/rsos.160417

Received: 22 June 2016
Accepted: 23 January 2017

Subject Category:
Cellular and molecular biology

Subject Areas:
biophysics/cellular biology/microbiology

Keywords:
cell size, bacteria, multi-fork replication, population variability, RecA, bacterial cell division

Author for correspondence:
Chaitanya A. Athale
e-mail: cathale@iiserpune.ac.in

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.3685441.

THE ROYAL SOCIETY PUBLISHING

Threshold effect of growth rate on population variability of *Escherichia coli* cell lengths

Manasi S. Gangan and Chaitanya A. Athale

Division of Biology, Indian Institute of Science Education and Research (IISER) Pune, Dr Homi Bhabha Road, Pashan, Pune 411008, India

MSG, 0000-0003-2980-7329; CAA, 0000-0002-9506-2153

A long-standing question in biology is the effect of growth on cell size. Here, we estimate the effect of *Escherichia coli* growth rate (r) on population cell size distributions by estimating the coefficient of variation of cell lengths (CV_L) from image analysis of fixed cells in DIC microscopy. We find that the CV_L is constant at growth rates less than one division per hour, whereas above this threshold, CV_L increases with an increase in the growth rate. We hypothesize that stochastic inhibition of cell division owing to replication stalling by a RecA-dependent mechanism, combined with the growth rate threshold of multi-fork replication (according to Cooper and Helmstetter), could form the basis of such a threshold effect. We proceed to test our hypothesis by increasing the frequency of stochastic stalling of replication forks with hydroxyurea (HU) treatment and find that cell length variability increases only when the growth rate exceeds this threshold. The population effect is also reproduced in single-cell studies using agar-pad cultures and 'mother machine'-based experiments to achieve synchrony. To test the role of RecA, critical for the repair of stalled replication forks, we examine the CV_L of *E. coli* Δ*recA* cells. We find cell length variability in the mutant to be greater than wild-type, a phenotype that is rescued by plasmid-based RecA expression. Additionally, we find that RecA-GFP protein recruitment to nucleoids is more frequent at growth rates exceeding the growth rate threshold and is further enhanced on HU treatment. Thus, we find growth rates greater than a threshold result in increased *E. coli* cell lengths in the population, and this effect is, at least in part, mediated by RecA recruitment to the nucleoid and stochastic inhibition of division.

1. Introduction

The size and shape of a cell is considered a characteristic feature of a given cell type, and quantifying its variability in a population provides information about the effect of fluctuations on a complex phenotype. *Escherichia coli* cells have typically been described as spherocylinders of length 2 µm and width 1 µm. Differences in sizes are primarily owing to cell length (L) and not so much the width [1–3]. Cell length frequency distributions show a positive skew owing to the presence of long cells ($L > 8$ µm), the proportion of which is increased by environmental factors such as low bacterial density at 22°C and 37°C or a shift to richer media [4]. In *Salmonella*, growth rate (r) alone has been shown to correlate with increased cell size and multiple nucleoids [5], but a microscopic study on *E. coli* has shown that cells grown at 22°C are shorter than at 37°C [4]. The effect of temperature and growth medium on cell size appears thus to suggest that growth rate might primarily regulate the cell size. However, the quantitative relationship and molecular mechanism by which growth could affect cell sizes remains unclear.

The growth rate of bacteria, in particular, *E. coli*, is regulated by numerous pathways that typically connect growth to nutrient availability [6–8]. Many genetic factors that link nutrient sensing to cell size regulation have been identified [9–11]. These pathways, however, link growth rate via pathways independent of replication to cell size. If DNA replication fails to complete and the bacterial nucleoid does not segregate, the nucleoid 'occlusion' response results in cell elongation [12–14]. Based on the BCD—birth (B), chromosome replication (C) and division (D)—cycle [15], growth rates exceeding one doubling per hour (doubling time, $t_d = 60$ min) result in insufficient time for the completion of the chromosome replication (C-period approx. 40 min) and cell division (D-period approx. 20 min). Cooper and Helmstetter postulated and experimentally demonstrated that *E. coli* undergoes simultaneous rounds of replication, multi-fork replication [16] to overcome the shortening of t_d in rapid growth. However, the role of multi-fork replication in cell size regulation has not yet been investigated.

Recent improvements in light microscopy image analysis have allowed quantification of bacterial morphology and growth dynamics with subpixel accuracy [17–19]. Combined with fluorescence microscopy of subcellular components [20,21], it has become possible to address single-cell dynamics of the bacterial cell division cycle. These advances now allow us to address the effect of population sizes and physical factors and probe the mechanisms that control cell sizes and cell size variability.

Theoretical studies have suggested that asymmetric cell division [2], lognormal distribution of growth rates [22] or stochastic partitioning of molecular components at cell division [23] could lead to a heterogeneity in cell sizes in the population. Recent single-cell bacterial growth kinetic data [19] combined with theoretical modelling have reopened the debate of whether cell-size robustness is determined by a 'timer' or 'sizer' mechanism [24] and currently the 'incremental' or 'adder' model appears to explain all available data [25,26]. However, the effect of molecular regulatory networks on cell size and the correlation of cell size variability with growth rate remain unclear.

RecA is a central regulator of the SOS response pathway, and deletion mutants of *E. coli* for the *recA* gene experience enhanced replication fork stalling [27]. Additionally, a *recA1* mutation is known to result in asynchronous replication and a reduction in the expected genome-copy numbers [15]. In previous work, we had found that a *recA1* mutation phenocopies typical cell septation defects, resulting in elongated cells containing multiple nucleoids and increased cell length variability [28]. While replication fork stalling and repair are important for DNA replication, as reviewed by Cox *et al*. [29], the artificial induction of replication stalling results in increased cell lengths [30]. The repair of stalled replication forks by RecA protein assembly on DNA [31] also triggers SulA-mediated cell division inhibition [32] via the SOS response pathway [33–35]. At the same time, the population growth rate affects the number of replication forks per cell in a step-wise manner [16]. As a result, the number of replication stalling events could be multiplicatively increased by growth rate and thus affect cell division. Therefore, we hypothesize that RecA might provide the molecular link between *E. coli* growth rate and cell length.

Here, we measure the correlation between cell length variability and growth rate from steady-state cultures, and test our method against single-cell agar-pad and microfluidic growth assays. We find that cell size variability remains unchanged for slow-growing cultures, but increases above a threshold growth rate. By increasing replication fork stalling with hydroxyurea (HU) in multiple mutant strains, we demonstrate that DNA replication fork dynamics can affect population cell size distributions in a RecA-dependent manner. From the growth-rate-dependent recruitment of RecA to the genome, we infer a molecular mechanism that links growth rate to cell size.

2. Material and methods

2.1. Bacterial strains and plasmids

Multiple *E. coli* strains were used: MG1655 (6300, CGSC), *ΔrecA* (JW26691, CGSC), *ΔsulA* (JW09411, CGSC), *ΔslmA* (JW56411, CGSC) and *E. coli* MG1655 with a GFP-tagged genomic copy of *recA* (*recA-GFP*) was grown in the presence of 25 µg ml^{-1} kanamycin as described previously [36] (gift from Dr G.P. Manjunath). Nucleoid segregation dynamics were followed in *E. coli* MG1655 with a pBAD24-hupA-gfp plasmid with 100 µg ml^{-1} ampicillin [37] (gift from Dr Josette Rouviere-Yaniv). We constructed two *recA* expression plasmids (i) *mCherry* tagged and (ii) arabinose-inducible, untagged. Two primer sets were used with complementary regions to the genomic RecA sequence and overhangs for restriction digestion for the p-recA-mCherry and pBAD-recA constructs (electronic supplementary material, table S1). The *recA* gene was PCR-amplified (Mastercycler proS, Eppendorf, Germany) using Taq polymerase and dNTPs (Bangalore GeNei, India) in recommended buffers. The template DNA, *E. coli* MG1655 genomic DNA, was extracted by a rapid extraction method that avoids polysaccharide contamination [38]. The *recA* amplicon for mCherry tagging and the p-mCherry plasmid were sequentially digested with *Sal*I and *Hind*III. The fragments were separated on an agarose gel, column-purified (QIAquick, Qiagen, Germany) and ligated using a T4 DNA Ligase (Bangalore GeNei, India). The plasmid p-mCherry was constructed by replacing the GFP sequence in a pGFP plasmid with mCherry from p-mCherry-N1 (both plasmids from Clontech, USA) by directional cloning using the restriction enzymes *Sal*I and *Eco*RI. The *recA* amplicon for arabinose-inducible expression was purified, and both the amplicon and pBAD24 digested sequentially by *Nhe*I and *Xba*I, and ligated as before. Plasmids were transformed using the $CaCl_2$ method [39] in *E. coli* DH5α cells. Plasmids were isolated using a spin column-based method (Miniprep Kit, Qiagen GmbH, Germany).

2.2. Growth media

For rapid growth, cells were grown in Luria–Bertani (LB) broth (HiMedia, Mumbai, India), while reduced growth rate was achieved using the reduced media yeast extract broth (YEB): 0.5% (w/v) yeast extract in 1% (w/v) solutions of NaCl and tryptone broth (TB): 1% (w/v) tryptone in a 1% (w/v) solution of NaCl. Additionally, M9 minimal salts medium [40] supplemented with 4 µg ml^{-1} thymidine were reconstituted with three different carbon sources (to result in successively slower growth rates): 0.4% (w/v) glucose or 0.9% (w/v) succinic acid or 0.5% (w/v) sodium acetate (all sugars from Sigma-Aldrich). All broths and media were made in deionized water and the pH was adjusted to 7.

2.3. Batch culture and growth rate estimation

Cells were grown at 37°C with shaking at 180 r.p.m. (Forma, ThermoScientific, USA) in 100 ml LB, YEB and TB using a 1% overnight inoculum. Identical conditions were used to grow *E. coli* MG1655 in M9 + sugars. Cell density was estimated by converting 1 $OD_{600\,nm} = 8 \times 10^8$ cells ml^{-1} [41]. To estimate the growth rate (r), the averaged OD with time curves were fit to the solution to the logistic equation by

$$N(t) = \frac{N(0) \cdot K}{N(0) + (K - N(0)) \cdot e^{-r \cdot t}}, \tag{2.1}$$

where $N(0)$ is the population at the time of inoculation, r is the growth rate (h^{-1}), K is the carrying capacity and t is time (electronic supplementary material, figure S1). Doubling time is $t_d = 1/r$ [42].

2.4. Continuous cell culture

A PDMS-based microfluidic device was used to grow *E. coli* MG1655 culture continuously based on the 'mother machine' design [19]. The device was designed as a two-layered micro-pattern mask in CleWin (WieWin Web, The Netherlands) and fabricated by using an approximately 100 nm layer of gold (for aligning the second layer) followed by spin-coating a 2 µm layer SU8-2 negative photoresist (Microchem, USA) onto a SiO_2 wafer using a spin coater model WS-400B-6NPP LITE (Laurell Tech. Corp., USA). The photoresist was cured by UV exposure with a mask (EVG, Austria) corresponding to the trench and dead-end channels. Unexposed photoresist was washed and a 20 µm layer of SU8-20 negative photoresist (Microchem, USA) spun and exposed to UV corresponding only to the trench, for

curing. The PDMS device was made by mixing elastomer : curing agent of 10 : 1 w/w (Sylgard 184, Dow-Corning, USA) and coating the wafer and heat-curing it in an oven at 60°C (Raut Scientific, Maharashtra, India) for 2 h. The cured PDMS membrane was cleaned with pentane (Sigma-Aldrich, Mumbai, India) and washed with acetone (Fisher Scientific, Mumbai, India). The air-dried device was then bonded in air using a plasma cleaner (Emitech K050X, Quorum Technologies, UK) under RF power of 70 W, washing time of 30 s and 1 mbar vacuum (Edwards Pumps, UK). The PDMS-glass device was integrated and channels passivated by passing 10 mg ml^{-1} of BSA (Sigma-Aldrich, Mumbai, India) for 1 h. *E. coli* MG1655 cells were infused into the device (OD ~ 1.0) and allowed to diffuse into the channels for 1 h at 37°C using a syringe pump (PHD Ultra, Harvard Apparatus, USA). The device was washed by flowing in either fresh LB medium or M9 + succinate at a constant flow rate of 0.3 ml h^{-1}. Cell growth and division were observed in DIC microscopy. To measure the effect of HU treatment in continuous culture, *E. coli* MG1655 and Δ*recA* strains expressing eGFP from a plasmid were introduced in the microfluidics device as before. Cells were grown under continuous flow in LB for 1 h (approx. three generations), followed by a change of the medium to the corresponding medium supplemented with 30 mM HU for 1 h ('treatment'). Subsequently, the medium without any drug (no HU) was once again replaced for 2.5 h of 'recovery'. Fluorescence time-lapse images were acquired in the GFP channel and analysed.

2.5. Hydroxyurea and trimethoprim treatment

Overnight cultures were grown from a single colony of *E. coli* MG1655, Δ*recA*, Δ*sulA*, Δ*slmA* and Δ*recA* + pRecA-mCherry. The cultures were diluted 1 : 100 (1% inoculum) into 5 ml of fresh LB and M9 + 0.9% succinate and grown at 37°C with shaking (180 r.p.m.). *E. coli* MG1655 with genomic RecA-GFP was similarly grown in LB, TB and YEB at 37°C with shaking. At $OD_{600\,nm}$ ~ 0.2, the cultures were incubated in 10–100 mM HU containing growth medium for three generations corresponding to 1 h in LB, 1.5 h in TB, 2 h in YEB, 3 h in M9 + succinate. Subsequently, cells were allowed to recover for another three generations. Similarly, *E. coli* MG1655 cells grown in LB and M9 + succinate were exposed to 1 µg ml^{-1} trimethoprim (Sigma-Aldrich, India) and allowed to recover for 1 and 3 h, respectively. After recovery, all cultures (treated and untreated) were washed, fixed and imaged.

2.6. Western blotting

Escherichia coli MG1655 and Δ*recA* cells grown in LB, YEB and TB media were grown for three generations and treated with HU as above. The OD at 600 nm was measured before treatment and after recovery, and 1 ml cell suspensions were diluted to result in comparable cell densities (electronic supplementary material, table S2). The cells were pelleted, washed in phosphate-buffered saline (PBS), resuspended in 50 µl of lysis buffer consisting of 10 µl of 5× SDS loading dye (250 mM Tris–Cl (pH 6.8), 10% SDS, 50% glycerol, 0.5% bromophenol blue and 500 mM DTT) and 40 µl PBS, heated to 95°C for 10 min with constant shaking at 700 r.p.m. (ThermoMixer, Eppendorf, Germany), and samples were centrifuged before loading on a 10% SDS–PAGE gel run at 120 V (Bio-Rad, USA). Proteins were transferred onto a PVDF membrane (Immobilon-P transfer membrane, EMD Millipore Corporation, USA) and the membrane blocked with 5% milk powder in TBST buffer (Tris–Cl-buffered saline (pH 7.4) and 0.1% Tween 20). Rabbit antiserum raised against *E. coli* RecA [43] (a gift from Dr K. Muniyappa) was diluted to 1 : 12 000 in blocking agent (5% milk powder in TBST buffer) and incubated with 100 ml of cell lysate of *E. coli* Δ*recA* for 12 h at 4°C to immunodeplete non-specific antibodies. The lysate was prepared by growing *E. coli* Δ*recA* cells to $OD_{600\,nm}$ approximately 2.0, resuspending the pellet in PBST (PBS with 0.05% (v/v) Tween 20) and lysis by pulse sonication for 2 min, 30 cycles. The membrane was incubated with this pre-treated serum at 4°C overnight, washed and hybridized with the secondary HRP-conjugated anti-rabbit antibody, 1 : 10 000 diluted (Jackson ImmunoResearch, USA). The blot was developed using a reagent Luminata Femto (Millipore Corporation, USA) and luminescence images acquired (LAS 4000, GE Healthcare, USA).

2.7. Cell immunostaining

Escherichia coli MG1655 were grown to the mid-log phase and cells fixed with 1.6% paraformaldehyde (PFA) and 0.01% glutaraldehyde, incubated for 1 h and washed three times with PBST. The cells were treated with GTE (50 mM glucose, 25 mM Tris–Cl, 10 mM EDTA) containing 5 µg ml^{-1} lysozyme and incubated at 37°C for 45 min, followed by 3× wash with PBST. The cell suspension was spread on poly-L-lysine-coated coverslips and air-dried for 1 h. Coverslips were washed three times with PBST, incubated

with 2% BSA (blocking agent) for 1 h, washed with PBST and incubated with the rabbit anti-RecA serum (1 : 1000 diluted in PBS) at 4°C for approximately 12 h. The coverslip was then incubated at 37°C for 1 h with Alexa647 conjugated anti-rabbit antibody (Thermo Fisher Scientific, USA) and mounted on slides.

2.8. Fluorimetry

E. coli recA-GFP was grown in 5 ml of LB, TB and YEB (1% inoculum) with 25 µg ml^{-1} kanamycin at 37°C with constant shaking. The cultures were treated with 30 mM HU as above. Cell suspensions (1 ml) were sampled at pre-treatment (pt) and recovery (r) stages, and r-samples were diluted in the respective growth medium based on the ratio of OD recovery : pre-treatment samples (LB: pt 0.23, r 1.3; YEB: pt 0.195, r 0.92; TB: pt 0.193, r 0.819). All samples were pelleted and resuspended in 50 µl PBS and fluorescence measured in a 96-well half-area round bottom black plate (Corning, USA) using 480 nm excitation and 510 nm emission in a Varioskan Flash multifunctional plate reader (Thermo Scientific, USA). Measurements were blank-subtracted by measuring *E. coli* MG1655 of the same density grown in LB, YEB and TB. Fluorescence per cell was estimated by dividing by the total cell numbers in 50 µl by using the conversion 1 $OD_{600\,nm} = 8 \times 10^8$ cells ml^{-1} as before.

2.9. Microscopy

Cells were sampled (200 µl) from the batch cultures at the mid-log phase and fixed in 4% PFA, stained with 0.1 µg µl^{-1} of DAPI (Sigma-Aldrich, India) and mounted, as has been described previously [28]. Fixed cells of *E.coli* MG1655 *recA-GFP* and *E. coli* Δ*recA* expressing RecA-Cherry were acquired using the 100× (Plan Apochromat N.A. 1.4, oil) objective of a Zeiss Axio Imager Z1 (Carl Zeiss, Germany) microscope in fluorescence and DIC channels. For live-imaging, cells were grown on 2% agar pads with 100 µg ml^{-1} ampicillin and induced for 2 h by 0.2% Arabinose (Sisco Research Labs, Mumbai, India) to express HupA-GFP and imaged on a Zeiss LSM780 confocal microscope (Carl Zeiss, Germany) simultaneously in fluorescence (Diode laser 405 nm, beam splitter MBS 405, pinhole 126.5 corresponding to 1 airy unit) and DIC modes using a 63× lens (Plan Apochromat NA 1.40, oil). Multiple positions were scanned as 512 × 512 pixel images (0.264 µm per pixel) with an image acquired every 2 min for approximately 2 h and at 37°C.

2.10. Image analysis

Cell lengths were automatically analysed from DIC images of fixed cells using a previously developed algorithm [28] in Matlab R2014b (MathWorks Inc., MA, USA). The source code has been released on a GPL basis and can be downloaded from a Github repository (https://github.com/athale/ecolilenDIC). The birth lengths and the division lengths of *E. coli* cells in the 'mother machine' and RecA puncta were interactively estimated using ImageJ (v. 1.50f) [44]. Kymographs of *E. coli* cells expressing plasmid-based eGFP grown in the 'mother machine' were generated by using 'multiplekymograph' plugin in ImageJ [44] based on a line of interest drawn along the growth channel. RecA puncta were quantified by selecting a segmented line corresponding to the length of the *E. coli* cell in DIC and used to generate an intensity profile in the RecA-GFP and DAPI channels. Co-localized peaks were used to score cells in the population and calculate the percentage cells showing such co-localization of RecA on the nucleoid. To follow the dynamics of nucleoid segregation, intensity profiles from timeseries of *pHupA-GFP*-transformed cells were plotted as a matrix to produce a kymograph (space–time plot), using the *imagesc* function in Matlab R2014b (MathWorks Inc.). Western blot intensity analysis was performed using 'gel analyser', an ImageJ plugin. The protein band area was obtained by using the 'label peaks' function and maximum normalized for comparison.

2.11. Data analysis

Cell frequency distributions were normalized by the sum of the area under the curve, fit to a lognormal distribution to obtain lognormal mean (μ) and variance (v) using *fitdist*, *lognpdf* and *lognstat* functions using Matlab with the Statistics Toolbox (MathWorks Inc.). The Kolmogorov–Smirnov (KS) test statistic

was calculated for the number of bins ($n = 44$) and significance level (α) of 0.01 to arrive at a test statistic ($D_{(\alpha,n)}$) given by [45,46]

$$D_{(\alpha,n)} = \frac{\sqrt{-\ln(\alpha/2)}}{2n}. \tag{2.2}$$

The cumulative distribution function (CDF) of observed and fit data was calculated for each bin (i) from the length–frequency distribution. The difference $|d_i| = |F_i - \hat{F}_i|$ between the observed (F_i) and expected ($\hat{F}_i$) values of the CDF was evaluated, and the maximum ($d_{\max}$) was found. The hypothesis that the fit to the data was good was accepted if $d_{\max} < D(\alpha,n)$. Variability in cell lengths was quantified by the coefficient of variation ($\mathrm{CV_L}$) using the expression $\mathrm{CV_L} = \sigma_\mathrm{L}/\mu_\mathrm{L}$, where σ_L and μ_L are the standard deviation and mean of cell lengths, respectively.

