FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Programa de Doutoramento em Engenharia Electrotécnica e de Computadores (3º Ciclo – Modelo de Bolonha)	**PhD Program in Electrical and Computer Engineering** (3rd Cycle – Bologna model)
Proposta de Plano de Tese	**Thesis Plan Proposal**

Generator of Simulated Time-Lapsed Microscopy Images of Bacterial Cells

Leonardo Pedro Donas-Boto de Vilhena Martins

Nº 43582

Supervisor:
Professor Doutor José Manuel Matos Ribeiro da Fonseca

Co-supervisor:
Professor Doutor André Sanches Ribeiro

Monte da Caparica, Março de 2015

List of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This section makes an introduction to the problem background, in which this research work aims to be developed, specifically a brief introduction to the problem and its open questions, and the main motivation behind this research work.

1.1. Introduction to the Problem and Main Motivation

Recent advances in microscopy imaging technology have prompted the detection of single molecules at the single live-cell level, due to biochemical techniques that are able to highlight the targets responsible for gene expression, such as the Deoxyribonucleic acid (DNA), Ribonucleic acid (RNA) and proteins but also due to the capability of acquiring multidimensional images with better quality and resolution. Specifically, the monitoring of the spatial and temporal distribution of protein aggregates and how they are segregated in prokaryotic cells (especially in *Escherichia coli*) has been important in the establishment of connections between cellular aging and these events.

The observation of the spatial distribution of these molecules at the single cell level, and studying the kinetics of segregation to the cell pole and partitioning during division can establish a correlation between those events and cellular aging

These new experimental studies heavily based on image analysis protocols have also led to the fusion of various multimodal microscopy images, allowing the integration of morphological and functional information.

All these advances have been accompanied by the development of state-of-the-art image processing techniques (such as image segmentation, image registration) and tools for statistical analysis have culminated in the increase of accuracy in the quantification of single-molecules at the single-cell level.

Most of the computational tools developed to analyse these microscopy images are still validated by using benchmark data of manually annotated images. In high-throughput experiments (with large amounts of data), manual validation is a very time-consuming and unfeasible task, which prompted the development of artificial image generators to create new “gold-standard” images.

Such artificial image generators need to be as close as possible of images acquired in the laboratory, so they should be based on mathematical models of the cell biophysical behaviour and empirical information acquired from experiments.

Looking at prokaryotic cells (and specifically to bacteria such as *Escherichia coli*) these biophysical models need to reproduce the cellular spatial and temporal organization by modelling the cell morphology (shape, size and spatial arrangement), cell growth, division, motility and internal functions and structure organization.

The main motivation for this work is to create a Generator of Simulated Time-Lapsed Microscopy Images that is capable of simulating the spatial and temporal organization of bacteria cells, namely *Escherichia coli*, in order to create gold-standard images to validate the computational tools that have been previously developed. This Generator can also be used to create a null-model that can be used to study the effects of removing or adding features to the model (e.g. to study the effects of the nucleoid by removing it from the cells) or sampling some parameters (e.g. evaluate the effects of adjusting the division rates to unrealistic values).

1.2. Open Questions in the Area

The development of Artificial Image Generators to create reliable gold-standard benchmarks that can be used to validate image processing tools is one of the emerging fields in Biomedical Engineering. Specifically looking at microscopy images, these Generators have focused on the simulation of the morphological features of the cellular biophysical models.

The morphological information can be enough to create multidimensional images, but is not enough to simulate time-lapsed functional images, where important time-dependent processes are present such as cell growth, cell division and cell movement. The main questions in this specific topic are: 'which models should be used?' and 'which processes are enough to create a realistic simulation of the cell spatial and temporal organization?'.

One of the open topics in the area is the development of a Generator of functional and morphological time-lapsed simulated images. Mathematical and empirical realistic biophysical models need to be implemented to create simulated images as close as possible of real images acquired in the Laboratory. Such a Generator should be made first on simple organisms such as bacteria and then could be fittingly adapted to more complex organisms.

Other open questions are related to the simulation of temporal and spatial modelling of bacterial response to external factors/stress conditions, such as temperature (heat-shock, cold-shock) and pH stress.

When a fully operational Generator of Time-Lapsed Microscopy Images is developed, new questions related to the creation of null models will arise and how they will be able to shed light about some biological processes.

Chapter 2

Research Question

This section makes an introduction to the main research question that will be addressed in this research work alongside its corresponding hypothesis and planned approach. Secondary research questions that arise from the main topic are also presented here together with a list of the topics that will be addressed.

2.1. Main Research Question

In the previous Section, the problem behind this research work was introduced alongside the emerging challenges in the area that are still open research questions. The main challenge emerges from the necessity of creating a benchmark ("gold-standard") of microscopy images for validation of newly developed image processing tools, as nowadays most benchmark data is produced manually, which is an unfeasible task for high-throughput experiments. A proper system for creating such a benchmark is to use simulated data, using realistic mathematical and empirical cell models, which need to be thoroughly studied and implemented into the simulator.

These models should be able to reproduce time-lapsed experiments by simulating time-dependent processes such as cell growth, cell division, and cell movement. In the initial simulation framework we will concentrate our attention to bacterial cells, more specifically using *E. coli* cells models. Simulation of temporal and spatial modelling of external factors/stress conditions can also be done to produce even more realistic results. Simulation of different acquisition systems should also be done in order to generate different the different features of morphological and functional microscopy images.

After validation of the image generation tool, we can begin the validation of image processing tools and expand the image generation to other bacterial models or to simple cell organisms such as yeast. Other studies could be made by developing null-models that can assist in research about biological processes, such as cellular aging.

From the above information, the main research question is subsequently written:

How to design a Generator of Microscopy Images that is capable of simulating bacterial time-lapsed experiments by reproducing realistic morphological and functional microscopy images?

In addition, four more detailed research questions are proposed with the aim of better solidifying the main research question, three of them related to the parameters that need to be simulated and one related to the applications that can be approached using the developed Generator.

The first question is related to the biological processes:

Which biological processes and environmental conditions are necessary and sufficient to create a realistic simulation of the cell spatial and temporal organization?

The second question is related to the image acquisition parameters:

Which image acquisition parameters are necessary and sufficient to simulate the realistic characteristics of each type of microscopy methodologies will be used?

The third question is related to the methodologies that will be used to validate the developed image generation tool:

Which methodologies should be used to validate the image generation tool?

Finally it is important to acknowledge other applications/topics that could also be interested in the cell modelling simulation:

What applications other than validation of image processing tools and the creation of time-lapsed microscopy image benchmarks can benefit from the simulation of bacterial time-lapsed experiments?

2.2. Hypothesis and Approach

The main hypothesis adopted for this work is:

An Artificial Image Generator capable of replicating realistic bacterial time-lapsed experiments can be developed if the produced morphological and functional images can emulate the characteristics of the images acquired in the laboratory, by reproducing the spatial and temporal cell morphological and functional features (implementing the existing mathematical or empirical models from the literature) and by simulating different image acquisitions systems and environmental conditions.

To answer the secondary research questions, we need to specify the biological processes and environmental conditions that need to be included in the image generator to recreate realistic experiments. To do this, we need to search in the literature for the existing mathematical and empirical models of cell modelling. These include the studies of the temporal and spatial features of bacterial growth and division (which is then linked to its morphological features of cell size and shape) and how this is linked to gene expression.

It is also important to study how these processes and cell motility are connected to the spatial arrangement into clusters. If some of these functions and connections cannot be described mathematically, it is possible to use machine learning techniques to reproduce the empirical data. In terms of external environmental conditions, we can consider the bacterial response to external factors such as temperature (heat-shock, cold-shock), pH stress, oxidative stress, nutritional stress or even exposure to antibiotics.

The image acquisition parameters that need to be simulated will depend on the illumination, contrast and florescence methods that will be considered. Simulation of even/uneven illumination, simulation of the primary sources of noise: CCD Read Noise, noise associated with dark charge electrons, Poisson noise (photon shot), noise due to electronic interference and simulation of other sources of image aberrations.

To test and validate this research hypothesis we will approach groups of microscopy and biotechnology experts to make a qualitative analysis of the generated simulated images, compared to the ones acquired in the laboratory. Quantitative analysis will be done by direct comparison with real *E. coli* images acquired in various image acquisitions systems and various environmental conditions. Then it is possible to compare the simulated distributions of the model parameters indicators such as cell sizes and shapes (distributions of bacterial spatial organization), motility velocity, division and growth rates (distributions of bacterial temporal organization), and also the production and localization of subcellular structures (fluorescent proteins, nucleoid, etc).

The main objective of developing this image generator is to create time-lapsed microscopy image benchmarks that can be used to validate image processing tools. There can be other applications to the image generator such as creating null-model that can be used to investigate how the removal or the insertion of features can affect the bacterial behaviour (e.g. to study the effects of the nucleoid by removing it from the cells or changing the bacterial size distribution) or sampling some parameters (e.g. evaluate the effects of adjusting the growth rates to unrealistic values).

Chapter 3

Literature Review

The following section gives a comprehensive literature review of the areas related to the planned research work. These areas show the relevance of microscopy images to biological studies, the development of tools for simulation of synthetic microscopy images and the visualization of microscopy images and finally the development of tools for cell segmentation and cell tracking in microscopy images. The spatial and temporal modelling of bacteria organization is also discussed in detail, in various subsections and finally it is also acknowledged how all the mentioned topics are tied into the main focus of the scientific project (Project SADAC – Study of the kinetics of asymmetric disposal of aggregates in cell division and its correlation to functional aging from in vivo measurements, one event at a time) where this study is also integrated.

