Physiological constraints and evolutionary trade-offs underlying bacterial aging, caloric restriction and longevity

Yifan Yang

To cite this version:

Yifan Yang. Physiological constraints and evolutionary trade-offs underlying bacterial aging, caloric restriction and longevity. Human health and pathology. Université Sorbonne Paris Cité, 2015. English. NNT : 2015USPCB158 . tel-01588777v1

HAL Id: tel-01588777

https://theses.hal.science/tel-01588777v1

Submitted on 17 Sep 2017 (v1), last revised 20 Sep 2017 (v2)

UNIVERSITÉ PARIS DESCARTES

DOCTORAL THESIS

Physiological constraints and evolutionary trade-offs underlying bacterial aging, caloric restriction and longevity

---

Author:
Yifan YANG

Supervisors:
Dr. François TADDEI
Dr. Ariel B. LINDNER

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

in the

INSERM U1001
Ecole Doctorale n°474 Frontières du Vivant

November 2016

“Physics is like sex: sure, it may give some practical results, but that’s not why we do it.”

Richard P. Feynman

UNIVERSITÉ PARIS DESCARTES

Abstract

Faculté de Médecine
Ecole Doctorale n°474 Frontières du Vivant

Doctor of Philosophy

Physiological constraints and evolutionary trade-offs underlying bacterial aging, caloric restriction and longevity

by Yifan YANG

The evolutionary theories of aging and the disposable soma theory in particular, have been the theoretical basis for a recent surge of animal aging research. Yet their central assumption about the physiology of cellular maintenance and repair has not been empirically tested. In this thesis, I analysed the physiology of *E.coli* aging under carbon starvation, as a model system to empirically validate evolutionary theories of aging. Microfluidic tools are used to isolate large populations of isogenic single *E.coli* cells, and to achieve homogenous carbon starvation. Despite sharing the same genetical background and environmental conditions, individual cells in the population exhibit significant variations in lifespans and causes of death. Distributions of lifespans exhibit typical features of the aging process, often seen in animal and human demographic studies. The rate of aging can be altered by mutations of the general stress response pathway. Resembling caloric restriction induced longevity, the general stress response pathway extends starvation lifespans of *E.coli* by attenuating the effect of aging at the expense of immediate needs of the cells. A quantitative model of this physiological trade-off is constructed and correctly predicted experimental observations. As a conclusion, I substantiate the disposable soma theory of aging with the physiological details of*E.coli* aging in starvation
.

Acknowledgements

I owe tremendous gratitude to my thesis advisors, Dr. Ariel Lindner and Dr. François Taddei, who have given me the intellectual environments, freedom, and material support for conducting research on subjects that truly intrigue me, and provided the guidance and motivations that helped me through difficult times.

This thesis work would not have been possible without the pioneering work of Dr. Luping Xu, who first explored both the idea and the technology of using mother machine to study *E.coli* chronological aging. I also have to thank various members of INSERM unit 1001, who have created stimulating intellectual environments and generously helped me with my experimental work. Among them, I particularly have to thank Dr. Ming Ni for teaching me microscopy and micro-fabrication techniques; Dr. Peipei Chen for micro-fabricating the templates used to make some of the microfluidic devices used in this work; Mrs. Chantal Lotton and Dr. Ana Luísa Santos for assisting me with the large workload of both plating and microfluidic experiments; Dr. Xiaohu Song for providing the image analysis software with which I extracted quantitative data from the mother machine experiments; Mrs. Marie-Florence Bredeche for constructing various fluorescent reporter strains used throughout my thesis work; Dr. Jake Wintermute, Mr. Antoine Decrulle and Dr. Ming Ni for various interesting scientific discussions. In addition to INSERM U1001, Dr. Johan Paulsson's group at Harvard Medical School has hosted me for 6 months, during which I worked on theoretical aspects of my work. Dr. Paulsson has provided important advice that contribute to the development of the work described in this thesis.

Last but not the least, the AXA chair on aging and longevity played an indispensable role throughout my thesis work. Besides the financial support I received the last 2 years of my thesis work, I also drew a lot of intellectual inspirations from the annual AXA workshops on aging and longevity.

Contents

List of Figures

Abbreviations

E.coli	***E**scherichia* ***coli***
CR	**C**aloric **R**estriction
OD600	**O**ptical **D**ensity at 600nm
GASP	**G**rowth **A**t **S**tationary **P**hase
PTFE	**P**oly**T**etra**F**luoro**E**thylene
AU	**A**rbitrary **U**nit
PI	**P**ropidium **I**odide
CFU	**C**olony **F**orming **U**nit
PEG	**P**oly**E**thylene **G**lycol
TVDiff	**T**otal **V**ariation Regularized Numerical **Diff**erentiation
MLE	**M**aximum **L**ikelihood **E**stimates
p.d.f.	**p**robability **d**ensity **f**unction
CWT	**C**ontinuous **W**avelet **T**ransformation
ROI	**R**egion **O**f **I**nterest
K-M	**Kaplan-Meier**
N-A	**Nelson-Aalen**

Chapter 1

Introduction

Aging research, although attracting far more than its fair share of public attention for obvious reasons, has been considered for decades as scientifically intractable, even suspicious as a field. Yet in recent years, this somewhat marginal status has changed dramatically, in large parts due to a convergence of three forces: genetic discoveries of lifespan regulations in model organisms, expansions of the biochemical and molecular biology toolboxes, and validation of some longstanding predictions from the evolutionary theories of aging. Despite making substantial empirical progress, the molecular biology agenda of aging research has rarely clarified the conceptual pictures underlying this inherently interdisciplinary field. Considering the complexity of biology, this is to be expected, just as Shermer (2012) wrote in reviewing Stuart Firestein's book on ignorance and science "... as the sphere of scientific knowledge increases, so does the surface area of the unknown ...". This ever expanding shroud of unknown is further compounded by the rush to apply whatever knowledge gained in a millennium-old hunt for the fountain of youth, causing much confusion not only for the public but also for researchers.

It might be illuminating to examine how an equally fundamental and mysterious biological phenomenon, inheritance, was tackled by the founders of the molecular biology revolution. The works of Delbrück, Luria and Hershey on the nature of mutations, and the work of the Jacob and Monod on gene regulation, share some striking features. There were the uses of quantitative experiments and modelling to address conceptual questions (Luria & Delbrück, 1943; Monod, 1958) and the convictions of the usefulness of simple model organisms such as phages and bacteria, as in Jacques Monod's famous quote "anything found to be true of *E. coli* must also be true of elephants".

This thesis describes my attempt to apply this approach to the study of aging. By using a quantitative physiological approach to study *E.coli* aging in carbon starvation, my aim is not to advance the frontier of human longevity, but to empirically explore the fundamental concepts in aging research. In this chapter, I will first describe as background, the brief history of the theories and concepts in aging research; and then, how these concepts fit into the particular lifestyle and physiology of *E.coli*. At the end of the chapter, I will introduce my conceptual framework, raise the central question of my thesis and outline the contents of the remain chapters.

1.1 Background

Biological aging, defined as the decline of physiological functions with age, is a multifacetted phenomenon. Genetics, environments, physiology, natural selection and even the way age is defined (i.e. age structure) all play important roles. Here, I outline the theories and concepts of aging relevant to this work, while presenting their limits and their possible over-interpretations. In doing so, I hope to make a case for an integrated, interdisciplinary approach to understand aging.

1.1.1 Evolutionary theory of aging

The evolutionary theory of aging has many versions. They all start from the same basic insights of Medawar and Hamilton, that selective pressure decrease with age. But they differ in the assumptions that they make about mutations' effect on aging. The mutation accumulation theory presumes mutations have additive, age-specific effects, while the antagonistic pleiotropy theory (Williams, 1957) presumes trade-offs between traits at earlier and later ages. Baudisch (2005) has shown theoretically that the declining selective force with age is not enough to explain aging patterns. Aging patterns are also shaped by the effects of mutations. In favour of the antagonistic pleiotropy theory, various empirical works have shown that trade-offs between older and younger ages are indeed true (Bouwhuis *et al.*, 2010; Bryant & Reznick, 2004; Gustafsson & Part, 1990; Hayward *et al.*, 2014). Many have argued for the validity of the antagonistic pleiotropy theory (Le Bourg, 2001). As an evolutionary theory, the subject matter is not the direct causes of aging, but how aging phenotypes could be possible in the face of fitness-optimising natural selection.

In order to develop a more useful theory, proponents of the antagonistic pleiotropy theory need to understand the quantitative relationship between young and old age fitness components. This requires knowledge about the physiological processes of mortality, aging and reproduction. One famous such attempt was by Kirkwood (1977), where he argued for translation "error catastrophe" being the physiological cause of cellular aging, and proposed the disposable soma theory presuming that trade-offs between young and old ages are the energetic cost of error repair in the soma. Reassessing the different parts in that original paper, the physiological cause of cellular aging are not believed to be translation "error catastrophe" (Gavrilov & Gavrilova, 2002); but whatever the cause of aging was, the idea that energetic cost of damage repair mediates trade-offs could still be valid. It has to be pointed out here that the disposable soma theory is in essence a particular version of the antagonistic pleiotropy theory, where the physiological nature of the young and old age trade-offs is assumed to be the competition for resources between cellular maintenance and more immediate needs.

Any evolutionary life-history theories could be potentially empirically tested by examining the plasticity and reaction norms of the life-history traits in question, i.e. how environmental factors modulate them (Stearns, 1989). The central assumption that organisms should adopt the optimal life-history strategy to maximise lifetime reproductive success should predict quantitatively the way organisms change their life-history strategies when environmental conditions fluctuate. For the disposable soma theory, one obvious environmental variable is the caloric intake since it is assumed to be the limiting resource. This is where the theory has found the most empirical support and potential applications. The phenomenon of caloric-restriction (CR) induced longevity are found in many model organisms, and the genetics of corresponding regulatory pathways have been the subjects of intense research efforts. However, to discern the quantitative effects of such pathways especially at older ages require longitudinal studies of large populations. Such studies face practical difficulties associated with the long timescales, and in controlling genetics backgrounds and environments of large populations (Nussey *et al.*, 2008). This makes laboratory studies of model organisms with short lifespans and controllable environments the ideal setting for testing the evolutionary theories of aging.

1.1.2 Laboratory studies of aging and their interpretations

Since the discovery of insulin/IGF-I pathway mutants that extend lifespans in nematodes (Kenyon *et al.*, 1993), similar mutants and pathways have been identified for multiple model organisms (Gems & Partridge, 2013). These nutrient-sensing pathways modulate aging in the way predicted by the disposable soma theory, and attract widespread attentions as potential targets of pharmaceutical intervention to slow aging.

However, there are some pitfalls in this genetics research agenda. Nutrients-sensing pathway mutants with prolonged lifespans in lab environments do not guarantee to slow the aging process through increased maintenance investments as the disposable soma theory predicts. First, the lifespan extensions could be due to the prevention of particular pathologies associated with lab conditions. This issue has been central to a lot of the recent controversies, ranging from microbiology (Burtner *et al.*, 2009) and primates research (Mattison *et al.*, 2012). In the former case, observed aging in budding yeasts were actually explained by the PH changes in the cell culture media. In the latter case of Rhesus macaque caloric restriction studies, different conclusions were reached because of differences in food availabilities in the control group.

The second unrelated but perhaps more central issue is that lifespans can be extended by a proportional reduction in mortality risks in all ages, without affecting the aging process. While the former is still good news as far as longevity is concerned, this type of pathways can not be argued to be modulating the aging process, even if it involves molecular repair or damage resistance. Thus if a certain nutrient-sensing, lifespan extending pathway only affect mortality in age-independent manners, it can not be considered to be evidence for the disposable soma theory of aging. In fact, the life-history trade-off assumed in disposable soma theory is between immediate needs for resources, and maintenance investments whose effects only manifest in older ages (Shanley & Kirkwood, 2000). This type of maintenance effects that have an age-dependent component is often presumed in models, but rarely observed or established empirically. Unfortunately, this central assumption is often misunderstood to mean trade-offs between reproduction and immediate stress resistance.

The difference between age-independent and age-dependent modes of lifespan extension not only have theoretical implications but gets at the heart of the physiological concept of aging. Aging is often understood to be due to the 'wear and tear' damages of molecular and cellular machinery, a process at the end of which is death. In this view, aging is simply the inescapable 'second law of thermodynamics' that eventual leads to death, and the intervention of aging is simply protection against damages (Gems & Partridge, 2013). However, if there are indeed pathways whose effects on mortality risks only manifest at older ages, and can be differentiated with age-independent protective mechanisms, then the aging process can be distinct from the different pathologies and physiological failures that lead to death. If validated for human, this view would provide the scientific basis for targeting aging itself for pharmaceutical interventions. In fact, a current effort is underway to gain governmental regulatory approval for clinical trials based on this point of view (Hayden, 2015).

1.1.3 Bacteria as models for aging

Our lab has shown previously that aging exists in even binary-fission bacteria such as *E.coli* (Stewart *et al.*, 2005). In this case, the age structure is that of the age of cell poles, since the two poles of one cell are synthesised at different cell divisions. It has been shown that accumulated molecular damages at old poles (Lindner *et al.*, 2008) contribute the aging process. Another possible age structure can be defined in growth arrested populations, and age is simply the time spent in these conditions (Nyström, 2003). It has been shown that individual *E.coli* cells can live as long as 150 hours (Wang *et al.*, 2010a). Although I do not expect the physiological aging processes associated with these two age structures to be the same as the aging process in multicellular organisms, the evolutionary theories of aging should still apply.

One of the main regulators of maintenance in *E.coli* is the general stress response pathway mediated by rpoS. The rpoS pathway respond to multiple environmental signals and especially starvations of various nutrients (Bougdour *et al.*, 2008; Peterson *et al.*, 2012). When activated, rpoS increases maintenance investments and stress responses. It is named the general stress response because once induced by one particular signal, it offers protection and prolonged survival under multiple conditions, i.e. cross protection. Being an alternative sigma factor for the RNA polymerase, the RpoS protein regulates maintenance investment through competition with the vegetative sigma factor RpoD for RNA polymerase holoenzymes (Nyström, 2004). Thus the general stress response coincide with many features of caloric-restriction induced longevity. Yet without quantitative characterisations, it is unclear whether the pathway offers protections to the cells in age-independent manners such as enhancing stress resistances or through slowing down the aging process as predicted by the disposable soma theory.

The rpoS pathway is also highly variable both across various environmental isolates of *E.coli* (Ferenci *et al.*, 2011) and also in laboratory evolution (Spira *et al.*, 2011). A genetic screen in our lab (Fontaine *et al.*, 2008) has identified a mutant of rpoS pathway, ΔrssB, with prolonged lifespan. These observations show that rpoS pathway is highly evolvable and is subject to natural selection based on its effects on life-history traits. Yet in almost all cases, despite lower stress resistance, loss-of-function mutants of rpoS pathway are selected in almost every conditions including stationary phase batch cultures and chemostats in extremely low dilution rates. This is understood to be due to Growth-At-Stationary-Phase (GASP) phenotypes (Finkel, 2006). It raises the technical issue of controlling environments in batch culture and elimination of social interactions between individual cells in a cohort. More curiously, it begs the question: what is the selective pressure that maintains a functional rpoS gene in the wildtype strain in the first place?

The short lifespan of *E.coli* (Wang *et al.*, 2010a), combined with a mature genetic toolkit and well-understood physiology, offers us a great opportunity to test the evolutionary theories of aging. In particular, assessing the way in which the general stress response pathway impacts lifespans regarding age will be a direct empirical test of the basic life-history assumptions and predictions of the disposable soma theory.

1.2 Questions, approaches and objectives

1.2.1 General questions

I decided to study with a quantitative approach *E.coli* conditional aging in growth arrests induced by carbon starvation. I focused on the role of general stress response in modulating lifespan and aging, both as an attempt to provide physiological and systems biology basis for life-history traits such as aging and longevity, and as a special case to test the evolutionary theory of aging.

The general questions I ask in this study are:

- In growth-arrested *E.coli* populations, are there trade-offs between growth and maintenance, between immediate energetic expenditure and future survival?
- How is the *E.coli* lifespan distributed under carbon starvation conditions? How is the distribution affected by the general stress response? Is the effect of general stress response age independent? Or it actually decreases the rate of aging?
- What is the life-history strategy associated with the observed lifespan patterns? Is such a strategy adaptive? If yes, to what environment?

1.2.2 Conceptual approach

Growth rate is possibly the most dominant trait for bacterial fitness, not only in terms of general physiology but also for the trade-off between maintenance and growth as well. On one hand, both growth and survival are major components of long-term population fitness. Divestment of resources into maintenance versus growth will directly impact the balance between the two. On the other hand, growth and division will dilute and segregate cellular damages thus change the probability of cellular death and survival. Thus the long-term optimal strategy in maintenance investment is dependent on growth rate. It is important to understand to understand general stress response's effect on growth to understand both its impact on overall fitness and cellular survival.

Maintenance investment impact growth through a mechanism known as 'sigma competition'. This is well understood in the literature, both mechanistically and quantitatively (Scott *et al.*, 2010). Whereas the relationship between mortality rates and maintenance investment is much less well known. Given both published data (Wang *et al.*, 2010a) and our own experimental results, mortality rate is assumed to be exponentially dependent on damage accumulation and maintenance investment.

Comparison between maintenance investment strategy derived from optimal fitness assumptions, and the actual behaviour of general stress response in *E.coli* MG1655 yields some interesting insights. The observation that the rpoS regulon is not highly induced under even modest growth conditions (doubling time $< 20h$) suggest that maintenance investment is only beneficial when cellular damage could not be effectively diluted by growth. The fact that specific stress response pathways such as oxyR or rpoH repress rpoS suggests that the general stress response of induced by rpoS is an insurance policy when growth is not possible, rather than dealing with specific stresses.

1.2.3 Experimental approach

Aiming to go beyond the stage of qualitative insights towards quantitative testing of the optimal fitness assumptions, I used microfluidic devices to measure lifespan distributions on a single-cell basis. There are several issues that these systems are designed to overcome.

First, we need a way to follow large populations of cells through their lifespans, to find correlations between life-history traits across different ages. Secondly, crosstalk and interactions through the media have to be prevented between the cells, since rpoS loss-of-function mutants are selected in growth competition in low nutrition environments due to the ability to utilise the nutrients released from the lysis of other cells (Vulic & Kolter, 2001). Both of these factors make a cell-array approach necessary, where large populations of individual cells are trapped in micro-fabricated arrays while exposed to fast flow of media.

The second issue, as discussed in the previous section, is that for the energetic cost of general stress response to be significant for the cells, they should be exposed to such low level of carbon source that they barely grow. Unfortunately for us, the fast flow of media designed to eliminate cell debris also provide carbon sources in the form of trace contaminants leached from lab plasticware. We need to reduce carbon source levels supplied by contaminants before carbon starvation can be studied in a controllable manner.

Both of these issues are resolved during my thesis work by making improvements on a pre-existing microfluidic device called the mother machine (Wang *et al.*, 2010a), documented in Chapter 2. I also provide detailed experimental protocols in Appendix A and Appendix B. With my experiment system, I measured the life-history traits of individual cells including growth, lifespans and transcriptional reporters of general stress response through time-lapse microscopy. With these data, we found that the Gompertz distribution, which is commonly used to model human and animal lifespan data, can also be used to model bacterial lifespans. We can also observe correlations among life-history traits across different ages. I measured the lifespan distributions of several mutants of the general stress response pathway to assess the quantitative effects of this pathway on lifespan distribution. Both these types of data provided consistent conclusions about the mode of action of the general stress response and supported the disposable soma theory of aging.

1.2.4 Outlining the remaining chapters

The following chapters described the experimental methods and implementation details, analysed *E.coli* carbon-starvation life-history traits on both the population and the individual level. These experimental results showed that the general stress response is indeed evolved to protect the cell through slowing the aging process. And finally I articulated the life-history strategy implemented by the general stress response through a toy model.

I documented the following achieved objectives in the remaining chapters:

- Developed microfluidic platforms that enable environmentally-stable, nutrition-controlled culture of individually-isolated large *E.coli* populations monitored by single-cell time-lapse microscopy, Chapter 2, Appendix A and Appendix B;
- Measured the lifespan distributions of aging *E.coli* cells under carbon-starvation conditions, Chapter 3;
- Assessed the origins of such distributions and the mode of action that maintenance investment have on the aging process, Chapter 4;
- Quantitatively re-constructed of the disposable soma theory with an experimentally validated physiological model of growth and maintenance in the context of *E.coli* starvation, Chapter 5.

Chapter 2

Carbon starvation in microfluidics

Before the discovery of *E.coli* aging associated with pole age in our laboratory (Stewart *et al.*, 2005), the only kind of bacterial aging assumed to exist are those found in growth arrested populations. When bacterial populations run out of nutrients, they enter into a physiological state usually called 'stationary phase'. This name indicates a stationary population density in culture, but it has been shown that cell physiology does deteriorate with time due to accumulation of damages on the molecular level (Nyström, 2003). While it is unclear whether the aging process associated with pole age is subjected to natural selection (Wang *et al.*, 2010a), the aging process associated with starvation and growth arrests is clearly a strong target for natural selection due to the 'feast-or-famine' lifestyle of gut bacteria. In fact, even in the day-to-day microbiology operations such as storage and transfer of bacterial strains, where efforts are taken to avoid evolution as much as possible, traits and genes associated with starvation still evolve (Spira *et al.*, 2011). For these reasons, I decided that the conditional aging in bacterial starvation presents the best opportunity for me to test the evolutionary theory of aging.

Previous investigations of single cell conditional aging under starvation are done in batch cultures with extremely high cell densities. In these conditions, media conditions as well as cell physiology can change with starvation time (Burtner *et al.*, 2009). Furthermore, deaths of a sub-population of cells in the batch culture could release debris and chemicals that either alter the media or directly interact with the remaining cells (Vulic & Kolter, 2001). In fact, stationary phase batch cultures are dynamic systems of interacting cells and their environment (Finkel, 2006). To overcome these drawbacks, we decided to use micro-fabrication and microfluidic techniques to isolate individual bacteria cells in separate micro-fabricated wells, while exposing them to a constant flow of fresh media. As far as single cells are concerned, this is the equivalent of an extremely diluted stationary phase culture. In addition, this setting allows time-lapse microscopy to track large numbers of individual cells. It is a longitudinal cohort study with large sample sizes.

In this chapter, I present the basic experimental setup with which most of the experimental results in later chapters are obtained. This includes both the basic dynamics of cell growth and starvation without carbon sources; and the microfluidic technology used to achieve such conditions. I show that starvation in our microfluidic devices are idealised conditions unachievable by traditional batch cultures. The main the technical challenge that I overcome in order to achieve such conditions experimentally is documented in Appendix B, while the detailed experimental protocols are documented in Appendix A.

2.1 Microfluidic devices

Our lab have developed several microfluidic platforms based on a common general principle described in (Wang *et al.*, 2010a). The shared general principle is to trap bacteria cells in μm scale dead-end channels that can only hold one cell each, with a main channel maintaining a constant flow removing dead cells and feeding fresh medium. The dead-end channels are micro-fabricated to fit the exact dimensions of cells used and shallow enough to contain as few cells as possible to limit cell-cell interactions. This is demonstrated in Figure 2.1. This device, named the mother machine, has since spread widely in the systems biology and microbiology community.