3. Results

3.1. Growth rate affects population cell length distributions of *Escherichia coli* MG1655

With the aim of measuring the effect of growth rate, *r*, on cell size, *E. coli* MG1655 cells were grown in LB, YEB, TB and M9 supplemented with glucose, succinate and acetate. As expected, the growth of cells was the fastest in LB and decreased for all other media with the slowest growth observed in M9 supplemented with acetate (figure 1*a*). Doubling time (t_d) values were obtained from logistic function (equation (2.1)) fit to the growth curves (electronic supplementary material, figure S1) and ranged between 33 and 273 min (table 1). Cells sampled from the mid-log phase of each culture (figure 1*a*) were imaged, analysed and the frequency distributions of cell lengths fit to a lognormal function (figure 1*b*). The goodness of the fit was validated based on the KS non-parametric test (electronic supplementary material, table S3). The number of cells analysed in each sample ranged between 10^2 and 10^3 cells, comparable to previous microscopic studies on population cell size distributions [2,47]. Corresponding to the decrease in growth rate, the cell length distributions also decreased in spread. The spread of the distribution was maximal in samples grown in LB and minimal in M9 + acetate, with intermediate growth rates (in YEB, TB, M9 + glucose, M9 + succinate), resulting in an intermediate spread of cell lengths. The growth rate thus appears to alter the quantitative nature of the cell length distribution, while leaving the qualitative nature (lognormal) unchanged. Dynamic imaging of a population of *E. coli* MG1655 cells expressing HupA-GFP to label the DNA demonstrated that most cells divided normally, whereas elongated cells arose rarely and were accompanied by hampered DNA segregation (electronic supplementary material, video S1). This hints at cell division failure and DNA replication–segregation coupling as a potential cause for the observed variability of cell lengths. However, because the population distributions analysed from fixed cell microscopy are taken from unsynchronized bulk cultures, we proceeded to examine if the cell-cycle stage does indeed affect our analysis, using live cells in continuous culture.

3.2. Single-cell analysis of lengths of newborn and dividing cells in microfluidics

The 'mother machine' microfluidics device described previously by Wang *et al*. [19] is ideally suited for single-cell analysis of rod-shaped cell growth dynamics. We capture birth and division events and estimate cell lengths (figure 2*a*) from timeseries of cells grown in LB (electronic supplementary material, video S2) and M9 + succinate (electronic supplementary material, video S3) at 37°C. The frequency distribution of the cell lengths from single-cell analysis also fit a lognormal distribution (figure 2*b–e*), similar to the fixed-cell data, with the goodness of fit validated by the KS Test (electronic supplementary material, table S3). While the variance of cell lengths in LB showed a difference between newborn cells (figure 2*b*) and cells just prior to division (figure 2*c*), the mean cell length of newborn cells was also smaller ($\mu = 2.7\,\mu\mathrm{m}$, $v = 0.9\,\mu\mathrm{m}^2$) than dividing cells ($\mu = 5.28\,\mu\mathrm{m}$, $v = 3.29\,\mu\mathrm{m}^2$). As a result, the normalized variability measured by the coefficient of variation of cell lengths ($\mathrm{CV_L}$) remained constant for cells grown in LB—0.3188 for newborn cells (arithmetic mean 2.73 µm, s.d. 0.87 µm) and 0.3129 for dividing cells (arithmetic mean 5.27 µm, s.d. 1.65 µm). This suggests that population cell length variability is independent of cell growth stage, based on the two extreme cases, i.e. newborn and dividing cells in the same growth medium. Compared with LB, cell length distributions of cultures grown in M9 + succinate have a narrower spread in data from both newborn (figure 2*d*) and dividing (figure 2*e*) cells. The $\mathrm{CV_L}$ of these cells is 0.156 (newborn) and 0.139 (dividing), twofold smaller than those measured in LB, confirming the qualitative impression. This suggests that the cell length variability as measured by $\mathrm{CV_L}$ from single-cell experiments is independent of the cell-cycle stage in a given medium, while

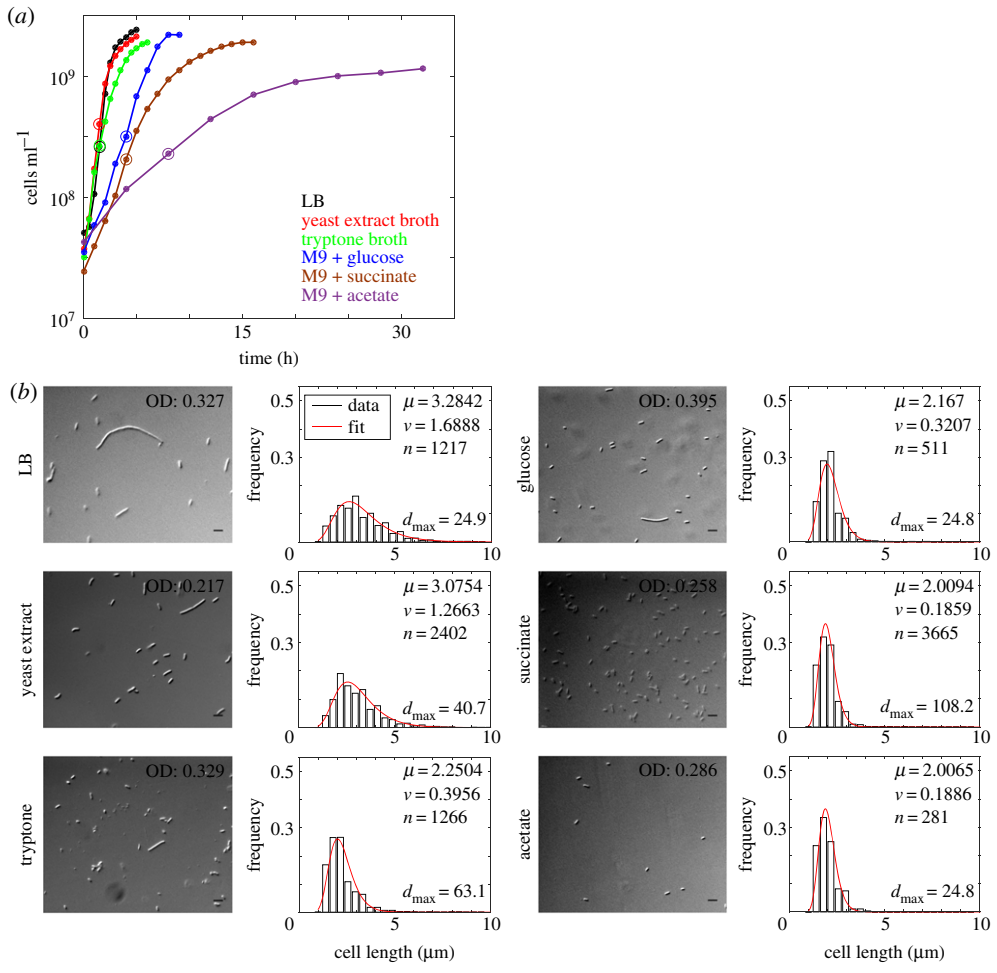


Figure 1. Growth rate and cell size distributions. (*a*) The growth ($\log_{10}$ cell density) of *E. coli* MG1655 cells was measured as a function of time in the following growth media: LB (black), yeast extract broth (YEB) (red), tryptone broth (TB) (green), M9 media supplemented either with glucose (blue), succinate (brown) or acetate (purple). Samples taken from the mid-log phase (circles) (*b*) (left) were examined in DIC microscopy (scale bar, 5 μm) and (right) the cell length distribution (bars) plotted. The distribution was fit to lognormal (red) where μ: mean cell length, v, variance; n, total number of cells. d_{max} is the Kolmogorov–Smirnov test value evaluated for the fit.

Table 1. The doubling times and growth rates of *E. coli* MG1655 were estimated from fitting the logistic equation to average ($n = 3$) OD measurements with time. Cultures were grown at 37°C with constant shaking.

growth medium	doubling time, t_d (min)	growth rate, r (h^{-1})
LB	33.11	1.814
yeast extract broth	39.87	1.505
tryptone broth	57.2	1.049
M9 + 0.4% glucose	63.13	0.9504
M9 + 0.9% succinate	131.98	0.4546
M9 + 0.5% acetate	273.35	0.2195

changing growth rates lead to measurable differences in the CV_L. To validate this finding, we also examine the growth of microcolonies, which form natural populations.

3.3. Growth-rate dependence of cell size variability in microcolonies

Agar-pad-based single-cell dynamics are routinely used to examine cell division dynamics in *E. coli*. We followed the growth dynamics of single cells, as they formed microcolonies at 37°C on LB agar

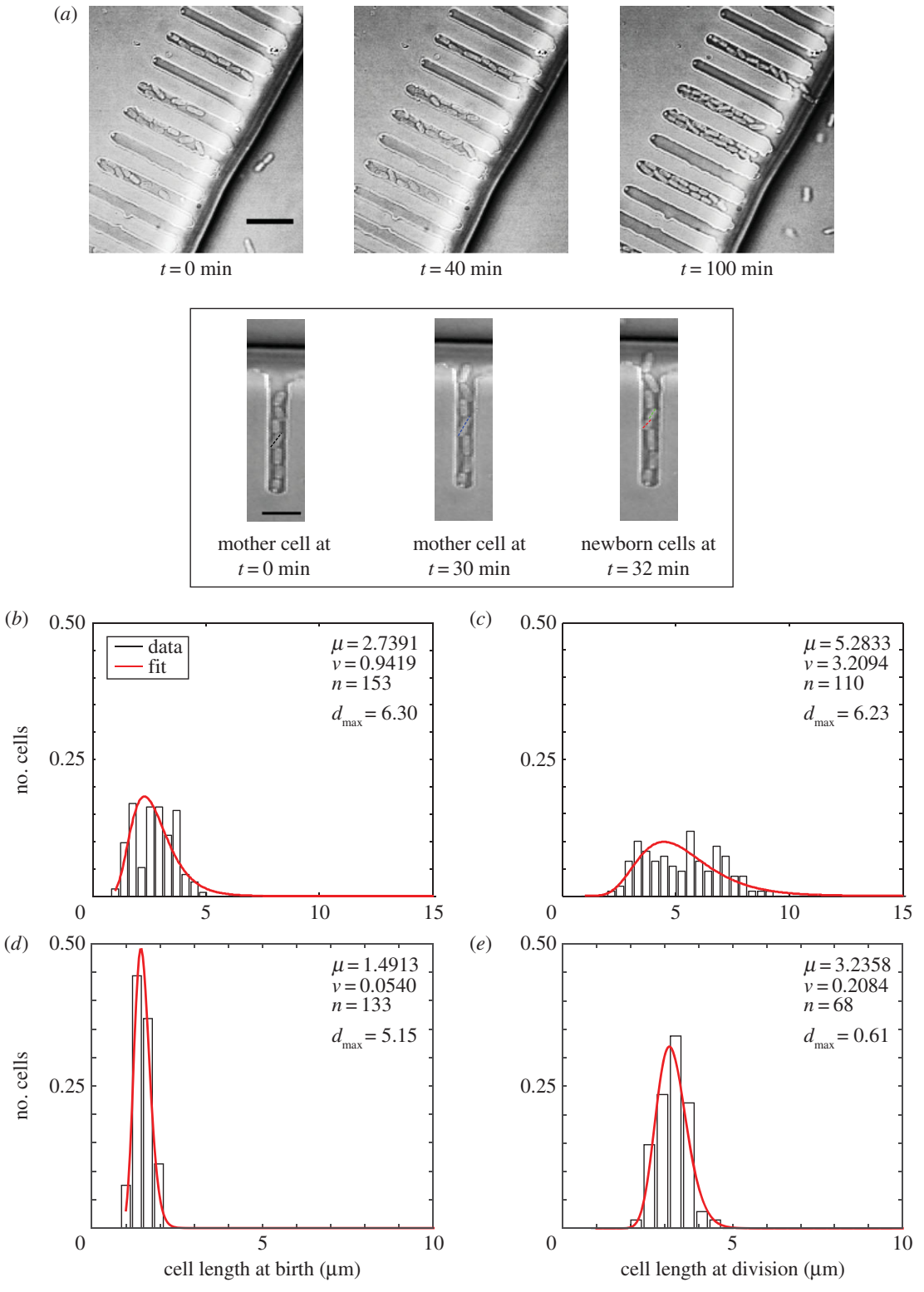


Figure 2. Cell length distribution of wild-type *E. coli* in continuous culture. (*a*) Representative DIC images of *E. coli* MG1655 cells grown in continuous culture in a 'mother machine' device with LB were recorded at 0, 40 and 100 min (also electronic supplementary material, video S3). Scale bar, 10 µm. Inset: a representative channel is marked to indicate a mother cell (black line), which grows in 30 min (blue line) and divides into two daughter cells at 32 min (green and red lines). The cell length distributions of cells grown in two growth media: (*b,c*) LB and (*d,e*) M9 + succinate were measured at birth (*b,d*) and division (*c,e*) and the frequency distributions are fit by lognormal distributions (red) with parameters μ and v and the goodness of fit was measured by d_{max} (evaluated for the KS test, electronic supplementary material, table S1).

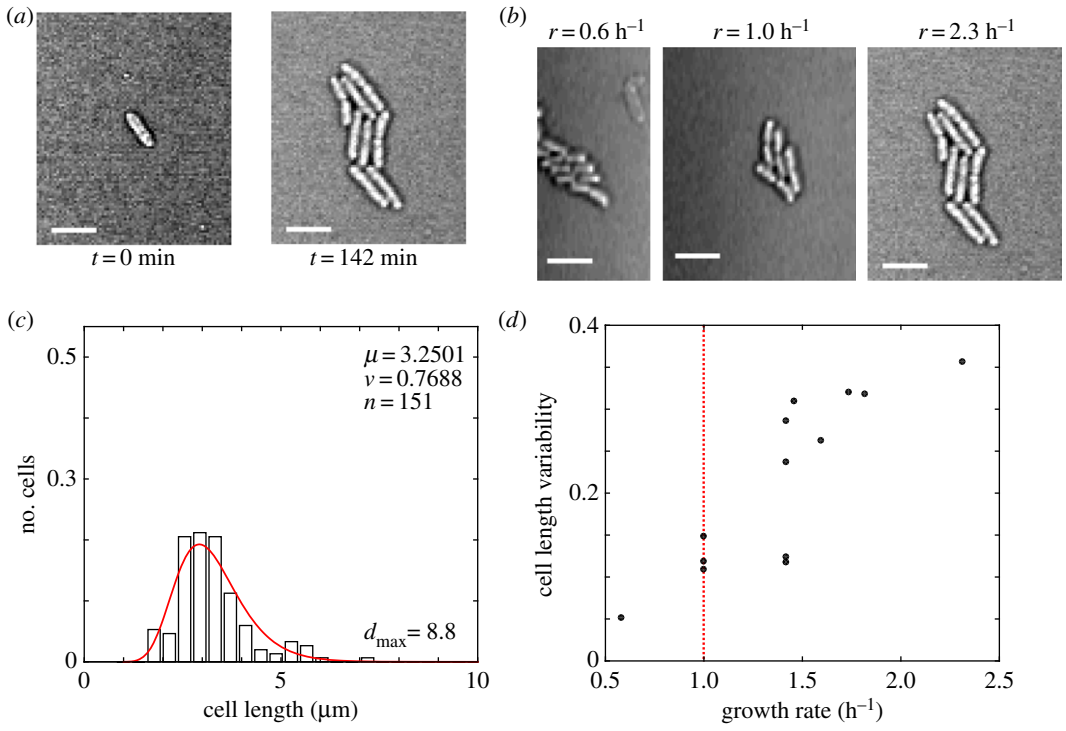


Figure 3. Population cell size variability in a microcolony. (*a*) The growth from a single cell to a microcolony at the end of 140 min is depicted. (*b*) Multiple microcolonies originating from a single cell after 140 min of growth are of different sizes, indicating differences in growth rates. Scale bar is 5 µm. (*c*) The cell length frequency distribution (bar) pooled from 13 microcolonies is fit by a lognormal distribution (red). μ, mean; v, variance; n, number of cells analysed, d_{max}, the KS test statistic measure. (*d*) The variability of cell lengths in a microcolony measured by the CV is plotted as a function of the average microcolony growth rate (h^{-1}). The dotted vertical line indicates a growth rate of $r_{mf} = 1\,h^{-1}$.

for 140 min (figure 3*a*; electronic supplementary material, video S4). Consistent with previous reports of growth rate heterogeneity in single cells [48], we find that growth rates vary in the range of 0.6–2.3 h^{-1} in the population (figure 3*b*). Each colony examined originates from a single cell, and hence at the end of 140 min when microcolony sizes vary, the cell–cell variation in growth rates is confirmed. The population cell length distribution of these microcolonies also fit a lognormal function (figure 3*c*), with the goodness of fit evaluated using the KS test (electronic supplementary material, table S3). To our surprise, the CV_L from individual microcolonies appeared to increase with increasing growth rate (figure 3*d*). Because the sample size in each CV_L measurement of a single microcolony is very small, and the growth-rate difference between single cells is difficult to control and is possibly the result of intrinsic stochastic variability, we instead proceeded to modulate average growth rate by the nutrient medium and analyse cell size variability in fixed cell microscopy, to take advantage of better population statistics.

3.4. Cell length variability affected by a combination of growth rate, *recA* and hydroxyurea

Cell length variability was quantified in the mid-log phase of cells grown in media resulting in growth rates ranging between 0.2 and 1.81 h^{-1} by using LB, the reduced media YEB, TB and M9 supplemented with sugars, to modulate growth rates (table 1). We find that cells grown in M9 supplemented with glucose, succinate and acetate are less variable ($CV_L < 0.25$), and the variability increases gradually with increasing growth rate (figure 4*a*). When the growth rate (r) exceeds 1 division per hour (in TB, YEB and LB), the CV_L appears to enter a second phase of a steeper increase. This inflection point also correlates with the growth rate threshold for multi-fork (mf) replication ($r_{mf} = 1\,h^{-1}$) [16]. Because multi-fork replication changes the genomic content per cell (G), we used a previously developed expression relating G with the BCD cycle [49] and doubling times, to estimate it as

$$G = \frac{t_d \cdot (2^{(C+D)/t_d} - 2^{D/t_d})}{C \cdot \ln 2}, \tag{3.1}$$

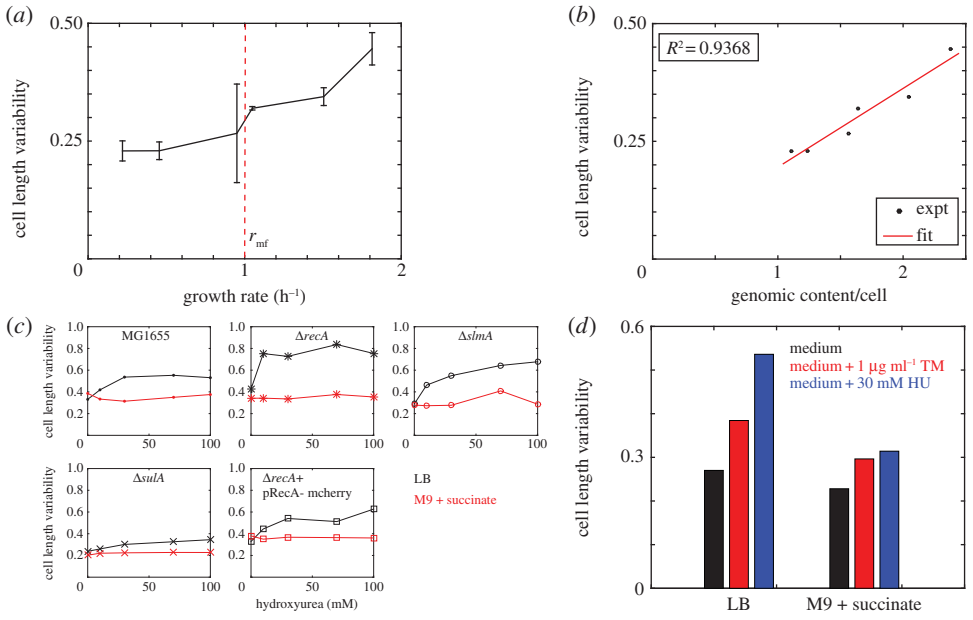


Figure 4. Effect of growth rate and replication stochasticity on cell lengths. (*a*) The cell length variability from the mid-log phase of cultures (*y*-axis) is plotted as a function of the growth rate (*x*-axis). The growth rate of $r_{mf} = 1\,h^{-1}$ (red line) is the multi-fork replication threshold based on Cooper & Helmstetter [16]. Error bars indicate s.d. (*b*) The measured cell length variability (*y*-axis) is plotted as a function of the expected total cellular DNA (*x*-axis) based on the model of Zaritsky *et al*. [49] (equation (3.1)). (*c*) Cell length variability of *E. coli* (*y*-axis) with increasing hydroxyurea (HU) concentration (*x*-axis) was estimated for the following strains: MG1655 (dots), Δ*recA* (asterisks), Δ*slmA* (circles), Δ*sulA* (x) and Δ*recA* + pRecA-mCherry (open squares). Cultures were grown either in LB (black) or M9 + succinate (red). (*d*) The cell length variability of *E. coli* MG1655 populations (*y*-axis) grown in LB or M9 + succinate (black) and compared with cells grown in the same medium but treated with 1 μg ml^{-1} of trimethoprim (TM) (red) or 30 mM HU (blue).

where t_d is the doubling time, *C* is the period of chromosome (DNA) replication and *D* is the time for cell division (septum formation). We combine the experimentally measured doubling times for different growth media (table 1) with an assumed *C*-period of 40 min and *D*-period of 20 min [16,50]. We find that the measured CV_L is positively correlated to increasing values of the estimated average genome content (*G*) per cell (figure 4*b*). This correlation demonstrates that genome content and cell size regulation could be coupled.