3.1. Microscopy Imaging

There has been an impressive progress over the recent years in the microscopy technology, resulting in multidimensional images with better quality and resolution. Associated with these progresses, the fusion of data coming from different microscopy techniques led to the development of several computational approaches in the last decade to deal and analyse image-based studies in Cell and Molecular Biology. Microscopy images and particularly sets from time-lapsed series can contain information about the cell dynamics, subcellular constituent distribution, such as the cell membrane, the cytoskeleton, genetic material (DNA and RNA) and various organelles (Danuser, 2011; Eils & Athale, 2003; Sung & McNally, 2011).

Three major improvements have been developed in parallel and driven the breakthroughs in the production of high quality microscopy images, which have played an important role in cellular studies (Danuser, 2011).

The first advancement, allowed the tracking of the activity of a great diversity of molecules using bright and genetically encoded fluorescent probes inside the cell. The second improvement was based on the optimization of the optical sensors and the usage of hardware controlled by feedback which permitted an efficient acquisition of large, high-quality microscopy image datasets. Third, an accelerated progress in the electronic detection technology enable the generation of high sensitivity datasets of microscopy images. In the latest years, single-molecule detection at the single-cell level grew into a conventional technique in the microbiology laboratories (Danuser, 2011).

Time-lapsed fluorescence microscopy imaging is also being used in live single bacterial cells to study the *in vivo* activity of transcription and translation, and also the protein interactions

using the previously described fluorescent probes (Meyer & Dworkin, 2007; Sung & McNally, 2011). These techniques have been used to study genetic circuits such as the Toggle Switch (Gardner, Cantor, & Collins, 2000) and the Repressilator (M B Elowitz & Leibler, 2000).

A revolutionary technique, capable of detecting and tracking single RNA molecules in *Escherichia coli* (*E. coli*) by fusing the RNA bacteriophage MS2 coat protein with GFP (Fusco et al., 2003; Golding, Paulsson, Zawilski, & Cox, 2005) have also been used recently by the Laboratory of Biosystem Dynamics (LBD) from Tampere University of Technology to produce time-lapsed microscopy images of *E. coli* cells. The RNA-MS2-GFP complexes were tracked to study to study the activity of the lar promoter (Kandhavelu et al., 2011), the activity of the arabinose promoter (Mäkelä et al., 2013) in *E. coli* and to study the partitioning of RNA (Lloyd-Price, Lehtivaara, et al., 2012) and proteins (Lloyd-Price, Häkkinen, et al., 2012) in cell division.

The main challenges in live-cell imaging can be divided between occurrences during the image acquisition and the post-acquisition processing (Coutu & Schroeder, 2013). For the first part, it is mostly related to the microscope components (e.g. shutter, lens, camera, stage) and can be solved by using better components, such as LED illumination, increasing the speed of the stage, using better filters and cameras. Improving all these solutions will make the system more expensive, so there is a necessity to compromise (Coutu & Schroeder, 2013).

The post processing limitations start with the data storage and archiving (see Figure 1-A for a detailed description of the limitations and solutions), which can be solved by using archiving software, having dedicated databases that can be easily accessed (Coutu & Schroeder, 2013). Using the safely stored data, there still needs be a correct cell tracking analysis in order to make accurate signal quantifications (figure 1-A), sometimes requiring image correction, image registration and other image processing techniques (Coutu & Schroeder, 2013).

Figure 1-B shows the typical workflow in a live-cell imaging setup, focusing on the image processing techniques, such as cell segmentation of external wall features and segmentation of internal cell structures, which with the advances in live-cell imaging can be at the single molecule level. Figure 1-A also shows intra-model and inter-model registration techniques that allow the fusion of multimodal microscopy images or the volume rendering of multi-dimensional data.

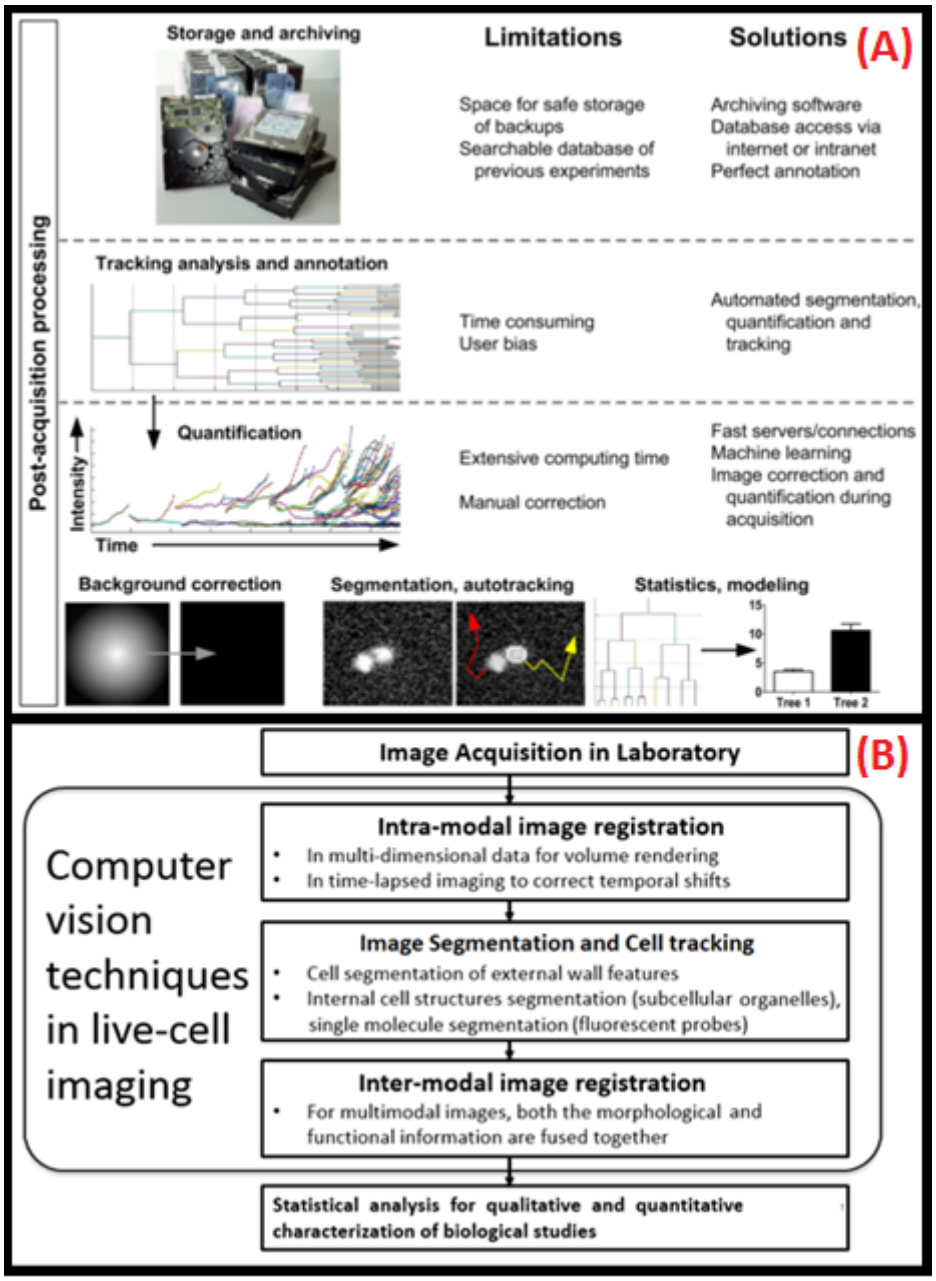

Figure 1 - (A) Limitations and challenges of the post-acquisition processing in Live-cell imaging Adapted from (Coutu & Schroeder, 2013). (B) A typical workflow in live-cell imaging, focusing on computer vision techniques related to the planned research work.

Both of these topics are included in the main trends in image microscopy, which will be presented in the next sub-sections with a detailed description and include the multimodality and multidimensionality in microscopy, microscopy image processing and statistical analysis and the simulation of microscopy images (Bonnet, 2004).

3.1.1. Multimodal and Multidimensional Microscopy

The establishment of novel biological studies that depended on the detection of fluorescent aggregates in live cells led to the development of data fusion techniques coming from different

microscopy techniques. has become a necessity in order to integrate and correlate functional (coming from fluorescent methods of microscopy) and morphological information (coming from illumination and contrast methods of microscopy), which can be combined to provide new information about biological processes (Stephens & Allan, 2003).

For live cell imaging, these microscopy modes include the use of bright-field and dark-field imaging, phase contrast, differential interference contrast, fluorescence microscopy, total internal reflection fluorescence microscopy, single and multiple photon excitation and even a multitude of super-resolution microscopy techniques, such as the stimulated emission depletion and scanning near-field optical microscopy (Bonnet, 2000, 2004; Glasbey, Buildings, Eh, & Martin, 1996; Stephens & Allan, 2003). Information integration based on the image fusion of multimodality microscopy images for the study of co-localization of internal cellular structures became a common strategy in many biotechnology studies (Bolte & Cordelières, 2006).

One of the first biotechnology applications of microscopy image fusion was the study of double labelled DNA via the fusion of dual colour fluorescence (specifically the red and green filter components) from a three-dimensional confocal microscopy in order to study the temporal and spatial organization of DNA inside the interphase nuclei of eukaryotic cells (E. M. M. Manders, Verbeek, & A, 1993; E. M. Manders, Stap, Brakenhoff, van Driel, & Aten, 1992).