We have several variants of the mother machine, designed for different experimental objectives and physiological conditions of *E.coli*. I mainly used two versions of this systems to achieve long-term, carbon-limited *E.coli* culture, monitored by time-lapse microscopy. They share the same design principle in Figure 2.1, but the arrays of dead-end channels are aligned either vertically or horizontally to the imaging plane. The left panels of Figure 2.2 demonstrate the operation of the two devices, named a) “trap array” and b) “mother machine”. The devices loaded with cells can be monitored by time-lapse microscopy. The right panels of Figure 2.2 are fluorescent images of cells in our devices. While "trap array" allows us to measure the mortality rate of $10^4 - 10^5$ single cells over time during starvation, "mother machine" allows us to measure gene expression and cell morphology more accurately, of a smaller population, around 10^3.

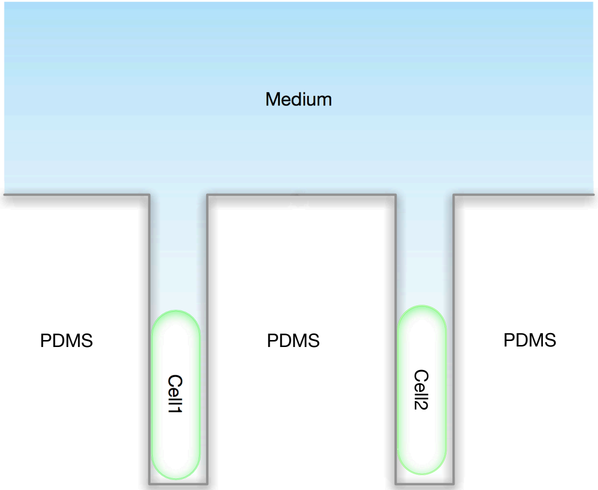

Figure 2.1: General principle of our microfluidic devices. Bacteria cells are trapped in μm scale dead-end channels as wide as single cells, with a main channel maintaining a constant flow removing dead cells and feeding fresh medium. The dead-end channels are micro-fabricated to fit the exact dimensions of cells used and shallow enough to contain as few cells as possible.

In Chapter 3, I mainly present the detailed single-cell life-history traits obtained in "mother machines". They are presented as both single-cell level description of the aging process, and a validation for the larger sample size studies obtained with "trap arrays". In Chapter 4, "trap arrays" are used to obtain lifespan statistics of large populations of different mutant strains. Both devices are operated with essentially the same experimental protocol. This protocol allows me to restrict carbon/energy supplies of cells while isolating them from each other to a degree unachievable by batch cultures. But in order to understand the purpose and limits of this protocol, we first have to examine carbon starvation in batch cultures, and compared them to microfluidic experiments.

2.2 Lower limits of carbon starvation in batch cultures

The obvious way to reduce cellular cross-feeding and social interactions through the media is to obtain extremely diluted cultures. In an ideal thought experiment, the most diluted batch culture would contain only one cell per container. Each cell would be isolated in their own test tubes and we would have enough test tubes to have a large

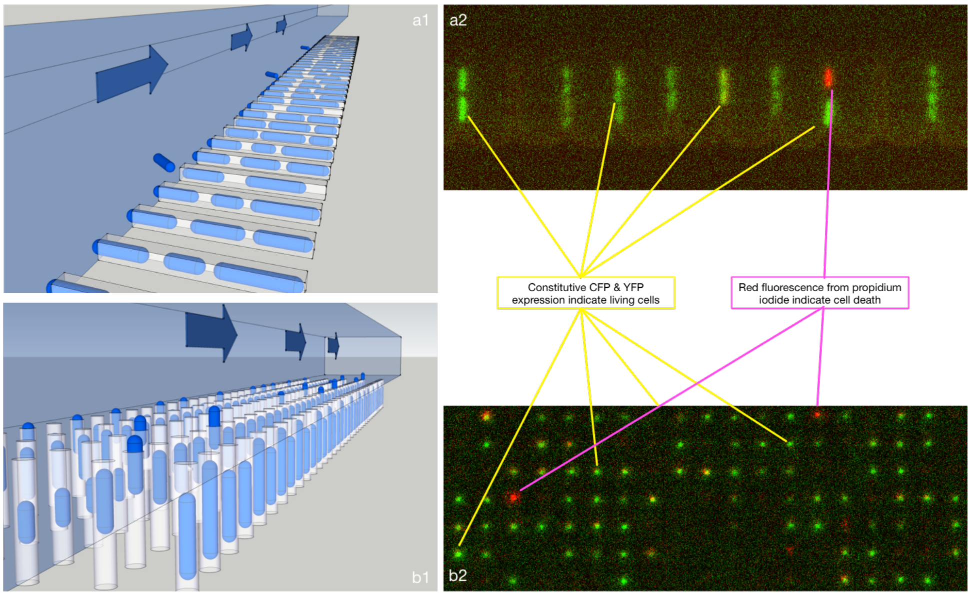

FIGURE 2.2: 3D demonstrations and fluorescent images of our microfluidic devices. (A1,B1) are the 3D cartoons of two variants of the mother machine. Cells are trapped inside dead end channels, and are growing as 1D colonies. Extra cells are pushed outside the dead end channels and removed by the flow in the main channel. The arrays of dead end channels are aligned either horizontally (A1,2) or vertically (B1,2) to the imaging plain. (A2,B2) are actual fluorescent images of these devices containing cells of E.coli constitutively expressing a variant of mVenus (shown in green). They are also exposed to propidium iodide in the medium, whose fluorescence (shown in red) is an indicator of cell death.

enough population to draw conclusions. The micro-wells in our microfluidic experiments can be thought of as extremely small test tubes. Although in reality, without the microfluidic devices, this extreme dilution is impractical, we could try to approach the ideally diluted experiment by simply washing and diluting harvested cells. I decided to test the limits of this approach in batch culture, as a way to test the potential benefits of the microfluidic systems and to calibrate them.

In order to carry out batch culture starvation experiments with low cell densities, instead of growing cells to stationary phase density of at least 10^9 cells per ml, I harvested, washed and diluted exponential phase cells into buffered minimum media (M9) without any carbon sources. These diluted cultures are then incubated and kept in physiological temperatures (37 °C) for up to two weeks while monitored by plating. The estimated cell densities are shown in Figure 2.3.

Despite the fact that the M9 buffer has no added carbon source, and the rigorous washing the initial populations of the cells went through, the low density starvation culture could still support a remarkable stable cell density of around 10^6 cells per ml. Different initial

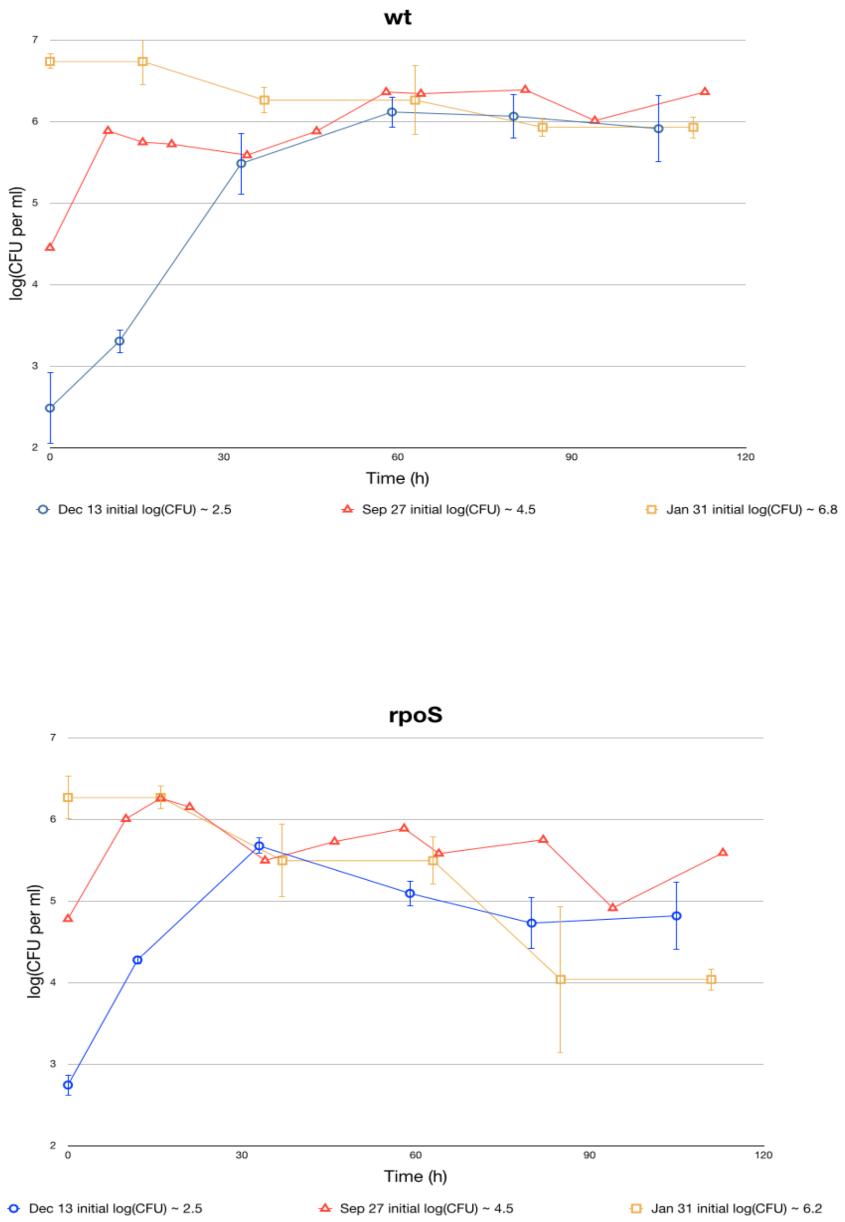

FIGURE 2.3: Carbon starvation of wildtype and ΔrpoS cells in diluted batch cultures. Cells are cultured in M9-glucose minimal media overnight from glycerol stock and then subcultured by at least 1000 fold dilution into fresh M9-glucose minimal media. Exponential phase cells are harvested at OD600˜0.2. Harvested cells are washed with M9 buffer without carbon source 3 times and then diluted into 100ml of the same buffer media to initial densities of $5{\times}10^2$, $5{\times}10^4$, $5{\times}10^6$ cells per ml. These starvation cultures are kept in 250ml Erlenmeyer flashs at 37 °C in the dark for up to two weeks. The cultures are monitored at regular interval by serial dilutions and platings, and estimated cell densities are plotted for 3 different starvation experiments. Cells used are MG1655 wildtype and ΔrpoS mutants, which has reduced stress resistance in stationary phase.

densities seem to have no effects on the steady state cell densities. The initially most diluted cultures grew around 1000 fold, from below 10^3 cells per ml to the same 10^6 cells per ml the other cultures have reached. This 'carrying capacity' suggests that the starvation media must have some kind of carbon sources at low concentrations. Further investigations indicate that the low concentration carbon sources are leached plasticisers from common plasticware used to prepare the media. This issue and its solution is detailed in Appendix B.

Compared to common stationary phase batch cultures that usually reach 10^{10} cells per ml, the dynamics of cell populations under starvation in these low density cultures are very similar (Finkel, 2006). After reaching the 'carrying capacity' of the media, the cell densities of the wildtype strain remain steady over the course of the experiments, while the cell densities of the ΔrpoS strain experience a light drop. These decreases are assumed due to increased cell deaths and inefficient use of resources (Vulic & Kolter, 2001). More interestingly, mutants with heterogeneous colony morphology start to appear in the culture after a week. This suggests that even in this relatively low density of 10^6 cell per ml, the GASP (Growth-At-Stationary-Phase) phenotypes previously described are still selected (Finkel, 2006; Vulic & Kolter, 2001).

Both the existence of low concentrations of carbon sources and the fact that GASP mutants are selected under these conditions, present significant challenges to the study of evolutionary trade-offs in both batch and microfluidic experiments. In order to isolate the effects of general stress response on maintenance from its effects on growth, starvation experiments must be done in conditions where growth is not possible even for ΔrpoS strains. The impacts of this constraint on microfluidic experiments and the resolution of this problem are discussed next.

2.3 Media infusion rate in microfluidic experiments controls the physiological regimes of trapped cells

If one compares the cells trapped in micro-wells in our microfluidic devices to the cells in batch culture starvation experiments in Figure 2.3, media refreshment rates controlled by the flow rates in the main channel determined essentially the initial cell dilutions. If the flow rate was infinitely fast, trapped cells would be essentially exposed to infinite volumes of fresh media. This situation corresponds to the condition at T_0 if the initial cell density approached zero in Figure 2.3. As we can see, if the media has a low concentration of carbon sources, as is the case in Figure 2.3, cells have non-zero growth rates. Since in microfluidic devices, extra cells are removed from the population, the population never reaches steady states and the trapped cells keep growing at a constant growth rate. This is easily experimentally observable using the same contaminated media and mother machine, as in Figure B.1. In fact, it can be shown that this growth rate is the maximal growth rate as defined in the exponential phase batch culture using the contaminant as carbon sources.

If we gradually decrease the flow rate from the very fast regime described in the previous paragraph, depending on the the concentration of the carbon source, at some point the cells near the inlets of the microfluidic devices will consume enough carbon sources to decrease the amount available for cells farther from the inlet. At this point, one can observe heterogeneity of growth rates between cells inside the device. When flow rate has decreased enough, the cells closes to the outlet will have stopped growing. The more one reduces the flow rate, the more cells enter growth arrests. Eventually, even the cells near the inlet stop growing. We call this range of flow rates under which heterogeneity of growth rate is observed transition flow rates. The transition flow rates is proportional to the concentrations of carbon source in the infusion media. However, there is a lower limit for flow rate reduction as well, since a fast enough flow is still needed to remove cell debris and counteract evaporation. The lower limit is usually $2\mu l$ per hour for our devices.

The main technical challenge that I have overcome in my thesis work, is essentially that I have reduced the level of carbon supplying contaminant in the media to such a low level that the lower bound of transition flow rate is significantly larger than $2\mu l$ per hour. To put it another way, a lot of technical work have went into cleaning up the media so that growth arrests can be induced in the mother machine while maintaining a flow rate larger than $2\mu l$ per hour. I have developed the experimental protocol that eliminated the majority of leached carbon sources from lab plasticware. In fact, the same problem has been repeatedly encountered by researchers who study bacteria in nutrient-poor conditions over the decades. In this protocol, the media is usually prepared with grade-I milliQ water, and only comes into contact with glass, stainless steel and PTFE surfaces. M9 media prepared with this protocol called carbon-free M9 hereafter. In Appendix B, I have documented in much more detail the sources of carbon-supplying contaminants, the method we identify them and the non-leaching equipments that replaced the leaching ones.

2.4 Typical experimental procedures of carbon starvation in microfluidic devices

Here I offer a minimalist description of our experimental protocols. For more details on media preparation, see Appendix B. For mother machine experimental setups, see Appendix A and Section A.3 in particular. For the image analysis methods after the experiments, see Section A.4.

Carbon-free M9 media is usually prepared before the experiments, filter-sterilised with PTFE filters and stored in glass bottles with PTFE caps. *E.coli* strains used for starvation experiments are usually are cultured in M9 minimal media overnight from glycerol stock and then subcultured by at least 1000 fold dilution into fresh M9 minimal media. Exponential phase cells are harvested at OD600˜0.2. Cells harvested from 50ml exponential phase culture are washed with carbon-free M9 3 times and then concentrated into 50 μl. These concentrated cells are injected into the microfluidic devices and loaded into the micro-wells by centrifugation. Microfluidic devices are then washed with carbon-free M9 and micro-bubbles to eliminate all extra cells not inside the micro-wells. Then the devices are connected to microfluidic infusion pumps and placed onto the stage of our temperature-controlled automated fluorescent microscopes. The experimental flow rate is determined beforehand to situate well below the transition flow rates that induce growth arrests, and is usually 10 μl per hour. The infusion bump maintain this constant infusion flow rate by injecting carbon-free M9 throughout the experiment, and cells are tracked by the automated fluorescent microscopes. Phase contrast and fluorescent images from the time-lapse microscopy are then analysed by a suite of in-house semi-automatic software to extract the life-history traits and gene expression profiles of the cells under starvation.

Chapter 3

Life-history traits of *E.coli* under carbon starvation

In this chapter, I will describe some basic observations of *E.coli* life-history under carbon starvation, in particular their lifespan distributions. Experiments in this chapter are done using the horizontal version of the mother machine, with the carbon starvation microfluidic platform and protocol established in Chapter 2, Appendix A and Appendix B. This enables us to monitor in detail the physiological changes of starving *E.coli* cells.

I established a quantitative method to monitor changes in cell membrane potentials, and use it as a proxy to infer cell viability. This method allows me to determine both the timing of cell death and the timescale of the dying process. Despite the clonal genetic backgrounds and homogenous environments, *E.coli* cells exhibit significant variations in both timing and timescale of deaths. Knockout mutation to the general response pathways significantly shortens the starvation lifespan. I then discuss the possible origins of lifespan variations and show that elongating cells in starvation tends to have shorter lifespans.

3.1 Time-series of a DNA-staining fluorescent dye as indicators of cell viability and membrane integrity

Propidium iodide (PI) staining has been established as a reliable way to assess the viability of *E.coli* cells (Ericsson *et al.*, 2000). For living cells, PI can not permeate through intact cell membranes. Upon cell death, PI stains DNA with a strong far-red fluorescence. To assess the impacts of chronic exposure to cell physiology, I incubated my strains in batch media with both PI and polyethylene glycol (PEG, used for surface treatment in microfluidic experiments), and used plating to measure population density. As in Fig 3.1, the addition of these additives do not change population dynamics in significant ways.

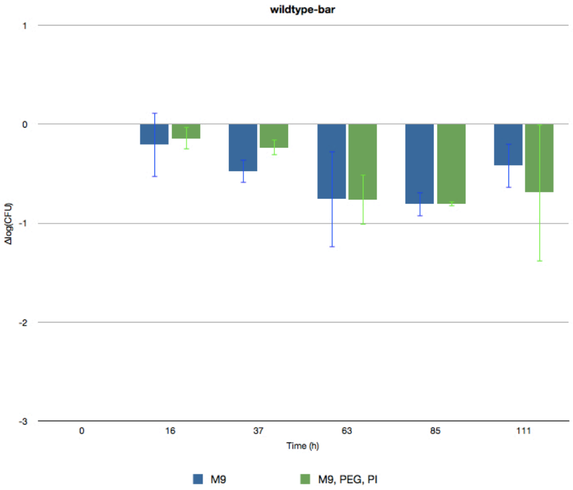

Figure 3.1: The survival of *E. coli* in bulk M9 medium with (green bars) and without PI and PEG (blue bars). *E. coli* were prepared in the same way as chip-based experiments and cultivated in Falcon® tubes at 37 °C without shaking. Samples were plated on LB agar plates to score viability by counting colony-forming units (CFUs) at the indicated time. Results shown are the average of three independent measurements.

We added PI into our infusion media in order to monitor the viability of cells in carbon starvation in real time. Fluorescent timelapse microscopy and semi-automatic image analysis allow us to track each cell individually and measure their viability via PI fluorescence, as described in Appendix A. Figure 3.2 presents the PI fluorescence time-series of all cells in a single experiment, in which a mix of two genotypes are subjected to carbon starvation for 120 hours. Each row is the time-series of one cell. See Appendix B for the equipments that allowed for such microfluidic carbon starvation possible.

These PI time-series have several features. First of all, for each time-series there are a low and a high phase, with the transition between the two defining the transition from life to death. We thus use the mid-point of the transition to define the moment of death. Cells in Fig 3.2 are sorted according to this measure of lifespan. Secondly, cells have different maximum PI intensities, presumably associated with the amount and state of DNA in each cell. It is not the absolute level of PI staining, but the relative to this maximum that defines the transition from life to death. Last but not the least, cells transit from being alive to being dead not only at different moments, but take different durations to do so.

PI fluorescence increases reflect losses of membrane integrity or at least membrane potentials. Based on the PI fluorescence data, I estimated for each cell the loss of membrane potential using a simple model of PI dynamics comprising of slow entry into cells and fast binding to DNA. Because most of the fluorescent signal is emitted by PI molecular bound to DNA, maximal fluorescent intensity gives an nice estimate of total DNA in the cell, while the difference between current and maximal fluorescence estimate the amount of DNA unbound to PI.

$$\frac{Fluo(t)}{dt} = k_1 e^{-\frac{E(t)}{k_B T}}[Fluo_{max} - Fluo(t)] - k_2 Fluo(t)$$

where $Fluo(t)$ is the observed PI fluorescence time-series, k_1, k_2 are reaction rate constants associated with PI entry and exit, $E(t)$ is the estimated membrane potential, k_B, T are the Boltzmann constant and temperature respectively. $-k_2$ can be estimated by the lower bound of $\frac{1}{Fluo(t)}\frac{dFluo(t)}{dt}$ for all t for all living cells because the minimal rate of PI entry is 0. Thus we can estimate the membrane potential by the rate of fluorescence increase:

$$\frac{E(t)}{k_B T} = -ln(k_1) + ln[\frac{Fluo(t)}{Fluo_{max} - Fluo(t)}] + ln[\frac{1}{Fluo(t)}\frac{dFluo(t)}{dt} + k_2]$$

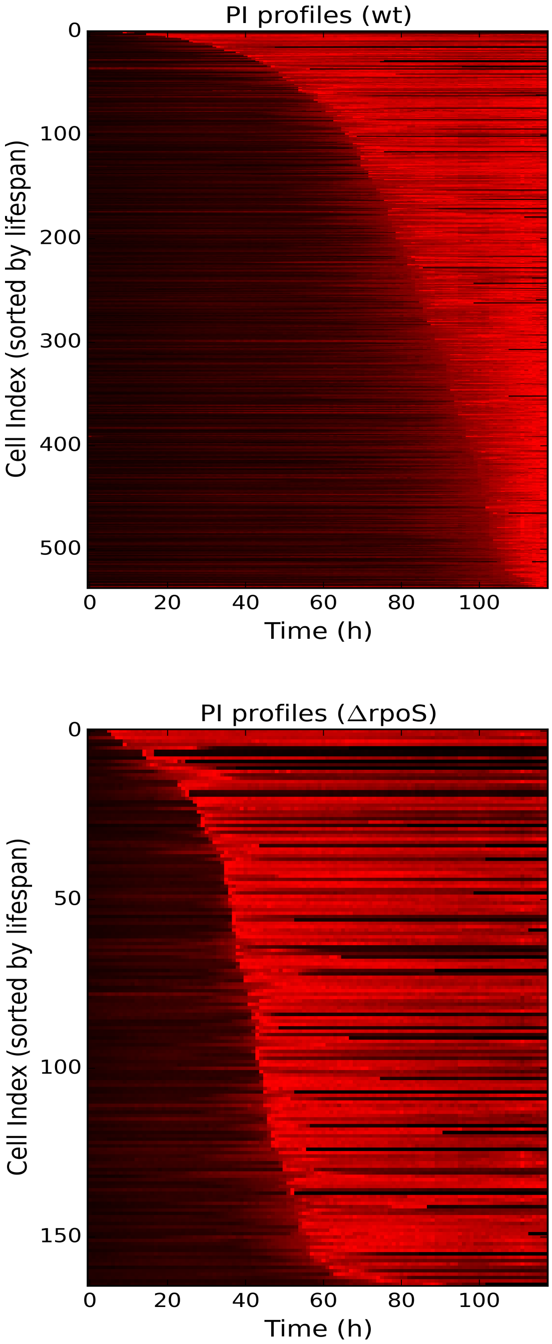

FIGURE 3.2: Propidium iodide (PI) fluorescent signals of wildtype and ΔrpoS cells in a single experiment. Each row correspond to the PI time-series of one cell. Cells are sorted according to their estimated time of death. As discussed in Section 3.1 and shown in Figure 3.3, PI only enter a cell, binds to DNA and fluorescence when cell membrane integrity is comprised. The transition from dark regions to bright regions indicate transition from living to death. Boundaries of such transitions describe the survival function of the populations. The dark parts at the ends of some time-series are the results of dead cells washed away from the chip.