Because the genome content dependence of CV_L does not, however, reproduce a biphasic, threshold dependence in cell length variability (figure 4*a*), we examined whether perturbing replication dynamics below and above the presumptive growth rate threshold could be used as a test of replication stochasticity as the underlying mechanism. HU is known to induce stochastic replication fork stalling [51,52] and the RecA protein is critical for restarting stalled replication forks [31]. Expectedly, cells mutant for *recA* have a reduced ability to recover stalled replication forks [35]. To our surprise, on treatment with HU, both wild-type and Δ*recA* cells showed an increase in CV_L when grown in LB, but not when grown in M9 + succinate (figure 4*c* and electronic supplementary material, figure S5). Thus, slow growth appears to protect cells from the cell-division defects of HU, but rapid growth induces an increase followed by saturation in cell length variability. *E. coli* Δ*recA* mutants, however, differ from wild-type, because the difference of CV_L between LB and M9 + succinate was more pronounced. On the other hand, minimal medium-grown cells lacking *sulA*, 'the effector' of RecA, are less variable when compared with wild-type and do not respond to HU treatment. The treatment of LB and minimal medium-grown Δ*slmA* and Δ*recA* cells expressing RecA-mCherry cells results in variability comparable to MG1655. The 'rescue' of the wild-type phenotype by expression of RecA in a mutant background and the variability of Δ*sulA* mutant suggest the specificity of the RecA-SulA mechanism for growth-rate-dependent regulation of cell size. Additionally, the increase and saturation of cell size variability on treatment with HU of wild-type cells only during rapid but not slow growth further validate the threshold growth rate dependence of cell length variability. The growth-rate-dependent increase in cell length variability with trimethoprim treatment (figure 4*d*), a drug known to increase replication fork stalling [53], reinforces replication stochasticity as a mechanism regulating cell size. Based on this, we hypothesize that, in addition to

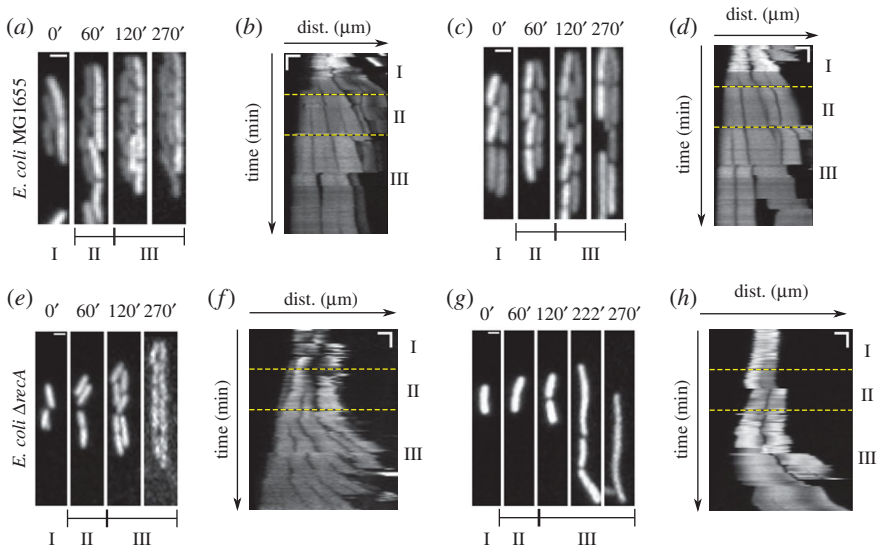


Figure 5. Effect of hydroxyurea (HU) on *E. coli* strain expressing eGFP grown in the 'mother machine'. (*a,c*) Images of *E. coli* MG1655 and (*e,g*) *E. coli* Δ*recA* mutants were both grown at 37°C in the 'mother machine' with three phases in the nutrient supplied I: LB ('pre-treatment'), II: LB + 30 mM HU ('treatment') and III: LB ('recovery'). Scale bar is 2 µm. (*b,d,f,h*) Kymographs of the preceding timeseries are used to follow cell division. The yellow horizontal lines mark the three phases. Time is in minutes. Horizontal scale bar is 3 µm and the vertical scale bar is 25 min.

its previously known roles, RecA recruitment to replication fork stalls [54] could mechanistically relate growth rate with cell size variability.

3.5. Single-cell dynamics of *Escherichia coli* MG1655 and Δ*recA* with hydroxyurea treatment

Based on this evidence from fixed cells, we expect that HU treatment should affect single-cell dynamics in a manner similar to the effect at a population level. To test this, *E. coli* MG1655 cells expressing eGFP from a plasmid were grown in the 'mother machine' in LB with the medium changed in three stages: 1 h pre-treatment (stage I), 1 h 30 mM HU treatment (stage II) and 2.5 h recovery from HU (stage III). Most cells continued to divide normally (figure 5*a*) and a few appeared to undergo moderate filamentation (cell lengths approx. 7 µm) at the end of the 'recovery' period (figure 5*c*). The kymographs suggest that the division of some cells is unaffected (figure 5*b*), while a few undergo filamentation (figure 5*d*). Most *E. coli* Δ*recA* cells continued to divide normally after recovery from treatment (figure 5*e*), whereas others became prominently filamentous, resulting in cell lengths of approx. 12 µm (figure 5*g*). The kymographs confirmed regular divisions of most cells (figure 5*f*), while a failure of division in some resulted in cell filamentation (figure 5*h*). This qualitatively corroborates our observations from population measurements that (i) cell division failure results in elongated cells, (ii) cell filamentation is probabilistic and (iii) the extent of cell filamentation is greater in Δ*recA* cells owing to a more extreme filamentation phenotype, when compared with *E. coli* MG1655. From population and single-cell dynamics, we hypothesize that cells that rapidly divide and are treated with HU are expected to have a higher frequency of filamentation and greater variability owing to increased RecA recruitment to the DNA. We proceed to test this hypothesis using microscopy.

3.6. Nucleoid localization of recA corresponds to increased cell length variability

To test the hypothesis of growth rate-dependent recruitment of RecA to the genome, an *E. coli* MG1655 strain expressing an endogenous RecA-GFP protein [36] was grown in three different media: LB, YEB and TB, to modulate growth rates as before. RecA foci co-localization with the nucleoid (labelled with DAPI) appeared to increase when cells were grown in LB when compared with YEB and TB (figure 6*a*). Treatment with HU resulted in increased co-localization when compared with untreated cells grown in the same growth medium (figure 6*a*). The proportion of cells with co-co-localization of the RecA protein on the nucleoid is low (less than 15% cells with RecA-DNA co-localization) during slow growth (TB), intermediate (approx. 15%) for cells in YEB and high (approx. 30%) in LB during

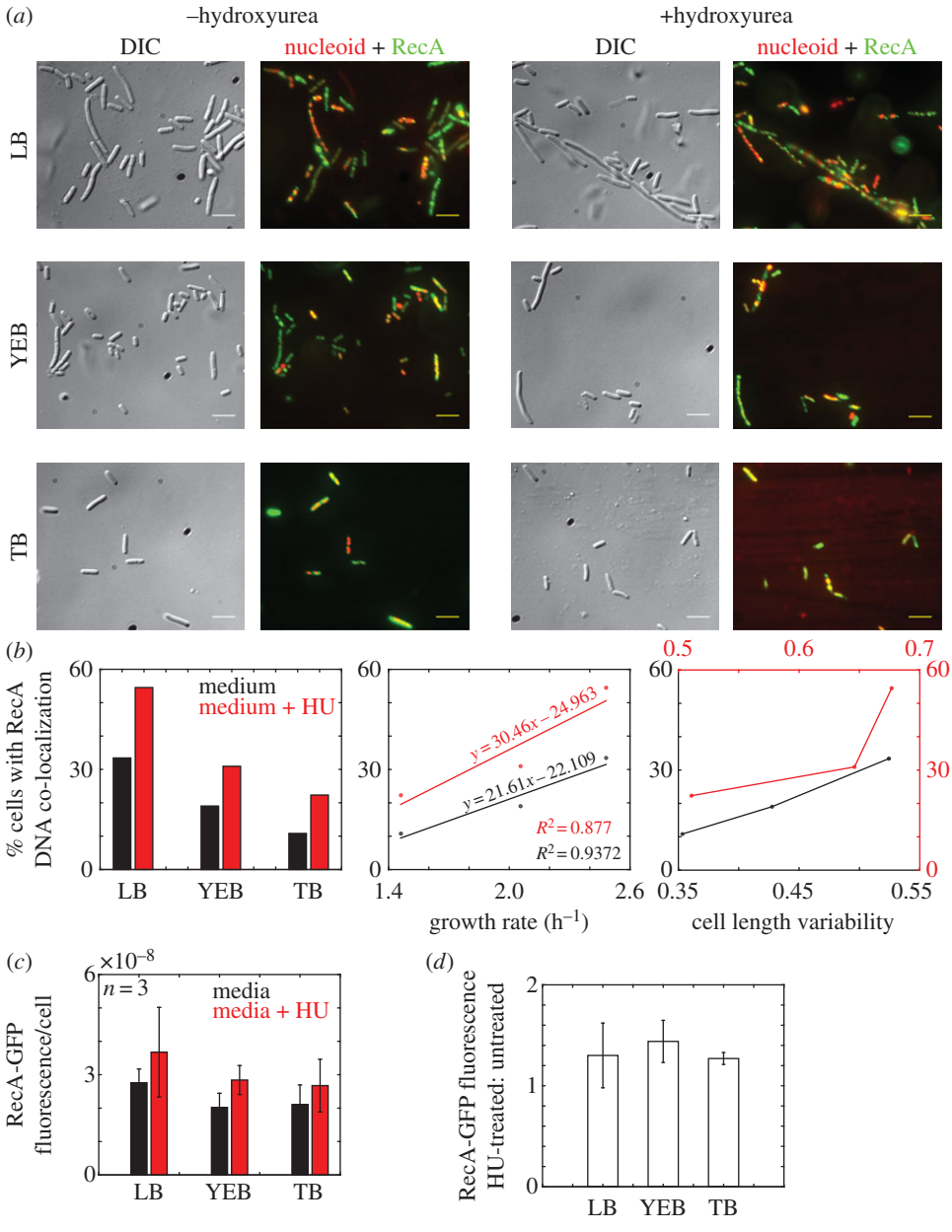


Figure 6. Growth rate dependence of RecA co-localization with nucleoids. (*a*) Mid-log cultures of *E. coli* expressing RecA-GFP (genomic) grown in LB, YEB and TB with (+) and without (−) 30 mM hydroxyurea (HU) were imaged in DIC (grey), and the fluorescence. The merged image of DAPI (red) and GFP (green) fluorescence with co-localization (yellow) is shown. Scale bar, 5 µm. (*b*) The proportion of *E. coli* cells in which RecA foci co-localize with nucleoids (*y*-axis) is plotted against three growth media (*x*-axis) with (red) and without (black) HU, growth rate (h^{-1}) (*x*-axis) and cell length variability (*x*-axis). (*c*) Mean (±s.d. from three samples) RecA-GFP fluorescence/cell from fluorimetry of liquid cultures is plotted as a bar chart for untreated (black) and HU-treated (red) cells grown in three media: LB, YEB and TB. (*d*) The ratio of RecA-GFP fluorescence of HU treated to untreated cells grown in LB, YEB and TB are compared.

rapid growth (figure 6*b*). The co-localization percentages are further increased on HU treatment. This is evidenced by the percentage cells with co-localized RecA-GFP and DNA increasing linearly with growth rate for untreated cells, with a 1.4-fold increase in the slope for HU-treated cells (figure 6*b*). The percentage co-localization of treated and untreated cells correlates positively with CV_L, consistent with our previous results of increasing cell length variability with growth rate and HU treatment. As before, the variability of HU-treated cells is higher for the same growth medium, when compared with untreated cells. To test if the co-localization of RecA with the nucleoid was not the result of an artefact of GFP tagging of RecA, the native RecA protein was immunostained in fixed cells of *E. coli* MG1655 and analysed for co-localization with DNA by DAPI staining in multiple fields of view

(electronic supplementary material, figure S6*a*). The percentage cells with RecA co-localized on the nucleoid matched the values from the RecA-GFP co-localization (electronic supplementary material, figure S6*b*). In addition, the measured cell length variability from the DIC images was also comparable between the RecA-GFP-expressing strain and the immunostained samples (electronic supplementary material, figure S6*c*). Alternatively, we also imaged the localization of RecA-mCherry protein expressed from a plasmid in *E. coli* Δ*recA* cells grown in LB (electronic supplementary material, figure S7*a*) and found the RecA-nucleoid co-localization to be comparable to anti-RecA and RecA-GFP-based quantification (electronic supplementary material, figure S7*b*). The CV_L of all three samples remained comparable (electronic supplementary material, figure S7*c*), suggesting that RecA protein tagging did not result in artefacts in either localization or cell length variability. To test if the increased co-localization could have resulted from RecA protein abundance instead of recruitment, we measured the RecA-GFP concentration per cell in fluorimetry and found that HU-treated cells had a slightly increased (approx. 1.25-fold) RecA-GFP expression per cell, for all three growth media tested (figure 6*c*). Western blotting using anti-RecA antiserum of cells grown in LB, YEB and TB with and without treatment also appear to show increased protein in cell lysates on HU treatment, independent of the growth medium (electronic supplementary material, figure S8*a*). Quantification showed an approximately 1.5- to threefold increase in RecA (electronic supplementary material, figure S8*b*) and RecA-GFP intensity after HU treatment, in all growth media tested, not owing to loading artefacts (electronic supplementary material, figure S8*c*,*d*). Additionally, the expression of RecA-GFP from the genomic locus also increased by 2.5-fold on HU treatment (electronic supplementary material, figure S8*e*,*f*), with non-specific protein content remaining comparable (electronic supplementary material, figure S8*g*,*h*). Taken together, the results suggest that RecA recruitment to the nucleoid could serve as one of multiple mechanisms that drive growth rate-dependent cell size variability in clonal populations, independent of RecA protein abundance.

4. Discussion

While the study of average values of cell size [55,56] and single-cell studies [19,26,48] have demonstrated that cell size is robust to environmental changes, understanding the population distribution remains important to the ecology of microbes and their survival in changing environments [57,58].

Here, we have examined the population variability of clonal cell sizes and their link to growth rate. We demonstrate that at a fixed growth rate, the cell length variability is constant and independent of the cell-cycle stage using a microfluidics-based continuous culture system. Additionally, the growth rate of microcolonies on an agar pad appears to correlate with cell-cycle synchronized variability in cell lengths. However, in our analysis, quantitative fixed cell microscopy from bulk cultures results in better statistics and a more robust control over growth rates. We find increasing growth rates increases the population variability in cell lengths in a bi-phasic manner, with the two phases separated by a growth rate threshold, r_{mf} (the growth rate of multi-fork replication). HU treatment, known to induce DNA replication fork stalling, increases the cell length variability of only those cells which are undergoing rapid growth. This HU-induced variability is further enhanced in a ΔrecA mutant when compared with wild-type. The linear increase in RecA co-localization on the nucleoid with growth rate and cell length variability suggests the involvement of the RecA protein in coupling increasing growth rates to higher cell length variability.

In general, stochastic partitioning of subcellular components has been shown to be a major source of cell phenotypic variability in a study combining theory and experiment [23]. Our observations on the role of stochastic replication dynamics potentially add to the potential contributors to phenotypic variability or 'noise'. We expect stochastic DNA replication–segregation effects on cell size, should result in cell elongation owing to incomplete segregation of DNA. In agar-pad growth experiments in DIC and fluorescence of *E. coli* MG1655 with nucleoids labelled by HupA-GFP, most cells divide to produce newborn cells with typical birth lengths of approximately 2 µm after successfully segregating their nucleoids in approximately 20 min (electronic supplementary material, figure S4*a* and video S1). On the other hand, rare cells become approximately 40 µm long, after their nucleoids fail to segregate even after 50 min (electronic supplementary material, figure S4*b* and video S1). Growing wild-type and mutant strains in multiple growth media with HU and trimethoprim, we show that cell length variability is greater when replication processivity is perturbed, depending on the growth rate.

Cell size is a complex phenotype and is influenced by multiple pathways such as nutrient sensing [9–11,59], the division site selection by the minCDE proteins [60,61], nucleoid occlusion to sense

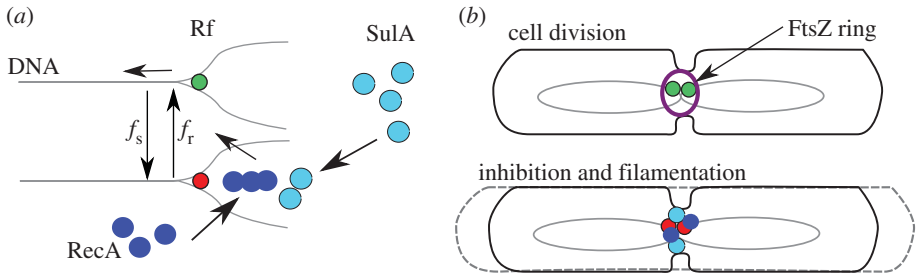


Figure 7. A proposed model of RecA recruitment and cell length regulation. (*a*) The stalled DNA-replication forks (red) result in the recruitment of RecA (dark blue) to the DNA and trigger SulA expression (light blue). The transition of recovered (green) and stalled replication forks is stochastic (determined by the frequency of stalling, f_s) and reversible (determined by the frequency of recovery, f_r). RecA filament assembly promotes increased recovery (f_r). (*b*) In normal cell division, replication forks are in a recovered state and FtsZ assembles at the septum. Recruitment of RecA results in SulA-based inhibition of division and cell filamentation (long cell).

incomplete replication [51,62] and the SOS response pathway [34,35]. However, this is the first study, to the best of our knowledge, that proposes a mechanism connecting growth rate with cell length variability based on multi-fork replication. It remains to be seen if a more direct method of replication fork tracking [30] can be used to test this proposed mechanism.

In previous studies on mammalian cell size regulation, the statistics of fixed cells were used to estimate the 'variability' [63]. While many older studies on bacterial cell length regulation [2,64] made use of such an approach, the advent of single-cell approaches have improved the robustness and accuracy of *E. coli* cell size and division measurements [18,19]. However, in the process, the population effects have been ignored. In this work, we attempt to bridge this gap. Additionally, we test our method of population variability measurement for artefacts that could result from a lack of cell cycle stage synchronization of the population. Indeed by measuring the growth of populations under different growth conditions, we find that the growth rate dependence of cell size variability is related to genome copy numbers per cell and is independent of synchronization.

We find the genome-copy number per cell, driven by growth rate, appears to positively correlate with cell size distributions in the population. An analogous study in yeast has examined the effect of the ploidy of specific genes on cell size regulation [65]. In the case of this study, while the genome copies per cell increase cell size variability, we have not estimated the possible role of specific genes and their ploidy on cell sizes. Additionally, our microcolony analysis once more reveals growth rate differences between clonal individuals under identical conditions (figure 3*d*). While these differences are not addressed in our study, it would be useful to extend our current analysis to the possible role of specific genes and proteins in single-cell growth rate variability.

The RecA protein, examined in this study for a relationship with growth-rate-dependent cell size variability, is an SOS response pathway protein. It has previously been shown to enhance the recovery of stalled replication forks [29,66]. At the same time, when RecA is recruited to DNA, it activates SulA, which sequesters FtsZ monomers [34], thus acting as a cell division inhibitor. Fast-growing *E. coli* are also known to initiate multiple replication forks [16,67]. DNA replication fork progression is known to be stochastic [68]. The increase in cell length variability that we observe as a function of growth rate in wild-type *E. coli* can thus be explained by a model where RecA recruitment to stalled replication forks (figure 7*a*) leads to an increase in the proportion of elongated cells owing to cell-division inhibition (figure 7*b*). The cause for the onset of this process during rapid growth, we hypothesize, results from the probabilistic replication fork stalling (figure 7*a*) and the multiplicative effect owing to multi-fork replication [16]. While a computer simulation of multi-fork replication in the *E. coli* cell cycle exists [49], an explicit model of replication stochasticity coupled to multi-fork replication dynamics could help further test our hypothesis.

Our observations of the role of RecA in cell size variability have two parts: (i) the effect of a functional copy of *recA* and (ii) the effect of a *recA* deletion. (i) We observed population cell length variability of *E. coli* to increase with growth rate (figures 1–3) and tested the hypothesis that increasing RecA localization on the genome with increasing growth rate correlated with this growth rate dependence (figure 6). We find that RecA-SulA-mediated cell division inhibition increases cell length variability in rapid growth in wild-type cells (figure 4). Additionally, in rapidly growing cultures of the wild-type (LB, 37°C), overexpression of the RecA protein from the pRecA-mCherry (electronic supplementary

material, figure S9*a*) and arabinose-inducible pBAD24-RecA constructs (electronic supplementary material, figure S9*b*) did not affect cell length variability (electronic supplementary material, figure S2*a*,*b*). In cells missing a copy of recA, treatment with 30 mM HU increased the variability of cell lengths (electronic supplementary material, figure S3*a*). The overexpression of plasmid-based copy of recA in these cells from pRecA-mCherry and pBAD24-recA plasmids reduced the spread of cell lengths to levels comparable to untreated MG1655 (electronic supplementary material, figure S3*b*). (ii) As reported previously, in the absence of RecA, the function of DNA replication fork stall rescue is hindered [31], resulting in replication defects [66] which are not repaired owing to the inability of the cell to induce an SOS response through LexA cleavage [69]. This results in cell division inhibition potentially owing to nucleoid occlusion [70,71] and additional RecA-independent pathways that detect incomplete replication [51]. However, distinct from previous work, we find greater variability in cell lengths in a Δ*recA* strain when compared with hydroxyurea-treated wild-type cells in LB (figure 4*c*). We also demonstrate that the phenotype can be rescued by plasmid-based expression of the RecA protein, both with and without a fluorescent tag (figure 4*c* and electronic supplementary material, figure S2*a*,*b*). In future, our data could form the basis of a mathematical model, extending a previously developed model of *recA* gene expression dynamics during UV-based damage [72], focusing instead on the effect of growth rates.