Multimodal image integration started with the fusion of fluorescence and phase contrast images, which was used in studies of the partitioning of F-plasmids molecules (see Figure 2-A) during the cell division of *Escherichia coli* (*E. coli*) cells (Niki & Hiraga, 1997) and the localization of DNA segments on the chromosome of *E. coli* cells (Niki, Yamaichi, & Hiraga, 2000). Other studies also used super imposed phase contrast images with fluorescence microscopy (see Figure 2-B) to investigate the stimulation of the proliferation and differentiation of endothelial progenitor cells by the erythropoietin darbepoetin alfa (Bahlmann et al., 2003). Similarly in another study, researchers explored the kinetic dynamics of the genetic circuit responsible for the utilization of lactose in *E. coli*, green fluorescence proteins were used to fuse fluorescence microscopy with inverted phase contrast images (see Figure 2-C) of the cells (Ozbudak, Thattai, Lim, Shraiman, & Van Oudenaarden, 2004).

The kinetic dynamics of protein production in *E. coli* cells at the single-molecule and single-cell level by fusing differential interference contrast and fluorescence microscopy images and using yellow fluorescence proteins (Yu, Xiao, Ren, Lao, & Xie, 2006) and to study the protein and mRNA copy numbers in *E. coli* (Taniguchi et al., 2010) by fusing fluorescence and phase contrast microscopy images (see Figure 2-F) .

To study the differential protein expression in *Colletotrichum acutatum* and its impacts in the pathogenicity of the strawberry, one group (Brown et al., 2008) used the superposition of differential interference contrast and fluorescence microscopy (using green fluorescent probes and Red Nile staining of lipid bodies) images (see Figure 2-D). Finally another group overlaid phase contrast and fluorescence microscopy (using green and cyan fluorescent proteins) in order to study the chromosome segregation in the bacterium *Caulobacter crescentus* by using a partitioning apparatus, similar to the existing spindles in eukaryotes (Ptacin et al., 2010).

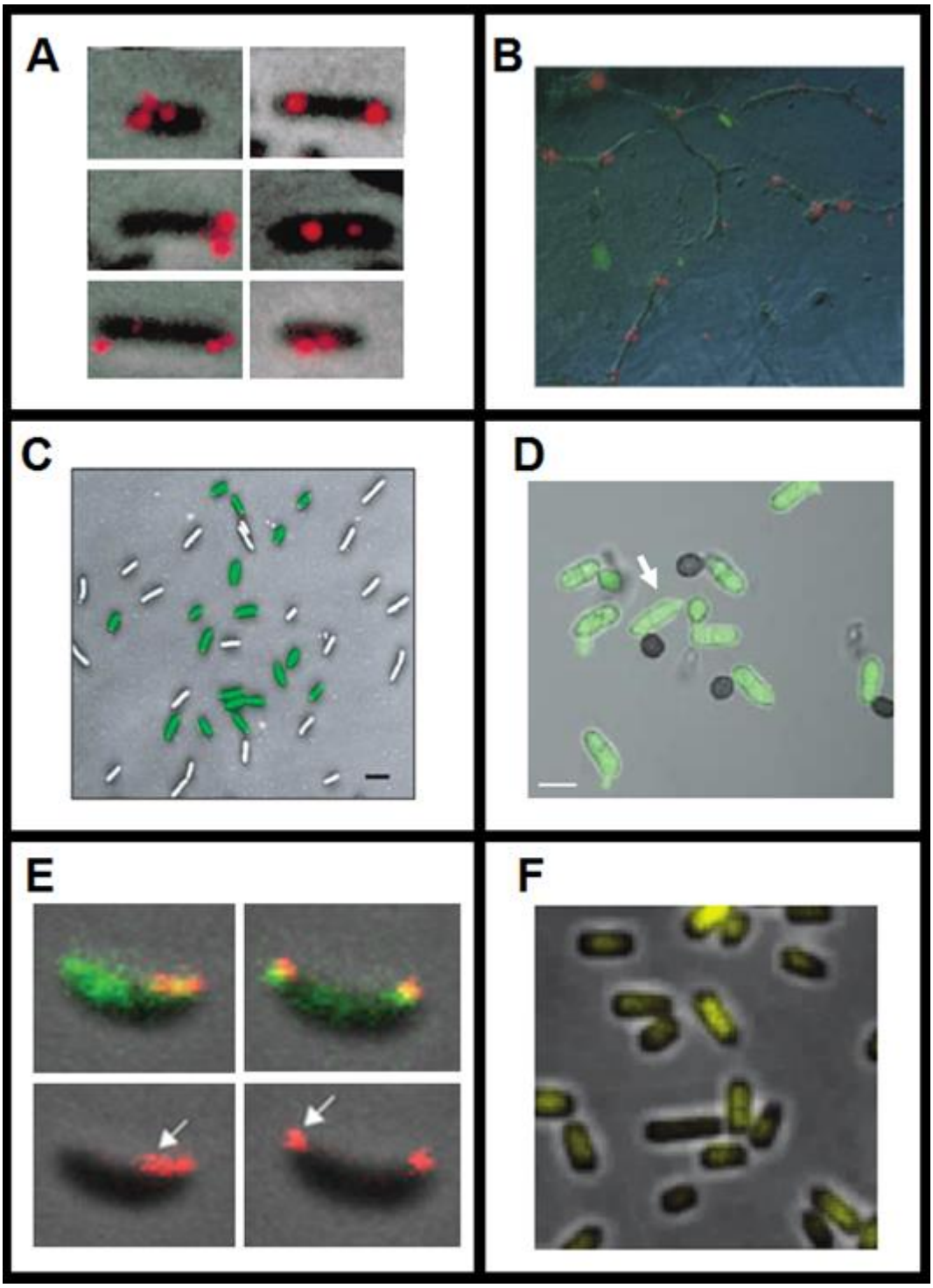

Figure 2 - Examples of multimodal image fusion. (A) Fluorescence and phase contrast images of *E. coli* cells. Adapted from (Niki & Hiraga, 1997). (B) Fluorescence and phase contrast images of endothelial progenitor cells. Adapted from (Bahlmann et al., 2003). (C) Fluorescence and inverted phase contrast images of *E. coli* cells. Adapted from (Ozbudak et al., 2004). (D) Differential interference contrast and fluorescence images of *Colletotrichum acutatum* cells. Adapted from (Brown et al., 2008). (E) Fluorescence and phase contrast images of *Caulobacter crescentus* cells. Adapted from (Ptacin et al., 2010). (F) Fluorescence and phase contrast images of *E. coli* cells. Adapted from (Taniguchi et al., 2010).

As can be seen in Figure 2, in some cases, the simple superposition of multimodal images will result in fused images where both images are correctly aligned (see Figure 2-B, C, D and F), while in some cases, possible registration misalignments can be observed, as several

fluorescent F-Plasmids, appear to be outside of the cell contours (see Figure 2-A and E), observable in the fused image, which is more than expected from diffraction effects. Intra-model registration can be used in images taken at different time frames, while inter-model registration can be used if images are taken from different sensors.

Since these type of misalignments can affect statistical analysis such as the calculation of the plasmids spatial distribution along the cell, various image processing algorithms started to be developed, such as image registration, image segmentation, which proved to be required to perform better statistical analysis for the qualitative and quantitative characterization of the processed data (Eils & Athale, 2003).

3.1.2. Microscopy Image Processing and Statistical Analysis

Usually, the simple superposition of multimodal images fusion will result in the aforementioned misalignment, particularly when the microscopy image acquisition is made by different sensors, which requires reliable image registration techniques to properly align the multimodal images (Deshmukh & Bhosle, 2011).

The process of image registration is done by properly overlaying two or more images of the same location taken at different time frames and/or from different viewpoints and/or by different sensorial devices. Image registration algorithms have been classified mainly based on modality, intensity or the methods used for registration (Deshmukh & Bhosle, 2011). More specifically they can also be separated on dimensionality, the domain and type of transformation, the type of data, source of features (intrinsic or extrinsic), automation level (Wyawahare, Patil, & Abhyankar, 2009).

Image registration methods have been divided in four steps: "(i) *feature detection* (ii) *feature matching* (iii) *transform model estimation* (iv) *image resampling and transformation*" (Zitová & Flusser, 2003). Based on these steps and the nature of the images the registration methods have been classified as area-based or feature-based.

Area based methods give more significance on the feature matching step rather than on their detection while featured based methods give more emphasis to the detection step (Zitová & Flusser, 2003). Examples of area-based methods are correlation-like methods (e.g. normalized cross-correlation and its modifications), Fourier methods (e.g. phase-correlation and its modifications to add rotation and scale factors to the transformation), mutual information methods and search techniques based on the sum of squared intensities (Chapnick et al., 1993; Wyawahare et al., 2009; Zitová & Flusser, 2003).

The aim of the feature-based methods is to find the pairwise correspondence, represented by control-points on the images (literal points, end points or centres of line features, centres of gravity of regions, etc.), using their spatial relations invariant descriptors (Torr & Zisserman, 1999). Other descriptors of features used as feature-based methods are the relaxation methods, pyramids and wavelets (Deshmukh & Bhosle, 2011). Multispectral/multisensory image registration can raise challenging problems due to different grey level characteristics,

making inadequate the application of simple techniques such as those based on area correlation (Deshmukh & Bhosle, 2011).