Figure 3.3 demonstrates the estimated dynamics of membrane potential of a few cells given their PI fluorescence time-series. Numerical differentiation are done with Total Variation Regularised Numerical Differentiation (TVDiff) (Chartrand, 2011). Moments of death are marked by sudden drops of membrane potentials, while different cell deaths share similar magnitude of membrane potential changes. This suggest that normalising the PI fluorescence with the maximum rather than using the absolute fluorescence is the correct way of determining cell death.

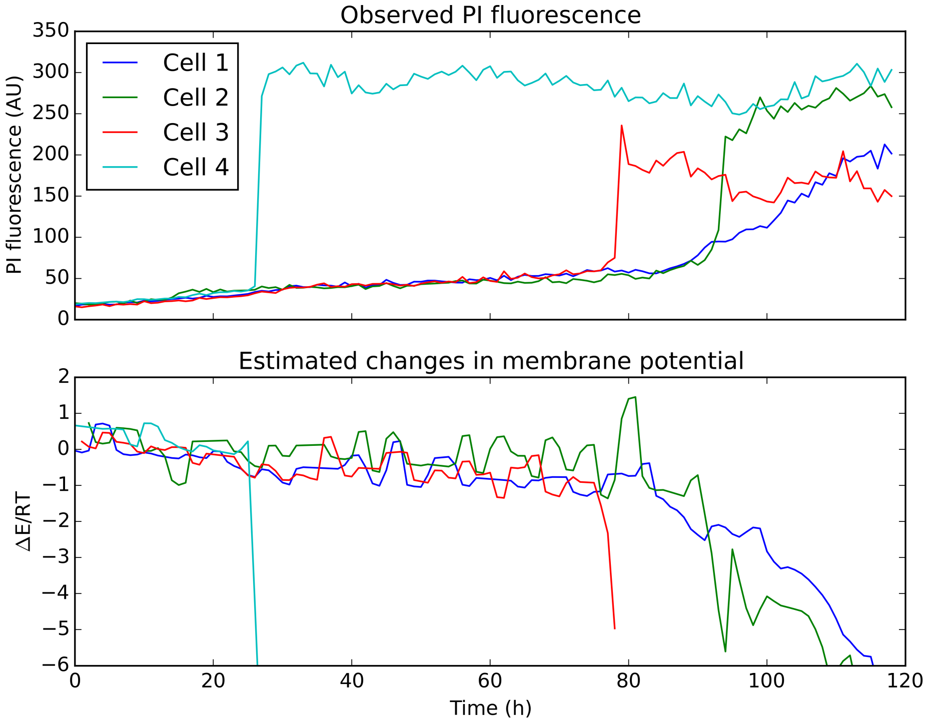

Figure 3.3: Estimating membrane potential from PI fluorescence. The upper panel shows the observed PI fluorescence of a few cells as examples. The lower panel shows the estimated membrane potential for these cells. The zip-zags of estimated potential around 0 are due to experimental and image analysis noise when fluorescence is low. The numerical differentiation (TVDiff) algorithm we used made it possible to extract genuine shifts among the noise in PI signal, but nonetheless allows some noise to pass through (Chartrand, 2011).

3.2 PI dynamics reveal both qualitative and quantitative individual differences in cell mortality

Despite sharing the same genetic background, being exposed to the same starvation media, and isolated from each other, cells die at different times. Figure 3.4 shows the survival curves of wildtype and ΔrpoS cells, while Figure 3.5 shows the corresponding lifespan distributions and hazard functions. Judging from the hazard functions of the two strains (log scale, right panel of Figure 3.5), the probability of death increases exponentially with starvation time. We are reminded of what is commonly known as the Gompertz law in demography and aging research, where the probability of death for humans increases exponentially with age. I used the Gompertz model to fit these distributions and obtained maximum likelihood estimates (MLE) of the parameters. Red and blue dashed lines in Figure 3.5 are the estimates for the probability density functions and hazard functions of the best fitted models to ΔrpoS and wildtype data respectively.

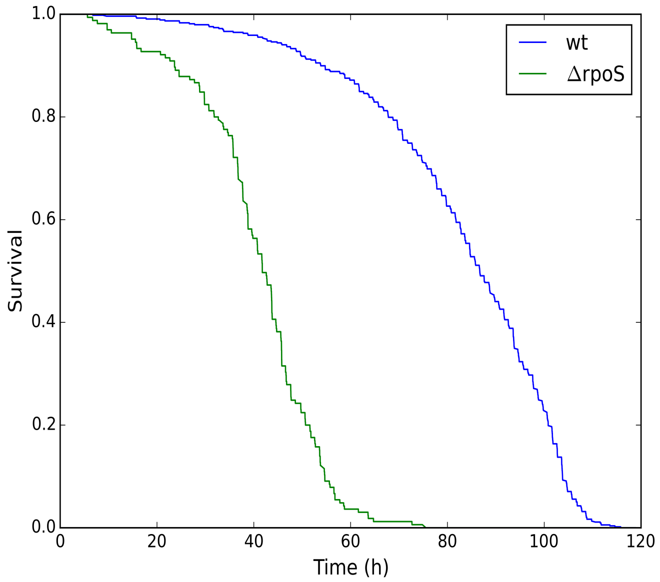

Figure 3.4: Survival curves of wildtype and ΔrpoS strains in the same carbon starvation experiment. Data are the same as those in Figure 3.2. Exact time of death is determined with the thresholding method, by taking the first time that PI signal are above half of the maximal strength reached in each time-series.

Not only do cells die at different times, but also in different manners. In order to quantify the different timescales that PI fluorescence transit from low to high, I used

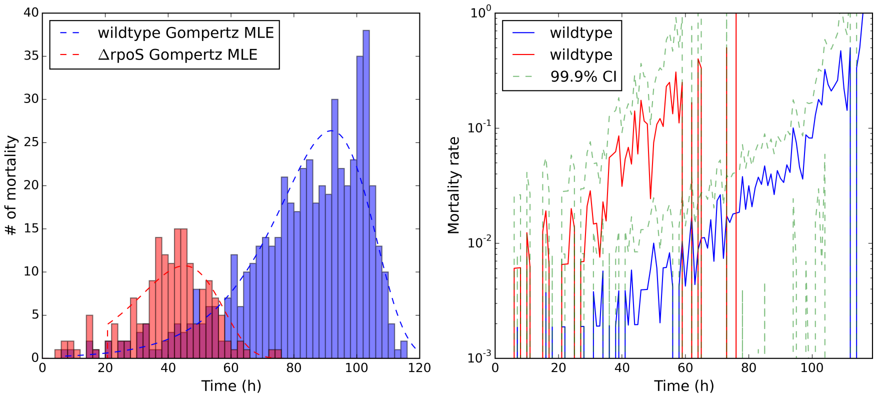

Figure 3.5: Estimated lifespan probability density functions and hazard functions of wildtype and ΔrpoS cells. Data are the same as those in Figure 3.4 and Figure 3.2. Number of deaths are counted for each 2-hour bins, and then are used to estimate the probability density and hazard functions. The hazard functions are estimated by $d_t / \sum_0^t d_t$, where d_t is the number of deaths at time t. Red and blue dashed lines in the left panel correspond to fitted Gompertz models using maximum likelihood method. Green dashes in the right panel describe the confidence interval according to Poisson sampling.

Continuous Wavelet Transformation (CWT) to locate the timing and timescale of PI fluorescence increases in the time-series. CWT essentially decompose a time-series into linear combinations of a family of orthogonal wavelets generated by translations and dilations of the same mother wavelet. Thus the wavelet with the biggest amplitude after CWT represents the most important feature of the time-series. This process is demonstrated in Figure 3.6. PI fluorescence data are first log transformed, and then transformed by CWT with a Ricker mother wavelet into a matrix of coefficients in the time-frequency domain, visualised in the lower panels of Figure 3.6. The largest coefficient (marked by white dots) correspond to the biggest increase in PI fluorescence. The scales (y-coordinates) and locations (x-coordinates) of the coefficients indicate the timing and timescale of the PI increases. What we are ultimately interested in from the CWT analysis are the timescales of PI increases, i.e. the y-coordinates of the largest CWT coefficients. The larger it is, the lower the PI fluorescence increases. Hereafter, this is defined as the timescale of morality events.

In Figure 3.7 I plotted the timescales of all mortality events in one experiment versus their lifespans. We can visually cluster two types of mortality events. One cluster is close to the x-axis, including those mortality events that have small timescales. They

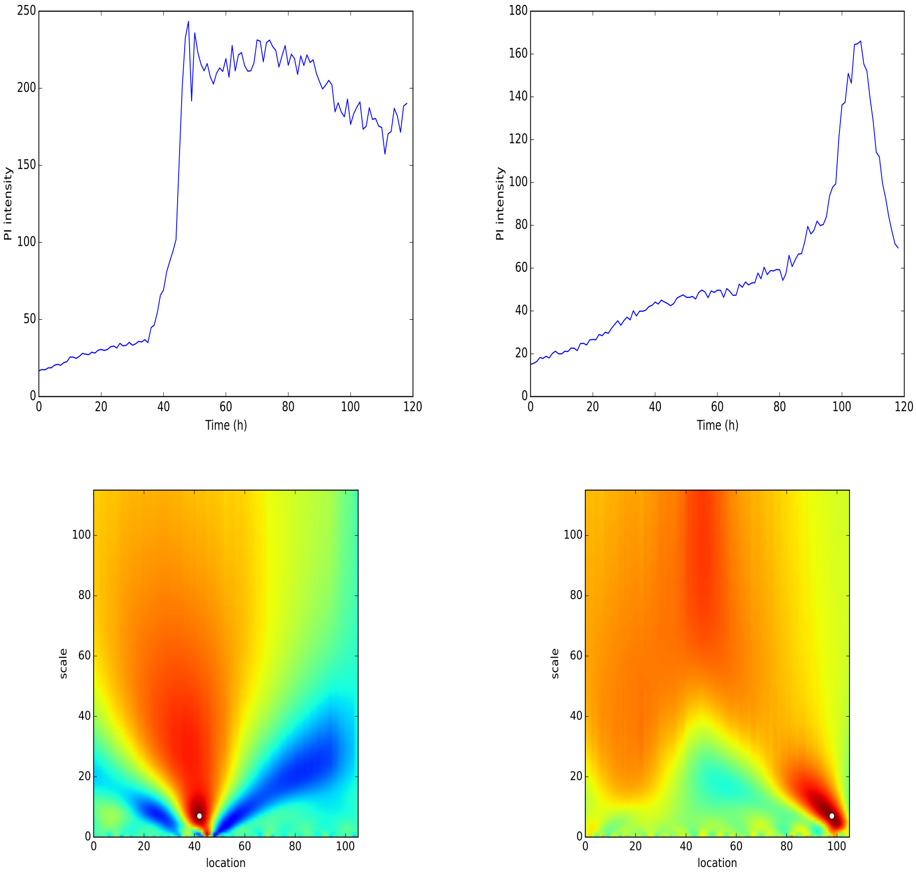

FIGURE 3.6: Examples of PI time-series (Upper row) and their continuous wavelet transformation (CWT) coefficients (Lower row). The upper row shows the PI time-series of 2 example cells, one with a sharp PI increase and one with a graduate PI increase. The CWT is aimed to extract both the time 'location' and 'scale' of those PI increases. Standard CWT algorithms are used with Ricker mother wavelet, because the signal we are interested in is a simple increase. The lower row are heat maps of the CWT coefficients. Red and blue correspond to positive and negative coefficients respectively, i.e. peaks and valleys. The most intense peak is labeled with the white circle, whose coordinates represent the time and timescale of the most prominent PI increase.

correspond to quick cell deaths and sharp PI increases where the cell membranes are permeated quickly. These cells are hereby named 'bursters'. The other cluster is around the diagonal line, which suggest a near equal relationship between lifespan and the timescale of mortality. Because we used log-transformed PI time-series in our CWT analysis, this suggest a slow but near exponential increase of PI fluorescence throughout the lifespan of the cell. Instead of abrupt membrane potential loss seen with the 'bursters', these cells lose small fractions of membrane potential throughout their lifespan, even before they lose most of their membrane potential. This pre-mortality losses of membrane potential in fact can be seen in Figure 3.2. These cells I call 'leakers'.

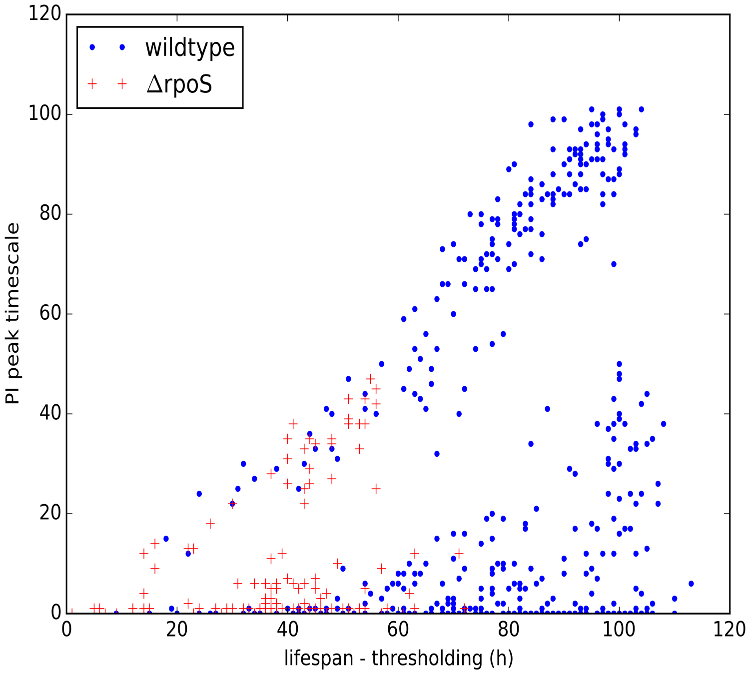

FIGURE 3.7: Two types of mortality events are differentiated by the timescale of PI peaks. These are the same cells as those in Figure 3.2. The x-axis is the lifespan of cells, determined as the time that PI signal has crossed half of the peak intensity, as in Figure 3.4. The y-axis is the timescale of cell PI peaks, determined by the continuous wavelet transformations analysis shown in Figure 3.6. Colors and shapes of the markers indicate the genotypes of the cells. The two type of mortality events are those closer to the diagonal line, indicating slow deaths, i.e. 'leakers'; and those closer to the x-axis, indicating quick deaths, i.e. 'bursters'.

3.3 Correlates and possible sources of lifespan variations

As discussed in Chapter 1, one main assumption of the evolutionary theory of aging is the growth-maintenance trade-off. Such a trade-off suggests that trying to grow despite starvation conditions is maladaptive because they shortened lifespans when growth is impossible. Yet this is hardly observed in bacteria in high-density batch culture experiments due to cross-feeding and bacterial cannibalism in those conditions (Vulic & Kolter, 2001). A original motivation for our microfluidic experiments is to test the growth-maintenance trade-off in true starvation conditions. We expect to do so by examining the longitudinal life-history traits obtained from time-lapse microscopy that tracks the same cells over their entire lifespans.

Despite the true carbon starvation condition achieved by our microfluidic setup (See

Appendix B) and the inabilities to grow over the entire course of experiments, cells do attempt to grow by elongation at the beginning of the experiments. I suspect that they do so with some reserve energy carried over from growing phases. Figure 3.8 shows the mean and variations of cellular elongation rates of the cell populations over the course of the experiments. Elongation rates are calculated for each cell based on cell length determined by image segmentation according to the definition of instantaneous growth rate: $\mu(t) = \frac{1}{ln2} \cdot \frac{dL(t)}{L(t)dt}$. Computationally, with the help of the TVDiff algorithm, we numerically differentiated the log transform of cell length to estimate $\mu(t)$ for each cell.

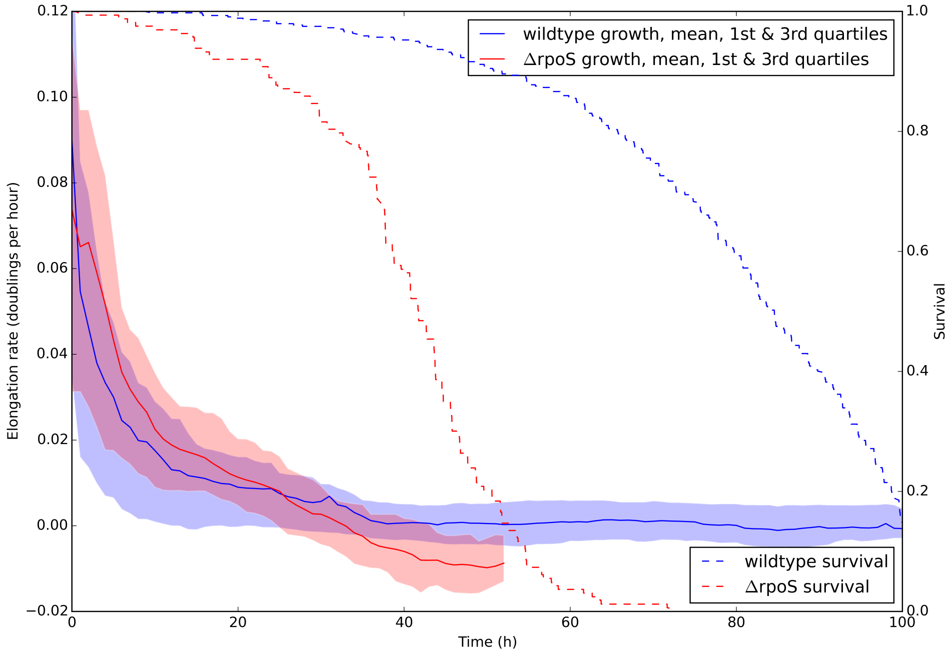

FIGURE 3.8: Instantaneous cell elongation rate and its variations in starvation experiments conducted in mother machines. The solid lines and shades plotted on the left axis are the mean and quartiles of elongation rates of wildtype and ΔrpoS cells, while the dotted line plotted on the right axis are the survival curves as a reference. Elongation rates are plotted for all time points with more than 15 surviving cells from the same genotype. Individual cell elongation rates are measured in units of doublings per hour, and are calculated for each cell based on cell length determined by image segmentation according to the definition of instantaneous growth rate: $\mu(t) = \frac{dL(t)}{ln2L(t)dt}$.

Elongation rates of cells quickly fall to near zero during the transition from growth to starvation. Consistent with the phenotypes described in the literature, ΔrpoS cells elongate more than wildtype earlier in starvation; but later elongate at negative rates, which is to say that on average they shrink.

In Figure 3.9, I plotted lifespans of the cells in starvation against their elongation rates. Elongation rates are averaged from instantaneous rates over the entire lineage consisting of each terminal cell and all her ancestral cells. Here the inverse of elongation rates measured in doubling time ($T_{dbl} = ln2/\mu(t)$) is plotted in the x-axis in log scale. Cells closer to the right and left extreme of the figure are barely growing or barely shrinking, while those to the middle are either growing or shrinking quickly, depending on the sign of the average elongation rate. Lifespan is negatively correlated with the absolute value of elongation rate. The fast growing and fast shrinking cells tend to die early, while those cells who live long tend to not change much in cell size throughout the starvation experiment.

This observed negative correlation between growth and lifespan is consistent with the growth-maintenance trade-off hypothesis in the evolutionary theory of aging. Under this theory, the optimal strategy for cells in true starvation is to invest in maintenance to counteract molecular damages and maintain homeostasis. Interpreting our results with this theory suggests that those fast elongating cells in Figure 3.9 are misallocating resources when growth is impossible, thus reducing their lifespan; while fast shrinking cells suffer from molecular damages without counteracting them.

If this interpretation is true, then individual differences in cell elongation rate and lifespan should be attributed to cell physiology, not external environment. The main source of environmental variations in our experiments is the depth that cells sit in the dead-end channels. A cell could be either protected or blocked by another cell sitting in between her and the main channel.

To see whether individual differences in elongation and lifespan could come from environmental differences, I redraw the data in Figure 3.9 while classifying cells according to their depths. Wildtype and ΔrpoS cells populations are separated into Figure 3.10 and Figure 3.11. Although I did not applied any statistical tests, these figures clearly indicate that depth is not a major determinant of either elongation rate or lifespan. Although we can not completely rule out some mysterious unaccounted for environmental sources, it is far more likely that this variation in life-history trade-off comes from physiological and epigenetic differences internal to the cells.

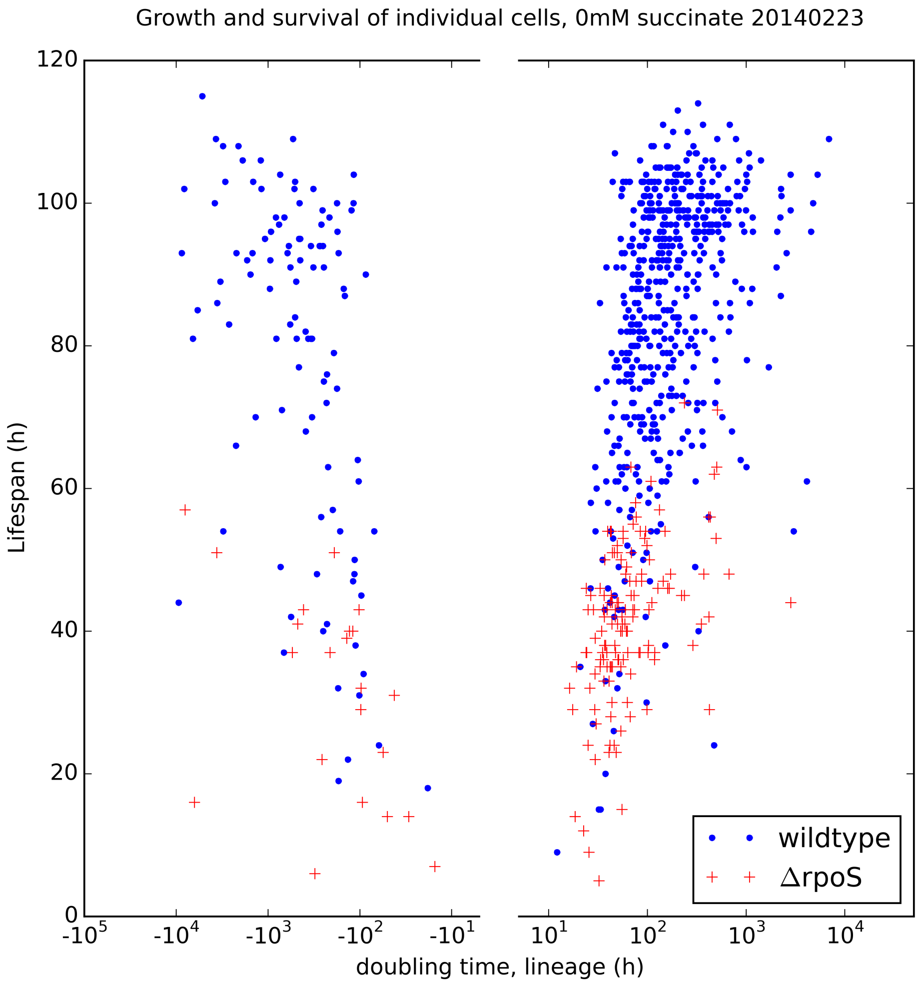

FIGURE 3.9: Correlations between cell elongation and starvation lifespan. Plotted are individual cell elongation rates and starvation lifespans as those in Figure 3.8. On the y-axis are starvation lifespans in hours as measured in Figure 3.4. On the x-axis in log scale are the inverse of average elongation rates, i.e. the doubling time ($T_{dbl} = ln2/\mu(t)$) in hours as well. For each cell, average elongation rate is calculated from instantaneous rates over the entire lineage, consisted of each terminal cell and all her ancestral cells. Lifespans are calculated the same way as those in Figure 3.4. Although cell elongation can be measured this way in doubling time, most cell have never actually doubled in size (see Figure 3.8). Negative doubling time indicate shrinking cells. Very large doubling time or very small doubling time are those cells that have barely changed in size. Colors and shapes of the markers indicate the genotypes of the cells.