HU treatment increases replication fork stalling and slows down DNA replication [73] and also results in cell division failure. Stalled replication forks that are not repaired, result in incomplete DNA replication [29] and result in elongated cells owing to cell division inhibition [12,13]. A direct measurement of replication fork dynamics of a population of dividing bacteria based on methods used to study the single molecule replication dynamics [30,74] could be used to potentially test our predictions at a subcellular level. Additionally, recently developed artificial 'replication roadblocks' in *E. coli* [30] and *Bacillus subtilis* [62] could be used in future to generate known number of replication fork stalling events and quantify their effect on population cell lengths, as a further test of the model.

The role of small proportions of outliers or 'tails' in phenotypic variability of a population has been shown to confer advantages to 'persister' cells, when the population undergoes selection [57,58,75]. However, a clear functional role for cell lengths is yet to be unambiguously determined. Suggestive evidence from clinical isolates of uropathogenic *E. coli* have implied that filamentous cells are harder for immune cells to clear than those of normal length [76]. In future, a study of the possible role of cell size and shape of not just *E. coli* in their natural environment could shed more light on the possible role in cell survival.

Our results suggest that an increased genome copy numbers in *E. coli* increases cell size heterogeneity. This is consistent with single-cell measurements in *E. coli* [30], but in contrast with *S. cerevisiae*, which shows that increased genome copy numbers lead to reduced 'noise' in cell size distributions [65]. To infer general principles from these results, the distribution of DNA replication origins in yeast and the concurrent nature of replication in bacteria will need to be taken into account. It can be presumed that the effect we report will only occur in organisms where rapid growth entails multiple simultaneous rounds of DNA replication.

5. Conclusion

In conclusion, we find that cell length variability of wild-type *E. coli* increases with increasing growth rate in a non-monotonic manner above a growth rate threshold for multi-fork replication (r_{mf}). This variability is independent of cell-cycle stage synchronization of the population. Increasing HU concentrations to modulate replication stochasticity only changes the cell length distributions of populations undergoing rapid growth, an effect amplified in Δ*recA* cells. The rescue of increased cell length variability by RecA expression in deletion mutants, the effect of replication stalling induction on cell lengths and recruitment of RecA to the DNA, all indicate a model of stochastic multi-fork replication involving SOS response proteins. This could provide a mechanistic explanation of how growth rate affects population cell length variability in *E. coli*.

Research ethics. The Institutional Biosafety Committee (IBSC) of the IISER Pune instituted under the aegis of the Department of Biotechnology, Government of India, has approved of the biological safety aspects of this work in 2015 and is valid.

Animal ethics. No human or animal subjects were used in this study.

Data accessibility. The software code developed as a part of this manuscript has been released as Open Source software and can be freely downloaded from https://github.com/athale/ecolilenDIC. The link to the program is also available

from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.2bs69. All experimental data are included in this manuscript and the accompanying electronic supplementary materials [77].

Authors' contributions. M.S.G. performed all the experiments, image analysis and statistics and made all the figures. C.A.A. designed the microfluidics chip, developed the image analysis code, designed the study and wrote the paper. Both authors gave their final approval for publication.

Competing interests. We have no competing interests.

Funding. M.S.G. is supported by a fellowship from ICMR 3/1/3/WLC/JRF-2011/HRD-156 (51550). IISER Pune and a grant from the Department of Biotechnology, Government of India (BT/PR1595/BRB/10/1043/2012) supported C.A.A. The Indian Nanoelectronics User Programme (INUP), IIT Bombay supported the microfluidics work with a short-term grant to C.A.A.

Acknowledgements. We acknowledge the gift of the HupA-GFP construct by Dr Josette Rouviere-Yaniv and the RecA-GFP-expressing strain provided by Dr G.P. Manjunath. The Coli Genomic Stock Centre (CGSC), Yale University provided the remaining strains. We are grateful to Prof. K. Muniyappa (IISC Bangalore, India) for the kind gift of the anti-RecA rabbit serum. Dr Suckjoon Jun (UCSD), Dr K. Nageshwari (IIT Bombay) and Dr Siddharth Deshpande (TU Delft) are acknowledged for help with the microfluidics. Hemangi Chaudhari and Shraddha Shitut were involved in the early stages of the project.

References

1. Grover NB, Woldringh CL. 2001 Dimensional regulation of cell-cycle events in Escherichia coli during steady-state growth. *Microbiology* **147**, 171–181. (doi:10.1099/00221287-147-1-171)
2. Cullum J, Vicente M. 1978 Cell growth and length distribution in *Escherichia coli*. *J. Bacteriol.* **134**, 330–337.
3. Kaya T, Koser H. 2009 Characterization of hydrodynamic surface interactions of *Escherichia coli* cell bodies in shear flow. *Phys. Rev. Lett.* **103**, 138103. (doi:10.1103/PhysRevLett.103.138103)
4. Trueba FJ, Spronsen EA, Traas J, Woldringh CL. 1982 Effects of temperature on the size and shape of *Escherichia coli* cells. *Arch. Microbiol.* **131**, 235–240. (doi:10.1007/BF00405885)
5. Schaechter M, Maaloe O, Kjeldgaard NO. 1958 Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *J. Gen. Microbiol.* **19**, 592–606. (doi:10.1099/00221287-19-3-592)
6. Mason MM. 1933 A comparison of the maximal growth rates of various bacteria under optimal conditions. *J. Bacteriol.* **29**, 103–110.
7. Powell EO. 1956 Growth rate and generation time of bacteria, with special reference to continuous culture. *J. Gen. Microbiol.* **15**, 492–511. (doi:10.1099/00221287-15-3-492)
8. Kubitschek HE. 1968 Linear cell growth in *Escherichia coli*. *Biophys. J.* **8**, 792–804. (doi:10.1016/S0006-3495(68)86521-X)
9. Yao Z, Davis RM, Kishony R, Kahne D, Ruiz N. 2012 Regulation of cell size in response to nutrient availability by fatty acid biosynthesis in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **109**, E2561–E2568. (doi:10.1073/pnas.1209742109)
10. Wang JD, Levin PA. 2009 Metabolism, cell growth and the bacterial cell cycle. *Nat. Rev. Microbiol.* **7**, 822–827. (doi:10.1038/nrmicro2202)
11. Weart RB, Lee AH, Chien A-C, Haeusser DP, Hill NS, Levin PA. 2007 A metabolic sensor governing cell size in bacteria. *Cell* **130**, 335–347. (doi:10.1016/j.cell.2007.05.043)
12. Veiga H, Jorge AM, Pinho MG. 2011 Absence of nucleoid occlusion effector Noc impairs formation of orthogonal FtsZ rings during *Staphylococcus aureus* cell division. *Mol. Microbiol.* **80**, 1366–1380. (doi:10.1111/j.1365-2958.2011.07651.x)
13. Norris V, Woldringh C, Mileykovskaya E. 2004 A hypothesis to explain division site selection in *Escherichia coli* by combining nucleoid occlusion and Min. *FEBS Lett.* **561**, 3–10. (doi:10.1016/S0014-5793(04)00135-8)
14. Wu LJ, Errington J. 2004 Coordination of cell division and chromosome segregation by a nucleoid occlusion protein in *Bacillus subtilis*. *Cell* **117**, 915–925. (doi:10.1016/j.cell.2004.06.002)
15. Allman R, Schjerven T, Boye E. 1991 Cell cycle parameters of *Escherichia coli* K-12. *J. Bacteriol.* **173**, 7970–7974. (doi:10.1128/jb.173.24.7970-7974.1991)
16. Cooper S, Helmstetter CE. 1968 Chromosome replication and the division of *Escherichia coli* B/r. *J. Mol. Biol.* **31**, 519–540. (doi:10.1016/0022-2836(68)90425-7)
17. Guberman JM, Fay A, Dworkin J, Wingreen NS, Gitai Z. 2008 PSICIC: noise and asymmetry in bacterial division revealed by computational image analysis at sub-pixel resolution. *PLoS Comp. Biol.* **4**, e1000233. (doi:10.1371/journal.pcbi.1000233)
18. Sliusarenko O, Heinritz J, Emonet T, Jacobs-Wagner C. 2011 High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Mol. Microbiol.* **80**, 612–627. (doi:10.1111/j.1365-2958.2011.07579.x)
19. Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, Jun S. 2010 Robust growth of *Escherichia coli*. *Curr. Biol.* **20**, 1099–1103. (doi:10.1016/j.cub.2010.04.045)
20. Locke JC, Elowitz MB. 2009 Using movies to analyse gene circuit dynamics in single cells. *Nat. Rev. Microbiol.* **7**, 383–392. (doi:10.1038/nrmicro2056)
21. Fisher JK, Bourniquel A, Witz G, Weiner B, Prentiss M, Kleckner N. 2013 Four-dimensional imaging of *E. coli* nucleoid organization and dynamics in living cells. *Cell* **153**, 882–895. (doi:10.1016/j.cell.2013.04.006)
22. Wakita J, Kuninaka H, Matsuyama T, Matsushita M. 2010 Size distribution of bacterial cells in homogeneously spreading disk-like colonies by *Bacillus subtilis*. *J. Phys. Soc. Jpn* **79**, 94002-1-5.
23. Huh D, Paulsson J. 2011 Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Gen.* **43**, 95–100. (doi:10.1038/ng.729)
24. Robert L, Hoffmann M, Krell N, Aymerich S, Robert J, Doumic M. 2014 Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism. *BMC Biol.* **12**, 17. (doi:10.1186/1741-7007-12-17)
25. Amir A. 2014 Cell size regulation in bacteria. *Phys. Rev. Lett.* **112**, 1–5. (doi:10.1103/PhysRevLett.112.208102)
26. Taheri-Araghi S, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, Vergassola M, Jun S. 2014 Cell-size control and homeostasis in bacteria. *Curr. Biol.* **25**, 385–391. (doi:10.1016/j.cub.2014.12.009)
27. Skarstad K, Boye E. 1993 Degradation of individual chromosomes in recA mutants of *Escherichia coli*. *J. Bacteriol.* **175**, 5505–5509. (doi:10.1128/jb.175.17.5505-5509.1993)
28. Athale CA, Chaudhari H. 2011 Population length variability and nucleoid numbers in *Escherichia coli*. *Bioinformatics* **27**, 2944–2948. (doi:10.1093/bioinformatics/btr501)
29. Cox MM, Goodman MF, Kreuzer KN, Sherratt DJ, Sandler SJ, Marians KJ. 2000 The importance of repairing stalled replication forks. *Nature* **404**, 37–41. (doi:10.1038/35003501)
30. Possoz C, Filipe SR, Grainge I, Sherratt DJ. 2006 Tracking of controlled *Escherichia coli* replication fork stalling and restart at repressor-bound DNA *in vivo*. *EMBO J.* **25**, 2596–2604. (doi:10.1038/sj.emboj.7601155)
31. Robu ME, Inman RB, Cox MM. 2001 RecA protein promotes the regression of stalled replication forks *in vitro*. *Proc. Natl Acad. Sci. USA* **98**, 8211–8218. (doi:10.1073/pnas.131022698)
32. Huisman O, D'Ari R, Gottesman S. 1984 Cell-division control in Escherichia coli: specific induction of the SOS function SfiA protein is sufficient to block septation. *Proc. Natl Acad. Sci. USA* **81**, 4490–4494. (doi:10.1073/pnas.81.14.4490)
33. Trusca D, Scott S, Thompson C, Bramhill D. 1998 Bacterial SOS checkpoint protein SulA inhibits polymerization of purified FtsZ Cell division protein. *J. Bacteriol.* **180**, 3946–3953.
34. Mukherjee A, Cao C, Lutkenhaus J. 1998 Inhibition of FtsZ polymerization by SulA, an inhibitor of septation in *Escherichia coli*. *Proc. Natl Acad. Sci USA* **95**, 2885–2890. (doi:10.1073/pnas.95.6.2885)

35. Aksenov SV. 1999 Induction of the SOS response in ultraviolet-irradiated *Escherichia coli* analyzed by dynamics of LexA, RecA and SulA proteins. *J. Biol. Phys.* **25**, 263–277. (doi:10.1023/A:1005163310168)
36. Renzette N, Gumlaw N, Nordman JT, Krieger M, Yeh SP, Long E, Centore R, Boonsombat R, Sandler SJ. 2005 Localization of RecA in *Escherichia coli* K-12 using RecA-GFP. *Mol. Microbiol.* **57**, 1074–1085. (doi:10.1111/j.1365-2958.2005.04755.x)
37. Wery M, Woldringh CL, Rouviere-Yaniv J. 2001 HU-GFP and DAPI co-localize on the *Escherichia coli* nucleoid. *Biochimie* **83**, 193–200. (doi:10.1016/S0300-9084(01)01254-8)
38. Chen W, Kuo TT. 1993 A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. *Nucleic Acids Res.* **21**, 2260. (doi:10.1093/nar/21.9.2260)
39. Sambrook J, Russell D. 2001 *Molecular cloning: a laboratory manual*, 3rd edn, 2344. Cold Spring Harbor, NY: Cold Spring Harbor Laboratories Press.
40. Chai NC, Lark KG. 1970 Cytological studies of deoxyribonucleic acid replication in *Escherichia coli* 15T-: replication at slow growth rates and after a shift-up into rich medium. *J. Bacteriol.* **104**, 401–409.
41. Sezonov G, Joseleau-Petit D, D'Ari R. 2007 *Escherichiacoli* physiology in Luria–Bertani broth. *J. Bacteriol.* **189**, 8746–8749. (doi:10.1128/JB.01368-07)
42. Schlegel HG. 1993 *General microbiology*. Cambridge, UK: Cambridge University Press.
43. Venkatesh R, Ganesh N, Guhan N, Reddy MS, Chandrasekhar T, Muniyappa K. 2002 RecX protein abrogates ATP hydrolysis and strand exchange promoted by RecA: insights into negative regulation of homologous recombination. *Proc. Natl Acad. Sci. USA* **99**, 12 091–12 096. (doi:10.1073/pnas.192178999)
44. Schneider CA, Rasband WS, Eliceiri KW. 2012 NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675. (doi:10.1038/nmeth.2089)
45. Birnbaum ZW, Tingey FH. 1951 One-sided confidence contours for probability distribution functions. *Ann. Math. Stat.* **22**, 592–596. (doi:10.1214/aoms/1177729550)
46. Zar JH. 2008 *Biostatistical analysis*, 4th edn. Delhi, India: Pearson Education Inc.
47. Trueba FJ, Woldringh CL. 1980 Changes in cell diameter during the division cycle of *Escherichia coli*. *J. Bacteriol.* **142**, 869–878.
48. Reshes G, Vanounou S, Fishov I, Feingold M. 2008 Timing the start of division in *E. coli*: a single-cell study. *Phys. Biol.* **5**, 46001. (doi:10.1088/1478-3975/5/4/046001)
49. Zaritsky A, Wang P, Vischer NOE. 2011 Instructive simulation of the bacterial cell division cycle. *Microbiology* **157**, 1876–1885. (doi:10.1099/mic.0.049403-0)
50. Skarstad K, Steen HB, Boye E. 1985 *Escherichia coli* DNA distributions measured by flow cytometry and compared with theoretical computer simulations. *J. Bacteriol.* **163**, 661–668.
51. Cambridge J, Blinkova A, Magnan D, Bates D, Walker JR. 2014 A replication-inhibited unsegregated nucleoid at mid-cell blocks Z-ring formation and cell division independently of SOS and the SlmA nucleoid occlusion protein in *Escherichia coli*. *J. Bacteriol.* **196**, 36–49. (doi:10.1128/JB.01230-12)
52. Cha RS, Kleckner N. 2002 ATR homolog Mec1 promotes fork progression, thus averting breaks in replication slow zones. *Science* **297**, 602–606. (doi:10.1126/science.1071398)
53. Slager J, Kjos M, Attaiech L, Veening JW. 2014 Antibiotic-induced replication stress triggers bacterial competence by increasing gene dosage near the origin. *Cell* **157**, 395–406. (doi:10.1016/j.cell.2014.01.068)
54. Simmons LA, Grossman AD, Walker GC. 2007 Replication is required for the RecA localization response to DNA damage in Bacillus subtilis. *Proc. Natl Acad. Sci. USA* **104**, 1360–1365. (doi:10.1073/pnas.0607123104)
55. Shehata TE, Marr AG. 1975 Effect of temperature on the size of *Escherichia coli* cells. *J. Bacteriol.* **124**, 857–862.
56. Trueba FJ, Neijssel OM, Woldringh CL. 1982 Generality of the growth kinetics of the average individual cell in different bacterial populations. *J. Bacteriol.* **150**, 1048–1055.
57. Kussell E, Kishony R, Balaban NQ, Leibler S. 2005 Bacterial persistence: a model of survival in changing environments. *Genetics* **169**, 1807–1814. (doi:10.1534/genetics.104.035352)
58. Gefen O, Gabay C, Mumcuoglu M, Engel G, Balaban NQ. 2008 Single-cell protein induction dynamics reveals a period of vulnerability to antibiotics in persister bacteria. *Proc. Natl Acad. Sci. USA* **105**, 6145–6149. (doi:10.1073/pnas.0711712105)
59. Hill NS, Buske PJ, Shi Y, Levin PA. 2013 A moonlighting enzyme links *Escherichia coli* cell size with central metabolism. *PLoS Genet.* **9**, e1003663. (doi:10.1371/journal.pgen.1003663)
60. Lutkenhaus J. 2007 Assembly dynamics of the bacterial MinCDE system and spatial regulation of the Z ring. *Annu. Rev. Biochem.* **76**, 539–562. (doi:10.1146/annurev.biochem.75.103004.142652)
61. Kerr RA, Levine H, Sejnowski TJ, Rappel W-J. 2006 Division accuracy in a stochastic model of Min oscillations in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **103**, 347–352. (doi:10.1073/pnas.0505825102)
62. Bernard R, Marquis KA, Rudner DZ. 2010 Nucleoid occlusion prevents cell division during replication fork arrest in Bacillus subtilis. *Mol. Microbiol.* **78**, 866–882. (doi:10.1111/j.1365-2958.2010.07369.x)
63. Tzur A, Kafri R, LeBleu VS, Lahav G, Kirschner MW. 2009 Cell growth and size homeostasis in proliferating animal cells. *Science* **325**, 167–171. (doi:10.1126/science.1174294)
64. Koppes LH, Woldringh CL, Nanninga N. 1978 Size variations and correlation of different cell cycle events in slow-growing *Escherichia coli*. *J. Bacteriol.* **134**, 423–433.
65. Di Talia S, Skotheim JM, Bean JM, Siggia ED, Cross FR. 2007 The effects of molecular noise and size control on variability in the budding yeast cell cycle. *Nature* **448**, 947–951. (doi:10.1038/nature06072)
66. Skarstad K, Boye E. 1988 Perturbed chromosomal replication in recA mutants of *Escherichia coli*. *J. Bacteriol.* **170**, 2549–2554. (doi:10.1128/jb.170.6.2549-2554.1988)
67. Nielsen HJ, Youngren B, Hansen FG, Austin S. 2007 Dynamics of *Escherichia coli* chromosome segregation during multifork replication. *J. Bacteriol.* **189**, 8660–8666. (doi:10.1128/JB.01212-07)
68. Tanner NA, Hamdan SM, Jergic S, Loscha KV, Schaeffer PM, Dixon NE, van Oijen AM. 2008 Single-molecule studies of fork dynamics in *Escherichia coli* DNA replication. *Nat. Struct. Mol. Biol.* **15**, 170–176. (doi:10.1038/nsmb.1381)
69. Little JW. 1983 The SOS regulatory system: control of its state by the level of RecA protease. *J. Mol. Biol.* **167**, 791–808. (doi:10.1016/S0022-2836(83)80111-9)
70. Woldringh CL, Mulder E, Valkenburg JAC, Wientjes FB, Zaritsky A, Nanninaga N. 1990 Role of the nucleoid in the top regulation of division. *Res. Microbiol.* **141**, 39–49. (doi:10.1016/0923-2508(90)90096-9)
71. Bernhardt TG, de Boer PAJ. 2005 SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over chromosomes in *E. coli*. *Mol. Cell* **18**, 555–564. (doi:10.1016/j.molcel.2005.04.012)
72. Krishna S, Maslov S, Sneppen K. 2007 UV-induced mutagenesis in *Escherichia coli* SOS response: a quantitative model. *PLoS Comp. Biol.* **3**, e41. (doi:10.1371/journal.pcbi.0030041)
73. Odsbu I, Skarstad K. 2009 A reduction in ribonucleotide reductase activity slows down the chromosome replication fork but does not change its localization. *PLoS ONE* **4**, e7617. (doi:10.1371/journal.pone.0007617)
74. Uphoff S, Reyes-Lamothe R, Garza de Leon F, Sherratt DJ, Kapanidis AN. 2013 Single-molecule DNA repair in live bacteria. *Proc. Natl Acad. Sci. USA* **110**, 8063–8068. (doi:10.1073/pnas.1301804110)
75. Rotem E, Loinger A, Ronin I, Levin-Reisman I, Gabay C, Shoresh N, Biham O, Balaban NQ. 2010 Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proc. Natl Acad. Sci. USA* **107**, 12 541–12 546. (doi:10.1073/pnas.1004333107)
76. Horvath DJ, Li B, Casper T, Partida-Sanchez S, Hunstad DA, Hultgren SJ, Justice SS. 2011 Morphological plasticity promotes resistance to phagocyte killing of uropathogenic *Escherichia coli*. *Microb. Infect.* **13**, 426–437. (doi:10.1016/j.micinf.2010.12.004)
77. Gangan MS, Athale CA. 2017 Data from: Threshold effect of growth rate on population variability of *Escherichia coli* cell lengths. Dryad Digital Repository. (doi:10.5061/dryad.2bs69)

PLOS ONE

RESEARCH ARTICLE

Automated Multi-Peak Tracking Kymography (AMTraK): A Tool to Quantify Sub-Cellular Dynamics with Sub-Pixel Accuracy

Anushree R. Chaphalkar, Kunalika Jain, Manasi S. Gangan, Chaitanya A. Athale*

Div. of Biology, IISER Pune, Pashan, Pune, India

* cathale@iiserpune.ac.in

Citation: Chaphalkar AR, Jain K, Gangan MS, Athale CA (2016) Automated Multi-Peak Tracking Kymography (AMTraK): A Tool to Quantify Sub-Cellular Dynamics with Sub-Pixel Accuracy. PLoS ONE 11(12): e0167620. doi:10.1371/journal.pone.0167620

Editor: Jinxing Lin, Beijing Forestry University, CHINA

Received: July 12, 2016

Accepted: November 17, 2016

Published: December 19, 2016

Data Availability Statement: All data is included in the paper. Software that forms an important part of the report has been released as open source and can be downloaded from the author's website (http://www.iiserpune.ac.in/~cathale/SupplementaryMaterial/Amtrak.html) and from GitHub at https://github.com/athale/AMTraK.