For the past decades numerous works have been conducted on segmenting cells and internal cellular structures (Meijering, 2012). Most of the developed solutions have been applied in isolated applications, where dedicated automatic methods are developed for a specific problem. One of the biggest challenges is to design methods sufficiently generic (automatic or semi-automatic) in order to attain a high specificity and sensitivity for an extensive range of cases (Meijering, 2012).

Various techniques have appeared over the years, as can be seen in Figure 3, such as intensity thresholding (in blue), feature detection (in red), morphological filtering (in green), region accumulation (in yellow), deformable model fitting (in violet) and techniques that were not classified as any of the former (in magenta).

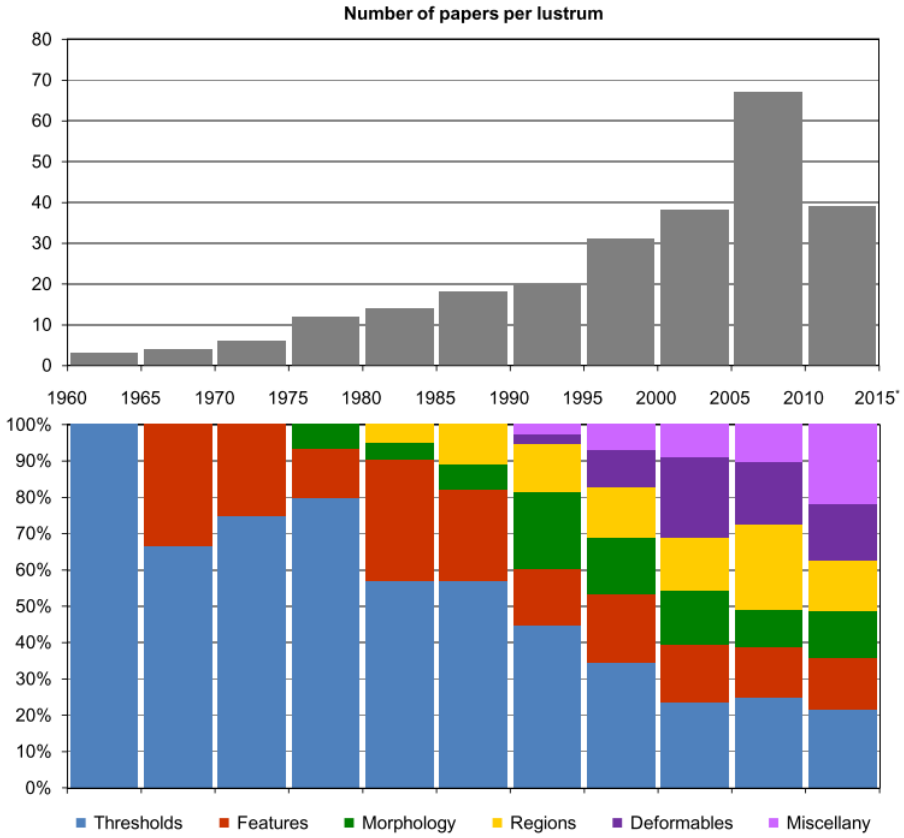

Figure 3 – Fifty year analysis of cell segmentation techniques. Top graphic shows the number of articles in the area per 5 years, and the bottom graphic shows the evolution of the used cell segmentation methods. These include intensity thresholding, feature identification, morphological filters, region accumulation, deformable model fitting and other miscellaneous approaches. Taken from (Meijering, 2012).

Nowadays, most of the recent techniques use a mixture of several approaches and started to be available in open-source platforms, in order to increase testability of all those methods described in the literature (Meijering, 2012).

Contests and open challenges, based on benchmark data (acquired by an independent laboratory or created by artificial image simulators) are being organized to prevent some abuses of method comparison in the literature, where research groups are trying to "prove" superiority over the previous methods (Meijering, 2012). One of these contests is the 'Cell Tracking Challenge', which is already in its 2nd Edition and is part of the 'Grand Challenges in Biomedical Image Analysis' – (http://grand-challenge.org/Home/), where a benchmark of artificial and real datasets was created in order to measure six segmentation and tracking algorithms (Maška et al., 2014).

In terms of microscopy images of bacteria, the usual techniques for image segmentation are based on thresholding (using a fixed or adaptive intensity value) and watershed transform (based on the image morphology) that, mimicking the idea of a water source on each regional minima flooding the image, being able to segment two or more touching cells (Hand, Sun, Barber, Hose, & Macneil, 2009; Malpica & Solorzano, 2002; Vallotton, Mililli, Turnbull, & Whitchurch, 2010), or even the merging of both techniques with probabilistic methods. The efficiency of these techniques is limited to high contrast images with well-defined cell wall limits and uniform illumination (Jung & Scharcanski, 2005). It is also known that watershed methods have a tendency for over-segmentation when acquisition conditions are not perfect, which is common in most of the microscopy modes in high-throughput experiences, due to the high variability of the illumination and contrast conditions, even within the same time series (Jung & Scharcanski, 2005).

With the development of novel image processing techniques, various computational toolboxes have been published and publicly available. Presently, one of those open source tools that generated a successful scientific impact (was cited in over 800 publications) is 'Cellprofiler' (Carpenter et al., 2006). An improved 'Cellprofiler' tool was developed in 2011 with new segmentation algorithms and features to facilitate high-throughput work, which also proved to be more robust and user friendly (Kamentsky et al., 2011). 'Cellprofiler' has been used to automatically identify and measure various eukaryotic constituents in images, including studies in human cells, mouse cells, yeast colonies, *C. elegans* colonies and other eukaryotic species (Hartwell et al., 2013; Lamprecht, Sabatini, & Carpenter, 2007; Wählby et al., 2013).

Cell segmentation in 'Cellprofiler' is performed in two steps. First, it uses a block-wise Otsu threshold (Otsu, 1979), followed by using a bilinear interpolation, which is applied to separate each cell colony from the background (Carpenter et al., 2006). The second steps used the intensity or shape as a feature for discrimination and segmentation of clumped objects (Carpenter et al., 2006).

Additional image analysis software was developed for segmentation of bacterial cells in microscopy imaging, based on DAPI staining and *fluorescence in situ hybridization* (FISH) images. The source-code of the software, named 'CellC', was publically released and was compared with an image processing and analysis automated tool developed in ImageJ, while both methods were validated with manual cell counting (Selinummi, Seppälä, Yli-Harja, & Puhakka, 2005).

Another of those tools is 'Cell-ID', which is optimized to bright-field images of yeast (mainly *Saccharomyces cerevisiae*) and other cell types (Gordon et al., 2007). In bright-field images

taken beneath the focal plane, the cell border pixels are darker than the image background and the cell internal pixels. A threshold cut-off level is determined independently and automatically for each image in order to segment the cells from the background, providing a method to study high throughput microscopy studies with a high sensitivity (Gordon et al., 2007).

Another publically available tool is the 'CellTracker' which was specially developed to track the movement of living cells and also to automatically segment cell boundaries and quantify the intensity of fluorescently tagged proteins (Shen et al., 2006). Cell borders were detected via thresholding and level setting and refined by detecting the cell edges based on an active contour algorithm (Kass, Witkin, & Terzopoulos, 1988; Shen et al., 2006).

Finally, it is also worth mentioning the 'Farsight' toolkit, a segmentation algorithm developed to study detailed biological microenvironments, such as histopathology images of the brain and other tissues (Al-Kofahi, Lassoued, Lee, & Roysam, 2010; Bjornsson et al., 2009).

The methods present in the 'Farsight' toolkit exploited graph-cuts-based algorithms for segmenting foreground signals from the image background. Then, the nuclear seed points are detected by a multiresolution edge detection method, based on Laplacian-of-Gaussian filters limited by an adaptive scale selection of distance-maps and refined by a second graph-cuts algorithm (Al-Kofahi et al., 2010).

The algorithms implemented in the above-mentioned tools use a variety of cell segmentation approaches and have been used mainly in eukaryotic cells. The major drawback of using the same segmentation algorithms in prokaryotic cells is that these cells are organized in large and dense clusters.

The main consequence of segmenting such clusters, is that accuracy of the algorithms will decrease as the clusters density increases, as its success depends in the initial marking and identification of cell boundaries, which can be a difficult task when cells are tightly clumped together, reducing the changes of portability of using those tools in bacteria segmentation (Q. Wang, Niemi, Tan, You, & West, 2010).

In order to make an adequate segmentation of live prokaryotic cells in time-lapsed microscopy, other approaches for segmentation methodologies have been developed.

Research done by Wang et al. described the usage of morphological methods for automatic segmentation of not only bacterial cells, but also yeast and human cells (Q. Wang et al., 2010). Their methods evolved from a previously developed tool, called 'CellTracer' (Q. Wang, You, & West, 2008).

Cell lineage in 'CellTracer' was constructed using a neighbourhood-based scoring methodology to track each bacterium along each frame in the time-lapsed microscopy (Q. Wang et al., 2010). Their methodology exhibited a satisfactory portability to various types of cells and integrating various types of microscopy imaging, namely phase-contrast, bright-field and fluorescent (Q. Wang et al., 2010).

Other segmentation methods were implemented in a toolbox called 'MicrobeTracker' and an accessory tool, 'SpotFinder' to do biological studies in *Escherichia coli* and *Caulobacter crescentus* (Sliusarenko & Heinritz, 2011).