3.4 Discussions and conclusions

In contrast to high density batch culture experiments (Vulic & Kolter, 2001), our carbon starvation experiments showed that the general stress response is beneficial in carbon starvation. This difference can be attributed to the high nutrient environments typically

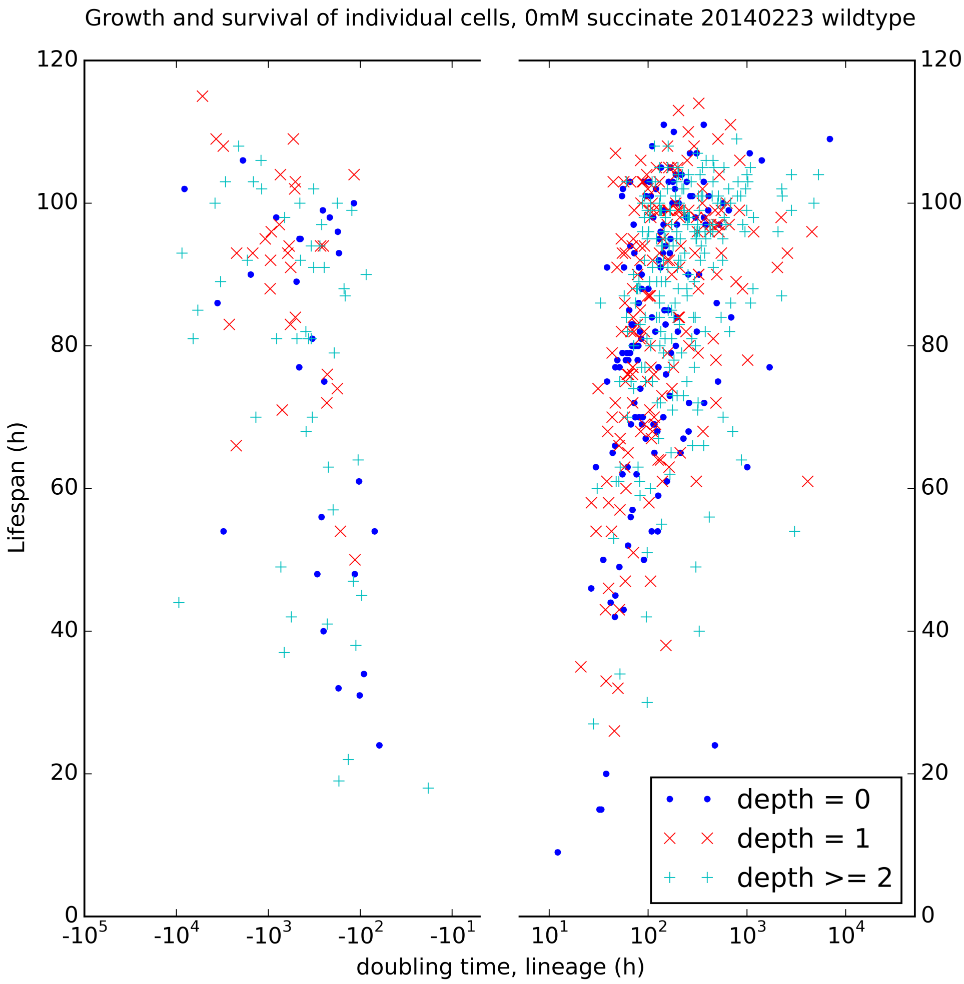

Figure 3.10: Wildtype cell elongation rates and lifespans, grouped by depth in dead-end channels. The data and axises are the same as the blue dots in Figure 3.9, while different markers represent cells at different depths in their respective dead-end channels in the mother machine, with depth = 0 being the cell closest to the main channel. Notwithstanding statistical testing, cell depth is not major source of variation in lifespans and elongation rates.

seen in batch cultures. These environments tilt the fitness trade-off between growth and maintenance towards the growth side, thus always select for strains with higher growth potential, even at the expense of maintenance. My work described in Chapter 2 and Appendix B in reducing carbon-supplying contaminants in the media has made it possible to experimentally examine the fitness benefits of the general stress response. When the nutrient capacity in the environment measured in lowest possible doubling time is brought closer to the timescale of starvation lifespan, the advantage of maintenance investment becomes observable. Our microfluidic devices are essentially mimicking the

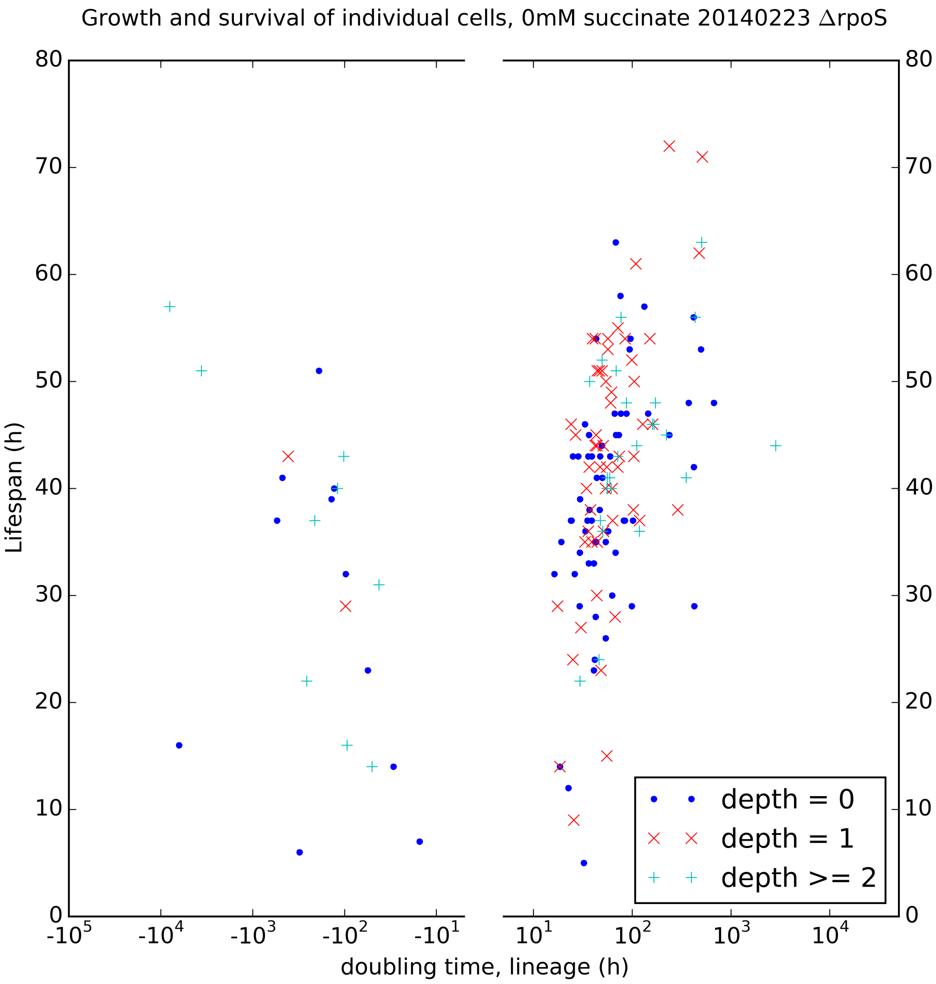

Figure 3.11: ΔrpoS cell elongation rates and lifespans, grouped by depth in dead-end channels. The data and axises are the same as the blue dots in Figure 3.9, while different markers represent cells at different depths in their respective dead-end channels in the mother machine, with depth = 0 being the cell closest to the main channel. Notwithstanding statistical testing, cell depth is not major source of variation in lifespans and elongation rates.

highly diluted environments where bacterial dispersal and starvation usually happen.

Quantitative individual differences in lifespan and qualitative differences in mortality are common in animal aging research, even in environmentally controlled cohorts of clonal populations. It has been speculated that unobserved heterogeneity in initial frailty, however defined, might determine such differences. Our results show that even environmentally-controlled clonal populations of simple organisms such as *E.coli*, still exhibit variations in lifespan, and even reproduce the similar statistical distribution seen in higher organisms. The origins of these distributions, may lie not in the details of molecular biology of specific organisms, nor in the particularities experimental environments, but in the statistics of system failures. Yet, individual differences in lifespan are biologically non-trivial, in the sense that at least they are partially correlated with other important life-history traits such as growth, and might have adaptive consequences.

In this chapter, I have established a reliable method of monitoring cell viability in carbon starvation in mother machines, and observed a few features of starving *E.coli* populations. Detailed measurements of *E.coli* lifespan reproduced lifespan distributions typical in the demographics of human and animals. I also confirmed the validity of a fundamental assumption of the evolutionary theory of aging, at least in the context of *E.coli* starvation. These results established *E.coli* carbon starvation as a model of chronological aging, and showed its value in understanding quantitative and evolutionary features of aging in general.

Chapter 4

The general stress response decreases the rate of aging

In Chapter 3, I have shown that a rpoS null mutant without general stress response has significantly shorter lifespan in carbon starvation than the wildtype. It is well known in the literature that rpoS null mutants have increased capability to grow on poor carbon substrates, which in turn might affect lifespans in those conditions through increased growth. To understand the combined effects of less maintenance investment and increased growth on both lifespan and overall fitness, the qualitative correlations described in Chapter 3 is not enough. We need a quantitative model for the general stress response's effects on life-history traits. To this end, we designed a vertical version of the mother machine, as described in Chapter 2. In this chapter, I will first described carbon starvation experiments in vertical mother machines, and then introduce a non-parametric estimator of mortality rates and rate of aging. Using these tools, I will show that the general stress response pathway controlled by rpoS indeed decreases the rate of aging.

4.1 Measuring carbon starvation lifespans of large populations of *E.coli* with vertical mother machines

Vertical mother machines trap and isolate individual cells in similar fashions to that of the mother machine used in the previous chapter, except that a much larger population can be tracked in a single experiment. Practically, we usually can obtain populations from 5×10^3 to 2×10^4; While 10^5 is easily reachable without changing our experimental setup if such a population is required. Multiplexed microfluidic chips allow for several strains to be imaged independently in parallel experiments. Experimental protocols are very similar to those used before, including the carbon starvation media preparation, and the use of PI to track cell deaths.

One disadvantage of the vertical chip is that we can not morphologically segment cells by the phase-contrast images. To ensure that we can locate all cells that are loaded into the array of dead-end channels, we inject the chip with ethanol at the end of experiments. Figure 4.1 demonstrates that despite different causes of death, cells that died during starvation and killed by ethanol have similar maximal PI fluorescences. This result suggests to us that we can reliably use PI fluorescence not only to determine lifespan as described in chapter 3, but also to locate those cells that have survived till the end of the experiments but are killed by the ethanol injection.

To analyse the time-lapse movies generated from this type of experiments, we first spatially aligned all the time-lapse images, and then project the fluorescence images at all times to a single 2D image. PI fluorescent dots on this t-projection image are identified as regions of interest (ROIs), and time-series of total fluorescence within these ROIs are extracted. These PI fluorescence time-series are analysed in similar fashion as in Chapter 3, and lifespans of individual cells are determined in similar ways.

4.2 Non-parametric mortality estimators and rate of aging

We used several different non-parametric estimators to analyse the lifespan distributions we obtained from carbon starvation experiments. For one experiment of wildtype in vertical mother machines, these estimators are shown in Figure 4.2. Intuitively, the survival curve (figure 4.2, upper left) similar to those in Figure 3.4 are estimators of survival functions of lifespan distributions. These are known as the Kaplan-Meier (K-M) estimators:

$$\hat{S}_{km}(t) = \prod_{i:t_i<t} (1 - \frac{d_i}{N_i}) \tag{4.1}$$

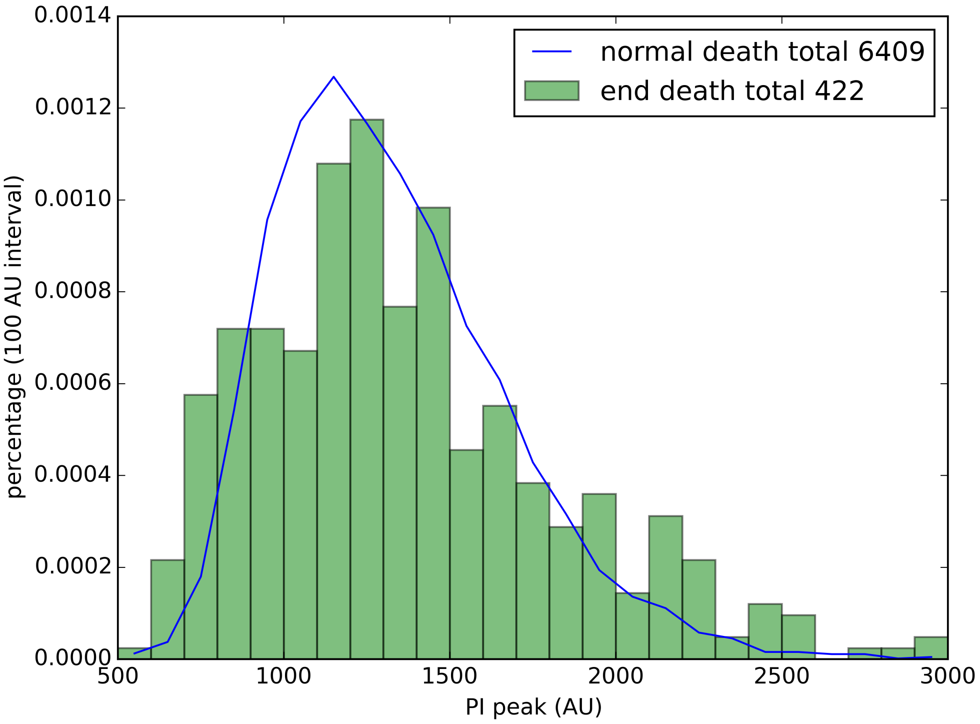

FIGURE 4.1: Histograms of maximal PI fluorescence in vertical mother machines. "Normal deaths" are mortality events whose peak fluorescence are reached during the starvation experiment, whereas "end deaths" are survivors of the starvation experiment, whose peak fluorescence are reached after the chip is injected with ethanol. The two types of death events give rises to similar PI fluorescence peaks, and could not be differentiated on the sole basis of maximal fluorescence.

where d_i are the numbers of deaths at time t_i, and N_i is the numbers of individuals at risk at time t_i. The Kaplan-Meier estimator has several advantages: it represents the overall experiment intuitively; the overall shape and trend are insensitive to irregular events within limited time intervals; and it does not require binning (i.e. uses exact time of death). However, the drawbacks are that it does not articulate events at specific time very well because of the relative insensitivity to in-continuities in d_i, and also do not reflect rate of aging very well, since the survival function is not linearly affected by the rate of aging. It affects the survival function at older ages much more than that at younger ages.

Because we have relatively large sample sizes in our experiments, we could estimate the probability density function (Figure 4.2, lower left) and hazard function (Figure 4.2, lower right) of starvation lifespan distributions, by binning the mortality events into fixed time intervals. In fact, the same binning procedure is done in Figure 3.5. If we suppose regular binning $T_b = b \times T_{int}$, where $b = 1, 2, 3 \ldots$, and T_{int} should be larger than the experimental time resolutions. We have the estimators for the probability density function $\hat{f}(T_b)$ and hazard function $\hat{h}(T_b)$:

$$\hat{f}(T_b) = \frac{\sum_{j:T_b<t_j<T_{b+1}} d_j}{T_{int} \sum_j d_j} \tag{4.2}$$

$$\hat{h}(T_b) = \frac{\sum_{j:T_b<t_j<T_{b+1}} d_j}{T_{int} \langle N_j \rangle_{j:T_{b-1}<t_j<T_b}} \tag{4.3}$$

We can easily identify mortality events and specific times of interests in the starvation process from these figures because as long as appropriate T_{int} is chosen, they are both sensitive to in-continuities in d_i. The hazard function estimator is particularly intuitive in understanding the aging process because the hazard function is essentially the conditional probability of death at specific time if cells have survives before that time. But these estimators depend on time binning, thus become artificially noisy because the timescale of the dying process is usually higher than the binning interval (see Chapter 3, Figure 3.6 and Figure 3.7), esp. for time intervals with small number of deaths.

To be able to intuitively visualise the rate of aging, while remain insensitive to binning and measurement noise, we introduce a cumulative hazard function estimator, known as the Nelson-Aalen (N-A) estimator:

$$\hat{\Lambda}_{na}(t_i) = \sum_{j=1}^{i} \frac{d_j}{N_j} \tag{4.4}$$

The cumulative hazard function $\Lambda(t)$ is defined as the time integral of the hazard function $\Lambda(t) = \int_{t_0}^{t} h(t)dt$. It also relate to the survival function by a simple exponential transformation: $S(t) = e^{-\Lambda(t)}$. Survival function estimator derived this way is called Breslow estimator and is asymptotically equivalent to Kaplan-Meier.

$$\hat{S}_{Breslow}(t) = e^{-\hat{\Lambda}_{na}(t)}$$

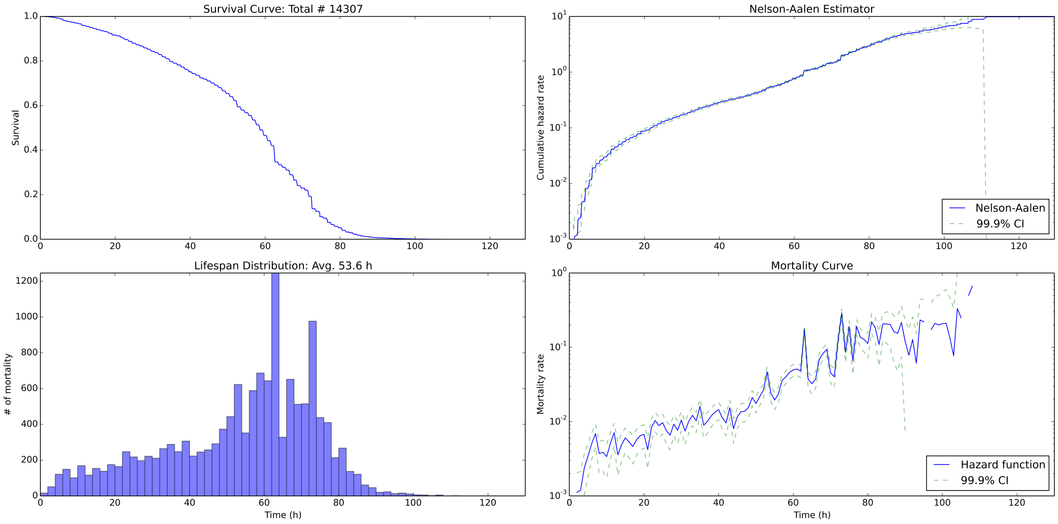

Figure 4.2: Four non-parametric statistical estimators of lifespan distributions. Upper left: Kaplan-Meier, Equation 4.1; Upper right: Nelson-Aalen, Equation 4.4; Lower left: Histogram with time binning estimating the shape of the probability density function, Equation 4.2; Lower right: hazard function with binning, Equation 4.3. The data is from a vertical mother machine carbon starvation experiment with wildtype MG1655 *E.coli* strain. Times of death are determined by thresholding PI fluorescence time-series as described in Chapter 3. Dashed line denote 99.9% confidence intervals. The upper panels (K-M and N-A estimators) do not require time binning and are thus much less noisier than the lower panels.

Except Breslow survival estimator, all four estimators introduced here are plotted in Figure 4.2. We use the K-M estimator and the density estimator (upper left, lower left) mainly to monitor experiments and to spot irregular perturbations in the data. As shown in Figure 3.5, our lifespan distributions can be modelled by the Gompertz law, i.e. exponential increase of hazard rate with time. The Gompertz slope parameter, which is understood as the rate of aging, can easily visualised as the linear slope of both the hazard function and the cumulative hazard function when they are plotted on log scale. Since the N-A estimator is much less noisier than the hazard function estimated by binning, as the lower left panel of Figure 4.2, the linear slope of the Nelson-Aalen estimator on log scale give the best estimate of the rate of aging.

Figure 4.3 shows the N-A estimators of the lifespan distributions of wildtype and ΔrpoS strains (same data as Figure 3.4 and Figure 3.5). Here, we can clearly identify the impact on rate of aging by the ΔrpoS mutation. This will be the focus of the next section.

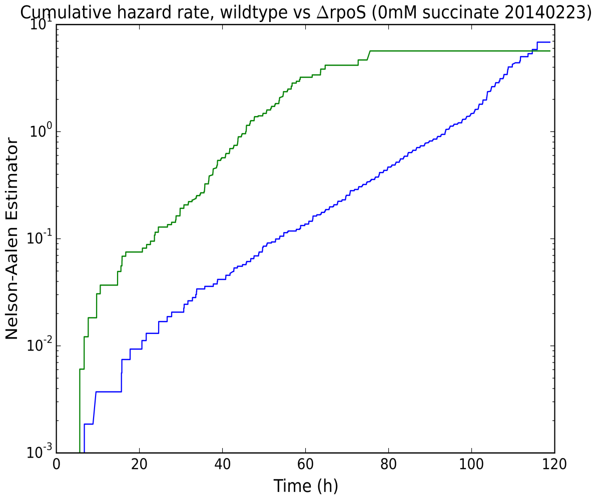

Figure 4.3: Cumulative hazard rates (Nelson-Aalen estimator) of wildtype and ΔrpoS populations in carbon starvation in mother machines. The y-axis is in log scale. Data shown are the same as those in Figure 3.4 and Figure 3.5, where populations of wildtype of ΔrpoS cells are in carbon starvation in the same mother machine. The increase in the Gompertz slope on indicate that the ΔrpoS mutation increases the rate of aging.

4.3 General stress response mutants have altered rates of aging in carbon starvation

In the previous section, Figure 4.3 demonstrated that mutants in the the general stress response pathway have altered rates of aging. To further explore this hypothesis, we systematically used the vertical mother machine to measure lifespan distributions of mutants from the Keio collection (Baba *et al.*, 2006). The Keio collection is a single gene knockout library based on the *E.coli* K-12 strain BW25113. The two most relevant mutants to the general stress response, are ΔrpoS and ΔrssB. ΔrssB is the anti-sigma factor that deliver σ_S for degradation by ClpPX protease. By knocking out ΔrssB, a larger proportion of cellular resources is devoted to the general stress response during starvation. Histograms of carbon starvation lifespans of the Keio collection wildtype, and these two mutants are shown in Figure 4.4.