Funding: This research was supported by The Department of Biotechnology (DBT), Govt. of India (BT/PR1595/BRB/10/1043/2012), University Grants Commission (F.2-14/2011 (SA-1)), Department of Science and Technology, Ministry of

Abstract

Kymographs or space-time plots are widely used in cell biology to reduce the dimensions of a time-series in microscopy for both qualitative and quantitative insight into spatio-temporal dynamics. While multiple tools for image kymography have been described before, quantification remains largely manual. Here, we describe a novel software tool for automated multi-peak tracking kymography (AMTraK), which uses peak information and distance minimization to track and automatically quantify kymographs, integrated in a GUI. The program takes fluorescence time-series data as an input and tracks contours in the kymographs based on intensity and gradient peaks. By integrating a branch-point detection method, it can be used to identify merging and splitting events of tracks, important in separation and coalescence events. In tests with synthetic images, we demonstrate sub-pixel positional accuracy of the program. We test the program by quantifying sub-cellular dynamics in rod-shaped bacteria, microtubule (MT) transport and vesicle dynamics. A time-series of *E. coli* cell division with labeled nucleoid DNA is used to identify the time-point and rate at which the nucleoid segregates. The mean velocity of microtubule (MT) gliding motility due to a recombinant kinesin motor is estimated as 0.5 μm/s, in agreement with published values, and comparable to estimates using software for nanometer precision filament-tracking. We proceed to employ AMTraK to analyze previously published time-series microscopy data where kymographs had been manually quantified: clathrin polymerization kinetics during vesicle formation and anterograde and retrograde transport in axons. AMTraK analysis not only reproduces the reported parameters, it also provides an objective and automated method for reproducible analysis of kymographs from *in vitro* and *in vivo* fluorescence microscopy time-series of sub-cellular dynamics.

Introduction

Kymographs, or space-time plots, have been extensively used to analyse sub-cellular microscopy time-lapse data with improvements in microscopy. It has been used in the past to characterize organelle transport, cell division and molecular motor motility as reviewed by Pereira et al. [1], and the wide-range of applications could be the result of the reduced spatial

Science and Technology (INSPIRE IF130394), and Indian Council of Medical Research (3/13/WLC/JRF-2011/HRD-156 (51550)).

Competing Interests: The authors have declared that no competing interests exist.

dimensions of complex microscopy time-series. Most often however, kymography has been used as a qualitative readout of movement or dynamics. In studies where kymographs have been quantified, most often this has been manual, as seen in the Multi Kymograph plugin for ImageJ [2]. Most of the existing tools such as the automated kymography tool [3] and 'guided' kymography [1] focus on automating the process of kymograph building. Few methods for the automated quantification of kymographs exist, such as 'Kymomaker' [4] and a curvelets based tool [5]. Both these tools automate quantification, but cannot deal with merging and spitting events. Despite the ubiquitous nature of merging and splitting events in typical sub-cellular processes, none of the existing tools for the automated quantification of kymographs include a feature to handle budding and coalescence.

Genome segregation is conserved across cellular systems and has been extremely well studied in the rod-shaped Gram-negative bacterium *Escherichia coli* [6,7]. However microscopic analysis of DNA segregation has only recently been made possible with improvements in microscopy and image-analysis [8–10]. Given the almost 1D geometry of segregation of the genome along the long axis of the cell, kymography is a convenient way to analyze the process of nucleoid DNA segregation. Recent studies using explicit 3D over time tracking have found compaction waves are associated with *E. coli* genome segregation [11]. Based on a reduction in dimensions to 1D over time, a quantitative kymograph-based analysis could be used to screen for changes and defects in segregation, without the need for more complex datasets and their analysis.

The process of microtubule transport by molecular motors reconstituted *in vitro*, referred to as a 'gliding assay' has been extensively used to examine the fundamental nature of multi-molecular transport of actin and microtubule filaments by motors [12–15]. Recent studies have also used 'gliding assays' to address microtubule mechanics based on the bending of filaments while undergoing transport [16]. Kymography of cytoskeletal filaments *in vivo* has been used to follow actin contractility and microtubule buckling dynamics [17]. However in most cases the use of kymography has been limited to visualizing the time-series in a single-image, as a compact form of data representation. A general tool that could use this information to objectively extract the measures of motility would hence be of some use to these multiple applications.

The assembly of proteins by 'recruitment' to structures is fundamental in multi-protein complex formation. The assembly of vesicles by budding off membranes and their fusion is critical for cellular function. For the assembly of coated pits with clathrin for endocytosis the site of assembly [18], sequence of binding events [19] and interactions of other proteins [20] is considered to be critical. Microscopy of *in vitro* reconstituted membrane bilayers has become a powerful tool to study the dynamics of protein assembly during vesicle formation [21,22]. Proteins such as epsin, which were reported to accelerate clathrin 'recruitment' [23] have been examined using kymography of the fluorescently labelled clathrin and the effect of mutant epsins on the process [24]. While such an approach lends itself to high-content screening, the analysis of the kymograph has been manual. Many other such 'recruiment' dynamics studies could benefit from an automated routine to quantify the kinetics of assembly through intensity measurements coupled to kymography.

Neuronal vesicles are transported in axons by the action of molecular motors. Microscopy of *in vitro* reconstituted [25] and the *in vivo* transport in cultured cells [26,27] has provided insights into both the components and forces regulating transport. Recent technical developments have allowed whole animal *in vivo* microscopy of sub-cellular vesicle movements in neurons [28]. In this and comparable studies, quantitative statistics have been obtained using manual detection of kymographs. This is possibly due to the complex nature of the time-series with cross-overs and the crowded *in vivo* environment. An approach that uses objective

criteria and automates the process of quantification could provide valuable improvements to our understanding of fundamental nature of vesicle transport as well as aid in the process of modeling vesicle transport.

Here, we have developed a novel tool to automatically quantify kymographs from fluorescence image time-series. We proceed to demonstrate the utility of the automated multi-peak tracking kymography (AMTraK) tool by quantifying dynamics from diverse sub-cellular fluorescence microscopy data sets. These include bacterial genome-segregation, microtubule (MT) motility of 1D filaments and 2D radial asters, membrane protein assembly dynamics and vesicle transport in axons.

Algorithm and workflow

The automated multi-peak tracking kymography (AMTraK) is open source software based on an algorithm that combines peak detection and distance minimization based linking to quantify dynamics of fluorescence image time-series. The source code has been released with a GPL license and can be accessed from: http://www.iiserpune.ac.in/~cathale/SupplementaryMaterial/Amtrak.html and https://github.com/athale/AMTraK

The program has a GUI front-end and is accompanied by a detailed help file. The algorithmic workflow (Fig 1A) is divided broadly into three steps:

1. Making the kymograph
2. Peak detection and tracking
3. Statistics

These steps in the workflow are reflected in the graphical user interface (GUI) layout (Fig 1B). The functioning of each of these steps is briefly described as follows:

(a) Making the kymograph. The user chooses an input image time-series with the "Open File" button. Image time-series are assumed to be uncompressed, multi-page TIF files (independent of bit depth). The user can choose to process either the whole or a subset of frames using the "Frame nos." text box. For example entering "2:2:8" will now result in only frames 2, 4, 6 and 8 being processed for further analysis. The text box "Save as sub-folder" takes a number input (default "1") indicating where the outputs will be stored (e.g.: "./amtrak-1"). The drop-down menu "Apply LOI" allows the user to either choose a line of interest (LOI) using the mouse ("Interactive") or apply a pre-existing LOI on a different channel (color) of the image time-series ("From file"). Once an interactively drawn LOI is selected, it is stored in the output sub-folder as "LOIselection.txt" (S1 Data). This LOI can subsequently be applied, to another channel or the same region of another dataset (e.g.: microfluidics channels) using the "From file" mode. For this, the user is required to load a separate TIF time-series using "Open File" and change the sub-folder number in order to prevent overwriting old data. The "LOI width (pixels)" allows a user to choose the width of the LOI, to compensate for occasional drift of the object, in a direction orthogonal to the LOI orientation. The choice widths- 1, 3 and 5 pixels- is centered around the selected LOI pixels, similar to that implemented in the ImageJ Multi Kymograph plugin [2]. The drop-down menu "Units" allows the user to select distance and time units, and the text boxes "Scaling factor" and "Time interval" are used to provide conversion factors per pixel and frame respectively. This results in scaling the pixels and frame numbers to physical units. The button "Make Kymograph" produces a maximum intensity projection image of the input time-series, if the user had chosen the "Interactive" mode (default) in the "Apply LOI" menu. The user is required to select the line of interest by drag-

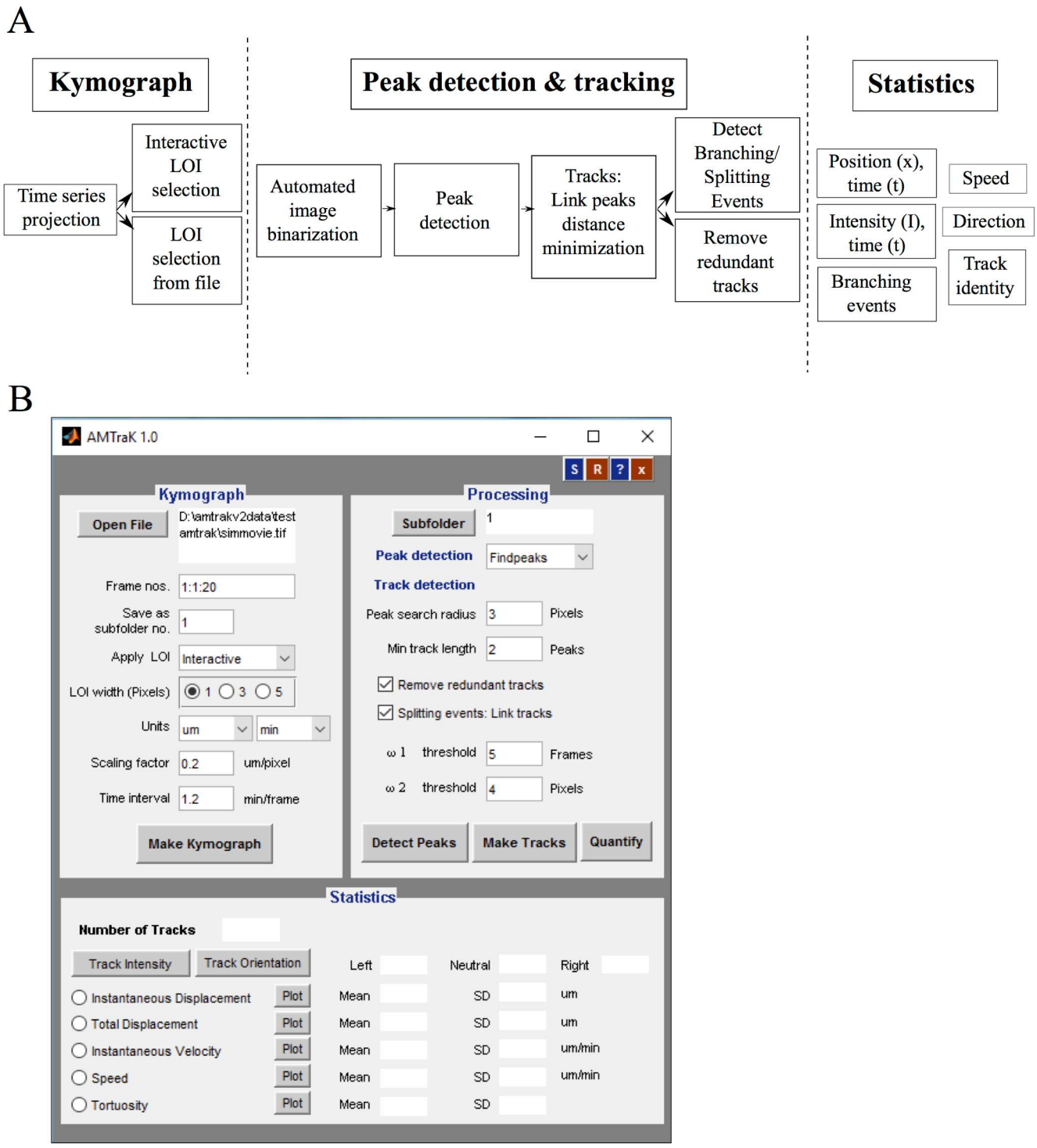


Fig 1. Algorithm workflow and user-interface. (A) The workflow of the algorithm involves three steps (1) kymograph generation, (2) peak detection and tracking and (3) quantification and the functions invoked by each part are elaborated. (B) The GUI is organized to reflect this workflow.

doi:10.1371/journal.pone.0167620.g001

clicking the mouse. Double-clicking ends the selection, and throws a dialog box, which prompts the user to choose to either select more LOIs or continue with the processing of the

one already selected. This generates file one or more "LOIselection.txt" files in the sub-folders. If the "From file" mode was selected, the program allows the user to select a pre-existing "LOI-selection.txt" from the directory structure. The program then generates kymographs based on these LOIs and stores the matrices corresponding to the LOIs in sub-folders numbered according to the sequence of LOI selection (e.g.: "/amtrak-1/", "/amtrak-2/" etc.).

(b) Peak detection and tracking. *Detecting peaks*: The button "Subfolder" allows the user to choose the kymographs to be processed using "Add", which adds the subfolders created earlier to the active list. Using this feature, a user can either process a single kymograph at a time, or process multiple kymographs using the same parameters. The kymograph is segmented row-wise using Otsu's method [29] and the resulting binary image is processed for "Peak detection". The user can choose between three alternative methods: (i) *findpeaks* [30] and (ii) watershed [31] to find central peaks, while (iii) Canny edge detection [32] is useful if the edge information is the most reliable descriptor of the dynamics. Typically *findpeaks* and watershed are ideal for spherical objects.

Linking: The list of peaks *P(t)* for each time point *t* is linked resulting in tracks, based on user input parameters of "Peak search radius" (λ_1) and "Min. track length" (λ_2). Peaks are linked if the minimal pair-wise distance dj(t,t+τ) between every j^{th} peak in successive rows (*t*, *t* +τ) satisfies the condition $\min(d_j(t, t+\tau)) \leq \lambda_1$, iteratively for the j^{th} peak in every subsequent time step (t+τ). If two or more peaks are equidistant, the peak that makes the largest angle (0 to π) with the existing track is chosen, similar to our previously developed branch detection method [33]. For the peaks in t = 1, the angle criterion does not hold true and equidistant peaks are resolved by user-input. Tracks are eliminated from further analysis if their number of peaks linked $\text{len}(P) \leq \lambda_2$, to avoid artifacts due to very short tracks.

Remove redundant: If the checkbox "Remove redundant tracks" is selected, each i^{th} track with η_i coordinates, is tested for intersections using the inbuilt *intersect* function. If the number of common coordinates η_c satisfies the condition $\eta_c \geq \eta_i/3$, it is eliminated as a redundant track.

Splitting and joining tracks: If the checkbox "Splitting events: Link tracks" is selected, events where two tracks merge are identified by a two-step process. First, all peaks (I(x,t)) are evaluated for the condition $I(x,t) = (d^te \leq \omega_1)$ AND $(d^xe \leq \omega_2)$, where d^te is the distance on the time-axis (t) and d^xe is the distance on the spatial (x) axis. Then, a peak with the minimal (Euclidean) distance is minimized for the distance to the end-point (e) coordinate J_m. The time and distance thresholds are set by the user in the text box for ω_1 (frames) and ω_2 (pixels) respectively.

The button "Detect Peaks" then outputs an image of the kymograph with the peaks overlaid in color, while invoking the button "Make tracks" links the peaks based on the input parameters. Lastly the button "Quantify" produces a text file corresponding to each track (S2 Data, S3 Data and S4 Data).

(c) Statistics. This section of the code produces both text-file outputs and plots of the dynamics estimated from the kymograph. The frequency distribution of "Instantaneous Displacement", "Total Displacement", "Instantaneous Velocity", "Speed" and "Tortuosity" (i.e. directionality) are plotted if the button "Plot" corresponding to these variables is pressed. Additionally the mean and standard deviation (s.d.) of these variables are also generated in the text boxes. Pressing the "Track Intensity" button plots the normalized (0–1) grey value intensity of each track as a function of the time. The button "Track orientation" triggers a recoloring the tracks in the kymograph based on the net direction of movement along the X-axis- blue (-ve, left), red (+ve, right) and green (stationary, neutral).

The outputs of the analysis are stored in multiple tab-delimited text files: "LOIselection.txt" with the LOI coordinates (S1 Data), "USER_TrackStats.txt" which reports track-wise mean

values (S2 Data), "USER_InstStats.txt" which reports the time-dependent statistics (S3 Data), "Tracklist.txt" which stores the grey-value intensities as a function of track number and time (S4 Data) and "Branchpoints.txt" which stores the position and time coordinates of detected branches (S5 Data). The user-inputs to AMTraK in terms of files, directories and parameters are all stored in "All_Parameters.txt" (S6 Data), to enable reproducible analysis.

Materials and Methods

Simulated test images

Simulated images of static beads were generated by creating 8 bit images with a black background (intensity: 0) with equally spaced single white pixels (intensity: 255) in MATLAB (MathWorks Inc., USA). To simulate bead motion, a simple 1D random-walk was implemented where each bead was moved randomly in each frame, with displacement drawn from a normally distributed random number with mean m = 0 and standard deviation (s). The standard deviation is a measure of the mean speed of motion. Both the static and mobile bead image time-series were filtered with a 5x5 disk filter and smoothed using a 3x3 averaging filter. The resulting convolved circular objects (S1A Fig) have intensity profiles that resemble point sources of fluorescence signal (S1B Fig). The time-series were saved as a multi-page TIF files. Noise was added to individual time-series in order to simulate increasing levels of image-noise using a Gaussian filter with increasing standard deviation (0–100) using ImageJ [34].

Bacterial growth and microscopy

E. coli MG1655 (CGSC, Yale, USA) expressing the pBAD24-hupA-GFP [35] were cultured in Luria Bertani (LB) medium (HiMedia, Mumbai, India) with 100 μg/ml Ampicillin (Sigma-Aldrich, Mumbai, India) at 37˚C with shaking at 170 rpm (Forma, ThermoScientific, USA). Nutrient 'agar-pads' with 0.2% arabinose (Sisco Research Labs, Mumbai, India) and 100 μg/ml ampicillin were imaged on a glass-bottomed Petri dish (Corning, NY, USA) at 37˚C using an inverted Zeiss LSM780 confocal microscope (Carl Zeiss, Germany) with a Plan Apochromat 63x (N.A. 1.40, oil) lens in DIC and fluorescence (excitation by 405 nm diode laser with a beam splitter MBS 405 and the emission collected between 487–582 nm) modes. Images were corrected for drift using the rigid body transformation in the StackReg plugin [36] for ImageJ.

Microtubule gliding assay

A 1:4 ratio of TRITC-labeled bovine and unlabeled porcine tubulin (Cytoskeleton Inc., USA) at a concentration of 20 μM were used to prepare taxol stabilized MT-filaments in general tubulin buffer as described by the supplier (Cytoskeleton Inc., USA). Into a double backed tape chamber, we sequentially flowed in 4.1 μg/μl of a 67 kDa recombinant human kinesin (Cytoskeleton Inc., USA), blocking buffer (5 mg/ml Casein) and MT filaments. The chamber was then washed with a casein-containing buffer and the reaction was started with 1 mM ATP with anti-fade mix (0.05 M glucose, 1% sucrose, 0.5 mg/ml catalase, 0.5 mg/ml glucose oxidase, 0.5% beta-mercaptoethanol (Cytoskeleton Inc., USA)). Time-series images were acquired every minute for 30 minutes on an upright epifluorescence microscope with a 40x (N.A. 0.75) EC Plan Neofluar lens mounted on a Zeiss Axio Imager.Z1 (Carl Zeiss, Germany) using filters for excitation (563 nm) and emission (581 nm) and an MRC camera (Carl Zeiss, Germany).