The segmentation methodology from 'MicrobeTracker' starts with a sequence that starts with inverting the image, then combines algorithms based on edge detection and watershed transform to start the segmentation (Sliusarenko & Heinritz, 2011). The initial cell contours are then refined, similarly to the 'CellTracker' algorithm (Shen et al., 2006), by separating the foreground objects from the background, separating the under-segmented clumped cells from each other, and discarding false positives that can occur due to noise in the images (Sliusarenko & Heinritz, 2011).

Another state of the art software tool is 'Schnitzcells' (J. W. Young et al., 2012). 'Schnitzcells' provides solutions for segmentation and tracking of Escherichia coli cells from images by confocal or phase contrast microscopy. 'Schnitzcells' cell segmentation is a multi-stepped process, starting with an operation of edge detecting in order to generate an initial cell segmentation (J. W. Young et al., 2012). The next step is to splits long or clustered cells. Finally, it considers too small objects as false positives and discards them (J. W. Young et al., 2012).

The main issue with the 'Schnitzcells' toolbox happens to be the large number of parameters that, without proper tuning, cause the accuracy of the segmentation to decrease notably (Queimadelas, 2012). Further, it has a limited scope of application, i.e. it only handles *E. coli* and *Bacillus subtilis* cells and often presents a significant number of false positives (Queimadelas, 2012).

An additional segmentation method, 'MAMLE' (Multi-Resolution Analysis and Maximum Likelihood Estimation), was proposed for detecting *E. coli* cells within dense clusters (Chowdhury, Kandhavelu, Yli-Harja, & Ribeiro, 2013).

'MAMLE' executes cell segmentation in two stages. The first relies on state of the art filtering technique, edge detection in multi-resolution with morphological operator and threshold decomposition for adaptive thresholding. From this result, a correction procedure is applied that exploits maximum likelihood estimate as an objective function. Also, it acquires morphological features from the initial segmentation for constructing the likelihood parameter, after which the final segmentation is obtained (Chowdhury et al., 2013).

To provide more solutions to the prokaryotic segmentation problem, the Laboratory of Biosystem Dynamics from Tampere in collaboration with the Computational Intelligence Research (CA3) Group from FCT-UNL initiated a project to develop automatic image processing toolboxes for cell segmentation and tracking. The first developed tool (called 'CellAging') was used applied in studies of the segregation and partitioning that occur in cell division (Häkkinen, Muthukrishnan, Mora, Fonseca, & Ribeiro, 2013).

A new tool ('iCellFusion') is being developed (Santinha et al., 2014), which continues the work of the 'CellAging', but focuses on the fusion of different microscopy methods to provide a better integration of functional and morphological information by fusing phase-contrast and fluorescence microscopy images.

Table 1 – Review of the publically available toolboxes of microscopy image processing

Toolbox	Availability	Reference
'Cellprofiler'	http://www.cellprofiler.org/	(Carpenter et al., 2006)
'Cell-C'	https://sites.google.com/site/cellcsoftware/Home	(Selinummi et al., 2005)
'Cell-ID'	http://lbms.df.uba.ar/	(Gordon et al., 2007)
'CellTracker'	http://www2.warwick.ac.uk/fac/sci/systemsbiology/staff/bretschneider/celltracker/	(Shen et al., 2006)
'Farsight'	http://farsight-toolkit.org/wiki/Main_Page	(Al-Kofahi et al., 2010)
'CellTracer'	https://stat.duke.edu/research/software/west/celltracer/	(Q. Wang et al., 2008)
MicrobeTracker'	http://microbetracker.org/	(Sliusarenko & Heinritz, 2011)
'Schnitzcells'	http://cell.caltech.edu/schnitzcells/	(J. W. Young et al., 2012)
'MAMLE'	http://www.cs.tut.fi/~sanchesr/CellSegment/index.htm	(Chowdhury et al., 2013)
'CellAging'	http://www.cs.tut.fi/~sanchesr/cellaging/	(Häkkinen et al., 2013)

Quantitative performance evaluation of the image processing techniques requires the use of reference datasets for validation. Manual processing techniques were the first gold standard for validation of the methods included in the automatic tools for cell parametric measurements (Cytometry). The ground-truth established by the manual processing can be challenged as they are expert-dependent (repeatability of results depends on the user, and even intra-user variability can be very high) and can become unfeasible for large data-sets (the case of high-throughput microscopy studies), due to becoming a non-viable and time-consuming task (Coelho, Shariff, & Murphy, 2009).

For image registration techniques, known gold standard registration methods include manual processing, by back projection of image features (mainly by implanting fiducial landmarks) or by using simulated data (van de Kraats, Penney, Tomazevic, van Walsum, & Niessen, 2005).

The use of fiducial landmarks in live-cell imaging has two major problems: (i) having several fiducials too close to cellular constituents will affect the imaging and (ii) the fiducials on the image might show divergent drift patterns, either due to movement within the sample or due to the variability of the drift within the sample (Pertsinidis, Zhang, & Chu, 2010; Shivanandan, Deschout, Scarselli, & Radenovic, 2014). Moreover, such fiducial marks might have a disruptive impact on the regular behaviour of bacteria.

Simulations of biological processes using computational modelling is a viable alternative to create a "ground truth" by producing artificial deformable images that can be used for quantitative evaluation of the image processing algorithms has been one of the growing trends in microscopy imaging in the last years (Bonnet, 2004).

Simulated images have already been used as benchmark with precisely known “Ground-truth” in biological studies, as for example phase-contrast image of a fish epidermal keratocyte was simulated to analyse the dynamics of cell motility (Ambühl, Brepsant, Meister, Verkhovsky, & Sbalzarini, 2012) and fluorescent images of human embryonic stem cells were also used to test four segmentation methods: k-means clustering, Expectation Maximization, Otsu’s threshold, and the Global Minimization of the Active Contour Model in order to evaluate both quantitatively and qualitatively each of the methods(Du & Dua, 2010). The data that was utilized in the previous study was obtained by the development of a synthetic image generator called ‘SIMCEP’ (Ruusuvuori et al., 2008).

Such image generators of biological systems have been helpful in the improvement of the quantitative measures of cell segmentation, including statistical analysis such as Sensitivity, Precision and F-Score (Powers, 2011).

The accurate measurement of the cell boundaries and their internal constituents is important to provide precise statistical distributions (spatial and temporal) of cellular internal functions, which is critical in increasing the present knowledge of transcription, translation, regulation, degradation and other genetic mechanism dynamics at the single-molecule level in a living cell (Kalisky, Blainey, & Quake, 2011).

3.1.3. Simulation of Microscopy Images

Allied with the development of tools capable of image processing and statistical analysis, computational tools for the generation of synthetic images started to be established for the purpose of serving as validation of the former tools, but also they are able to create “null-models” (Gotelli & Bldg, 1996), that can be used to study the statistical patterns in the absence of a particular mechanism (e.g. it could be used to study how the nucleoid affects the production of RNA molecules target for fluorescent proteins by removing the nucleoid from the cell).

To create such simulation tools, realistic biological models need to be developed using data coming from theoretical and experimental knowledge that arise out of the statistical distributions of cellular geometry (Xiong, Wang, Ong, Lim, & Jiang, 2010) and spatial and temporal information (Kruse, 2012). Those models include the cell shape and size, the location of subcellular structures, kinetic and spatial models of cell growth, cell division cell migration and internal cell functions (such as gene expression).

Regarding just tools of microscopy image simulation, most of them only focused on the spatial information of the cell, producing just a single frame of the desired synthetic image model. Chronologically, the image simulation tools started by producing simple objects, such as the point-like objects of Fluorescence in situ hybridization (FISH) spots (Grigoryan, Hostetter, Kallioniemi, & Dougherty, 2002), which were randomly placed in a 3D space, as can be seen in Figure 4.

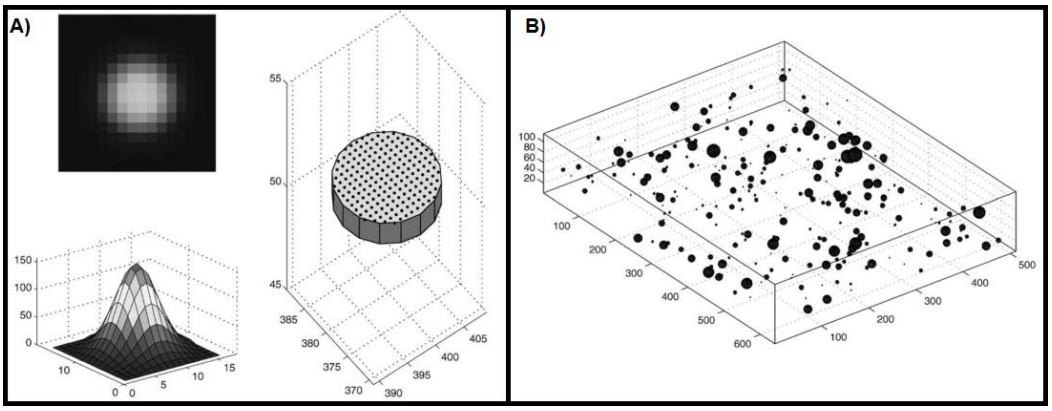

Figure 4 – (A) Slice of a simulated point-like object (based on the Fluorescence in situ hybridization spots); (B) Random 3d spots in a Rectangular Prism. Adapted from (Grigoryan et al., 2002).