In Figure 4.5, I plotted the Nelson-Aalen estimators of ΔrpoS and ΔrssB strains, compared to the wildtype strain of the Keio collection. While expectedly, ΔrpoS and ΔrssB have shorter and longer lifespans respectively, the rate of aging is also increased by ΔrpoS and decreased by ΔrssB. This is a qualitatively non-trivial observation since that lifespans could conceivably be shortened or prolonged without changing the rate of aging, by raising or lowering proportionally the hazard rate across all ages. In fact, the proportional hazard model is the default assumption in many survival analysis studies with small sample sizes. Yet in our case of the general stress response pathway, the proportional hazard model is clearly rejected in favour of the accelerated life model.

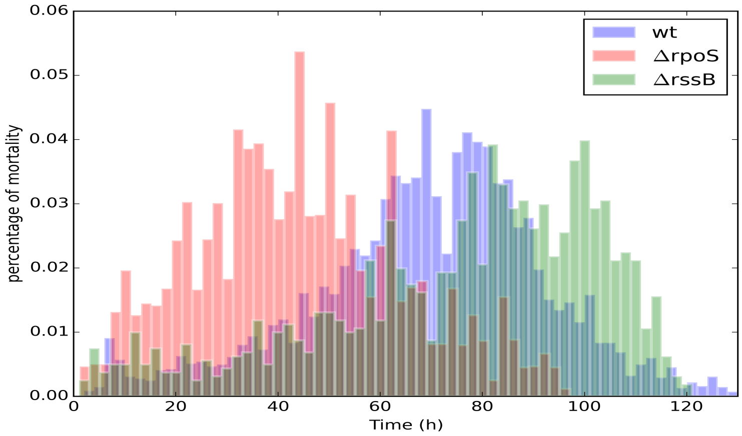

FIGURE 4.4: Lifespan distribution of Keio collection strains in carbon starvation: wildtype, ΔrpoS, ΔrssB. The colour bars are semi-transparent to allow side-by-side comparison of 3 different strains. When the different color bars overlap, they show composite colour. Data shown are from carbon starvation experiments in vertical mother machines. Number of cells to construct the histograms: 6012 for Δrpos = 6012; 6831 for wildtype; 1605 for ΔrssB.

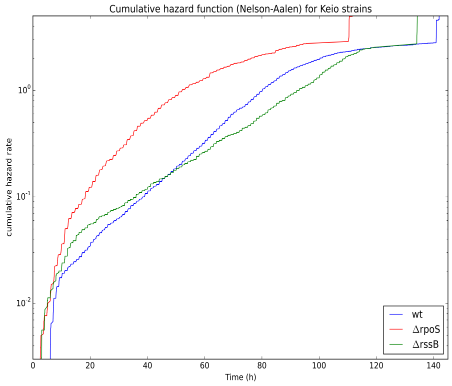

FIGURE 4.5: Cumulative hazard rates (Nelson-Aalen estimator) of general stress response mutants from the Keio collection. ΔrpoS abolish the general stress response, while ΔrssB have enhanced general stress response. Data shown are from carbon starvation experiments in vertical mother machines, same as those in Figure 4.4. N-A estimators are calculated according to Equation 4.4.

To put the clearly shifted aging rates in Figure 4.5 in context and to assess the impacts of changing genetic backgrounds, in Figure 4.6, I also plotted the cumulative hazard rate of wildtype MG1655, compared to the wildtype and ΔrssB strains of Keio collection. BW25113 is a lab K-12 strain with a slightly different culturing history and genetic background compared to MG1655, which is used in the previous chapters. Interestingly, the two wildtype strains share similar rates of aging, and clearly different from that of the ΔrssB strain. This suggests that the rate of aging is inherent to the genotype of the general stress response, further confirming the titular hypothesis of this chapter. Whether the different intercepts of the two wildtype strains, i.e. proportional hazard differences, can be attributable to the few background mutations between the two strains, requires further studies.

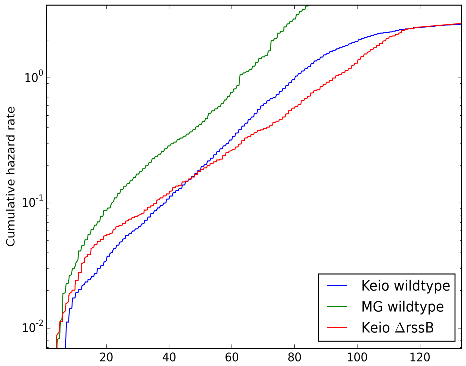

FIGURE 4.6: Cumulative hazard rates (Nelson-Aalen estimator) of MG1655 wildtype strain compared to two Keio collection strains. The Keio collection is based on strain BW25113, which is a strain of *E.coli* K-12, as is MG1655. The two wildtype strains follow similar Gompertz slopes. ΔrssB is shown for comparison. The different initial cumulative hazards does not indicate true biological differences, but are results from the different number of deaths in the first 15h, which came from the initial shock of cell loading. Data shown are from vertical mother machine carbon starvation experiments, and for the two Keio collection strains are the same as those strains in Figure 4.5.

4.4 Conclusion

Here, I have shown that by using the vertical version of the mother machine, we can measure the rate of aging with high quality lifespan data. The lifespan distributions of general stress response pathway mutants have shown that the general stress response prolong cellular lifespan by decreasing the rate of aging. Because the benefit of decreased rate of aging is mostly seen at older ages, i.e. further into starvation, the general stress response should be considered not just a protection mechanism against immediate stresses, but also as a hedge against time spent in starvation. This observation have important implications for understanding the physiological and evolutionary role of the general stress response.

Chapter 5

Physiological and evolutionary trade-offs of the general stress response

In the previous two chapters, I have presented experimental results on both the individual and population level, indicating that the general stress response in *E.coli* decreases the rate of aging at the expense of growth potentials. The main qualitative conclusions from these experiments are:

- In carbon starvation, *E.coli* cells have significant individual variations in the manner and timing of death;
- The lifespans of *E.coli* cells in carbon starvation exhibit distributions typical of aging processes;
- At the expense of growth potentials, the general stress response promotes long-term survival in starvation through decreasing the rate of aging;
- The trade-offs between long-term maintenance and short-term growth persist on both the population and individual level.

In this chapter, in order to quantify the overall fitness impacts of the general stress response and understand the selective forces driving the evolution of this pathway, I present a physiological model integrating some of these empirical observations with the well understood bacterial physiology of growth. In the model, I represent the phenotypes of the cells with two proteome sectors, one in charge of growth and the other in charge of maintenance. While the environment controls the overall energy budget of the cells, through adjusting the synthesis rates of the two proteome sectors, the genome could strategically adjust the ratio between growth and maintenance investments. Given environmental histories typically consisted of both feast and famine phases, the fitness of a particular genotype is defined and calculated as the long-term population growth rate. I will attempt to use this model to explain some of my experimental observations and those reported the the literature regarding the evolution of the general stress response pathway.

5.1 Distinction between aging and death suggests vitality-based mortality models

The qualitative observations of the PI dynamics described in Chapter 3 suggest that the physiological processes underlie aging and cell death are on different timescales. The aging process is characterised by the Gompertz slope, i.e. the exponential exponent of the mortality risk increase; while the dying process is characterised by the variable timescales of PI dynamics indicated by the CWT analysis, reflecting the different ways cells lose their membrane potentials. Although we do not know definitively the biochemical natures of aging and death under starvation, the different timescales suggest that they might be distinct processes. Genetic results in Chapter 4, where the rate of aging can be altered by mutations, are consistent with the distinction between aging and death.

The distinction between aging and death has significant implications for modelling the aging process. A large class of aging models sometimes called ‘redundancy theory of aging’ assumes that aging is a stochastic damage accumulation process ending in death (Gavrilov & Gavrilova, 2001; Weitz & Fraser, 2001). They tend to have only one timescale, i.e. the timescale of damage accumulation or redundancy reduction. In the case of my results on *E.coli* carbon starvation, aging might still be due to ‘redundancy reduction’, but there are at least two timescales associated with a slower cumulative aging process and faster mortality processes. While the mortality process is clearly stochastic, the aging process could very well be deterministic.

Strehler & Mildvan (1960), in trying to explain the Gompertz law for human populations, proposed an aging model with exactly these features. They proposed both a linearly declining 'vitality' as the driving dynamic variable behind aging, and a memory-less mortality model. The mortality model simply compares 'vitality' with stochastic challenges characterised by Maxwell-Boltzmann distributions: the likelihood of death at a given age equals to the proportion of challenges bigger than the current 'vitality'. While the assumptions of Maxwell-Boltzmann distribution and linearity of 'vitality' decline of Strehler & Mildvan (1960) do not have to be true for *E.coli* carbon starvation, the general framework seems to fit my empirical observations.

While it is hard to directly identify physiological basis or biomarkers for 'vitality' in humans and most model organisms, it is not far fetched to presume that vitality is associated with depletion of energy reserves in the case of *E.coli* carbon starvation. While more experimental evidences are needed to confirm this hypothesis, it leads to a relatively complete picture connecting the physiology of starving *E.coli* cells to the exponential increase in mortality risks characterised by the Gompertz distribution. It has already been shown experimentally that starving *E.coli* cells continuously synthesise proteins (Gefen *et al.*, 2014) from energy and nutrients reserves. The energy and nutrients reserves are in the forms of recycled amino acids from degraded proteins (Kuroda *et al.*, 2001), thus relate to the overall biomass of cells. The eventual depletion of resources will lead to a reduction of the number of functional proteins inside the cell. Zhang *et al.* (2012) proves that in certain simple stochastic models, mortality risks defined as the inverse of mean first passage time scale exponentially with system size. If this model applies to *E.coli*, the overall mortality risks scale exponentially with the biomass of starving cells, while different pathways of cell death could still exist. The general stress response decreases the rate of aging by decreasing the rate of biomass and resource depletion. We could test this hypothesis experimentally by examining the rate of aging when the cells are exposed to weak decouplers, which decrease the oxidative-phosphorylation efficiency and increase the respiratory rate.

The connection between resource depletions and mortality increases is an interesting subject worth further experimental investigations, but it is not necessary for the purpose of building a toy model to understand the overall fitness trade-offs. After all, phenomenologically we already have the lifespan distributions and the accelerated life model through which the general stress response modulate these distributions. They already allow me to calculate long-term fitness in the next section. But still, the conceptual link between energy reserve depletions and mortality rate increases lends itself very well to the resource trade-off hypothesis. In the model described in the next section, we could easily translate the accelerated life model of general stress response to the speed of energy consumption.

5.2 Modelling *E.coli* life-history trade-offs by two proteome sectors

The basic trade-off that starving *E.coli* cells face is where to invest the limited energy/nutrients reserves. On one hand, the general stress response decreases the rate of aging thus the chance of long-term survival; On the other hand, investing in the general stress response does not help the cells in immediate terms and decreases the availability of resources for other purposes such as metabolising alternative nutrients available in the environment. The strategies that *E.coli* cells could theoretically adopt are on the conservative end a 'sit-and-preserve' strategy of high general stress response, or a risky exploratory growth strategy of low general stress response. The latter strategy is one adopted by the GASP strains. The fitness and optimality of the different strategies depend on environmental parameters, such as the length of starvation or availabilities of alternative resources. In this section I present a toy model following this outline.

The model considers the energy budget of the cell, consisted of J^+, the energy income flow, and J^-, the energy out flow or the energetic cost of living. When J^+ is bigger than J^-, i.e. the conditions are favourable for growth, the energy surplus is consumed by protein synthesis; when J^+ is smaller than J^-, the energy deficit is paid by autophagy, i.e. amino acids catabolism supplied by protein degradation. We consider two types of proteins: vegetative proteins D, and maintenance proteins S. The total amount of protein in the cell is P, while the amount of D and S are P_D and P_S.

The connection between vegetative protein synthesis and bacterial growth rate is well understood quantitatively in the literature (Scott *et al.*, 2010). In this model, the function of vegetative proteins, Ds, are to generate energy income through catabolising nutrients in the environment. We have:

$$J^{+} = \Gamma^{prod}_{E/P} \cdot P_D \cdot q(env), \tag{5.1}$$

where $\Gamma^{prod}_{E/P}$ is the rate constant of energy production due to enzymatic activities of D, and $q(env)$ is the quality of energy production substrate controlled by the environment, with $q(env) = 0$ indicating complete energy starvation and $q(env) = 1$ indicating the availability of the best energy source (usually glucose).

On the other hand, resource depletion in starvation is represented by J^{-}. As I mentioned at the end of last section, vitality-based mortality models connect J^{-} with the aging process described by the increasing likelihood of cell death, i.e. the Gompertz slope. And since we know from Chapter 4 that the general stress response decreases the Gompertz slope, I assume that the function of S protein in the model is to decrease J^{-}. We can also infer that J^{-} scale with P. Since we do not know the exact functional form of $J^{-}(P_S)$, let us make the simplest assumption for now. Therefore, we have:

$$J^{-} = f(P - P_S) = \Gamma^{maint}_{E/P} \cdot P_D, \tag{5.2}$$

where $\Gamma^{maint}_{E/P}$ is the rate constant of maintenance energy needed for protein D.

Now, at steady states, the specific growth rate of the bacterial population is:

$$\begin{aligned} \mu \equiv \frac{\dot{P}}{P} &= \frac{(J^{+} - J^{-})Y_{P/E}}{P} && (5.3)\\ &= \frac{(P_D\Gamma^{prod}_{E/P}q(env) - P_D\Gamma^{maint}_{E/P})Y_{P/E}}{P} && (5.4)\\ &= p_D(\Gamma^{prod}_{E/P}q(env) - \Gamma^{maint}_{E/P})Y_{P/E}, && (5.5) \end{aligned}$$

where $Y_{P/E}$ is the protein synthesis yield per unit of energy, and p_D is defined as the proportion of D proteins: $p_D \equiv \frac{P_D}{P}$. p_D is controlled by the sigma factor competition between RpoS and RpoD (Nyström, 2004).

Interestingly, p_D scale linearly with the absolute value of specific growth rate under these assumptions. In a cyclic environment with feast ($\Gamma_{E/P}^{prod} q(env) - \Gamma_{E/P}^{maint} > 0$) and famine ($\Gamma_{E/P}^{prod} q(env) - \Gamma_{E/P}^{maint} < 0$) phases, p_D essentially plays the role of time accelerator. This reminds us the results in Chapter 4: $p_S \equiv 1 - p_D$ decrease the rate of aging through accelerated life model, in which p_S plays the role of time decelerator.

5.3 General stress response evolution is shaped by the trade-off represented by the model

Wildtype *E.coli* cells employ a complex signal transduction network to determine the outcome of this competition between RpoS and RpoD. Through adjusting p_D, the signalling network controls both the phenotype of the cells and their current strategy in the maintenance-growth trade-off.

If the signalling network controlling sigma competition is understood, we could define p_D as either constants or functions of some environmental or physiological states. Absolute fitness could be calculated as the long-term population growth rate given a certain environmental regime such as constant growth, constant starvation or alternative cycles of growth and starvation. Relative fitness could be calculated by comparing the populations of different genotypes. These fitness measures could be either analytically derived if environments are simple enough, or determined by numerical simulations under any given environmental histories.

The simulation of evolutionary dynamics is outside of the scope of this thesis. Here I list the following non-trivial empirical observations from this study and the literature that can be directly tested by simulating the evolutionary dynamics:

- When external resources are unavailable, cells that elongate more tend to die earlier, unless they are able to reach cell division (see Chapter 3);
- Long-term population increases select for aggressive growth strategies characterised by higher p_D while long-term population decline selects for the conservative strategies characterised by low p_D (Maharjan *et al.*, 2006; Vulic & Kolter, 2001);
- Growths in lower quality nutrients generate higher selective pressures for high p_D, i.e. stronger selection for GASP phenotypes (King *et al.*, 2004);

- Higher external stresses in long-term growth select for higher p_D. (Zhang *et al.*, 1998).

5.4 Conclusions

Through quantifying the effects on life-history traits, I showed that the general stress response is a microbial example of caloric-restriction induced longevity. The basic assumptions of the disposable soma theory of aging are shown to be true, in the case of *E.coli* carbon starvation. I integrated the physiologies of *E.coli* starvation with the evolutionary theories of aging into a simple model, and provided some testable predictions from my experimental results and the literature. Through this model, the selective forces associated with the evolution of aging rates can be thoroughly understood. Although the physiological causes of aging might not be the same in other model organisms and human, this study have provided an example of how a course-grained physiological model could provide the basis for understanding the origins of evolutionary trade-offs and their roles in shaping the evolution of aging and longevity..

Appendix A

Book chapter in *Methods in Microbiology*: Time-lapse microscopy and image analysis of Escherichia coli cells in mother machines

Manuscript submitted as a chapter in the upcoming book “Imaging Bacterial Molecules, Structures and Cells Methods” (December 2016) edited by Grant Jensen and Colin Harwood, volume 43 in the book series **Methods in Microbiology**, published by Elsevier.

Title: Time-lapse microscopy and image analysis of Escherichia coli cells in mother machines

Authors: Yifan Yang1,2,3, Xiaohu Song1,2 and Ariel B. Lindner1,2,3

Affiliations: 1. INSERM U1001, 2. Faculté de Médecine, Université Paris Descartes, Paris, France 3. Center for Research and Interdisciplinarity, Paris , France

Abstract

High-throughput longitudinally measurements of single cells by automated time-lapse microscopy, combined with dynamic system modelling, have in recent decades advanced our understanding of microbial physiology, such as cell size control, gene expression and regulation, sub-cellular organisation and organelle dynamics, development and cell fate decision. However, the sheer large numbers of cells resulting from exponential growth often limits both the experimental duration and numbers of single cells quantified. To address these issues, microfluidic devices are specifically designed to spatially track and align single cells or lineages of cells. In this chapter, we use the example of tracking Escherichia coli (E.coli) cells in mother machines to introduce methods and potential issues in the design, execution and image analysis of such experiments. For experimental design, we highlight microfluidic and time-lapse microscopy considerations important for the successful interfacing between the two technologies and for obtaining easily quantifiable images. Then we provided typical protocols in executing such experiments. Finally, we devote a large portion of this chapter to introduce image analysis methods automating the segmentation, tracking and quantification of cell images obtained from mother machine experiments.

Keywords: Time-lapse microscopy, microfluidic cell culture, single cell life-history, mother machine, image analysis automation

A.1 Introduction

Quantitative longitudinal measurements of single cells by time-lapse microscopy have been a popular technique in systems biology and microbial physiology in recent decades. Studies employing these types of experiments have advanced our understanding of cell size control (Campos *et al.*, 2014; Taheri-Araghi *et al.*, 2015), gene expression and regulation (Cai *et al.*, 2008; Golding *et al.*, 2005; Locke *et al.*, 2011), sub-cellular organisation and organelle dynamics (Babic *et al.*, 2008; Parry *et al.*, 2014), cell fate determination (Suel *et al.*, 2006).

However, the exponential nature of microbial growth often limits the number of generations cells can be followed through microscopy. For instance, in experiments where micro-colonies are followed, as the single layers of cells are expanding exponentially over time, tension and friction between cells inevitably push them into multiple layers. Furthermore, the post-experimental image analysis process limits the number of cells whose relevant traits can be quantified. In each image frame, cells have to be segmented from each other. The same cells and their progenies have to be tracked through consecutive frames to form lineages. Consequently, the data generated from a single overnight experiment could take weeks to analyse.

These experimental constraints could be overcome by the applications of microfluidic devices specifically designed to spatially isolate and align single cells or lineages of cells. Examples of such devices are the mother machine, designed to track old pole cells and their immediate progenies of coliform bacteria (Wang *et al.*, 2010b), and U-shaped traps (Nagarajan *et al.*, 2014; Rowat *et al.*, 2009) designed to isolate mother cells of budding yeasts. These devices contain geometrically fitting microstructures to restrict cells to spatially regular patterns. These geometrical restrictions greatly facilitate the tasks of image analysis. In addition, constant microfluidic flows remove most of the newborn cells and focus the analyses on constant numbers of cells within ever-expanding populations.

This chapter aims to use the example of tracking E.coli cells in a device known as the mother machine to demonstrate the general principles of such spatially structured microbial time-lapse microscopy experiments. We introduce common methods and potential issues in the design, execution and image analysis of such experiments. Since such experiments involve micro-fabrication, microscopic, microfluidic and image analysis methods, each of which could be covered in their own chapters, we chose to focus this chapter on the interplay between them.

A.2 Experimental designs

A.2.1 General principles of the mother machine

The mother machine is a good example of microfluidic devices used to isolate and align single bacterial cells in microscopic studies. It was originally designed to track the old pole cells of elongating and dividing lineages of E.coli, hence the name. Due to its simplicity of design, ease of use and the ability to track single cells for up to one hundred and fifty generations, its use spreads quickly and has been successfully used to culture and study coliform bacteria such as Escherichia coli and Bacillus subtilis (Norman *et al.*, 2013; Wang *et al.*, 2010b).

The general principle of the mother machine is very simple: an array of single-cell-wide channels with dead ends on one side, and a much larger main flow channel on the other supplying fresh medium and removing newborn cells. After cells are loaded into the dead-end channels and fed with fresh medium, they grow into one-dimensional 'colonies'. Once the dead-end channels are completely filled, cells on the open side of the colony will be pushed into the main flow channel and washed away, as shown in Figure A.1a. By constantly supplying fresh medium through the main channel, the growth conditions reach steady states quickly and can be kept in the exponential regime for hundreds of generations (Wang *et al.*, 2010b). In these conditions, the dead end channels contain the inner most cells of the whole lineage, i.e. the mother cells and their immediate progenies.

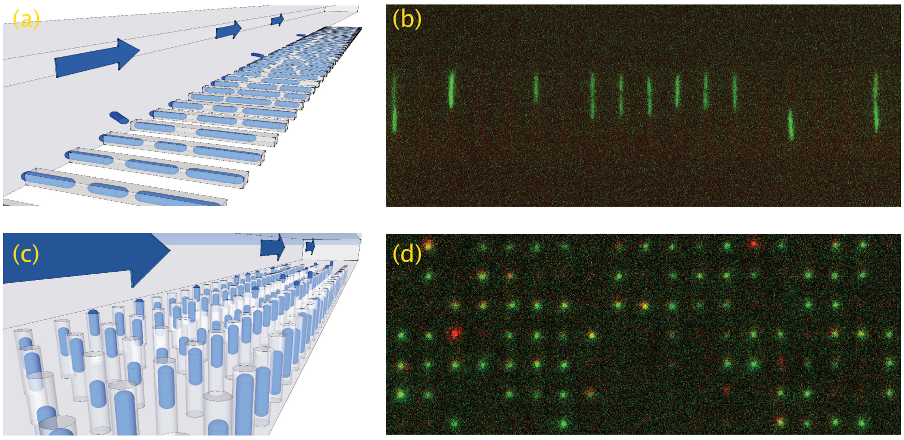

Figure A.1: The general principle and design variations of the mother machine. Shown are the 3D cartoons (a,c) of two variants of the mother machine. Cells are trapped inside dead end channels, and are growing as 1D colonies. Extra cells are pushed outside the dead end channels and removed by the flow in the main channel. The arrays of dead end channels are aligned either horizontally (a) or vertically (c) to the imaging plain. (b,d) are actual fluorescent images of these devices containing cells of E.coli constitutively expressing a variant of mVenus (shown in green). They are also exposed to propidium iodide in the medium, whose fluorescence (shown in red) is an indicator of cell death.

The obvious benefit of using the mother machine is that cells could be followed for days in exponential growth, without either change in media conditions or obscuration due to over-crowding. This type of long term longitudinal data could very useful if the phenomena of interests are rare or takes more than 10 generations to develop, such as ageing (Wang *et al.*, 2010b) and stochastic developmental decisions (Norman *et al.*, 2013). Another feature of the mother machine and similar microfluidic systems is the ability to easily and quickly shift culturing conditions. Coupled with time-lapse microscopy, single-cell level dynamic behaviour in physiology and gene expression can be easily observed (Izard *et al.*, 2015).