Image processing

The acquired time-series and movies taken from published data were converted to uncompressed TIF time-series using ImageJ (Schneider et al., 2012) and online converters for MOV

files. MT-gliding assay images were de-noised using a median filter in ImageJ. For manual analysis of kymographs of MT-gliding, a program was written in MATLAB (MathWorks Inc., USA) to generate a kymograph from the time-series, interactively draw a segmented line along the edges and extract coordinates to calculate velocities. The automated multi-peak tracking kymography (AMTraK) code was implemented in MATLAB R2014b (MathWorks Inc., USA) in combination with the Image Processing (ver. 7.0) and Statistics (ver. 7.3) Toolboxes and tested on Linux, Mac OSX and Windows7 platforms. Vesicle transport image time-series in *C. elegans* from supporting material of published work [28] were calibrated based on the width of the axon from the same report.

Data analysis

All data analysis and plotting was performed using MATLAB 2014b (MathWorks Inc., USA). Fitting of custom functions was performed using either the Levenberg-Marquardt non-linear least square routine or the Trust-Region method, implemented in the CurveFitting toolbox (ver. 3.5) of MATLAB.

Results

Accuracy of detection

To test the positional detection accuracy of the algorithm, we have created simulated image time-series of circular objects that represent typical fluorescence images of circular objects (Fig 2A), comparable to images of sub-cellular structures in pixels (S1A Fig). Since the time-series consists of the same image, the objects are perfectly static as seen in the resulting kymograph (Fig 2B) output from running AMTraK on the data. Intensity variations are a result of the noise from the spatial filter (s.d. 40). The difference between the position of the detected tracks (x_D) and the simulated position (x_S) is used as an estimate of the limit of accuracy in position detection, $\Delta x = |x_S - x_D|$. The normalized frequency distribution of Δx can be fit to an exponential decay function to obtain a mean accuracy $<\Delta x> = 1/b$ from the fit, in pixel units (Fig 2C). For all images with noise of s.d.< 40, the mean error (from fit) in detection $\langle \Delta x \rangle < 1$ pixel. For higher values appears to saturate between 2–3 pixels (Fig 2D). Using the arithmetic mean as an estimate of the accuracy for a given noise s.d. appears to result in an underestimate that does not change with increasing noise s.d. (Fig 2D), and hence the mean from the exponential decay of the frequency of Δx was taken to be more representative of the central tendency. To test if motility affected the positional accuracy, we also evaluated the positional accuracy of particles undergoing a random walk (as described in the Materials and Methods section) with a fixed image noise (noise s.d. 30). By increasing the s.d. of the random walk we estimated the effect of increasing velocity on Δx (Fig 3A). The accuracy of positional detection using both the arithmetic and exponential mean error ($<\Delta x>$) as before, is less than 1 pixel for the chosen range of velocities of the random walk (Fig 3B). At higher velocities, the tracking errors accumulate, suggesting image noise is the major limiting factor for the positional accuracy of detection, independent of particle motility. Thus, while AMTraK analysis can result in sub-pixel accuracy of position detection, it is essential that the input data have low-noise. We proceeded to test our method on the multiple experimental datasets to examine the utility of this program involving bacterial DNA segregation, microtubule motility and vesicle assembly and transport dynamics.

Detecting splitting events in bacterial DNA-segregation

A time-series of growing *E. coli* is acquired in fluorescence (Movie A in S1 Video) and DIC (Movie B in S1 Video) to follow the nucleoid segregation dynamics of HupA-GFP labeled

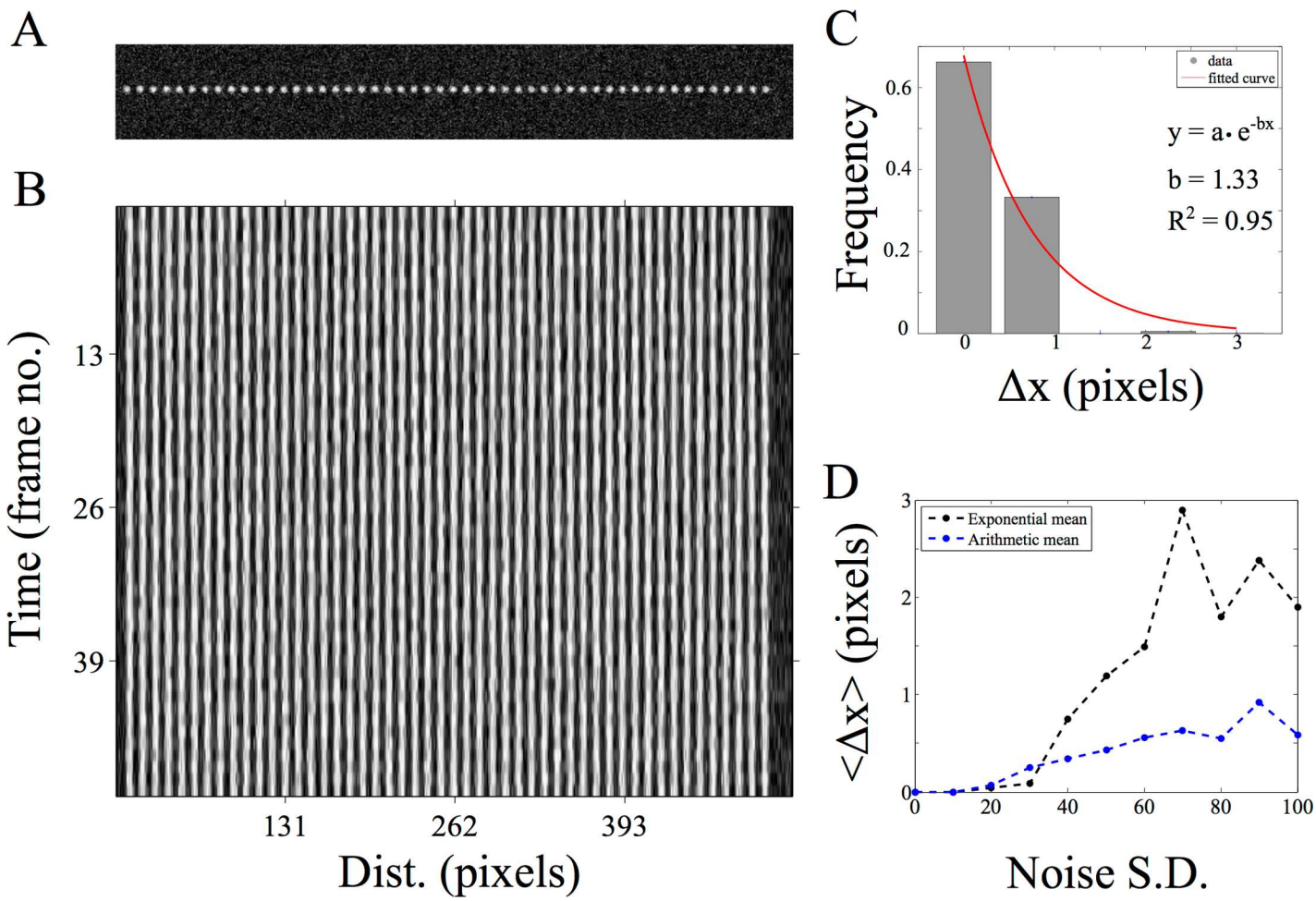


Fig 2. Estimating positional accuracy. (A) A single frame of a 2D image time-series of static spheres (with a peak intensity of 1) with Gaussian noise (mean = 0, s.d. = 40) is analyzed using AMTraK (B) resulting in a kymograph. (C) The frequency distribution of the error in position detection (Δx) by AMTraK (bars) is fit by an exponential decay (red). The mean error obtained is 0.75 pixels (goodness of fit $R^2 = 0.95$) for a representative time-series with noise s.d. = 40. (D) The mean error of detection (y-axis) from the exponential fit $\langle \Delta x \rangle = 1/b$ (black) is compared to the arithmetic mean (blue) in pixel units, plotted as a function of increasing noise s.d. (x-axis). The noise generates random intensities drawn from a Gaussian distribution with mean 0 and the specified s.d. being added to the image (based on the "Specified Noise" function in ImageJ).

doi:10.1371/journal.pone.0167620.g002

DNA (Fig 4A). Using the maximum intensity projection produced from AMTraK, the LOIs are chosen (Fig 4B) and used to generate and analyze two kymographs (Fig 4C and 4D). The segregation of the genome is captured by the branched structures of the tracks marked in the kymographs. Additionally we can evaluate both the instantaneous velocity for time-dependence (Fig 4E) and average statistics (Fig 4F). The mean nucleoid transport velocity is 0.103±0.12 μm/min (arithmetic mean ± standard deviation). Based on the form of the frequency distribution of instantaneous velocities, we also fit an exponential decay function to obtain the exponential mean velocity $v_{ex} = 0.104$ μm/min. These values of nucleoid movement speed from *E. coli* MG1655 (wild-type) cells are comparable to a previous report in which nucleoids were tracked in 3D over time [11]. While nucleoids form a diffraction-limited spot in microscopy images, un-branched cytoskeletal filaments form typical 1D structures and dynamics of transport on them and of the filaments themselves, are ideally suited for kymography.

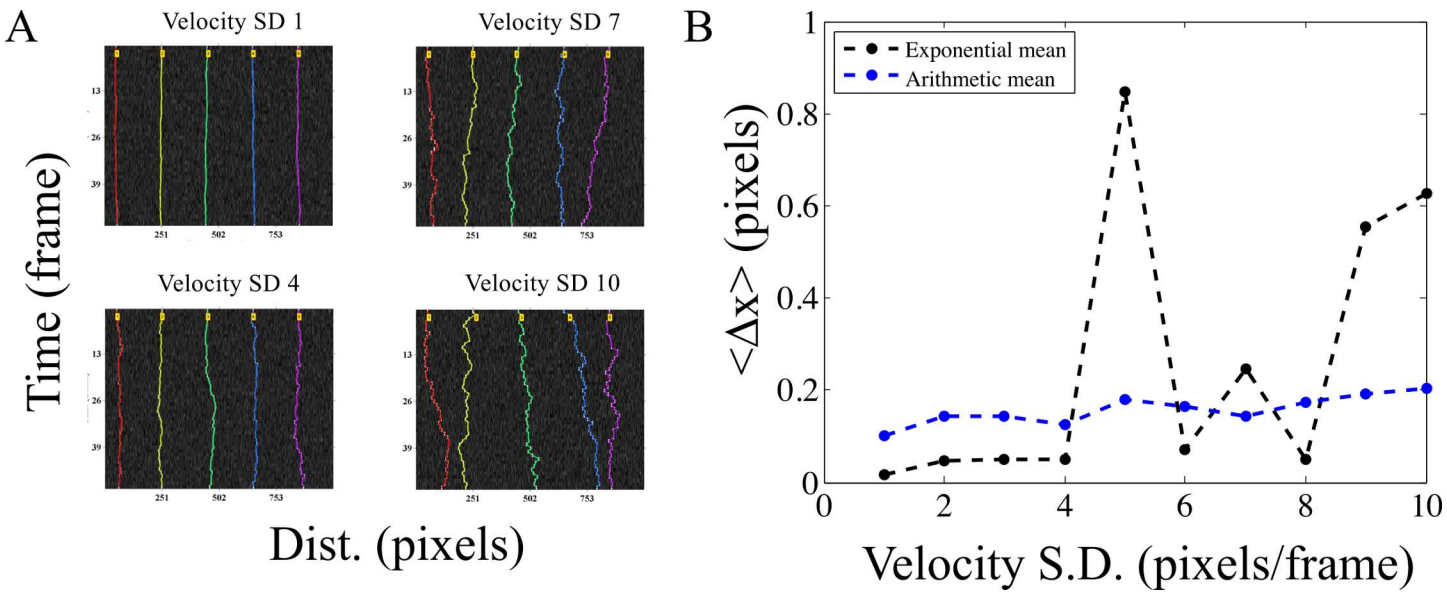


Fig 3. Positional accuracy of tracking simulated motility. (A) Kymographs of time-series of spheres undergoing a 1D random walk with Gaussian noise (s.d. = 30) were tracked. The colors indicate the detected tracks. (B) The arithmetic mean (blue) and exponential mean (black) of error in position detection (Δx) (y-axis) over 3 iterations of the time-series is plotted for increasing velocity of the random-walk (x-axis) as inferred from the standard deviation (s.d.).

doi:10.1371/journal.pone.0167620.g003

Microtubule transport: filament edges, centers and time-dependence of velocity

The transport of microtubule (MT) filaments by surface-immobilized molecular motors in the presence of ATP and buffers is referred to in the literature as 'gliding assay' or 'collective transport assay'. Here, we analyze the gliding motility of MT on kinesin, as described in the methods section, using AMTraK. The analysis of a representative kymograph using either peak- (Fig 5A) or edge-detection (Fig 5B) successful traces the centroids and edges respectively. The mean velocity estimates for collective motor transport show variations between individual filaments. The centroid and edge velocity estimates of multiple MT filaments (n = 10) are strongly correlated as evidenced by the straight line fit with slope ~ 1 (Fig 5C and 5D), as expected. However, the linear correlation of edge-based velocities has a slope of ~0.9 (Fig 5E), suggesting small deviations from the ideal slope, within the range of the average positional detection error (Fig 2C). While typical kymograph analysis of cytoskeletal transport averages the edge information (movement of the tips over time), correlating edge-velocities could potentially be used to estimate small alterations in the filament geometry such as bending and length change. The mean velocity of 0.5 μm/min obtained from our analysis of the assay (Fig 5F) is consistent with previous reports for the same construct [37,38]. While the transport of effectively 1D MT filaments lends itself to kymography, we proceeded to investigate if 2D radial MT structures or asters can also be analyzed by kymography.

Fusion of MT asters

In recent experiments by Foster et al. [39] they examined the spontaneous contraction dynamics of radial MT arrays or asters labeled with Alexa647-tagged tubulin, in *Xenopus* egg extracts. We have taken a time-series of such asters from published data (kindly shared by the author Peter J. Foster) and analyzed coalescence events using AMTraK (Fig 6A) The projection of the time-series for selecting the LOI enables us to reduce the complex movements of such 2D

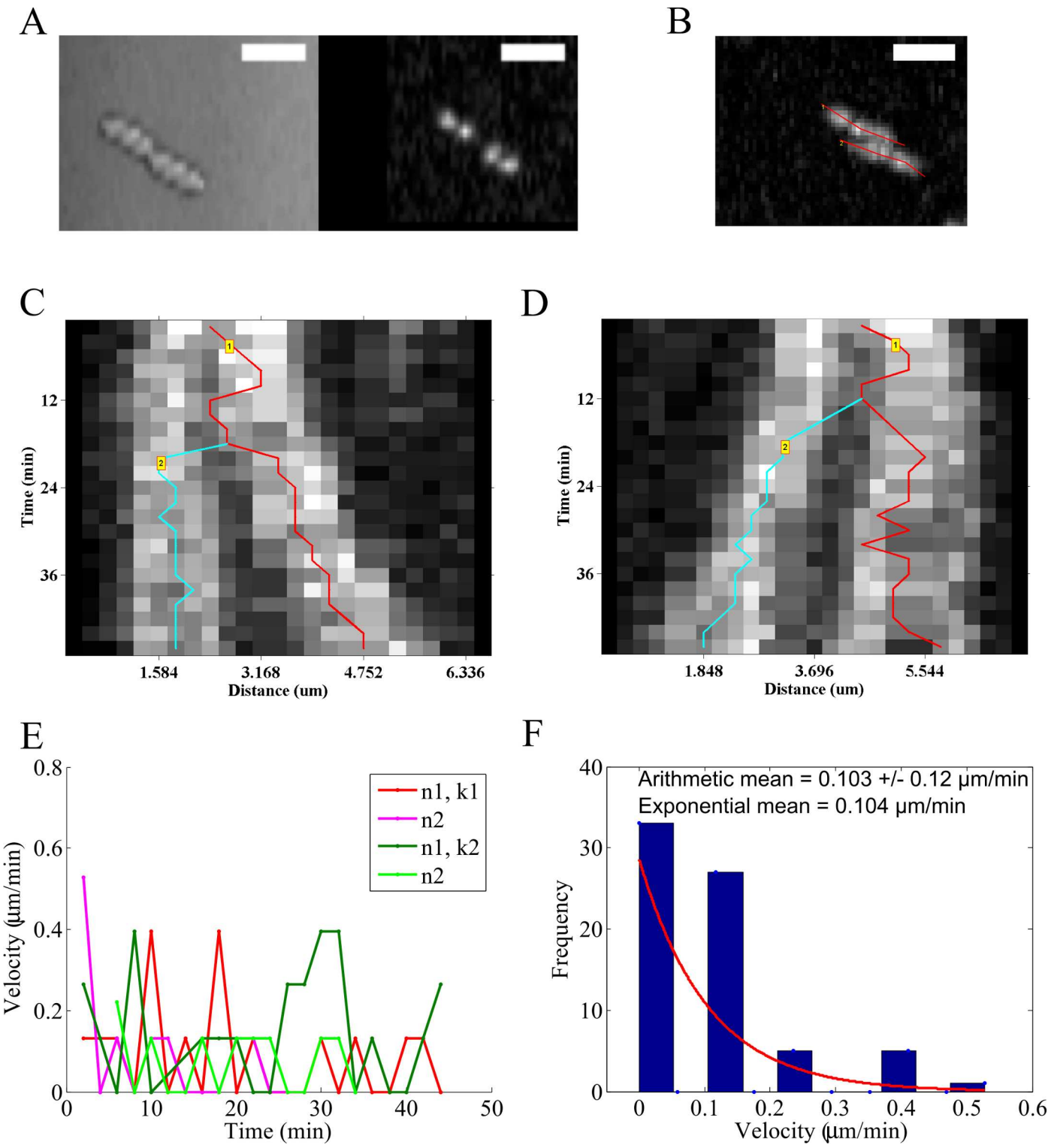


Fig 4. Nucleoid segregation dynamics of *E. coli*. (A) Image time-series of *E. coli* MG1655 grown on agar pads and imaged in DIC (left) and fluorescence based on HupA-GFP (right) are analyzed using AMTraK. (B) AMTraK generates a maximum intensity projection on the basis of which user-selected lines of interest (red lines) are used by the program to generate kymographs. The kymographs based on (C) LOI 1 (k1) and (D) LOI 2 (k2) were tracked resulting in branched tracks (colored lines). (E) The instantaneous velocities of nucleoids 1 and 2 (n1, n2) from kymographs 1 (k1) and 2 (k2) are plotted as a function of time (colors indicate nucleoids n1, n2 each from the kymographs k1, k2). (F) Mean velocities are estimated using both the arithmetic mean (±s.d.) and v_{ex}, the mean of the exponential decay ($y = e^{-1/vex}$) that was fit (red line) to the frequency distribution of instantaneous velocity (bars). Scale bar 4 µm.

doi:10.1371/journal.pone.0167620.g004

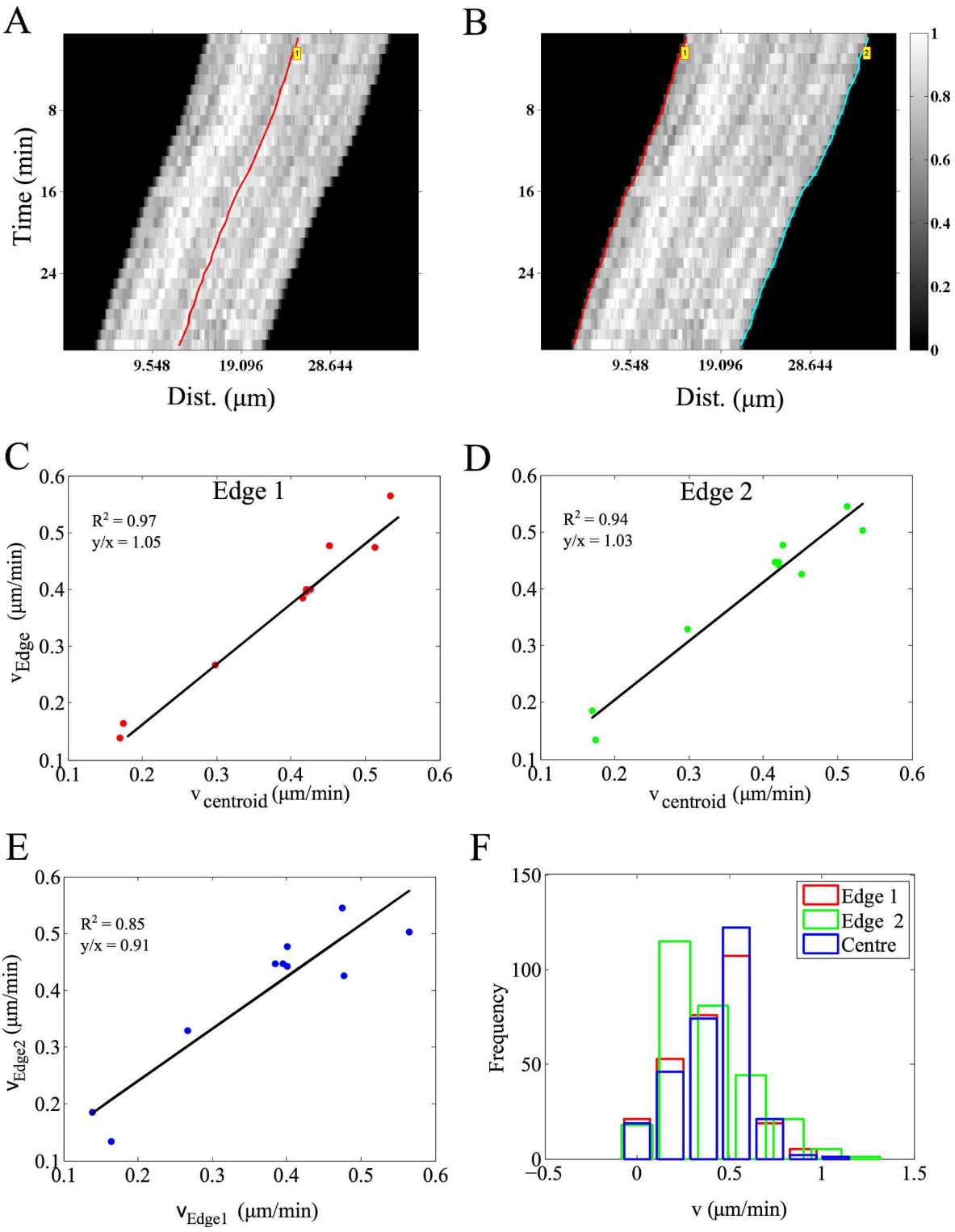


Fig 5. Microtubule (MT) gliding motility on kinesin motors. MTs gliding on kinesin (images acquired every 1 minute for 30 minutes) were analyzed using AMTraK by either detecting **(A)** the centerline (red) or **(B)** the two edges the filament, edge 1 (red) and 2 (cyan). Color bar: gray scale image intensity normalized by the maximal value for the bit-depth. **(C, D)** The velocity estimates from the centroid-based velocity estimates and the two edges and **(E)** the velocity estimated from each edge are correlated. **(F)** The frequency distribution of the instantaneous velocity estimates using the centroid (blue) is compared to edge-based estimates. r^2: goodness of fit, y/x: slope of the linear fit. Number of filaments analyzed, n = 10.

doi:10.1371/journal.pone.0167620.g005

structures to a 1D over time process. The movement of the smaller aster as it merges with the larger one is rapid. The fluorescence intensity following the merger fluctuates, but does not increase, which we interpret to mean tubulin density at the center of the new aster does not increase (Fig 6B). While the coalescence appears not to result in a compaction of the aster, it demonstrates the utility of the code for 2D MT array transport. On the other hand, intensity measurements are expected to change during processes such as molecular 'recruitment' of sub-cellular structures, so we proceed to test the tool on this process, which had previously been studied using manual kymography.