A more complex simulator (Lehmussola, Selinummi, Ruusuvuori, Niemisto, & Yli-Harja, 2005) was designed to produce a simulated image of large eukaryotic cell populations by creating a parametric model for each individual cell geometric contour (including size, shape and texture). This simulator also included common errors obtained in the image acquisition systems, such as uneven lighting and optical aberrations (Lehmussola et al., 2005), as can be seen in the workflow presented in Figure 5-1. In Figure 5, it is also possible to observe the creation of the nuclei (Figure 5-2-A) and the cytoplasm (Figure 5-2-B) and the simulation of the overlap of both structures (Figure 5-2-C). This simulator (Lehmussola et al., 2005) evolved towards the development of a standalone and publically available toolbox called 'SIMCEP' (Lehmussola, Ruusuvuori, Selinummi, Huttunen, & Yli-Harja, 2007).

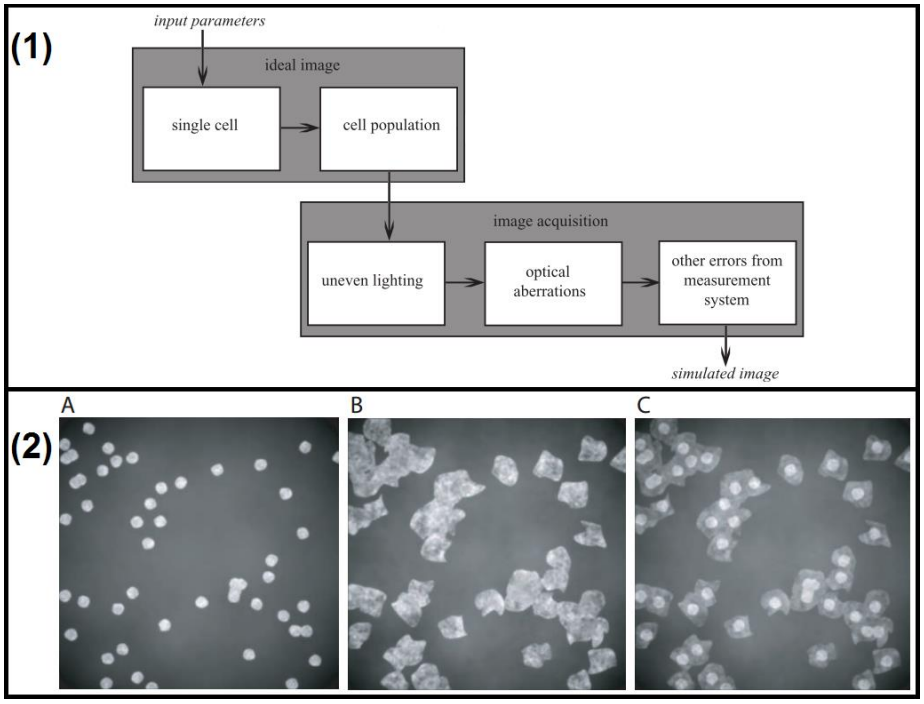

Figure 5 - (1) Workflow of the simulator; (2-A) nuclei images; (2-B) cytoplasm images; (2-C) Overlapped image of the simulated nuclei and cytoplasm. Adapted from (Lehmussola et al., 2005)

The 'SIMCEP' simulator (Lehmussola et al., 2007) provided a framework to validate and test various image processing tools, such as the previously mentioned 'CellProfiler' (Carpenter et

al., 2006) and 'CellC' (Selinummi et al., 2005), an open-source and Java-based image processor (ImageJ v1.36b) and a commercially available software (MCID Analysis from Imaging Research Inc., Catharines, ON, Canada; Evaluation ver. 7.0). A benchmark dataset of synthetic images along with manually labelled images was produced to be used as ground-truth to tune and test other image processing tools (Ruusuvuori et al., 2008). Various images were produced with different cell parameters, such as probability of clustering, cell radius, and cell shape and image parameters such as background noise and illumination disturbance (Ruusuvuori et al., 2008). Eleven methods for segmentation of subcellular constituents, which have a spot-like structure, were tested both microscopy images from real experiments coming from human and yeast cells but also validated their comparison by using simulated microscopy images from 'SIMCEP' (Ruusuvuori et al., 2010). Their results showed that no algorithm outperformed the others in every situation, as was shown that the selection of spot detection algorithms should consider each situation (the type of cells, the quantity and quality of images, the quantity of spots, etc) and choose the method accordingly (Ruusuvuori et al., 2010).

The same group from the 'SIMCEP' simulator started working on developing parameterized models of different bacterial populations (Lehmussola, Ruusuvuori, Selinummi, Rajala, & Yli-harja, 2011). The five proposed models can be observed in Figure 6-A:

(1) bacilli-type bacteria (rod-shaped), like the *Escherichia* and *Salmonella* genera;
(2) cocci-type bacteria (spherical-shaped), like the *Streptococcus and Micrococcus* genera;
(3) spirochaetes-type or spirilla-type bacteria (corkscrew-shaped), like the *Treponema, Borrelia* and *Brachyspira* genera;
(4) coccobacilli-*type* bacteria (intermediate shape between spheres and rods), like *Haemophilus* and *Chlamydia* genera;
(5) vibrio-type bacteria (curved rods or comma-shaped) like the *Caulobacter* and *Vibrio* genera (Lehmussola et al., 2011).

More information on bacteria shapes will be given in Sub-Section 3.2.1 - Bacterial Cell Morphology - Shape, Size and Spatial Arrangement. A simulation of a population set containing both *Escherichia coli* and *Micrococcus luteus* cells is shown Figure 6-B, using the first and second models respectively (Lehmussola et al., 2011).

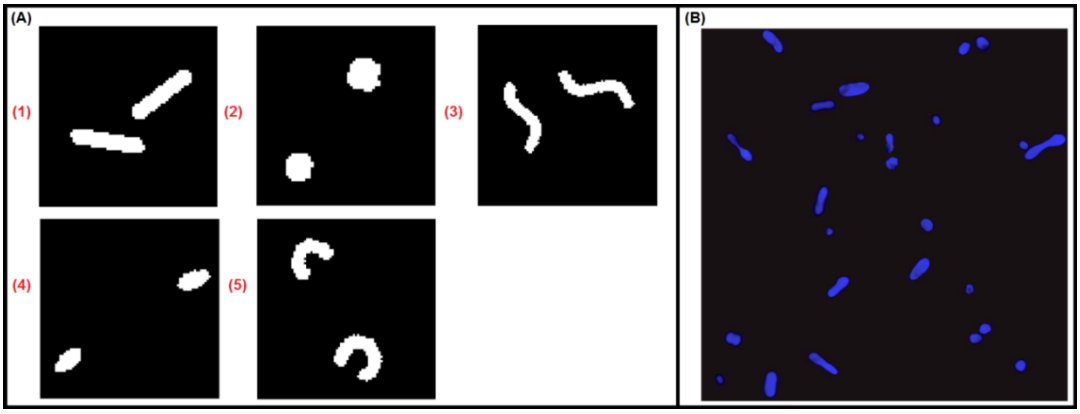

Figure 6 - (A) Parametric bacterial shape models; (B) Population of 40 cells sampled from models learned for E. coli and M. luteus bacteria. Both synthetic cell types show variation in cell sizes and shapes. Adapted from (Lehmussola et al., 2011).

This simulator ('SIMCEP') is capable of creating population with similar characteristics (but each cell parameters are drawn from a random variable), but also sub-populations with specific characteristics, for example, stress response to drugs and gene knock-downs (Lehmussola et al., 2011). As this toolbox was designed only to simulate 2D images (and its extension to 3D was not straightforward), another toolbox was developed to extend that model to a higher dimension, although it limited the maximum number of generated cells (Svoboda, Kasik, Maska, & Hubeny, 2007).

Generation of the object (e.g. HL-60 Nucleus) starts by creating an ellipsoid in black and white (Figure 7-1-a), then the ellipsoid is deformed using partial differential equation-based methods (Figure 7-1-b) and then texture is created by defining the internal structure and adding the nucleus (Figure 7-1-c) (Svoboda et al., 2007; Svoboda, Kozubek, & Stejskal, 2009). A final image of the cell passing through the optical system is finally simulated with the added nucleus (Figure 7-1-d) (Svoboda et al., 2009).

The generated artificial images were validated using four different methodologies (Svoboda et al., 2009):

(1) by visual comparison of real images acquired in a laboratory;
(2) comparison of the log intensity histograms of the artificial and real images;
(3) comparison of the computed descriptors, namely the entropy and the second to sixth central moments using Quantile—Quantile plots from real and synthetic data;
(4) computing the 3D Haralick texture features, such as angular second moment, contrast, correlation and variance.

The full workflow of the simulator is shown in Figure 7-2, where it was also added a digital camera simulator module (Svoboda & Ulman, 2012), and it was designed to reproduce not only the spatial information, but also the temporal information by simulating motion of selected biological objects and generating an artificial time-lapse observation (Ulman & Hubeny, 2007), as can be observed in Figure 7-3 (Svoboda & Ulman, 2012).

The simulator has been published online, as a toolbox called 'CytoPacq', equipped with three different modules:

- The first module, '3D-cytogen' is the digital cell phantom generator (Figure 7-2) (artificial objects that mimic the cell structure and behaviour). '3D-cytogen' can generate microspheres, granulocytes, HL-60 Nucleus and images of Colon Tissue;

- The second module, '3D-optigen' is the optic system simulator (Figure 7-2) and simulates the transmission of the signal through the lenses, the objective, the excitation filter and the emission filter (various sets of equipment can be simulated for each part);

- The last module, '3D-acquigen' is the digital CCD camera simulator of the phenomenon's that occur during image capture (noise, sampling, digitization) by changing the camera selection, the acquisition time, the dynamic range usage and the stage movement in the z axis.