One advantage of the mother machine we find especially relevant for this chapter is the fixed locations and geometry of cells and colonies within the mother machine. Instead of having to segment and track hundreds of individual cells from crowded and moving 2D colonies (Ni *et al.*, 2012), one only need to identify cell divisions and boundaries between cells in one dimension. This made image processing much simpler by reducing manual intervention and error correction, and improving the quantification throughput from time-lapse movies.

The rest of this section discusses in turn the design considerations of both the mother machine itself, and time-lapse imaging process. The next section will demonstrate the operational procedure of mother machine time-lapse experiment using our experimental protocol. The last section is dedicated to the image analysis methods of these experiments.

A.2.2 Device design considerations

With a number of reasons for applying the mother machine to time-lapse experiments, the specific objectives of the study determine the exact designs and dimensions of the mother machine itself. The first design question to ask is that, in an exponentially expanding lineage of cells, which population does the experimentalist want to observe? Since the original mother machines will only trap the old pole cells and their immediate progenies, one has to consider whether tracking this sub-population will either serve the objectives or bias the conclusions. If the objective calls for tracking different sub-populations, the basic design of the mother machine has to be altered, for example by opening up the dead ends by connecting them to another channel.

Once the question of which cells within the exponential lineages settled, the second design decision involves the question of how many cells to follow and the length of the dead end channels. Dynamic physiological and developmental events often span multiple generations (Izard *et al.*, 2015; Norman *et al.*, 2013), and sometimes correlations between the cousins rather than immediate progeny have to be examined (Hilfinger & Paulsson, 2015). For time-lapse experiments designed to measure these phenomena, the tracked sub-population has to include every individual within several generations. This requires the dead end channels to be longer than $L \times 2^N$, where L is the average length of cells, and N is the number of generations needed to be followed.

The length of dead end channels is limited by two factors. The first is the heterogeneity of the growth conditions along the length of dead end channels. The small cross-sectional area within the channel may limit the diffusion of the media, and waste resulting from metabolic activity may accumulate. A variation of the original mother machine designed to relieve the diffusion problems (Norman *et al.*, 2013), involves the enlargement of the upper halves of the dead end channels. The other limiting factor is the surface friction between cells and the device itself, caused by growth and cell displacement. Since one end of the one-dimensional colony is spatially fixed, the other end has to be displaced at a velocity proportional to the total length of the colony and average growth rate.

This friction may cause the misalignment of colonies in a way that is counter-productive to the image analysis process. Thus the dead end channels should not be longer than necessary. Depending on the requirements of the experiments, growth homogeneity and the correct colony alignment needs to be confirmed experimentally (Wang *et al.*, 2010b).

The width of dead end channels is also important in aligning the cells in ways that are convenient for image analysis and quantification. Specifically, if the dead end channels are significantly wider than the cells in the intended culture conditions, the friction mentioned above may push neighbouring cells into overlapping configurations. Irrespective of whether the overlap is perpendicular to the imaging plain or along the imaging plane, as is the case in Figure A.2, it would significantly complicate the segmentation step in image analysis. Obviously, the dead end channels should not be too small such as to prevent the cells from entering them. The best approach for determining the correct width of the dead end channels would be to first measure accurately the cell width under the intended culture conditions, micro-fabricate a series of mother machines with a range of dead end channel widths at or around that the measured cell width, and then determine experimentally those with the most suitable width, as described in Section A.3.

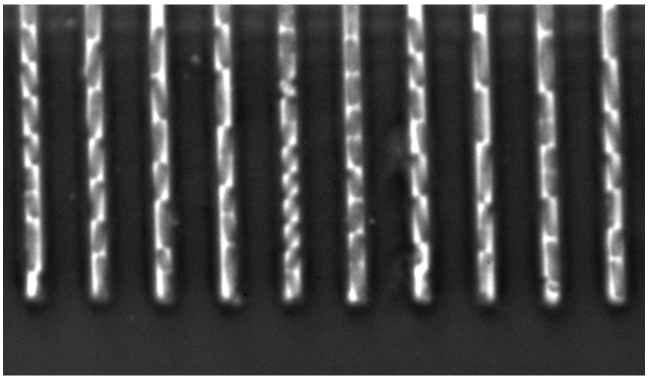

Figure A.2: Phase-contrast image of the dead end channels of a mother machine filled with narrower E.coli cells. Due to the disparities between the widths of the cell and the dead end channels, the bacterial cells are misaligned and might cause segmentation problems in image analysis.

Last but not the least, different layouts of the main flow channels and dead end channels allows for either multiplexing of independent experiments, or increasing of population size in a single device, depending of course on the requirements of the experiment. For multiplexing, the standard approach of having identical units of main channels lying parallel to each other is suffice. This allows independent microfluidic experiments and culture populations be recorded in the same device and time-lapse microscopy session. For increased population size, Figures A.1c and A.1d demonstrate a variant of the mother machine, where dead end channels are aligned vertically to the imaging plane and the main flow channel lies on top. This configuration increases the number of cells that can be monitored in a single field of view at the cost of lineage information, and is used in our laboratory to study E.coli mother cell lifespan distributions.

A.2.3 Time-lapse considerations

In general, time-lapse microscopy can be conducted on mother machines or similar microfluidic devices without major alterations or additional equipment. We conducted all of our mother machine experiments on conventional inverted epifluorescence microscopes from major commercial brands. Besides the microscopes and illumination sources, our platforms are fitted with temperature control chambers, motorised and automatically controlled stage and focusing, fluorescent filter wheels and shutters, and CCDs. The systems are managed by commercial microscope automation software such as MetaMorph® installed on dedicated PCs. Yet, when the mother machine is employed, the time-lapse experiments are usually designed to last over very long durations. To achieve the best results, certain aspects of time-lapse microscopy need to be specially considered during the design process.

An inherent trade-off for multi-positional time-lapse microscopy is between sample size and time resolution. This trade-off can be expressed as:

$$k_{pos} = \frac{N_{sample}}{\langle n \rangle} \leqslant \frac{T_{\text{int}}}{\langle t_{img} \rangle},$$

Where T_{int} is the imaging interval, $\langle t_{img} \rangle$ is average time of moving to, focusing and imaging one position, $\langle n \rangle$ is average number of cells or lineages per position, N_{sample} is the total sample size, and k_{pos} is the number of imaging positions.

For quantitative experiments aiming to measure dynamic variables such as growth rate and promoter activity, T_{int} should be much shorter than the mean generation time. Even if these measurements are not required, it is necessary to have a minimum of 4 frames per division time in order to correctly track and segment cells. This poses an upper boundary for T_{int}. And since N_{sample} is determined by the feature layout in a microfluidic device and cell loading efficiency, the maximum T_{int} also set upper boundaries for k_{pos} and N_{sample}. Before recording the imaging positions and starting the time-lapse loop, the experimentalist should estimate $\langle t_{img} \rangle$ and $\langle n \rangle$ experimentally and set an appropriate k_{pos} and T_{int} regime to achieve appropriate sample size. If appropriate N_{sample} values cannot be met, aspects of the experiment need to be changed, for example by increasing loading efficiency or shortening the focusing and imaging procedure at each position. In order to maximise overall sample size, it is important to minimise imaging time at each position and only use complex focusing procedures when required, as we will discuss next.

Microfluidic devices, including the mother machine often complicate the focusing procedures of automatic time-lapse microscopy. This is because the system has to automatically identify the cells within the device among multiple material interfaces and micro-fabricated structures. Software-based auto-focusing methods such as those maximising image contrasts may fail because they choose larger and more salient device structures rather than the cells. This problem can be avoided by using hardware-based methods such as PFS from Nikon or Definite Focus from Zeiss, which use LEDs to track the bottom of the device. However, these systems often require manual resetting when the stage travels large distances and outside their narrow operational z-ranges. In our experience, the most robust and efficient approach is to use software-based auto-focusing as a complement to LED-based systems: when they report out-of-range errors, we use contrast maximisation to locate the narrow z-range in which both the micro-fabricated structures and cells are positioned, thus automatically moving the focus to within the operational range of the hardware systems. This approach can be automated with scripts in MetaMorph® or other microscope automation software packages.

A.3 Experimental procedures

In this section, we present our experimental procedures for setting up E.coli mother machine experiments for inverted fluorescence microscopy. This protocol is one that is used routinely by non-specialists in our laboratory. The protocols for these types of experiments are necessarily complex, consisting of several phases, including the preparation of the mother machine, cell cultures, fluidic systems and time-lapse microscopy. Instead of detailing every possible alternative method for each phase of the protocol, we leave the detailed discussions of each phase to specialists. We hope to demonstrate the general process of the experiments with an emphasis on their successful integration.

A.3.1 Making polydimethylsiloxane (PDMS) mother machine devices

We use a soft-lithography approach to fabricate our devices, by casting PDMS structures out of pre-fabricated master negatives on silicon wafers. The lithography methods used to fabricate these master negatives are outside the scope of this chapter. We only want to emphasise that the choice among different lithography methods depends on the degree of precision required to manufacture the dead end channels, which in turn depends on natural variations of cell morphologies.

To cast PDMS layers out of the master negatives, we pour uncured PDMS mixes onto the wafers, de-gas, spread out with either gravity (in case of mm-level design thickness) or spin-coating (in case of µm-level design thickness), partially cure with heat. We cut out the PDMS slabs with the micro-fabricated area from this partially cured PDMS, and separate it from the wafer. To provide fluidic inlets and outlets for the main fluidic channels we punch holes in the PDMS slabs with 22-gauge Luer Stubs, entering from the front side where the ends of the main channels are located, and exiting through the backside of slabs. 25 gauge Luer Stubs are used to remove any PDMS dust particles from the inlets and outlets. Then PDMS layers are bounded with each other or onto cover glasses after both surfaces are treated by room air plasma for 90s. The bounded device is then immediately fully cured at 80 ℃ overnight. These soft-lithography steps should be done in a clean room or at least in an equivalent compartment in which the air has been cleaned by HEPA filters.

Before mother machines can be used in cell culturing experiments, two types of surface treatment may be necessary. First, due to the small dimensions of the dead end channels, the naturally hydrophobic PDMS surfaces need to be made temporarily hydrophilic to enable the channels to be wetted. Secondly, the main channel surfaces need to be modified to prevent bacterial adhesion and biofilm growth. To achieve both ends, we treat the devices with room air plasma for 90s and then immediately inject a wetting solution consisting of the cell culture buffer (in our case M9 base medium) and 1.5% polyethylene glycol 400 (PEG 400). Depending on the applications, surfactants such as polysorbate 20 (Tween® 20) can also be added to the wetting solution. As a general rule, before injecting any solutions into the device, including the wetting solution, they need to be pre-filtered to prevent blockage by dust particles or crystal sediments. Before it can be loaded with cells, the device needs to be treated by the wetting solution for at least 30min by maintaining a flow rate of at least 0.2ml per hour.

A.3.2 Setting up the fluidic system

We use pulseless syringe pumps with high-precision stepper motors to drive our fluidic systems. Since most of the mother machine applications only require infusions at constant flow rates, pulseless syringe pumps are chosen for their simplicity over alternative solutions such as pressure-based flow control systems. Peristaltic pumps should not be used due to their lack of precision and tendency to generate flow pulses.

To prepare the flow system, we first fill several syringes with the required culture media. If non-covalent surface modifying molecules such as PEG 400 or Tween® 20 are used in the wetting solution, they should also be added to the culture media to the same concentrations. The media should be filter-sterilised before being loaded into sterile syringes. After the syringes are filled, we also cap them with sterile 0.2 µm syringe filters to prevent cross contamination before adding 23-gauge blunt end needles for insertion into tubing. We then load all syringes onto syringe pumps, and set the pump to the correct syringe diameters.

We use a semi-flexible thin-walled polytetrafluoroethylene (PTFE) tubing (I.D. X O.D. = 0,56mm x 1,07 mm) to connect the syringes and the microfluidic devices. Traditional rigid PTFE tubing is too inflexible to allow the range of motions needed for multi-positional time-lapse microscopy. Yet, the super-flexible PVC tubing traditionally used for flexible infusion applications should also be avoided because they have been found to leach phthalate plasticisers into the media (McDonald *et al.*, 2008), that, in our laboratory, was found to have physiological effects on the bacteria. Similar semi-flexible plasticiser-free tubing can also be used, such as TYGON® formulation 2001, can also be used. 23 gauge metal couplers are used to connect tubing segments to each other and to the PDMS device.

A.3.3 Cell culture and loading

Cell cultures can be conducted with conventional methods. Because of their physiological and morphological homogeneity, we usually load harvested exponential phase cells into the mother machine. Successful cell loading requires highly dense cell suspensions, concentrated at least 1000-fold from exponential phase cultures. We usually start with 50-ml culture volumes per main channel. Since the final cultures will be injected into the microfluidic device, we usually filter sterilise culture media to remove dust particles. Once the cultures reach the desired optical densities, cells are harvested by centrifugation at physiological temperatures. If the medium in the mother machine is different from those in exponential cultures, cells should be washed by repeated re-suspension and centrifugation.

After the last centrifugation step, minimal volumes of filtered media, usually less than 50 µl, should be used to gently re-suspend the cell pellets. Then these dense but homogenous cell suspension should be injected into the microfluidic device. The main channels are considered to be completely filled when the cell suspensions are visible in both the inlet and outlet tubing. After the removal of the inlet and outlet tubing and couplers, the device is ready for cell loading by centrifugation. The device should be secured to the centrifuge rotors with the dead end channels pointing towards the direction of the centrifugal force. In the original mother machine design, cell loading can be done with 15min centrifugation at 200 x g.

Successful cell loading should be confirmed by microscopy. In the ideal case, more than 80% of the dead end channels should be each filled with at least one cell. If loading is unsuccessful, a mother machine with wider dead ends may be required. If loading is successful, cell suspensions left in the main channel should be washed away immediately to avoid cell adhesion and biofilm formation. In our experience, this is done most efficiently by manually injecting and passing 4–6 small air packets quickly through each of the main channels. The water-to-air surfaces of the air packets are mostly responsible for removing the cells from the surfaces of the main channels. Passing each air packet should be done in seconds to avoid cells inside the dead end channels from drying.

A.3.4 Microfluidic and time-lapse setup

The last steps of the protocol involve immobilising the microfluidic device on to the microscope stage while connecting to the fluidic system. We use scotch tape to fix both the microfluidic device and inlet/outlet tubing onto the microscopy stage adapters, either supplied by the microscope manufactures or custom made. Then we connect the inlet tubing with the syringe pumps, and outlet tubing to waste collection. Immediately after, we start media infusion at a high flow rate, usually 2ml per hour but depending on the resistance of the device, to eliminate air packets that formed at the tubing connection sites. Once all media inside the device and the tubing are refreshed, flow rate is set to that required by the in-device cell culture.

Having stabilised the flow system stabilised, the time-lapse microscopy automation cycles can be started. We use the Multi Dimensional Acquisition module of MetaMorph® to manage the automation process. Besides the imaging positions where cells of interests are located, we also add a few positions along the main flow channel to the imaging queue to monitor the overall condition inside the device.

At this stage, the experiments should be left to run on its own. Both temperature and flow rate monitoring can be installed and data logged along with the time-lapse images. The 16-bit images recorded by our CCDs can take up large amount of digital storage space. We therefore have written custom scripts to compress them for uploading onto a centralised data storage facility.

A.4 Image analyses for lineage construction and single-cell traits

A.4.1 Pre-processing

The goal of image preprocessing is to enhance the visual appearance and improve the manipulations of the datasets. Microfluidic microscopy images often suffer from various difficulties:

- Images may be noisy as a result of limited light intensity;
- Images may suffer from uneven illumination;
- The cell channels may not be aligned vertically;
- The distance between two channels is unknown

To remove the noise and equalise illumination, we use a Fourier transformation based band-pass filter on the spectral space to eliminate the unwanted high frequency signals and flatten the background by filtering out the low frequency signals. Instead of keeping time frames within ImageJ stacks, dead end channels are cut out of each frame. And then for each individual dead end channel, frames are displayed chronologically from left to right on a single image, see Figure A.3. This greatly simplifies the tracking process and save space for restoring the images. For example, a 500–1000 frame time-lapse movie of a mother machine with 4 µm between dead end channels can be transformed into 16–19 images that include only cells within dead-end channels. In order to achieve this, the exact distance between channels in terms of pixels and the rotation of the channel should be known. By transforming the edge images obtained by Canny filter (Canny, 1986) into parameter Hough space (Hough, 1962), the localisations of these lines are defined parametrically within a distribution of points. The average rotation and distance of these lines can be easily computed based on the most frequently represented points in the transformed parameter space. As a result, rotating, cutting and connecting of the channels can easily be automated.

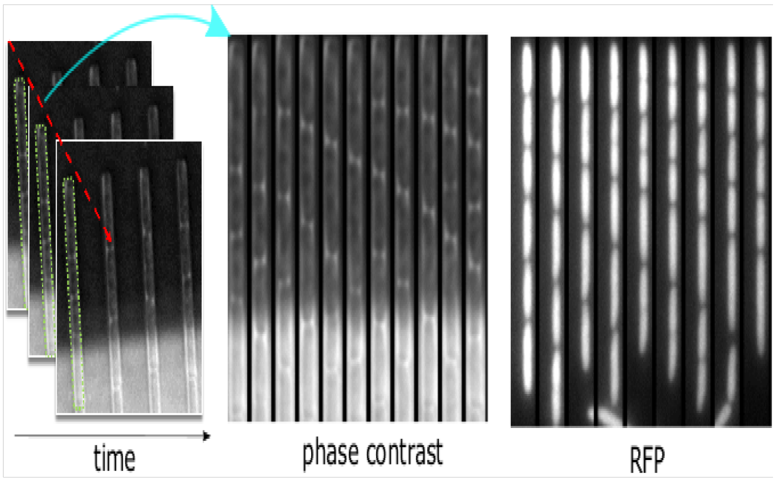

FIGURE A.3: Time-lapse microscopy images automatically rotated and cut into a time series image.

A.4.2 Segmentation approaches

It is clear that segmentation accuracy directly affects cell tracking. It therefore makes sense to approach the design of the segmentation method from the point of view of one-dimensional cell sequences inside dead-end channels: The microfluidic device is simpler than the 2D micro-colonies, with key information already provided because the cells are restricted to grow in a vertical channel. This type of structure can help to determine what methods of segmentation are truly useful compared to those used to segment 2D microcolonies. With the exemption of simple techniques such as thresholding, segmentation algorithms require the examination of intensity profiles along the centre of the channel. Based on the type of images that the experiments generate, we will discuss two different segmentation approaches for fluorescent or phase-contrast images. Our results show that the segmentation of fluorescent images can be realised by dilatation-like region growing process (Primet *et al.*, 2008) and the segmentation of phase-contrast image can be implemented on an estimated intensity profile. We introduce these very different approaches and compare results using experimental data.

A.4.2.1 Phase contrast image

The benefit of segmenting one-dimensional cell sequences is that the segmentation can be realised by peak detection on the intensity profile. In this way segmentation can act to prevent undesirable results. Instead of detecting the peak on the original image's intensity profile, it is more useful to apply this peak detector to a more textured image, defined by structure tensor (Bigün & Granlund, 1987). Its purpose is to replace each pixel of the image with their eigenvalues within a predefined scale. Thus, the intensity profile will be extracted from the so-called Eigen image. Figure A.4 clearly shows that the resulting structures are more remarkable and smooth than the original image. The original image's intensity profile caused it to keep a noisy peak and jettison a real peak, simply because the non-uniformed intensity inside the cell. The robustness of this method in decreasing false positive peak positions and the effect of noise by Eigen image has been confirmed experimentally, see Section A.4.4.

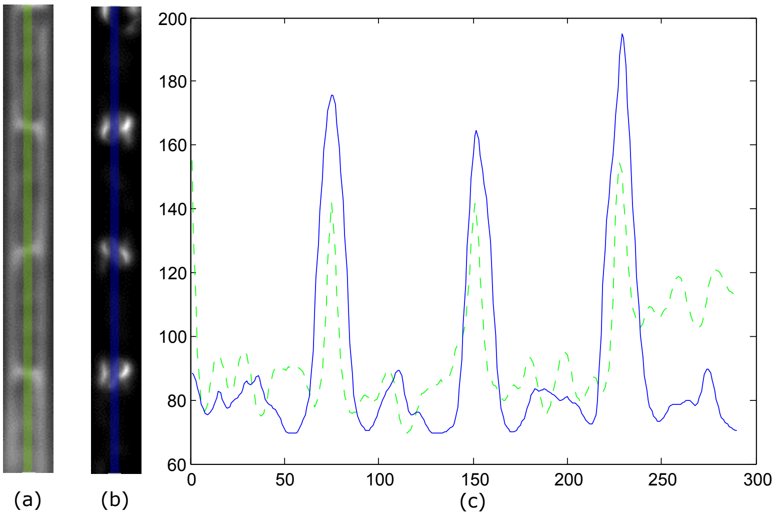

FIGURE A.4: (a) Original image and (b) structure tensor transformed image. (c) Representation of the intensity profiles through the original image (blue) and Eigen image (green).

A.4.2.2 Fluorescent image

The use of fluorescent microscopy provides us additional tools to record the time-lapse microscopy images. To capture images of cells in microfluidic devices, fluorescent proteins are often expressed so that cells can be detected with much less noise. As a result it is important to find segmentation method based on such images. One of the most powerful tools for doing this is to non-uniformly dilate from predefined seed points in the image and then to recursively add the most similar point as determined by its intensity and geometric distance to the seeds to form labelled regions (Primet *et al.*, 2008). Our challenge is to discover ways to precisely detect the seed points and support region growing to each seed points, including using automatic local thresholding methods. Although the simple thresholding method is widely used in defining seeds, it makes no use of the geometric information present in the image and consequently the results are limited. Therefore, we need to carefully reconstruct the seed points' detector from observed cell intensity distribution. In the method developed by Aguet and colleagues (Aguet *et al.*, 2013), the point-spread function model is used to configure the statistical p-value of the estimated intensity and the background noise and, as a result, the significance of each candidate signal, detection sensitivity and selectivity are improved over existing single-molecule detection methods. In the fluorescent image shown in Figure A.5, the cell intensity distribution function is an approximate Gaussian function and the cell intensity profile follows a Gaussian distribution. As a result, cell intensity can be approximated by the sum of estimated Gaussian amplitude, the background intensity and the noise, as described in (Aguet *et al.*, 2013).

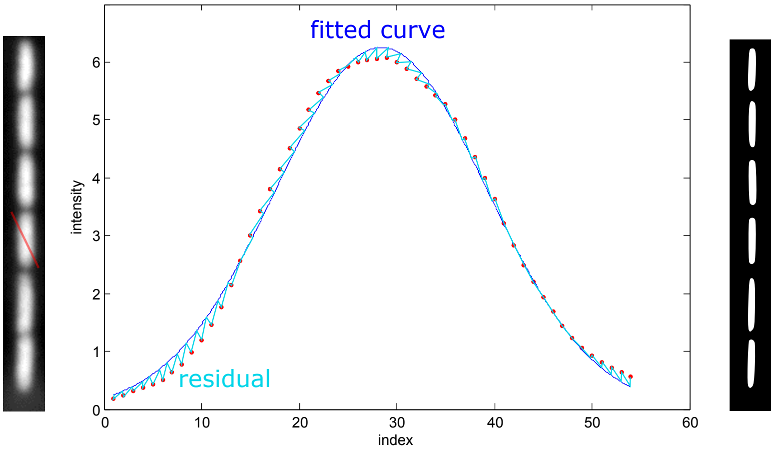

FIGURE A.5: An example of the intensity profile through a selected cell. The fitted curve is performed by an approximation of Gaussian function. The most statistically significant pixels (seed points) are selected by performing a one-sided, two-sample t-test of the fitted amplitude and the noise, estimated by the residual sum of squares (RSS). The right image presents the mask of the seed points that are defined as $m[k] := p[k] < \alpha$, as described in (Aguet *et al.*, 2013).