Kinetics of clathrin assembly during in vitro vesicle formation

We proceed to quantify the assembly kinetics of clathrin on membranes from an *in vitro* reconstitution assay of clathrin assembly on vesicle precursors reported previously by Holkar et al. [24]. This process has been analyzed using kymography due to its effectively 1D spatial extent and the multiple simultaneous events of assembly. The published time-series of fluorescently labeled clathrin assembly kinetics in the presence of wild-type epsin (supplementary movie 3 in [24]) and L6W mutant epsin (supplementary movie 5 in [24]) in the form of 16 bit TIF images were provided by the authors (Sachin Holkar, personal communication). AMTraK was used to analyze this data without any pre-processing, resulting in tracked kymographs of assembly kinetics with wild-type (Fig 7A) and mutant epsin (Fig 7B). The software outputs a text-file of grey-value intensities normalized by the bit-depth (maximum normalized, between 0–1) (S4 Data), which when multiplied by the bit-depth of the input images, produced intensity profiles of clathrin assembly in grey-values with time in the presence of wild-type (Fig 7C, S3A Fig) and mutant epsin (Fig 7D, S3B Fig). These intensity profiles were fit to a single phase exponential function $y = a+(b-a)^*(1-e^{-c^*t})$, where y is the intensity which increases with time t, and depends on three fit parameters, a, b and c, the same function as used by Holkar et al. [24]. A large proportion of the assembly events were successfully tracked and most showed

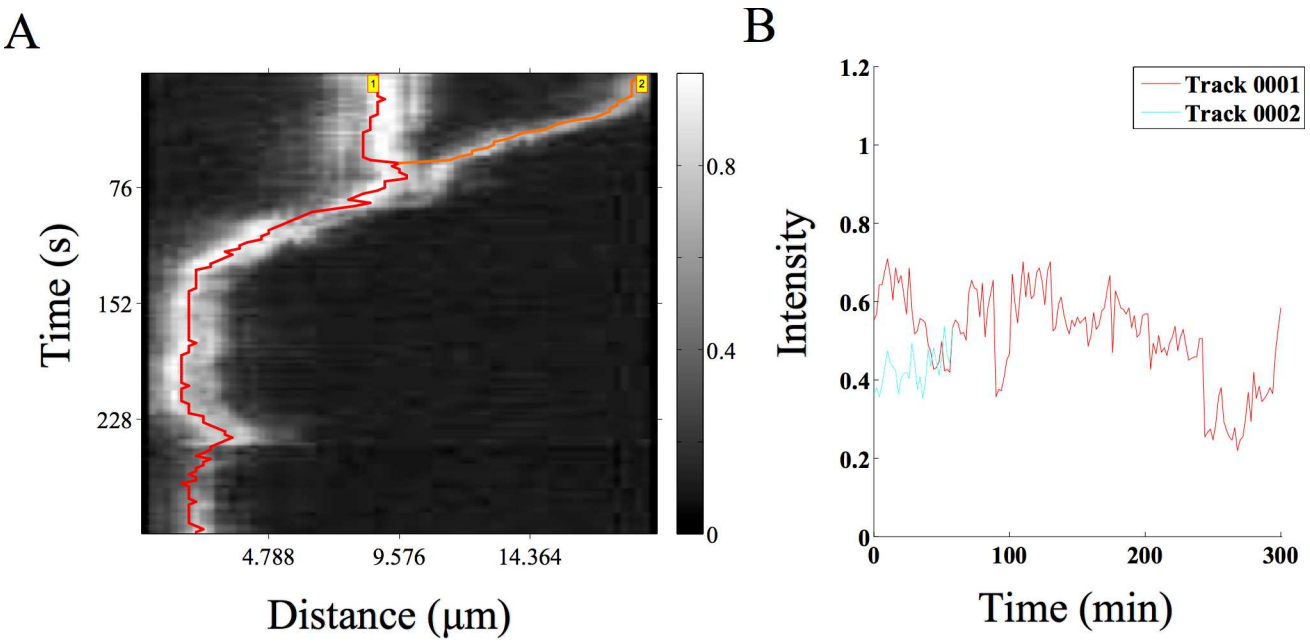


Fig 6. MT aster coalescence. (A) A time-series of MT asters undergoing fusion (time-series taken from previous work by Foster et al. [39]) was analyzed using AMTraK. The grey scale bar indicates normalized fluorescence intensity of Alexa-647 labeled tubulin. **(B)** The relative intensity over time of the two coalescing asters is plotted.

doi:10.1371/journal.pone.0167620.g006

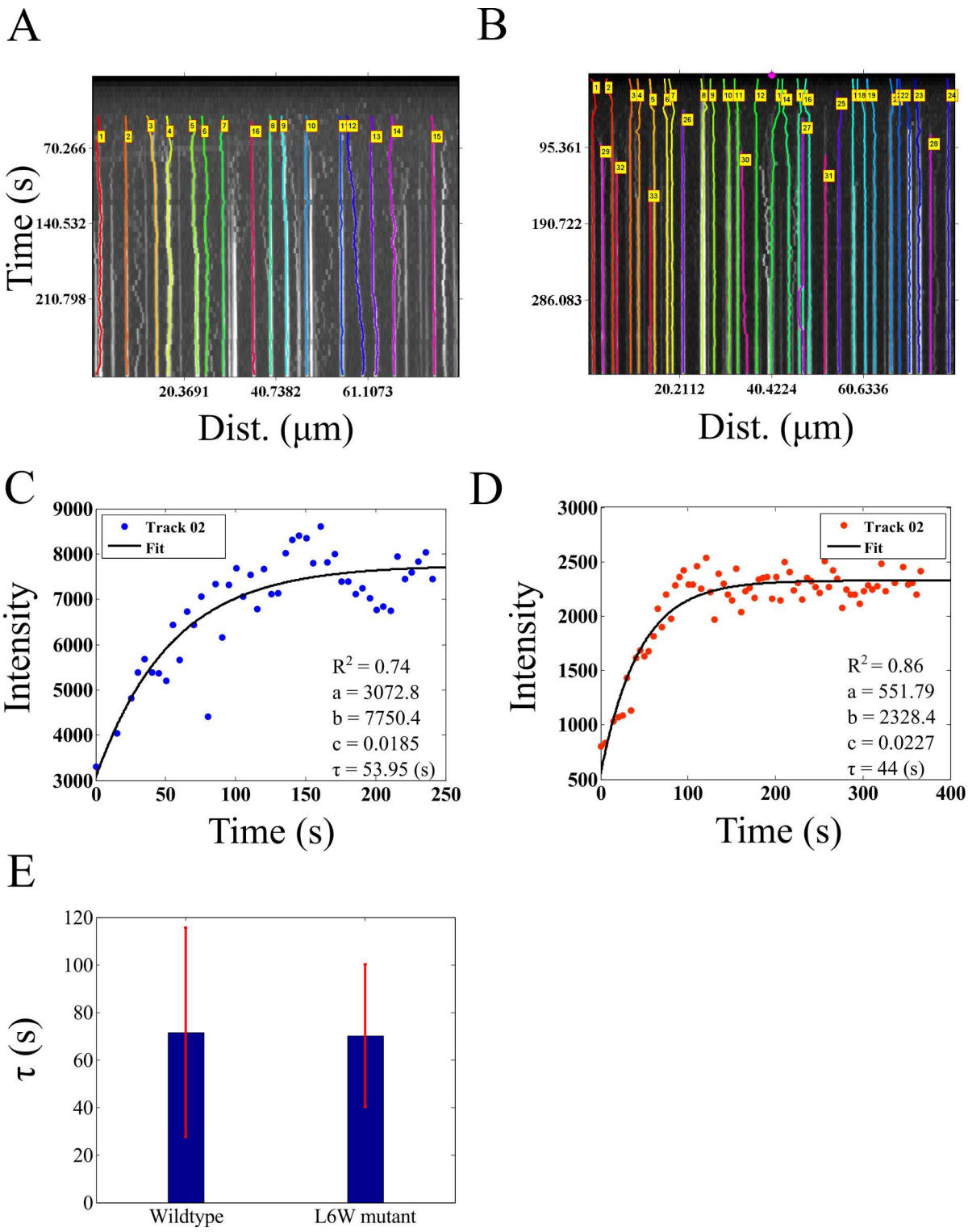


Fig 7. Dynamics of clathrin assembly. (A, B) Microscopy time-series taken from Holkar et al. [24] of fluorescently labeled clathrin assembly in the presence of **(A)** wild-type and **(B)** mutant epsin were analyzed using AMTraK. Colored lines in the kymographs indicate detected tracks. **(C, D)** The change in intensity as a function of time based on AMTraK detected tracks from **(C)** clathrin + w.t. epsin and **(D)** compared to clathrin + (L6W) mutant epsin. The intensity kinetics plots are fit to a single-phase exponential function, $y = a+(b-a)*(1-e^{-c*t})$ to obtain the time constant of assembly $\tau = 1/c$ (red). R^2: goodness of fit. **(D)** The mean values (error bar represents s.d.) of the time constant of assembly of clathrin (τ) in the presence of wild-type and mutant epsin are compared.

doi:10.1371/journal.pone.0167620.g007

saturation kinetics that were fit by curves with $R^2>0.7$ (S3 Fig). While the parameters *a* and *b* are scaling factors, *c* determines the characteristic clathrin polymerization time, $\tau = 1/c$. In our analysis the clathrin assembly time in presence of wild-type epsin is $<\tau> = 71.49\pm44.09$ s while with mutant epsin $<\tau> = 70.16\pm29.89$ s. In our estimate of the mutant assembly time is indistinguishable from wild-type, consistent with the previous report, which used manual quantification of the kymograph [24]. We proceed to examine if our tool, which appears to work successfully on *in vitro* data with low background noise, can also be used for the quantification of *in vivo* dynamics inside the crowded environment of an intact cell.

Axonal vesicle transport: Characterizing directional switching

Synaptic vesicles in *Caenorhabditis elegans* mechanoreceptor neurons labeled with GFP-Rab3 have been recently studied by Mondal et al. in a whole-animal microfluidics device, providing retrograde and anterograde vesicle transport statistics [28]. Such *in vivo* data is complex, involves multiple crossovers and has many objects close to each other. AMTraK based analysis of the published data could detect up to 17 different tracks (Fig 8A). Vesicles that were not detected have typically low intensity or were out of focus and were not segmented. The spread of the distribution of instantaneous velocities (left-ward: negative, anterograde; right-ward: positive, retrograde, non-motile: paused) shows that the GFP-Rab3 vesicles are equally likely to be anterograde and retrograde in their transport (Fig 8B). Based on the shape of the frequency distribution of the non-zero velocities in anterograde (Fig 8C) and retrograde (Fig 8D) directions, an exponential decay fit to the frequency distribution was used to estimate mean velocities (goodness of fit, $R^2 = 0.99$). To enable comparison with the arithmetic means reported in literature [28], we also estimate the average. The mean velocity from the exponential fits of anterograde transport is 0.625 μm/s (n = 425, arithmetic mean±s.d.: 0.77±0.53 μm/s) while the mean retrograde velocity is 0.714 μm/s (n = 540, arithmetic mean±s.d.: 0.854±0.67 μm/s). In this case, both means are comparable since only non-zero values were the analyzed. Velocities in both directions are of comparable order of magnitude to the published values obtained by manual detection [28], but 1.5-fold lower, due to a (non-zero) threshold velocity used by the authors to define pauses (as personally communicated by the author, Sudip Mondal). Thus, AMTraK can be reliably used to quantify transport and assembly dynamics from both *in vitro* and *in vivo* fluorescence microscopy data, as seen from the quantification, which is consistent with literature.

Discussion

In this report, we have described a novel tool for automatic detection and quantification of kymographs from fluorescence microscopy time-series. Using simulations we have demonstrated sub-pixel position detection accuracy of our proposed method, in conditions of low Gaussian noise. The program quantifies position, motility, and brightness intensity of fluorescence signal and fusion/splitting events. The utility of the code is tested on *in vitro* and *in vivo* fluorescence time-series ranging from *in vitro* assays of MT gliding assays with kinesin, coalescence dynamics of MT-asters, clathrin assembly kinetics on lipid tethers to *in vivo* axonal synaptic vesicle transport. The measures of average transport and kinetics of these diverse data types are consistent with published data and provides opportunities for improved statistics of individual events from a dynamic time-series, which were not as easily accessible with current methods.

Manual quantification of kymographs [2] depends typically on reliable edge detection. As a result, quantification varies between individuals and requires prior information or experience [40]. Yet, manual kymography is widely reported in cell-biological literature for the analysis of

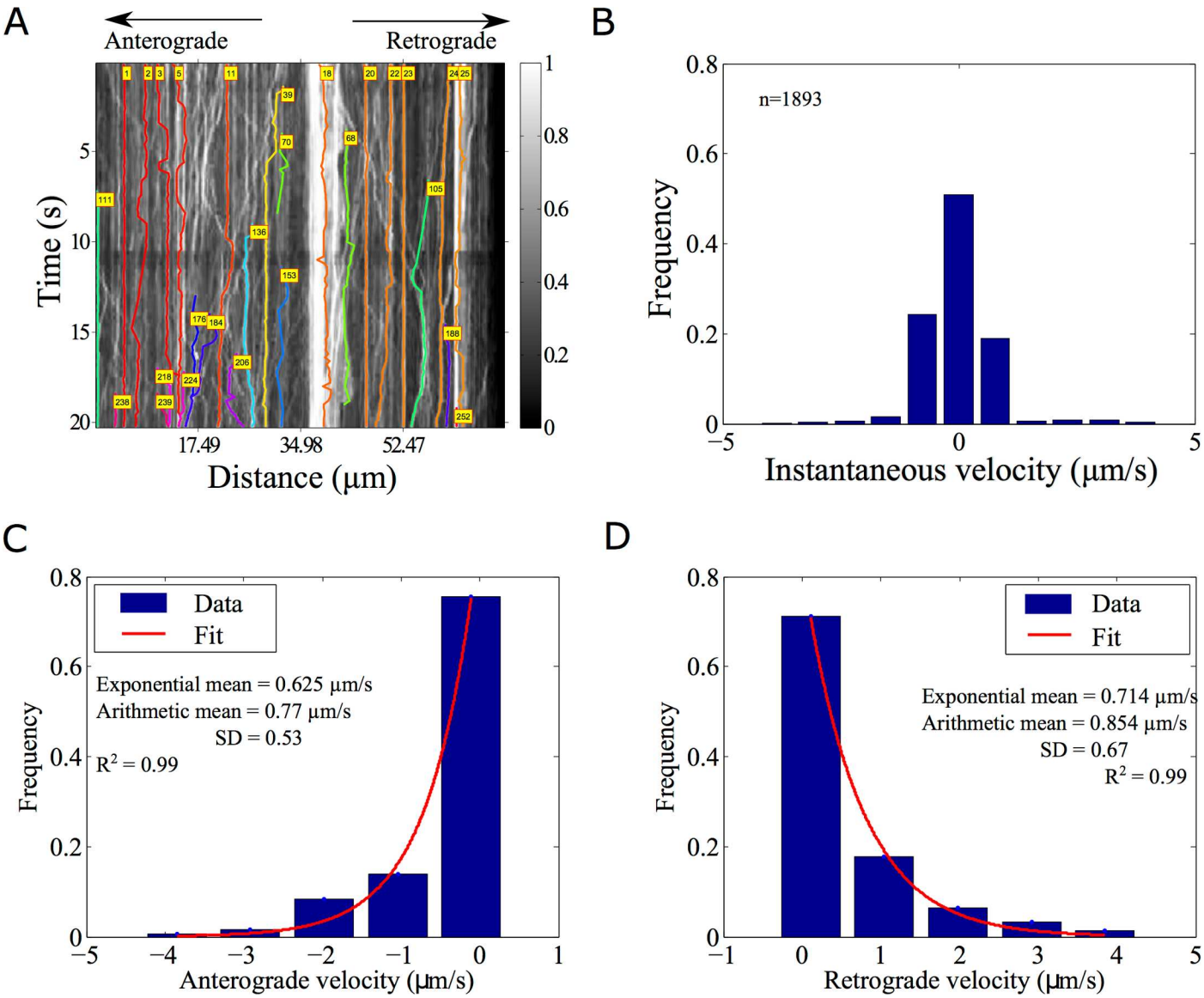


Fig 8. Analysis of synaptic vesicle transport. (A) GFP-Rab3 tagged vesicles from posterior touch cell neurons in *C. elegans* (experimental data from taken from supporting movie S1 Movie from [28]) were analyzed using AMTraK. Colored lines with index numbers indicate tracks. **(B)** The frequency distribution of instantaneous velocities of the vesicles (n = 1592) is plotted using AMTraK (mean: 0.49 µm/s, s.d. 0.88). **(C, D)** The frequency distribution of non-zero velocities are fit with an exponential decay function $y = A*e^{-x/m}$ (red line), where A: scaling factor and *m*: mean. **(C)** The mean anterograde velocity from the fit is 0.625 µm/s with arithmetic mean 0.77±0.53 µm/s (n = 425) and **(D)** the mean retrograde velocity from the fit is 0.714 µm/s with arithmetic mean 0.854±0.67 µm/s (n = 540). Arithmetic means are reported ± standard deviation (s.d.). R^2 indicates the goodness of the fit.

doi:10.1371/journal.pone.0167620.g008

dynamic processes, possibly due to the heterogeneity of the data types and the absence of a single standard method or even criterion, which to make the process less interactive. While developing AMTraK, we tested global (whole-image) methods of edge-segmentation (contour-, watershed- and gradient-based), but found them to be inadequate for the task. Possible reasons include the time-dependent brightness and contrast changes of the sample resulting from either bleaching or intrinsic dynamics. We find that for some applications such as vesicle transport and protein recruitment, the detecting and tracking peaks is ideal, while for microtubule gliding assays edge detection is better. As a result our code allows the user to choose

amongst three different methods of segmentation based on the nature of their data (a) peak detection by *findpeaks* and (b) watershed and (c) edge detection using the Canny edge detector.

Typical problems in peak or edge detection arise when the data has poor signal to noise. This is also seen in our error analysis with increasing noise amplitude (Fig 2D). One solution is to background subtract the image, which can be easily done in multiple tools. The occasional loss of some particles in a time-series such as synaptic vesicles (Fig 8A), despite being visible to the eye, results from a failure in detection or a 'pruning' step used to remove spurious and redundant tracks. Such pruning however was found to be necessary to ensure robustness of the code for handling multiple data types and is simple to trouble-shoot due to the limited number of adjustable parameters. While intensity matching did not improve the percentage vesicles tracked, in future additional features like those used in pattern-matching for tracking [41] could be used further improve the detection percentages. Our test with increasing Gaussian random image noise (Fig 2) also suggests that increases of fluorescently tagged proteins (for instance due to expression level increases *in vivo*), could result in reduced spatial contrast. Such data would then be difficult to automatically quantify using AMTraK. The data would require pre-processing with something similar to an anisotropic diffusion filter [42] to preserve edge information but reduce non-specific signal. In future, multiple data pre-processing routines could be implemented in a separate module, to add to the functionality of the program.

Our quantification of the frequency distribution of synaptic vesicle transport in anterograde and retrograde directions (Fig 8C and 8D) suggests the instantaneous velocities are exponentially distributed. While the arithmetic mean suffices for comparison with experimental reports [28], the quantification of the precise nature of the distribution of velocities could be used as a test of theoretical models. Such a comparison has been made in previous work on synaptic vesicle precursor trafficking [43]. Such models are relevant for both neurophysiology as well as understanding of collective effects in molecular-motor driven vesicle transport *in vivo* [44,45].