The generation of artificial time-lapsed microscopy images marks a very important step for the validation of automatic tools used in live cell imaging, as a time series extends the observation from a unique time-point (just 1 frame) to the observation various frames containing cellular dynamics, such as measuring protein or RNA levels or even observing cell migration, cell division and cell growth (Coutu & Schroeder, 2013; Sung & McNally, 2011).

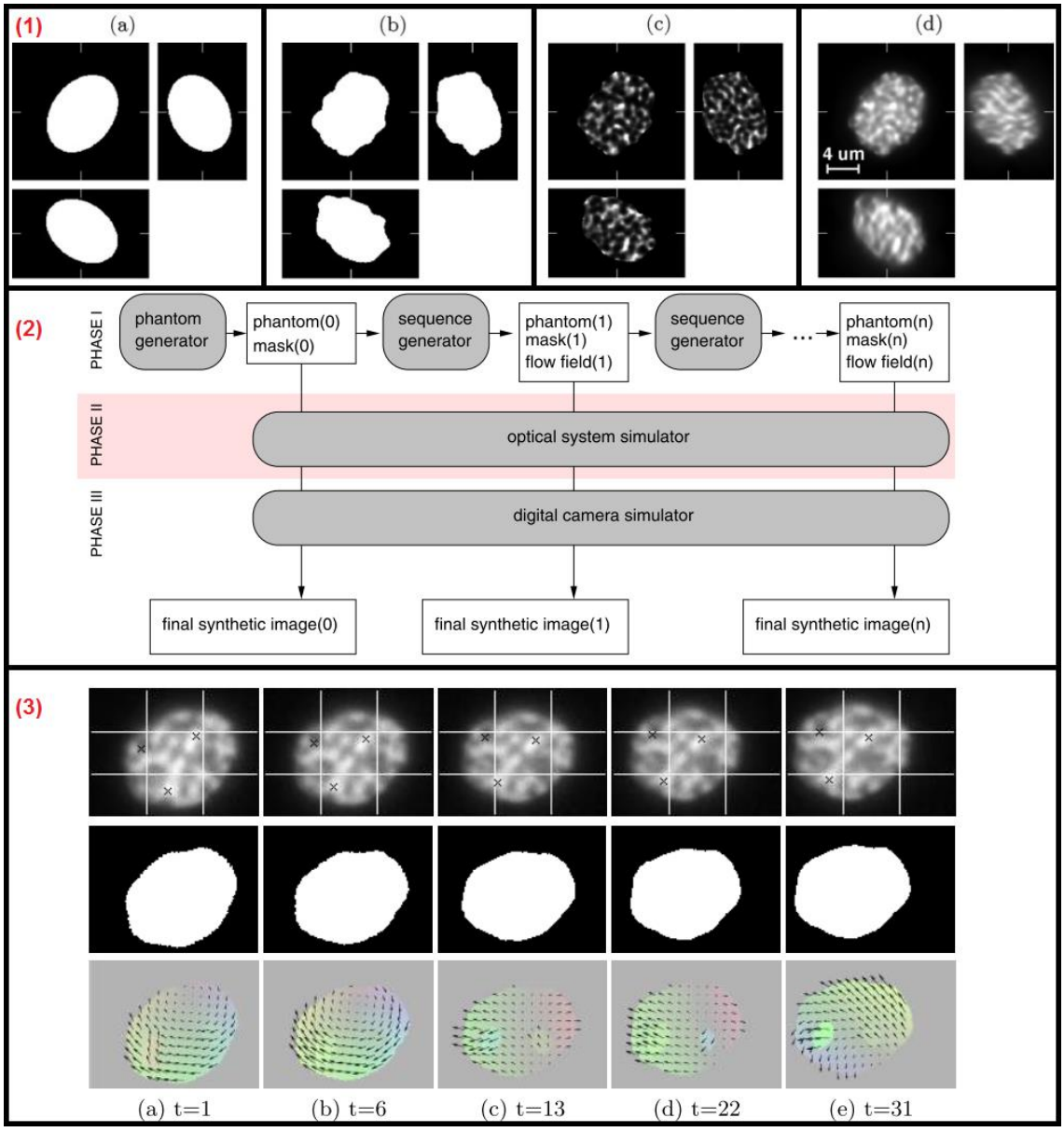

Figure 7 - (1) Steps for the artificial image generation of a HL-60 Nucleus (Svoboda et al., 2009); (2) Workflow of the CytoPacq toolbox (Svoboda & Ulman, 2012); (3) Artificial time-lapse observation of a generated HL-60 Nucleus (Svoboda & Ulman, 2012).

A recently developed toolbox called 'SimuCell' (Rajaram; Satwik, Benjamin, Nicholas, Steven, & Lani, 2012), capable of generating synthetic microscopy images with a heterogeneous cellular population and diverse cell phenotypes. In Figure 8-a, the workflow of 'SimuCell' is presented, showing that it allows for the development of novel phenotypes by creating new Plug-ins and that the algorithm development cycle is dependent of the comparison between the full ground truth and the result analysis (Rajaram; Satwik et al., 2012).

Each cell and their internal organelles can be modelled with different shapes and have distinct distribution of markers over each shape.

One of the biggest innovations of this tool is that it allows that the distribution of biomarkers inside the cell to be affected by the cell's microenvironment (See Figure 8-b), making the placement of each cell an important task in 'SimuCell', which can be in clusters, or placed near existing cells, or randomly placed but also allowing the overlap of cells in the images. Figure 8-c shows four examples of different images with different phenotypes and with different densities. Examples of simulated cellular organelles include the nucleus; nuclear body; cytoplasm and lipid droplet. Each object can be rendered using its own specific Plug-in. The 'SimuCell' toolbox can also simulate image artefacts that occur during the process of image acquisition, such as adding basal brightness, or adding linear or radial image gradient. It can also simulate cell artefacts, such as adding cell staining or misfocusing some cells.

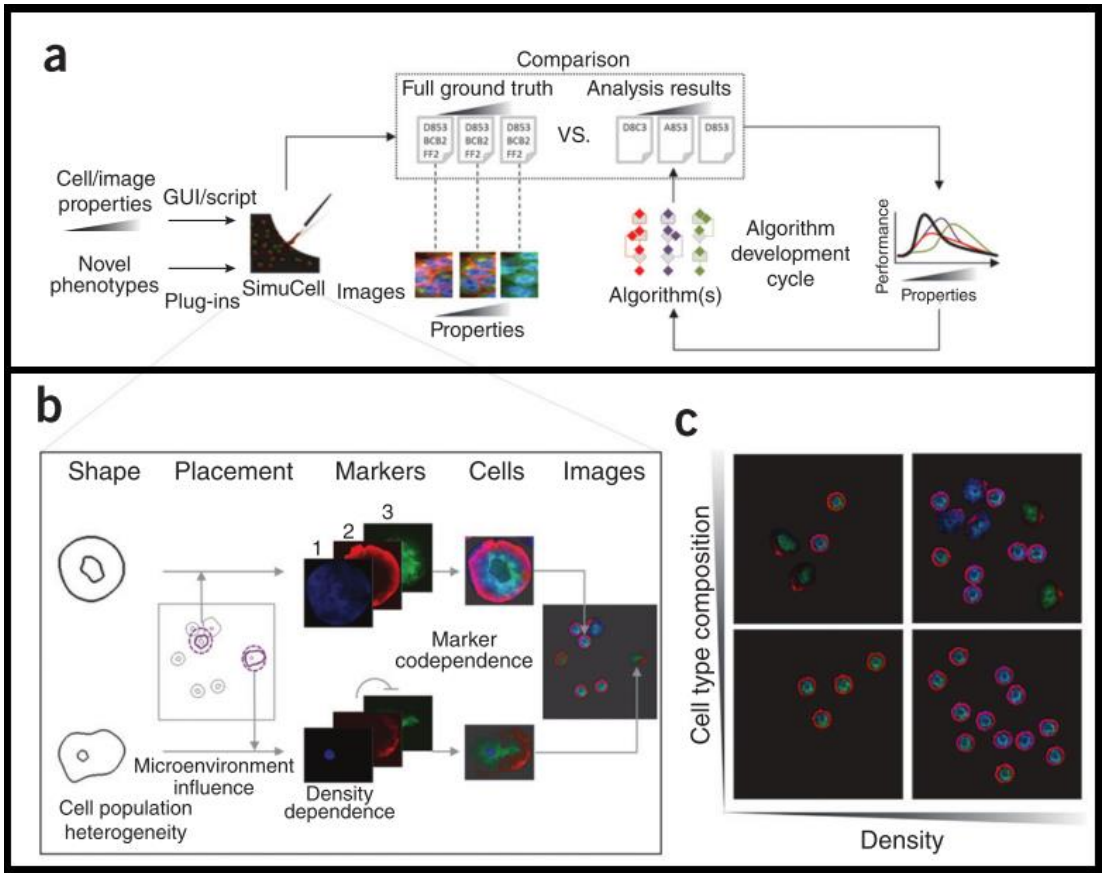

Figure 8 - (a) Workflow of 'SimuCell'; (b) Observation of the cell population heterogeneity and creation of different phenotypes in the same image; (c) Examples of images with different densities and different cell type composition. Taken from (Rajaram Satwik, Benjamin, Nicholas, Steven, & Lani, 2012)

In addition, a recently developed image generator focused on extending the range of microscopy methodologies, extending its simulation capabilities to include bright-field and phase-contrast techniques (Wiesmann, Sauer, Held, Palmisano, & Wittenberg, 2013). Their approach included the possibility of creating a realistic multi-channel simulation of cells (such as protoplasts and macrophages) in order to create unbiased and objective ground-truth data that can be used to validate the image processing tools, which have a significant advantage

over manually established ground-truth data, as it is expert-dependent and suffer from observer variance (Wiesmann et al., 2013).