A.4.3 Lineage approaches

For time-lapse microscopy images, the most important biological information is the life-histories of individual cells, e.g. cell lineages captured through either manual observation or image analysis. After each imaging interval, a cell in the previous frame normally ends up in two states, cell continuation $C \rightarrow$ and cell division $C\%$, each of which needs to be defined as an assignment mapping of a cell at time t and t+1 (Figure A.6). Because of the design of the mother machine, a third cell state $C \downarrow$, is marked to indicate that either a cell is pushed off due to the growth of other cells or a cell without any ancestor or descendant. In addition, it is also possible for a cell to explode and disappear; in this case a virtual cell C_v is used to replace the invisible cell to avoid tracking errors. In this section, we will discuss the three main steps for building cell lineage, (1) cell tracking, (2) tracking error detection and (3) semi-automatic error correction.

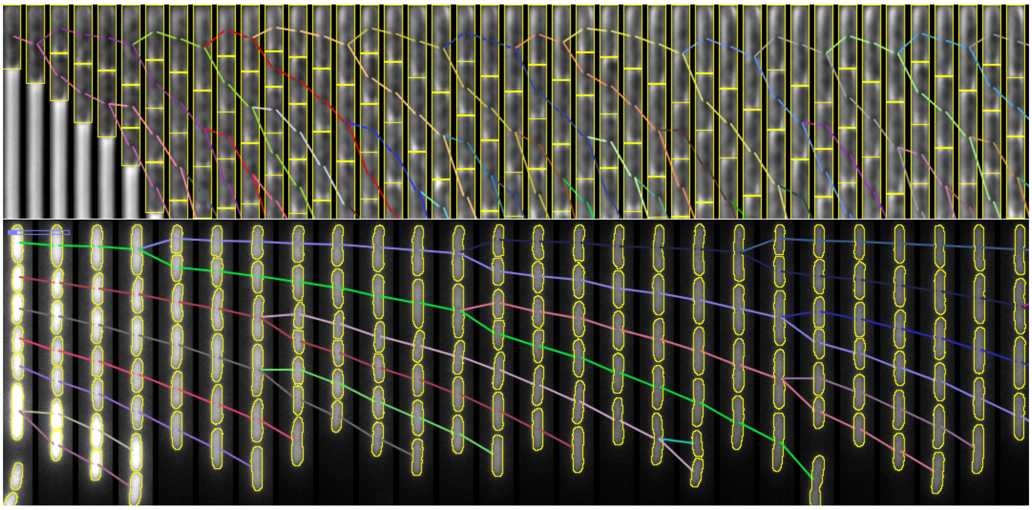

Figure A.6: Cell lineage expressed by linked arrows assigned to cell segmentation calls. The protocol for tracking cells is the same for the phase-contrast and fluorescent images, and therefore no additional segmentation information is needed.

A.4.3.1 Cell tracking

To determine the life-history of a cell from the first generation until its ultimate fate, cell tracking is required to quantify cell behaviour. Tracking considered as an assignment model (Jiang *et al.*, 2007), has been studied widely and transformed as a problem of global optimisation. Instead of giving a definite segmentation of a cell, Jug and colleagues consider all possible combinations of over-segmentations (Jug *et al.*, 2014). Based on certain physical and geometrical penalties, the most plausible segmentation was chosen via linear programming optimisation, with several constraints related to the simple biology of the system. However, the definition of the penalty, which is at a low level, is incomplete and has developed rather arbitrarily. Moreover, since the intensity profile is often determined on the basis of noisy un-delineaged noisy images (Figure A.4), there is a high risk that the hypothesis of the combined segmentations at the first stage is incomplete. In our work, we are primarily concerned with expressing the tracking observations in a concise assignment hypothesis due to the high accuracy of the segmentation process. 1, $C_{a(t)} \to C_{b(t+1)}$ (continuation): If cell $C_{a(t)}$ has an ancestor and has not assigned a descendant, meanwhile, if $C_{b(t+1)}$ has not yet an ancestor and if size of $C_{a(t)} < \lambda_1 C_{b(t+1)}$, then $C_{b(t+1)}$ is assigned as $C_{a(t)}$'s descendant, and similarly $C_{a(t)}$ is assigned as $C_{b(t+1)}$'s ancestor. In our experience, λ_1 is a fixed value that is slightly bigger than 1 ($\lambda_1 = 1.055$), since some cells gradually shrink, especially when starved.

In cases where the size of $C_{b(t+1)}$ is beyond the size of $\lambda_2 C_{a(t)}$ ($\lambda_2 = 1.6$), then $C_{b(t+1)}$ will be assigned also as the descendant of the bottom cell (named $C_{a\_bottom(t)}$ if existed) if and only if the size of $C_{a(t)} + C_{a\_bottom(t)} < 1.2\ C_{b(t+1)}$. By mapping $C_{b(t+1)}$ as descendants twice, this assignment reflects an error in segmentation. We use Random Forest discussed in the next section to detect this kind of error. 2, $C_{a(t)}$ % $C_{b(t+1)}$ (division): If cell $C_{a(t)}$ has an ancestor and has not been assigned a descendant , meanwhile, if $C_{b(t+1)}$ has not yet an ancestor and if size of $C_{a(t)} > \lambda_1 C_{b(t+1)}$, then $C_{b(t+1)}$ is assigned as $C_{a(t)}$'s daughter cell, $C_{a(t)}$ is assigned as $C_{b(t+1)}$'s ancestor. In the case where $C_{b(t+1)}$'s bottom cell existed (named $C_{b\_bottom(t+1)}$) and its size is $1.55 C_{a(t)} < C_{b(t+1)} + C_{b\_bottom(t+1)}$, then $C_{b\_bottom(t+1)}$ is assigned as $C_{a(t)}$'s younger daughter. $C_{a(t)}$ is assigned as $C_{b\_bottom(t+1)}$'s ancestor as well. 3, $C \downarrow$: If from step 1 to 2 the cell is not assigned an ancestor nor a descendant. C_v is assigned to the descendant of the cell if the actual cell has no overlap with the previous cell.

These are effective assignment definitions, not only because the parameters are easy to set up, but more importantly because they provide more visual and easy-to-comprehend feedback to users. Because these assignment definitions remain largely independent of the cell's image content (either phase-contrast or fluorescence) it connects, the segmentation and tracking error detection processes can be performed without requiring the presence of the original images. We will discuss the error detection and correction in the next sections.

A.4.3.2 Error detection

The graphics with which cells are connected, in some respects, can influence the identification of segmentation errors: cell-tracking with visible linkages such as arrows intuitively reveals the segmentation errors more than the unmarked cells. Nevertheless, to help discriminate between the correct and wrong segmentations based on these linkages, far more efforts needs to be taken. In our work, we make use of a Random Forest classification method of (Ho, 1995) to train a classifier of segmentations with respect to its high accuracy and efficiency. By learning the relevant segmentation and tracking information from manually curated training images, Random Forest repeatedly generates decision trees, from bagging to Random Forests, to categorise segmentations based on the weighted trees. It then uses the "majority vote" to predict the categories of tracking errors. The results showed that the effect of feature selection by the Random Forest was markedly inferior to traditional classification methods. Despite this, it is more appropriate to only give the algorithm the necessary features, rather than expecting it to learn to ignore the irrelevant ones. In our work, we have chosen the following measurements as the key features for the Random Forest. For segmentation: mean, standard deviation, minimum value, maximum value, median, mode. For tracking: estimated growth rate, presumed growth rate, difference of cell lengths and cell centre coordinates, Fourier open curve descriptor. For the descriptor, each cell is considered as a training data point composed of a vector of segmentation and tracking features. For the tracking features, the coefficients of Fourier series were used to describe a lineage around a cell that, in our case, corresponds to an open curve. We adapt the type-P Fourier descriptors using the slope information (Uesaka, 1984) to achieve the invariance of translation, rotation, scale and the start point.

The accuracy of this classification on the test data set was assessed in the context of segmentation classification in Section A.4.4. We show that, in most cases, using only tracking descriptors provides very good results compared to segmentation features. Once again, results show that the tracking process helps to detect segmentation errors.

A.4.3.3 Error correction

Our studies have demonstrated a strong connection between the segmentation and tracking. Incorrect segmentation always leads to the wrong assignment of cell lineages. Visibly connecting cells by arrows has greatly increased the detection of segmentation error. In our work, one solution to automatically correcting of segmentation errors, the voting mechanism, is discussed below.

Because of the efforts we have put into the segmentation step, around 99.5% of the segmentation calls we make are correct. Incorrect segmentation calls often occur when the cells have exploded, are out of focus, have low light intensity, or have aggregates inside. Fortunately, we can use the consensus among the correct segmentation calls to help correct the wrong segmentation calls by what we call the "voting mechanism". The ancestors and descendants of a segmented cell will all be considered as voting parties. The majority vote taken by this mechanism will decide if the cell is over or under segmented and this process is illustrated in Figure A.7. The voting mechanism benefits from high initial segmentation accuracy so that the participation of correction segmentation calls in the voting procedure greatly increases the judgement of wrong segmentation calls.

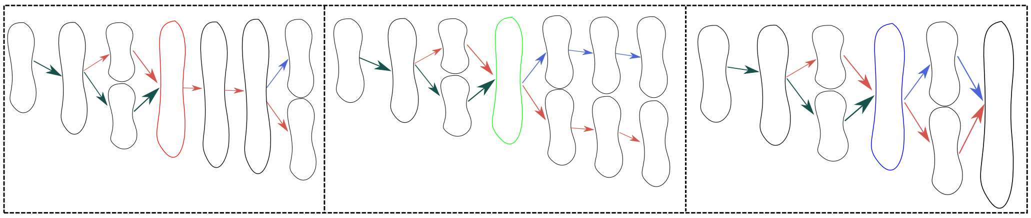

FIGURE A.7: An illustration of the voting mechanism. There are two states of wrong segmentation calls: oversegmentation and undersegmentation. The voting mechanism is used to determine which type of segmentation error is more likely to be the case considering all the evidence present in the time lapse frames, and then correct the errors accordingly. The voting score is determined by the difference in the number of descendants and the number of ancestors within the same generation. If the score equals 0, no action is taken until intervention from the operator. From left to right, the difference in the voting score is: (3–1) > 0, (1–4-n) < 0 and (1–1) = 0. Therefore the ancestor of the red segmentation is merged with its bottom cell. The green segmentation is split into two cells. For the blue segmentation, no correction is taken. The voting score is computed iteratively because the value will be changed after every correction process.

A.4.4 Image analysis performances and results

We propose to assess segmentation power and tracking accuracy though two types of image, phase-contrast and fluorescent. To this end we performed five datasets for each type of image to evaluate the performance of segmentation and tracking. Each dataset contains 101–280 images, each image contains 17–19 channels, and each channel contains 4–6 cells. Hence, one position contains at least 101x17x4 (8k) cells. The accuracy of the segmentation and tracking is computed on these datasets. We will also discuss the segmentation error detection rate by Random Forest and their correction ability by the voting mechanism. The software is written as a plugin of ImageJ.

Figure A.8 represents the Recall-Precision curve for different imaging conditions across increasing variations of the original image. Each plot corresponds to a Recall-Precision curve of our segmentation algorithm performed on the image transformed from the original image. For evaluating the robustness of our algorithms, we manually change the imaging condition by adding Gaussian noise and Gaussian blur. The segmentation process labels example cells as either positive or negative. The result can be represented in a structure as a confusion matrix that has four categories: True positives (TP) are cell segmentations correctly labelled as positive. True negatives (TN) refer to the wrong segmentations correctly labelled as negative. False positives (FP) correspond to the wrong cell segmentations incorrectly labelled as positive. Finally, the false negatives (FN) are the good cell segmentations labelled as negative.

	Positive	Negative
True	TP	TN
False	FP	FN

The recall and precision are obtained from the formula $Recall = \frac{TP}{TP+FN}$, $Precision = \frac{TP}{TP+FP}$. In our work, we notice that a very high recall rate was obtained even through very difficult imaging conditions. This means that the average missed cell segmentations rate is 0.07%. In addition, the resulting precision rate confirms the high accuracy rate of cell segmentation recognition. Compared with conventional methods, the results show that the proposed method has advantages of a high and robust ability to segment cells correctly, and can be used for the various imaging conditions. For the fluorescence image, the mean segmentation accuracy is 99.97% tested on five datasets with 361 images. The background illumination is often non-uniform, so we do not alter the imaging conditions manually.

Despite the fact that the segmentation process gives a highly correct accuracy rate of cell recognition, we still get some erroneous segmentation calls due to the noise, uneven illumination and cell aggregation, and therefore error identification is still necessary. This step is performed by Random Forest which is extracted from the Application Programming Interface (API) of a java machine learning package WEKA (Smith & Frank, 2016). The training feature vector is composed of four tracking descriptor and eight Fourier coefficients, as described in the section on error detection. In our work, we use 10% correct segmentations and 30% error segmentations as positive and negative examples to train Random Forest. We obtained a recall rate of 99.97% and a precision rate of 99.85%. This means our detection algorithm is able to determine the category of the extracted features in predicting the segmentation error. Cell segmentation error correction is implemented by the voting mechanism that allows the neighbouring segmentation to vote for the forthcoming action. In our work, we successfully corrected 71.5% of the segmentation errors.

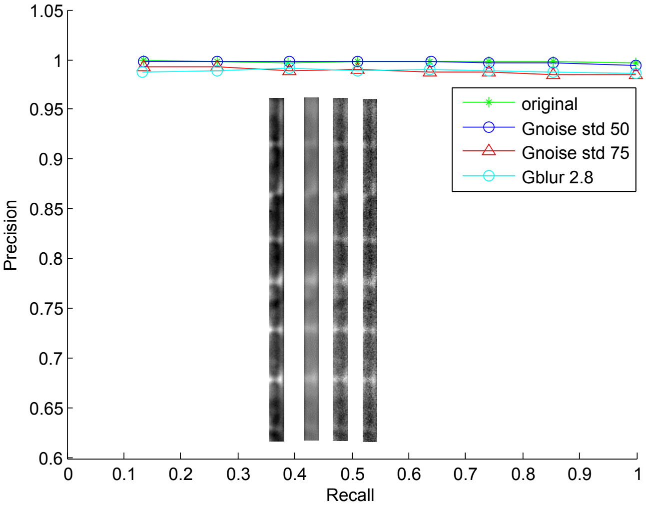

FIGURE A.8: In the graph above, the four Recall-Precision curves represent the performance of our segmentation algorithms. For evaluating algorithm robustness, the original image was transformed into images with a white Gaussian noise (standard deviation from 50 to 70) and into a blurred image with a Gaussian smoothing by scale 2.8. The used dataset contains 101 images with 17 channels. Therefore around 8k (101x17x5) cells were tested.

A.4.5 Image analysis summary

We have developed an approach for microfluidic bacterial automatic analysis of microscopy image sequences. This approach starts from preprocessing to the cell recognition. We use Hough Transform to detect the channel rotation and the internal distance between two channels. Based on the structure of microfluidic devices, a robust tensor structure-based segmentation method is used for the segmentation of phase-contrast images. We also introduce a statistical analysis, based on point spread function (PSF) to detect the seeds for cell segmentation of fluorescence images. Lineage information is then put into the learning system. Although a high precision rate of segmentation was obtained from the proposed method, there is still a need to detect the erroneous segmentation calls. Using the features of Fourier open curve coefficients, we successfully detected cell segmentation errors that were subsequently corrected using a new voting mechanism algorithm.

Appendix B

Preventing leachable contaminants in *E.* coli carbon starvation experiments

B.1 Introduction to the problem

B.1.1 Background

Bacteria is mostly well-known among the general audience for their small size, and abilities to metabolise and grow on common nutrients. However, exponential growth is hardly the mode of existence for bacteria. No matter culturable or not, pathogenic or commensal, most bacteria's life-histories are dominated by long periods of starvation and dispersal. Both the genomes and phenotypic traits of bacteria could not be understood purely in the context of growth. For an organism as well studied as *E. coli*, a large percentage of genes in its genome are not well characterised functionally and are only activated in non-growing conditions. In experimental evolution studies, *E. coli* mutants with dramatically different physiology arose in adaptation to continuous growth in chemostats (Maharjan *et al.*, 2006). These observations suggest that the wildtype *E. coli* genome and physiology are shaped by selective forces associated with starvation and dispersal. If one wishes to learn not only the basic biochemistry and molecular biology of bacterial growth, but also the subtlety of bacterial physiology, behaviour and evolution on its own terms, it is crucial to study the starvation an as important part of bacterial life-history.

Most of our contemporary knowledge of bacterial genetics and physiology came from exponential growth on high concentrations of nutrients. This bias, in my opinion, does not originate from intentional agenda on the part of the pioneers of microbiology, but reside in the practicalities and technologies of experimental investigations of microbial life. It is the adoption of agar plates and the invention of the chemostat that dominated the history of microbiology. Indeed, most bacterial species known to human are isolated because they can be grown on select media in the first place. With exponential growth, the researchers could control macroscopically the behaviour of the bacterial populations by changing the most abundant chemical species in the medium, and the steady state reached during the exponential phase related directly to the traits of the bacteria strain in question and the media composition. No such convenience is afforded to the studies of bacterial starvation. Most studies of bacteria starvation are done in either stationary phase batch cultures, or continuous cultures with very low dilution rates. In the case of stationary phase, no steady states are reached: both the medium and the population are shifting throughout the process (Zinser & Kolter, 2004). In both batch and continuous cultures, bacterial populations quickly become heterogenous (Maharjan *et al.*, 2006; Vulic & Kolter, 2001), further complicating any macroscopic analysis that only concern population averages.

The adoption of microfluidic technology in microbial laboratories offer us the tools to potentially overcome these problems. In particular, the mother machine (Wang *et al.*, 2010a) present us with an existing tool that can in principle avoid both issues mentioned above. The continuous flow of fresh medium through the microfluidic chip essentially maintains steady state environment for the cells inside the chip. In addition, since each cells are spatially restricted in their own individual compartments, individual cell physiology can be easily tracked and analysed by automated time-lapse microscopy and image analysis.

Despite these advantages, the application of microfluidic devices to the study of bacterial starvation presents its own problems. In the cases of batch cultures, no matter the chemical integrity and exact compositions of the media, growth arrests are natural occurring consequences of exponential growth, due to exhaustion of nutrients and accumulation of metabolic wastes. In mother machines, there are no such guaranteed growth arrests. By continuously refreshing the media, we subject a small number of cells to a large volume of media, thus making their metabolism and physiology much more sensitive to trace amounts of chemical in the media. It is thus imperative to maintain the chemical integrity of media when using mother machines to study bacterial physiology during starvation.

B.1.2 Leachable contaminants can be used by E. coli as carbon sources

The first indication that the chemical integrity of the culture media might be comprised, is that even carbon-source-free M9 minimal media can sustain a population of around 10^6 cells per ml, described in and Figure 2.3 and Section 2.2. In batch cultures, this background carrying capacity of the media may be overlooked because in most cases when carbon sources are added, cell density are hundreds to thousands times much higher.

In typical mother machine experiments, there are at maximum, 2×10^4 cells per channel. If we inject the input media at 2ml per hour, which is the usual rate for growth experiments, we estimated the effective population density to be much smaller than 10^6 per ml. This suggests that even without adding carbon sources, if we use the same M9 media used in Figure 2.3 as input media in mother machine experiments, we expect it to be able to achieve carbon-limited growth using whatever carbon sources present in the media. This is indeed the case after some testing in mother machine experiments, shown in the upper panels of Figure B.1, from (A1) to (A6).

B.1.3 Potential sources of carbon-supplying contaminants

Because M9 is a medium buffered by phosphate and is only consisted of inorganic salt, *E. coli* can only derive its carbon and energy sources from trace amounts of organics leached in the media during the media preparation process in our laboratory. It is a well reported yet often ignored fact that disposable plasticware leach a variety of contaminants into the solution that comes into contact with them. (McDonald *et al.*, 2008). Thus we suspected that it is these leachables that are metabolised by *E. coli* cells and used as carbon sources in experiments such as the in the upper panels (A1-A6) in Figure B.1.

One of the most well-known example of leaching plasticware is plasticiser leaching from

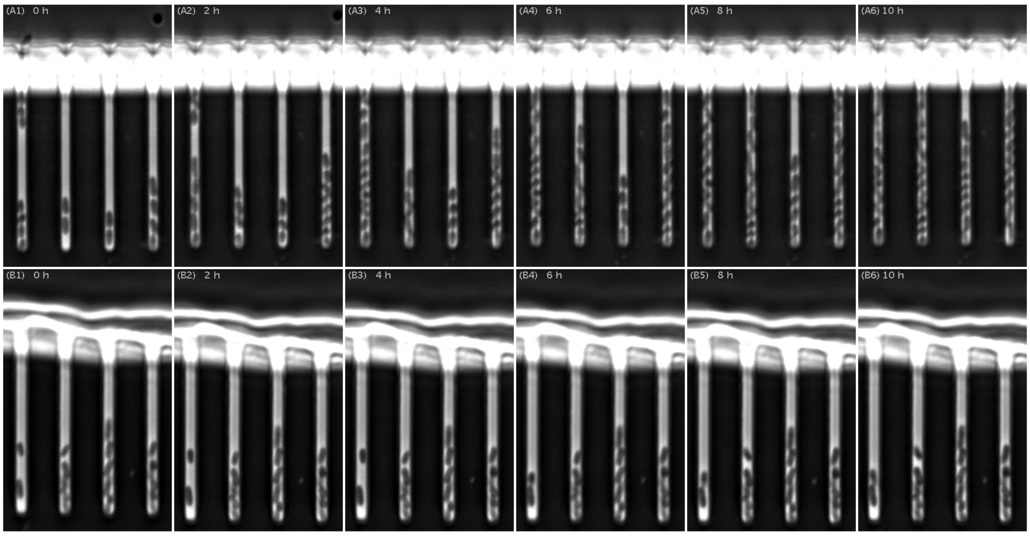

Figure B.1: Time-lapse phase-contrast images of *E. coli* cells in mother machines. (A1-A6) In M9 media without added carbon source prepared with common lab plasticware. (B1-B6) In M9 media without added carbon source prepared with improved protocol described in this appendix.

flexible disposable PVC (Polyvinyl chloride) tubings, which were used in as the media inlets and outlets to our microfluidic chip. Relatively large quantities of phthalate plasticisers such as those shown in Figure B.2 are used as additives in order to increase the flexibility of PVC polymers. Notice the non-polar part the phthalate molecules are consisted of relatively long alkoxy groups, which could serve as energy rich carbon sources when hydrolysed.

In fact, in order to conveniently prepare and store large volumes of sterile media, a variety of disposable plasticware are used in contemporary microbiology labs. Additives similar to phthalate plasticisers could leach from almost every disposable plastic components in the media processors. In fact, fatty acids and fatty alcohols esters are chosen as plasticisers for the physical properties that their long hydrocarbon chains bring to the composite material. Anything from the polypropylene disposable syringes used to infuse media, to the polystyrene disposable filter casings that are used to sterilised the media, even to polypropylene bottle caps that are used to store solutions, could leach contaminants that can potentially serve as carbon sources for our bacteria. Our goal is to identify the sources of contamination that serve as carbon sources in our bacteria starvation experiments, find chemically-inert substitutes and eliminate the leaching material from our protocols.