The collective motor velocity of human kinesin driven gliding of MTs has been well characterized in previous work [12,46,47]. Many of these studies have shown that the MT length and kinesin density do not affect the mean speed. However, the time-series of individual filaments show small time-dependent variations (Fig 4A and 4B), possibly a result of the local inhomogeneity of motor distributions. This information could be of some use when mixed-motor populations are used [48]. Recent studies of filament motility have used a filament-tracking approach based on a MATLAB program FIESTA [49], with a positional accuracy of 30 nm. We find the distribution of time-averaged velocity of gliding calculated using AMTraK match closely the distribution obtained from analysis using FIESTA (S2 Fig). This suggests that while complex transport dynamics in 2D are indeed better analyzed using tracking tools, for those data sets that are amenable to kymography analysis, AMTraK results are comparable to those obtained from tracking tools with sub-pixel accuracy.

While the dynamics of multiple particles can be simultaneously quantified using AMTraK, the selection of LOIs remains manual. However, once an LOI has been selected, the program can also be used in the "From file" mode to apply a pre-existing LOI to quantify kymographs in other channels (e.g.: bright field, fluorescence) and other fields of view with similar sample geometries. Potentially, LOIs could be generated independent of AMTraK too, provided they are compatible with the input format. The multiple bright-field and fluorescence correlative analysis tools for bacterial image analysis [10,33,50,51] are an example in case. More recent developments in image-analysis software to systematically extract data from microfluidics experiments automatically output channel information [52], which could also form the basis for the LOIs for multiple fields of view. These approaches could in future further increase the throughput our analysis tool.

Table 1. A comparison of features in kymography tools described in literature and commonly in use for cellular and sub-cellular scale images.

Feature / Tool	AMTraK	Multi-kymograph	Makekymograph	Icy-Kymograph Tracker	Kymomaker	Points from Kymograph	Kymograph mt2	KymographClear and KymographDirect
LOI selection	Manual	Manual	Semi-automated	Manual	Manual	No	Manual	Manual
Multiple LOIs	Yes	Yes	Yes	Yes	Yes	No	No	No
Automated track detection	Yes	No	No	Semi-automated	Yes	Semi-automated	No	Yes
Quantification	Yes	Separate	No	Separate	No	XY-coordinates	No	Separate
No. of adjustable parameters	8	1	1	7	13	-	-	-
Split and merge detection	Automatic	No	No	No	No	No	No	Manual
Open source	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Programming language	MATLAB	ImageJ macro	Java (ImageJ plugin)	Plugin for Icy	-	Java (ImageJ plugin)	Java (ImageJ plugin)	ImageJ macro and LabView
Reference	This report	[57]	[58]	[5]	[4]	[59]	[60]	[53]

doi:10.1371/journal.pone.0167620.t001

Multiple software tools for kymography have been described in the recent past in literature and their features are summarized in Table 1 for comparison. While most tools including this one, require user inputs for the process of kymograph generation, only AMTraK and Kymograph Direct [53] automates the detection and connection. However, certain features of AMTraK make it unique, being absent in other comparable tools, such as automated branch-point detection, an integrated quantification module and sub-pixel positional accuracy accessible with an easy to use GUI front-end. In addition, since the code is open source and written in MATLAB, it is more likely to be used in an existing microscopy analysis workflow, due to the increasing spread of MATLAB as a data analysis platform in quantitative cell-biology research [54,55]. Thus, AMTraK could serve as a tool for the rapid quantification of image time-series of transport and assembly kinetics from microscopy. This has become particularly relevant in the context of high-content screening [56], where the spatial interaction patterns are becoming just as important as bulk kinetics measured in traditional high throughput screens.

We have developed an automated tool for the quantification of kymographs. Our approach detects peak and edge information and utilizes a distance minimization approach to link them. We demonstrate the wide utility of our tool by quantifying microtubule transport dynamics, clathrin polymerization kinetics and vesicle transport. Combined with a user-friendly interface, objective detection criteria and open source code, we believe AMTraK can be used to extract more and reproducible statistics from microscopy of sub-cellular dynamics.

Supporting Information

S1 Data. The LOI coordinates generated are stored in the file "LOIselection.txt" when the user chooses the "Interactive" mode of LOI selection at the stage of generating a kymograph. This provides the 2D image coordinates (X and Y) in pixel units, as indicated by the columns labels.
(TXT)

S2 Data. The average statistics for all trajectories are stored in a file "USER_TrackStats.txt". It reports in a column-wise manner the track number, time over which it is tracked (in

user-provided units), speed, net-velocity (displacement/time), tortuosity (displacement/path-length), average of instantaneous velocity and the standard deviation of the average instantaneous velocity. All column headers are labeled for clarity.
(TXT)

S3 Data. The instantaneous (time-dependent) statistics of each track are stored in "USER_InstStats.txt" with track number, time interval to the previous frame in units provided by the user, displacement magnitude, positive/negative displacement (leftwards: negative, rightwards: positive), instantaneous velocity (displacement/time interval), signed-velocity (leftwards: negative, rightwards: positive), and cumulative time (adding up time intervals in units provided by the user). All column headers are labeled for clarity.
(TXT)

S4 Data. The file "Tracklist.txt" stores the time-dependent intensity statistics of each track. This provides the track number, position in distance from the origin (upper-left corner) in pixels, time-frame (frame number), normalized grey-value intensity (divided by the bit-depth of the image) and normalized time-frame (setting the first time-frame to 0). All column headers are labeled for clarity.
(TXT)

S5 Data. The branch-points detected by the code are stored in a file "Branchpoints.txt" which is generated when the user chooses to detect "Splitting events" (check-box) with an appropriate parameter choice. It contains the track-number that splits off from or joins another track, the 1D distance (from the origin at the left edge) and it's time point both in terms of user-provided units. The column headers describe the variables.
(TXT)

S6 Data. User provided values are stored in "All_Parameters.txt". This includes the name and path of the input TIF image time-series, scaling factors (distance, time) and parameters for the detection, tracking and splitting-events.
(TXT)

S1 Fig. The simulated image. (A) The simulated bead image used to estimate the accuracy of the code. A profile through the image (yellow line) is used to generate **(B)** an intensity profile through the three beads.
(PDF)

S2 Fig. Comparing kymography to filament tracking. The frequency distribution of instantaneous velocities obtained after analyzing time-series of MTs gliding on kinesin using AMTraK (red bars) and the high-precision filament-tracking tool, FIESTA (blue bars) are plotted.
(PDF)

S3 Fig. Kinetics of clathrin endocytosis. The fluorescence intensity in grey values (colored circles) as a function of time in seconds estimated from multiple detected tracks after AMTraK analysis (Fig 7A and 7B) of clathrin assembly kinetics in the presence of **(A)** wild-type and **(B)** mutant (L6W) epsin (based on data from Holkar et al. [24]). A single-phase exponential function (the same as in Fig 7C and 7D) is used to fit the data (black line) and the parameters are listed for each fit, with τ indicating the time-constant of assembly in seconds.
(PDF)

S1 Video. Time-series of division and genome-segregation in *E. coli* MG1655 is followed **(A)** in fluorescence with nucleoids labeled by HupA-GFP (grey) and **(B)** DIC is used to follow cell

morphology. Scale bar: 4 μm. Time indicated in minutes.
(ZIP)

Acknowledgments

A grant from Dept. of Biotechnology, Govt. of India in Basic Biology (BT/PR1595/BRB/10/1043/2012) financed part of this work. ARC, KJ and MSG are supported by fellowships from the University Grants Commission (UGC) India (F.2-14/2011 (SA-1)), DST-Inspire Fellowship and Indian Council of Medical Research (ICMR) respectively. We are grateful to Peter J. Foster for sharing the original MT aster time-series data and Sachin Holkar for sharing the original image time-series data and calibration values for the Clathrin assembly kinetics. We are grateful to Sudip Mondal for an explanation of the vesicle transport analysis.

Author Contributions

Conceptualization: CAA ARC.

Data curation: ARC.

Formal analysis: ARC.

Funding acquisition: CAA.

Investigation: ARC KJ MSG.

Methodology: ARC KJ MSG.

Project administration: CAA.

Resources: CAA.

Software: ARC CAA.

Supervision: CAA.

Validation: CAA ARC KJ MSG.

Visualization: ARC CAA.

Writing – original draft: CAA.

Writing – review & editing: CAA ARC.

References

1. Pereira AJ, Maiato H (2010) Improved kymography tools and its applications to mitosis. Methods 51: 214–219. doi: 10.1016/j.ymeth.2010.01.016 PMID: 20085815
2. Rietdorf J, Seitz A (2008) Multi Kymograph. Available: http://fiji.sc/Multi_Kymograph.
3. Chetta J, Shah SB (2011) A novel algorithm to generate kymographs from dynamic axons for the quantitative analysis of axonal transport. J Neurosci Methods 199: 230–240. Available: http://www.ncbi.nlm.nih.gov/pubmed/21620890. doi: 10.1016/j.jneumeth.2011.05.013 PMID: 21620890
4. Chiba K, Shimada Y, Kinjo M, Suzuki T, Uchida S (2014) Simple and direct assembly of kymographs from movies using KYMOMAKER. Traffic 15: 1–11. doi: 10.1111/tra.12127 PMID: 24102769
5. Chenouard N, Buisson J, Bloch I, Bastin P, Olivo-Marin J-C (2010) Curvelet analysis of kymograph for tracking bi-directional particles in fluorescence microscopy images. IEEE International Conference on Image Processing (ICIP). http://icy.bioimageanalysis.org/plugin/KymographTracker.
6. Skarstad K, Steen HB, Boye E (1985) Escherichia coli DNA distributions measured by flow cytometry and compared with theoretical computer simulations. J Bacteriol 163: 661–668. PMID: 3894332

7. Nordström K, Dasgupta S (2006) Copy-number control of the Escherichia coli chromosome: a plasmidologist's view. EMBO Rep 7: 484–489. doi: 10.1038/sj.embor.7400681 PMID: 16670681
8. Nielsen HJ, Ottesen JR, Youngren B, Austin SJ, Hansen FG (2006) The Escherichia coli chromosome is organized with the left and right chromosome arms in separate cell halves. Mol Microbiol 62: 331–338. doi: 10.1111/j.1365-2958.2006.05346.x PMID: 17020576
9. Spahn C, Endesfelder U, Heilemann M (2014) Super-resolution imaging of Escherichia coli nucleoids reveals highly structured and asymmetric segregation during fast growth. J Struct Biol 185: 243–249. doi: 10.1016/j.jsb.2014.01.007 PMID: 24473063
10. Sliusarenko O, Heinritz J, Emonet T, Jacobs-Wagner C (2011) High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. Mol Microbiol 80: 612–627. doi: 10.1111/j.1365-2958.2011.07579.x PMID: 21414037
11. Fisher JK, Bourniquel A, Witz G, Weiner B, Prentiss M, et al. (2013) Four-Dimensional Imaging of E. coli Nucleoid Organization and Dynamics in Living Cells. Cell 153: 882–895. doi: 10.1016/j.cell.2013.04.006 PMID: 23623305
12. Howard J, Hudspeth AJ, Vale RD (1989) Movement of microtubules by single kinesin molecules. Nature 342: 154–158. doi: 10.1038/342154a0 PMID: 2530455
13. Nitzsche B, Bormuth V, Brauer C, Howard J, Ionov L, et al. (2010) {S}tudying kinesin motors by optical 3D-nanometry in gliding motility assays. Methods Cell Biol 95: 247–271. doi: 10.1016/S0091-679X(10)95014-0 PMID: 20466139
14. Leduc C, Ruhnow F, Howard J, Diez S (2007) Detection of fractional steps in cargo movement by the collective operation of kinesin-1 motors. Proc Natl Acad Sci U S A 104: 10847–10852. doi: 10.1073/pnas.0701864104 PMID: 17569782
15. Toyoshima YY, Kron SJ, McNally EM, Niebling KR, Toyoshima C, et al. (1987) Myosin subfragment-1 is sufficient to move actin filaments in vitro. Nature 328: 536–539. doi: 10.1038/328536a0 PMID: 2956522
16. Martin DS (2013) Measuring microtubule persistence length using a microtubule gliding assay. Methods Cell Biol 115: 13–25. doi: 10.1016/B978-0-12-407757-7.00002-5 PMID: 23973063
17. Bicek AD, Tuzel E, Demtchouk A, Uppalapati M, Hancock WO, et al. (2009) Anterograde microtubule transport drives microtubule bending in LLC-PK1 epithelial cells. Mol Biol Cell 20: 2943–2953. doi: 10.1091/mbc.E08-09-0909 PMID: 19403700
18. Ungewickell E, Branton D (1981) Assembly units of clathrin coats. Nature 289: 420–422. PMID: 7464911
19. Avinoam O, Schorb M, Beese CJ, Briggs JA, Kaksonen M (2015) Endocytosis. Endocytic sites mature by continuous bending and remodeling of the clathrin coat. Science 348: 1369–1372. doi: 10.1126/science.aaa9555 PMID: 26089517
20. Skruzny M, Desfosses A, Prinz S, Dodonova SO, Gieras A, et al. (2015) An organized co-assembly of clathrin adaptors is essential for endocytosis. Dev Cell 33: 150–162. doi: 10.1016/j.devcel.2015.02.023 PMID: 25898165
21. Pucadyil TJ, Schmid SL (2010) Supported bilayers with excess membrane reservoir: a template for reconstituting membrane budding and fission. Biophys J 99: 517–525. doi: 10.1016/j.bpj.2010.04.036 PMID: 20643070
22. Neumann S, Pucadyil TJ, Schmid SL (2013) Analyzing membrane remodeling and fission using supported bilayers with excess membrane reservoir. Nat Protoc 8: 213–222. doi: 10.1038/nprot.2012.152 PMID: 23288321
23. Kalthoff C, Alves J, Urbanke C, Knorr R, Ungewickell EJ (2002) Unusual structural organization of the endocytic proteins AP180 and epsin 1. J Biol Chem 277: 8209–8216. doi: 10.1074/jbc.M111587200 PMID: 11756460
24. Holkar SS, Kamerkar SC, Pucadyil TJ (2015) Spatial Control of Epsin-induced Clathrin Assembly by Membrane Curvature. J Biol Chem 290: 14267–14276. doi: 10.1074/jbc.M115.653394 PMID: 25837255
25. Allan VJ, Vale RD (1991) Control of Microtubule-based Membrane Transport and Tubule Formation In Vitro. J Cell Biol 113: 347–359. PMID: 2010466
26. Welzel O, Knörr J, Stroebel AM, Kornhuber J, Groemer TW (2011) A fast and robust method for automated analysis of axonal transport. Eur Biophys J 40: 1061–1069. doi: 10.1007/s00249-011-0722-3 PMID: 21695534
27. Hill DB, Plaza MJ, Bonin K, Holzwarth G (2004) Fast vesicle transport in PC12 neurites: velocities and forces. Eur Biophys J 33: 623–632. doi: 10.1007/s00249-004-0403-6 PMID: 15071760

28. Mondal S, Ahlawat S, Rau K, Venkataraman V, Koushika SP (2011) Imaging in vivo Neuronal Transport in Genetic Model Organisms Using Microfluidic Devices. Traffic 12: 372–385. doi: 10.1111/j.1600-0854.2010.01157.x PMID: 21199219

29. Otsu N (1979) A Threshold Selection Method from Gray-Level Histograms. IEEE Trans Syst Man Cybern SMC-9: 62–66.

30. Brookes, Mike (2005) Voicebox. http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html.

31. Meyer F (1994) Topographic distance and watershed lines. Signal Processing 38: 113–125.

32. Canny J (1986) A Computational Approach to Edge Detection. IEEE Trans Patt Analy Mach Int 8: 679–698.

33. Athale CA, Chaudhari H (2011) Population length variability and nucleoid numbers in Escherichia coli. Bioinformatics 27: 2944–2948. doi: 10.1093/bioinformatics/btr501 PMID: 21930671

34. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9: 671–675. PMID: 22930834

35. Wery M, Woldringh CL, Rouviere-Yaniv J (2001) HU-GFP and DAPI co-localize on the Escherichia coli nucleoid. Biochimie 83: 193–200. PMID: 11278069

36. Thévenaz P, U.E. R, Unser M (1998) Thévenaz P1, Ruttimann UE, Unser M. IEEE Trans Image Process 7: 27–41.

37. Cytoskeleton (2005) Kinesin Motility Assay Biochem Kit BK027. Denver, CO, USA.

38. Stewart RJ, Thaler JP, Goldsteint LSB (1993) Direction of microtubule movement is an intrinsic property of the motor domains of kinesin heavy chain and Drosophila ncd protein. Proc Nat Acad Sci USA 90: 5209–5213. PMID: 8506368

39. Foster PJ, Fürthauer S, Shelley MJ, Needleman DJ (2015) Active contraction of microtubule networks Active Contraction of Microtubule Networks. Elife 4:e10837: 1–21. doi: 10.7554/eLife.10837 PMID: 26701905

40. Welzel O, Boening D, Stroebel A, Reulbach U, Klingauf J, et al. (2009) Determination of axonal transport velocities via image cross- and autocorrelation. Eur Biophys J 38: 883–889. doi: 10.1007/s00249-009-0458-5 PMID: 19404633

41. Miura K (2005) Tracking Movement in Cell Biology. In: Rietdorf J, editor. Adv. Biochem. Engin/Biotechnol. Springer Berlin / Heidelberg. pp. 267–295.

42. Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. IEEE Trans Pattern Anal Mach Intell 12: 629–639.

43. Maeder CI, San-Miguel A, Wu EY, Lu H, Shen K (2014) In vivo neuron-wide analysis of synaptic vesicle precursor trafficking. Traffic 15: 273–291. doi: 10.1111/tra.12142 PMID: 24320232

44. Bridgman PC (1999) Myosin Va movements in normal and dilute-lethal axons provide support for a dual filament motor complex. J Cell Biol 146: 1045–1060. PMID: 10477758

45. Hendricks AG, Perlson E, Ross JL, Schroeder HW, Tokito M, et al. (2010) {M}otor coordination via a tug-of-war mechanism drives bidirectional vesicle transport. Curr Biol 20: 697–702. doi: 10.1016/j.cub.2010.02.058 PMID: 20399099

46. Howard J (2001) Mechanics of Motor Proteins and the Cytoskeleton. Sunderland: Sinauer Associates.

47. Gibbons F, Chauwin JF, Despósito M, José J V (2001) A dynamical model of kinesin-microtubule motility assays. Biophys J 80: 2515–2526. doi: 10.1016/S0006-3495(01)76223-6 PMID: 11371430

48. Ikuta J, Kamisetty NK, Shintaku H, Kotera H, Kon T, et al. (2014) Tug-of-war of microtubule filaments at the boundary of a kinesin- and dynein-patterned surface. Sci Reports 4: 5281.

49. Ruhnow F, Zwicker D, Diez S (2011) Tracking Single Particles and Elongated Filaments with Nanometer Precision. Biophys J 100: 2820–2828. doi: 10.1016/j.bpj.2011.04.023 PMID: 21641328

50. Guberman JM, Fay A, Dworkin J, Wingreen NS, Gitai Z (2008) PSICIC: noise and asymmetry in bacterial division revealed by computational image analysis at sub-pixel resolution. PLoS Comp Biol 4: e1000233.

51. Ducret A, Quardokus EM, Brun Y V (2016) MicrobeJ, a high throughput tool for quantitative bacterial cell detection and analysis. Nat Microbiol 1: 1–7.

52. Sachs CC, Grünberger A, Helfrich S, Probst C, Wiechert W, et al. (2016) Image-Based Single Cell Profiling: High- Throughput Processing of Mother Machine Experiments. PLoS One 11: e0163453. doi: 10.1371/journal.pone.0163453 PMID: 27661996

53. Mangeol P, Prevo B, Peterman EJG (2016) KymographClear and KymographDirect: two tools for the automated quantitative analysis of molecular and cellular dynamics using kymographs. Mol Biol Cell 27: 1948–1957. doi: 10.1091/mbc.E15-06-0404 PMID: 27099372

54. Shekhar S, Zhu L, Mazutis L, Sgro AE, Fai TG, et al. (2014) Quantitative biology: where modern biology meets physical sciences. Mol Biol Cell 25: 3482–3485. doi: 10.1091/mbc.E14-08-1286 PMID: 25368426

55. Howard J (2014) Quantitative cell biology: the essential role of theory. Mol Biol Cell 25: 3438–3440. doi: 10.1091/mbc.E14-02-0715 PMID: 25368416

56. Zanella F, Lorens JB, Link W (2010) High content screening: seeing is believing. Trends Biotechnol 28: 237–245. doi: 10.1016/j.tibtech.2010.02.005 PMID: 20346526

57. Rietdorf, Jens (FMI B, Seitz A (EMBL H (2008) Kymograph (time space plot) Plugin for ImageJ. http://www.embl.de/eamnet/html/body_kymograph.html.

58. Hallman S (2007) Make_Kymograph plugin. http://www.ics.uci.edu/~fowlkes/sam_hallman/doc/

59. Das M, Drake T, Wiley DJ, Buchwald P, Vavylonis D, et al. (2012) Oscillatory dynamics of Cdc42 GTPase in the control of polarized growth. Science (80-) 337: 239–243.

60. Siebrasse JP, Kaminski T, Kubitscheck U (2012) Nuclear export of single native mRNA molecules observed by light sheet fluorescence microscopy. Proc Natl Acad Sci USA 109: 9426–9431. doi: 10.1073/pnas.1201781109 PMID: 22615357