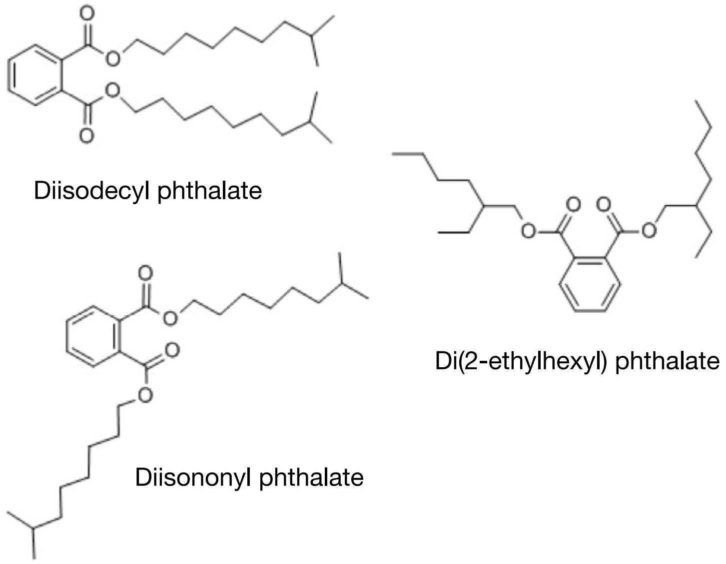

Figure B.2: Chemical structures of common phthalate plasticisers

B.2 Problem solving strategy

This section of the appendix aims to document my solution of maintaining chemical integrity in preparation and storage of sterile media and in microfluidic experiments. The hope is that it is generally applicable to other physiological experiments of bacterial starvation, not just to the particular work of bacterial aging in this thesis.

Our strategy of solving the leaching problem is essentially one by trial and error. One-by-one we replaced the components we suspected of leaching carbon-supplying contaminants, such as the PVC tubing mentioned in the previous section, with functionally equivalent alternatives made of chemically inert materials. Finding such alternatives are not as straight forward as it seems, because chemically inert material usually possess quite different mechanical, geometrical, thermal properties than the original leaching material, and are usually much more expensive and not disposable. The experimental protocols of media preparation, storage and microfluidic time-lapse experimental setup very often have to be modified in order to accommodate the new component.
After each component substitution, we tested the new protocol in mother machine starvation experiments similar to the one shown in Figure B.1. The main goal of such

testings is to see whether by substituting the component is question, whether we have hopefully completely eliminated sources of carbon-supplying contaminants, or at least the extent to which the contaminants has been reduced.

Protocols for setting up such experiments are documented in appendix A. After the cells are loaded and both the imaging and fluidic systems have stabilised, we first adjusted the infusion flow rate to the lower bound of flow rate, just enough counteract evaporation inside the chip, usually at $2\mu l$ per hour. We stay at this flow rate for 1 day and use the time-lapse images to estimate cellular growth rates. Then we raise the flow rate by 5-fold each day, until either the flow rate reach the maximum allowed by the fluidic system (usually 2 ml per hour for our system), or the estimate growth rates stop increasing with flow rates.

The estimated growth rates at different flow rates can inform the next step in multiple ways. Most importantly, the leaching problem could be considered operationally solved, if there is a reasonable large range of flow rates (4-fold) at which the cells could not grow. Secondly, the maximum growth rate reached at the maximum growth rate, i.e. the μ_{max}, could be used to estimate the quality of the best carbon-supplying contaminant left in the media, thus giving us clues about its source.

B.3 Identified leaching plasticware & replacements

We carried out the experimental strategy outlined above, and identified at least 6 plastic components that leach carbon-supplying contaminants. Figure B.3 illustrated some of these components and their replacements.

The PVC tubing is by far the worst offender: even at the lowest flow rate, the contaminant is enough to sustain continued growth at doubling time of about 2 hours. We know that a large portion of the tubing mass is consisted of phthalate plasticisers. In fact, we incubated PVC tubings filled with M9 buffer without added carbon sources for 3 days, and then inoculated it with bacteria. The concentration of plasticiser is so high such that the culture could reach visual turbidity. After continuously infusing M9 buffer for more than 1 week, the PVC tubings have leached so much plasticisers that the tubings have become more rigid and fragile.

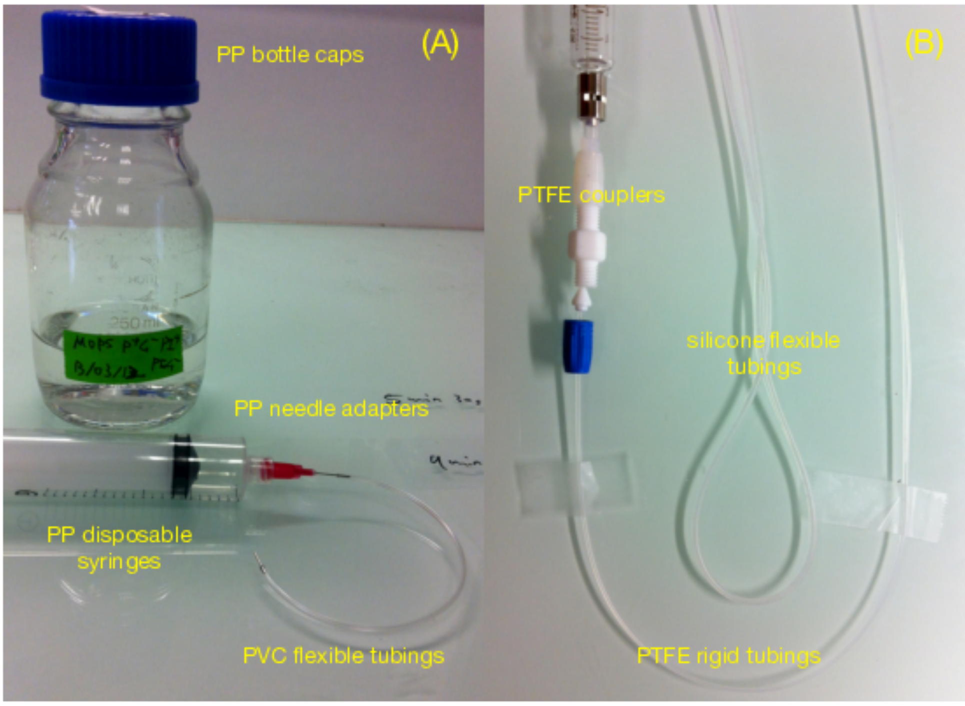

FIGURE B.3: Leaching plasticware and their replacements.

The fact that phthalates leaching from PVC plasticware is well known in the medical community, where PVC tubings were commonly used (Sathyanarayana, 2008). It was used for a reason in hospitals and our microfluidic systems: in these applications, tubing flexibilities are really desirable. In our case, whatever tubing that substituted PVC, besides needs to be chemically stable, has to 1) securely connect to the microfluidic chip through 23 gauge metal couplers; 2) be flexible enough to connect to the chip on the automatised microscope stage and inside the temperature-controlled chamber. Traditional Perfluoroalkoxy alkane (PFA) and Polytetrafluoroethylene (PTFE) tubings are too rigid for either requirements.

We tested serveral alternatives tubings, including those made of silicone, Polyethylene (PE), a non-leaching PVC formulation named Tygon® 2001 and semi-rigid thin-walled (0.56mm x 1.07mm) PTFE tubings. Silicone tubings tend to evaporate too much and limit the lower boundary of flow rates that can be used. We can not be sure whether plasticisers are used in the PE tubings we have. We found the best solution to be the thin-walled PTFE tubing. It is both more flexible and less expensive than Tygon® 2001.

After the PVC tubings are removed from the system, there are still contaminants in the media that can support carbon-limited growth, although at a much longer doubling time of about 4 hours. Furthermore, the flow rates at which cells are able to growth vary from experiments to experiments, indicating varying concentrations of contaminants. We found that the contaminant concentrations tend to be lower for newly prepared media. This leads us to suspect the bottles and syringes that are in contact with the media, which are mostly made of Polypropylene (PP) hard plastics are also leaching carbon sources into the media, although at much slower rates and lower concentrations than PVC tubings.

Glass solution bottles are used instead of disposable Falcon® centrifuge tubes made of PP. We also replaced the disposable syringes made of PP barrels and rubber plungers used to infuse media into the microfluidic system, with chromatography-grade syringes with glass barrelled and PTFE-coated plungers. The connections between syringes and tubings are switched to either metal needles with metal or fluoropolymer hubs, or PTFE couplers in case of rigid tubings. Even the usual PP bottle caps for glass solution bottles, while not disposable and are in contact with the media through repeated vaporisation and condensation, are experimentally confirmed to leach carbon sources. We replaced them as well with a premium bottle cap made of fluoropolymer.

Table B.1 lists all these components that are confirmed by mother machine experiments to be leaching carbon-supplying contaminants, and their replacements we settled on. There may be less expensive non-leaching plastic components such as those made of high density Polyethylene (HDPE), but we decided to opt for the safety options to save time.

Another experimentally confirmed source of contaminants is from autoclaving solutions. In autoclaves, volatile organic compounds might migrate between the water tank and the autoclaved solutions through steam. Since the water reservoirs in autoclaves are often changed after sterilising complex media rich in nutrients, they are a major source of carbon-supplying contaminants. We made sure to sterilise all our solutions with filter sterilisation using glass vessels and hydrophilic PTFE membrane filters. We do use dry autoclaving to sterilise our equipments. We also clean and refresh the water tank of our autoclave before we use it.

Table B.1: List of plasticware that leach carbon-supplying contaminants and their non-leaching replacements.

Leaching component	Material	Function	Replacement	Additional notes
Flexible tubing	PVC	Fluidic system	PTFE tubing, I.D. X O.D. = 0.56 x 1.07mm	Tubings have to thin-walled and seal well with metal couplers/needles
Disposable sterile syringes	PP barrel	Fluidic system	chromatography grade glass/PTFE syringes	Non-disposable, have to be properly cleaned/sterilised
Disposable sterile syringes	PP barrel	Cell loading	Manual glass syringes	Non-disposable, have to be properly cleaned/sterilised
Disposable Luer Stubs	PP hub	Fluidic system	Kel-F® or metal hub needles, 23 gauge	Non-disposable, have to be properly cleaned/sterilised
Disposable Falcon® centrifuge tubes	PP	Cell loading	Pyrex® glass centrifuge tubes	
Disposable Falcon® centrifuge tubes	PP	Media storage	Glass solution bottles	
Media storage bottle caps	PP	Media storage	Duran® TpCh260 bottle caps and pouring ring with PTFE coated silicone seal	
Sterile vacuum filter units	PS flask	Media prep.	All-Glass filter holder assembly with fritted base	Non-disposable, have to be properly cleaned/sterilised
Sterile vacuum filter units	PES membrane	Media prep.	0.2µm hydrophilic PTFE disc membrane filter	Needs to be sterilised before use

If the component is consisted of parts made of different materials, the material column describe the material most likely to be the culprit in leaching. PVC = Polyvinyl chloride, PTFE = Polytetrafluoroethylene, PP = Polypropylene, TpCh260 TZ is a material similar to Perfluoroalkoxy alkane (PFA), Kel-F® is the trade name for Polychlorotrifluoroethylene (PCTFE), PS = Polystyrene, PES = Polyethersulfone

In addition to those documented in table B.1, we also made other deliberate equipment choices to avoid leachable contaminants in media preparation, storage and cell culture. They are not listed in table B.1 because we have not experimentally confirmed that they are necessary, but we made them out of precautions. In general, we chose equipments whose wetted surfaces are made of glass, metal or fluoropolymer whenever possible. If not, we chose HDPE options, and decreased the time media is in contact with the plastic as much as possible.

Bibliography

Aguet F, Antonescu CN, Mettlen M, Schmid SL, Danuser G (2013) Advances in Analysis of Low Signal-to-Noise Images Link Dynamin and AP2 to the Functions of an Endocytic Checkpoint. *Developmental Cell* **26**: 279–291

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**: msb4100050–E1–msb4100050–E11

Babic A, Lindner AB, Vulic M, Stewart EJ, Radman M (2008) Direct visualization of horizontal gene transfer. *Science* **319**: 1533–6

Baudisch A (2005) Hamilton's indicators of the force of selection. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 8263–8268

Bigün J, Granlund GH (1987) Optimal Orientation Detection of Linear Symmetry. *Proceedings of the IEEE First International Conference on Computer Vision* **6**: Pages 433–438

Bougdour A, Cunning C, Baptiste PJ, Elliott T, Gottesman S (2008) Multiple pathways for regulation σS (RpoS) stability in Escherichia coli via the action of multiple anti-adaptors. *Molecular Microbiology* **68**: 298–313

Bouwhuis S, Charmantier A, Verhulst S, Sheldon BC (2010) Individual variation in rates of senescence: natal origin effects and disposable soma in a wild bird population. *The Journal of animal ecology* **79**: 1251–1261

Bryant MJ, Reznick D (2004) Comparative studies of senescence in natural populations of guppies. *The American naturalist* **163**: 55–68

Burtner CR, Murakami CJ, Kennedy BK, Kaeberlein M (2009) A molecular mechanism of chronological aging in yeast. *Cell cycle (Georgetown, Tex)* **8**: 1256–1270

Cai L, Dalal CK, Elowitz MB (2008) Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature* **455**: 485–90

Campos M, Surovtsev IV, Kato S, Paintdakhi A, Beltran B, Ebmeier SE, Jacobs-Wagner C (2014) A constant size extension drives bacterial cell size homeostasis. *Cell* **159**: 1433–46

Canny J (1986) A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell* **8**: 679–98

Chartrand R (2011) Numerical differentiation of noisy, nonsmooth data. *ISRN Applied Mathematics* **2011**: 1–11

Ericsson M, Hanstorp D, Hagberg P, Enger J, Nyström T (2000) Sorting out bacterial viability with optical tweezers. *Journal of bacteriology* **182**: 5551–5555

Ferenci T, Galbiati H, Betteridge T, Phan K, Spira B (2011) The constancy of global regulation across a species: the concentrations of ppGpp and RpoS are strain-specific in Escherichia coli. *BMC Microbiology* **11**: 62+

Finkel SE (2006) Long-term survival during stationary phase: evolution and the GASP phenotype. *Nature Reviews Microbiology* **4**: 113–120

Fontaine F, Stewart EJ, Lindner AB, Taddei F (2008) Mutations in two global regulators lower individual mortality in Escherichia coli. *Molecular microbiology* **67**: 2–14

Gavrilov LA, Gavrilova NS (2001) The reliability theory of aging and longevity. *Journal of theoretical biology* **213**: 527–545

Gavrilov LA, Gavrilova NS (2002) Evolutionary theories of aging and longevity. *TheScientificWorldJournal* **2**: 339–356

Gefen O, Fridman O, Ronin I, Balaban NQ (2014) Direct observation of single stationary-phase bacteria reveals a surprisingly long period of constant protein production activity. *Proceedings of the National Academy of Sciences* **111**: 556–561

Gems D, Partridge L (2013) Genetics of Longevity in Model Organisms: Debates and Paradigm Shifts. *Annual Review of Physiology* **75**: 621–644

Golding I, Paulsson J, Zawilski SM, Cox EC (2005) Real-time kinetics of gene activity in individual bacteria. *Cell* **123**: 1025–36

Gustafsson L, Part T (1990) Acceleration of senescence in the collared flycatcher Ficedula albicollis by reproductive costs. *Nature* **347**: 279–281

Hayden EC (2015) Anti-ageing pill pushed as bona fide drug. *Nature* **522**: 265–266

Hayward AD, Mar KU, Lahdenperä M, Lummaa V (2014) Early reproductive investment, senescence and lifetime reproductive success in female Asian elephants. *J Evol Biol* **27**: 772–783

Hilfinger A, Paulsson J (2015) Systems biology: Defiant daughters and coordinated cousins. *Nature* **519**: 422–3

Ho TK (1995) Random decision forests. In *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, vol. 1. pp. 278–282 vol.1

Hough PVC (1962) Method and means for recognizing complex patterns

Izard J, Gomez Balderas CD, Ropers D, Lacour S, Song X, Yang Y, Lindner AB, Geiselmann J, de Jong H (2015) A synthetic growth switch based on controlled expression of RNA polymerase. *Mol Syst Biol* **11**: 840

Jiang H, Fels S, Little JJ (2007) A Linear Programming Approach for Multiple Object Tracking. In *2007 IEEE Conference on Computer Vision and Pattern Recognition.* pp. 1–8

Jug F, Pietzsch T, Kainmüller D, Funke J, Kaiser M, van Nimwegen E, Rother C, Myers G (2014) *Optimal Joint Segmentation and Tracking of Escherichia Coli in the Mother Machine.* Cham: Springer International Publishing, pp. 25–36

Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R (1993) A C. elegans mutant that lives twice as long as wild type. *Nature* **366**: 461–464

King T, Ishihama A, Kori A, Ferenci T (2004) A regulatory trade-off as a source of strain variation in the species Escherichia coli. *J Bacteriol* **186**: 5614–5620

Kirkwood TBL (1977) Evolution of ageing. *Nature* **270**: 301–304

Kuroda A, Nomura K, Ohtomo R, Kato J, Ikeda T, Takiguchi N, Ohtake H, Kornberg A (2001) Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in E. coli. *Science* **293**: 705–708

Le Bourg E (2001) A mini-review of the evolutionary theories of aging. *Demographic Research* **4**: 1–28

Lindner AB, Madden R, Demarez A, Stewart EJ, Taddei F (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proceedings of the National Academy of Sciences* **105**: 3076–3081

Locke JC, Young JW, Fontes M, Hernandez Jimenez MJ, Elowitz MB (2011) Stochastic pulse regulation in bacterial stress response. *Science* **334**: 366–9

Luria SE, Delbrück M (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**: 491–511

Maharjan R, Seeto S, Notley-McRobb L, Ferenci T (2006) Clonal adaptive radiation in a constant environment. *Science* **313**: 514–517

Mattison JA, Roth GS, Beasley TM, Tilmont EM, Handy AM, Herbert RL, Longo DL, Allison DB, Young JE, Bryant M, Barnard D, Ward WF, Qi W, Ingram DK, de Cabo R (2012) Impact of caloric restriction on health and survival in rhesus monkeys from the NIA study. *Nature* **489**: 318–321

McDonald GR, Hudson AL, Dunn SM, You H, Baker GB, Whittal RM, Martin JW, Jha A, Edmondson DE, Holt A (2008) Bioactive contaminants leach from disposable laboratory plasticware. *Science* **322**: 917

Monod J (1958) *Recherches sur la croissance des cultures bactériennes, deuxième édition, thèse de 1942.* Paris V: Hermann, 6, Rue de la Sorbonne

Nagarajan S, Kruckeberg AL, Schmidt KH, Kroll E, Hamilton M, McInnerney K, Summers R, Taylor T, Rosenzweig F (2014) Uncoupling reproduction from metabolism extends chronological lifespan in yeast. *Proc Natl Acad Sci U S A* **111**: E1538–47

Ni M, Decrulle AL, Fontaine F, Demarez A, Taddei F, Lindner AB (2012) Pre-disposition and epigenetics govern variation in bacterial survival upon stress. *PLoS Genet* **8**: e1003148

Norman TM, Lord ND, Paulsson J, Losick R (2013) Memory and modularity in cell-fate decision making. *Nature* **503**: 481–6

Nussey DH, Coulson T, Festa-Bianchet M, Gaillard JM (2008) Measuring senescence in wild animal populations: towards a longitudinal approach. *Functional Ecology* **22**: 393–406

Nyström T (2003) Conditional senescence in bacteria: death of the immortals. *Molecular microbiology* **48**: 17–23

Nyström T (2004) Growth versus maintenance: a trade-off dictated by RNA polymerase availability and sigma factor competition? *Molecular Microbiology* **54**: 855–862

Parry BR, Surovtsev IV, Cabeen MT, O'Hern CS, Dufresne ER, Jacobs-Wagner C (2014) The bacterial cytoplasm has glass-like properties and is fluidized by metabolic activity. *Cell* **156**: 183–94

Peterson CN, Levchenko I, Rabinowitz JD, Baker TA, Silhavy TJ (2012) RpoS proteolysis is controlled directly by ATP levels in Escherichia coli. *Genes & development* **26**: 548–553

Primet M, Demarez A, Taddei F, Lindner AB, Moisan L (2008) Tracking of cells in a sequence of images using a low-dimension image representation. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. pp. 995–998

Rowat AC, Bird JC, Agresti JJ, Rando OJ, Weitz DA (2009) Tracking lineages of single cells in lines using a microfluidic device. *Proc Natl Acad Sci U S A* **106**: 18149–54

Sathyanarayana S (2008) Phthalates and Children's Health. *Current Problems in Pediatric and Adolescent Health Care* **38**: 34–49

Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T (2010) Interdependence of cell growth and gene expression: origins and consequences. *Science (New York, NY)* **330**: 1099–1102

Shanley DP, Kirkwood TB (2000) Calorie restriction and aging: a life-history analysis. *Evolution; international journal of organic evolution* **54**: 740–750

Shermer M (2012) Philosophy: What we don't know. *Nature* **484**: 446–447

Smith TC, Frank E (2016) *Introducing Machine Learning Concepts with WEKA*. New York, NY: Springer New York, pp. 353–378

Spira B, de Almeida Toledo R, Maharjan R, Ferenci T (2011) The uncertain consequences of transferring bacterial strains between laboratories - rpoS instability as an example. *BMC Microbiology* **11**: 248+

Stearns SC (1989) The Evolutionary Significance of Phenotypic Plasticity. *BioScience* **39**: 436–445

Stewart EJ, Madden R, Paul G, Taddei F (2005) Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS biology* **3**: e45+

Strehler BL, Mildvan AS (1960) General Theory of Mortality and Aging. *Science* **132**: 14–21

Suel GM, Garcia-Ojalvo J, Liberman LM, Elowitz MB (2006) An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* **440**: 545–50

Taheri-Araghi S, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, Vergassola M, Jun S (2015) Cell-size control and homeostasis in bacteria. *Curr Biol* **25**: 385–91

Uesaka Y (1984) A new fourier descriptor applicable to open curves. *Electronics and Communications in Japan (Part I: Communications)* **67**: 1–10

Vulic M, Kolter R (2001) Evolutionary cheating in Escherichia coli stationary phase cultures. *Genetics* **158**: 519–526

Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, Jun S (2010a) Robust growth of Escherichia coli. *Current Biology* **20**: 1099–1103

Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, Jun S (2010b) Robust growth of Escherichia coli. *Curr Biol* **20**: 1099–103

Weitz JS, Fraser HB (2001) Explaining mortality rate plateaus. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 15383–15386

Williams GC (1957) Pleiotropy, Natural Selection, and the Evolution of Senescence. *Evolution* **11**: 398–411

Zhang A, Altuvia S, Tiwari A, Argaman L, Hengge-Aronis R, Storz G (1998) The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *The EMBO journal* **17**: 6061–6068

Zhang Y, Ge H, Qian H (2012) One-dimensional birth-death process and Delbrück-Gillespie theory of mesoscopic nonlinear chemical reactions. *Studies in Applied Mathematics* **129**: 328–345

Zinser ER, Kolter R (2004) Escherichia coli evolution during stationary phase. *Research in microbiology* **155**: 328–336