Another group also developed a generator of artificial images of a neural stem cells' line, derived from the human umbilical cord blood (Korzynska & Iwanowski, 2012). Their system also took into account the simulation of the optical system and the camera, as other groups did, namely the 'CytoPacq' toolbox (Svoboda & Ulman, 2012). The main objective of this framework is also to validate the image segmentation tools produced by their group (Iwanowski & Korzynska, 2010) and estimate its segmentation error. Their methodology allowed the generation of various types of cells, as shown in Figure 9 (Korzynska & Iwanowski, 2012).

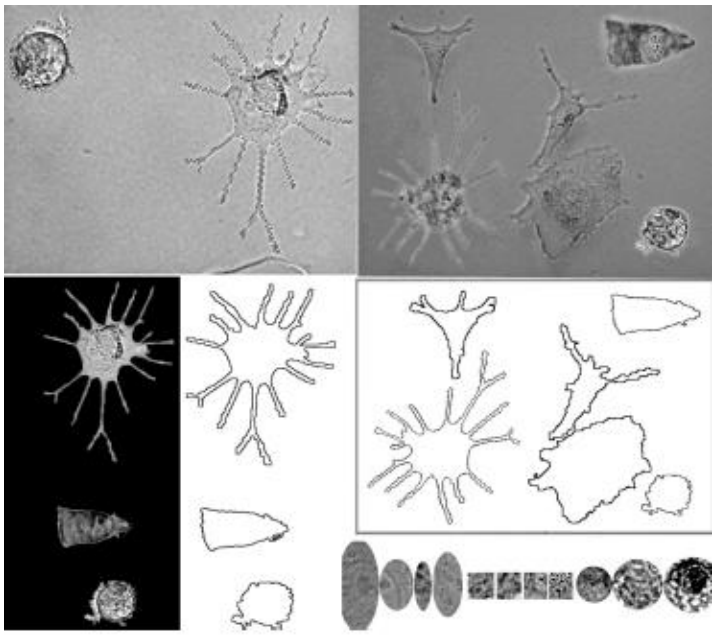

Figure 9 – Six types of artificially generated cells from a neural stem line (Korzynska & Iwanowski, 2012)

Using a different approach, another group developed a tool using machine learning trained from real data. The tool was capable of generating the whole cell, including structures like the nucleus, proteins, cell membrane and cytoplasm components such as microtubules. Although the model was capable of extracting a very precise shape model from real image data, the model could not be described in precise mathematical terms (Zhao & Murphy, 2007).

A description of the models as Bayesian networks, showing how the objects are created is shown in Figure 10-1 (Zhao & Murphy, 2007). The developed work later induced the development of a publically available toolbox called 'CellOrganizer' (Murphy, 2012). An overview of the modelling approach for the generation of synthetic microtubules is shown in Figure 10-2. Examples of a simulated 2D image of a lysosome (see Figure 10-3) and 3D images (see Figure 10-4) of a Lysosome (a), a Mitochondria (b), a Nucleolus (c) and an Endosome (d) are shown in Figure 10. The nuclear DNA (Nuclear Shape) is represented in red, the cell membrane (Cell Shape) is represented in blue and different protein location patterns are represented in green.

An example of a 2D slice generated from of a synthetic 3D microtubule image model is represented in Figure 10-5-A, while another example of the same model with additional free tubulin is represented in Figure 10-5-B.

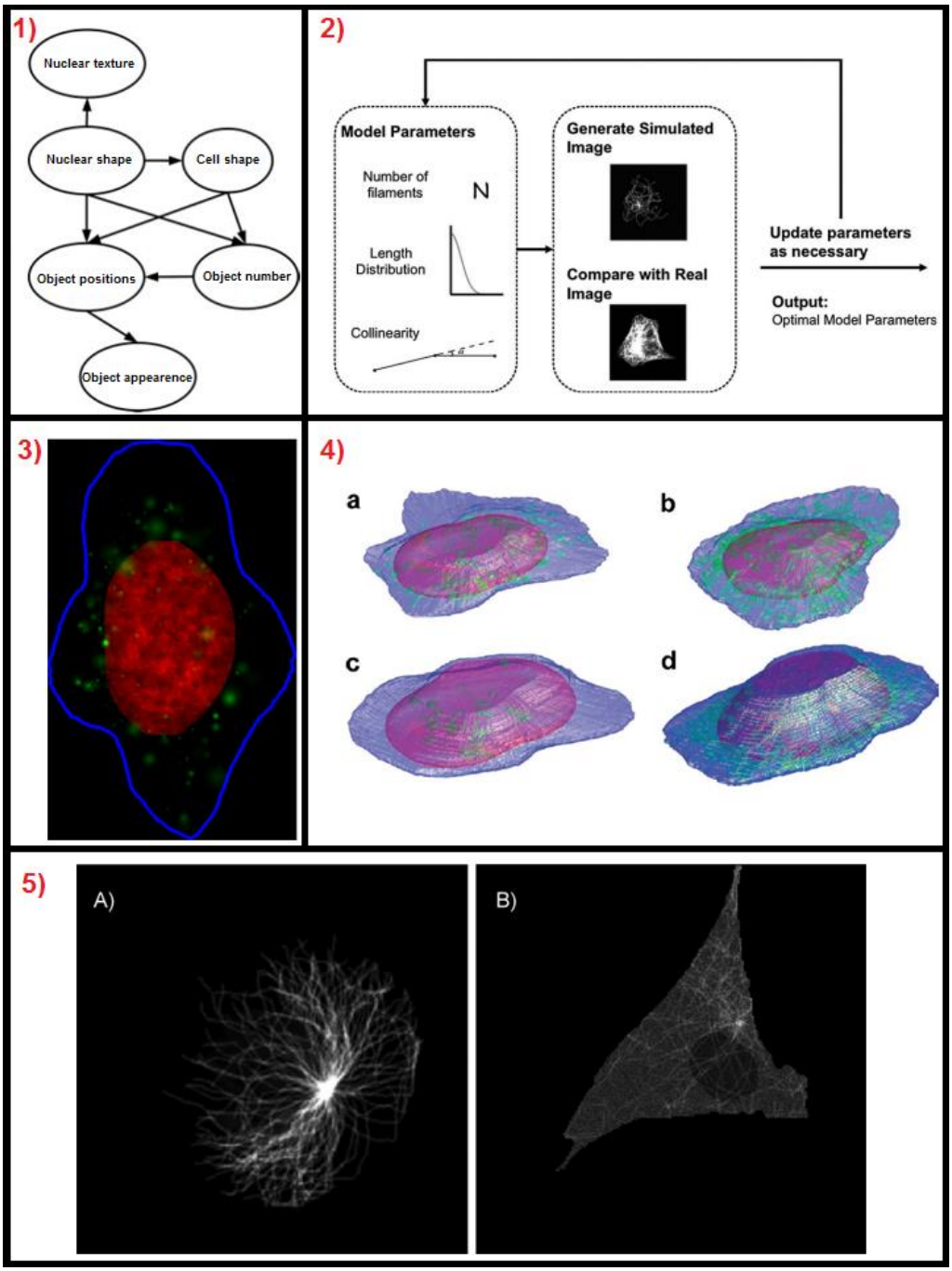

Figure 10 - (1) Description of the models as Bayesian networks (Zhao & Murphy, 2007). (2) Overview of inverse modelling approach for estimating parameters of the microtubule generative model (Murphy, 2012). (3) Example synthetic image generated by a 2D model learned from images of the lysosomal protein LAMP2. The DNA distribution is shown in red, the cell outline in blue, and the lysosomal objects in green. (Murphy, 2012) (4) Synthesized 3D images displayed in pseudo color surfaces for different protein location patterns (green), with nuclear (red) and cell shapes (blue). (a) Lysosome, (b) Mitochondria, (c) Nucleolus, and (d) Endosome. (Peng & Murphy, 2011) (5) A 2D slice example with the maximum plane intensity from generated 3D image using microtubule model. B: A 2D slice example with the maximum plane intensity from generated 3D image using microtubule model plus free tubulin model (Buck, Li, Rohde, & Murphy, 2012).

Another toolbox ('CompuCell3D') capable of simulating tissue development, homeostasis or even diseases over a timeframe was develop to aid the experimental studies in this area (Swat et al., 2012). The graphic user interface of 'CompuCell3D' is presented in Figure 11-A, showing the cell drawing tool. The proposed cell modelling is based on a multi-cell and open source Monte Carlo algorithm (Graner & Glazier, 1992) and is capable of simulating a cell-sorting model (shown in Figure 11-B), where the less-cohesive cells (lighter grey) envelop the more cohesive (dark gray) and condensing cells (forming a central cluster domain), simulating vascular tumor growth (as shown in Figure 11-C) or simulating angiogenesis models (Graner & Glazier, 1992).

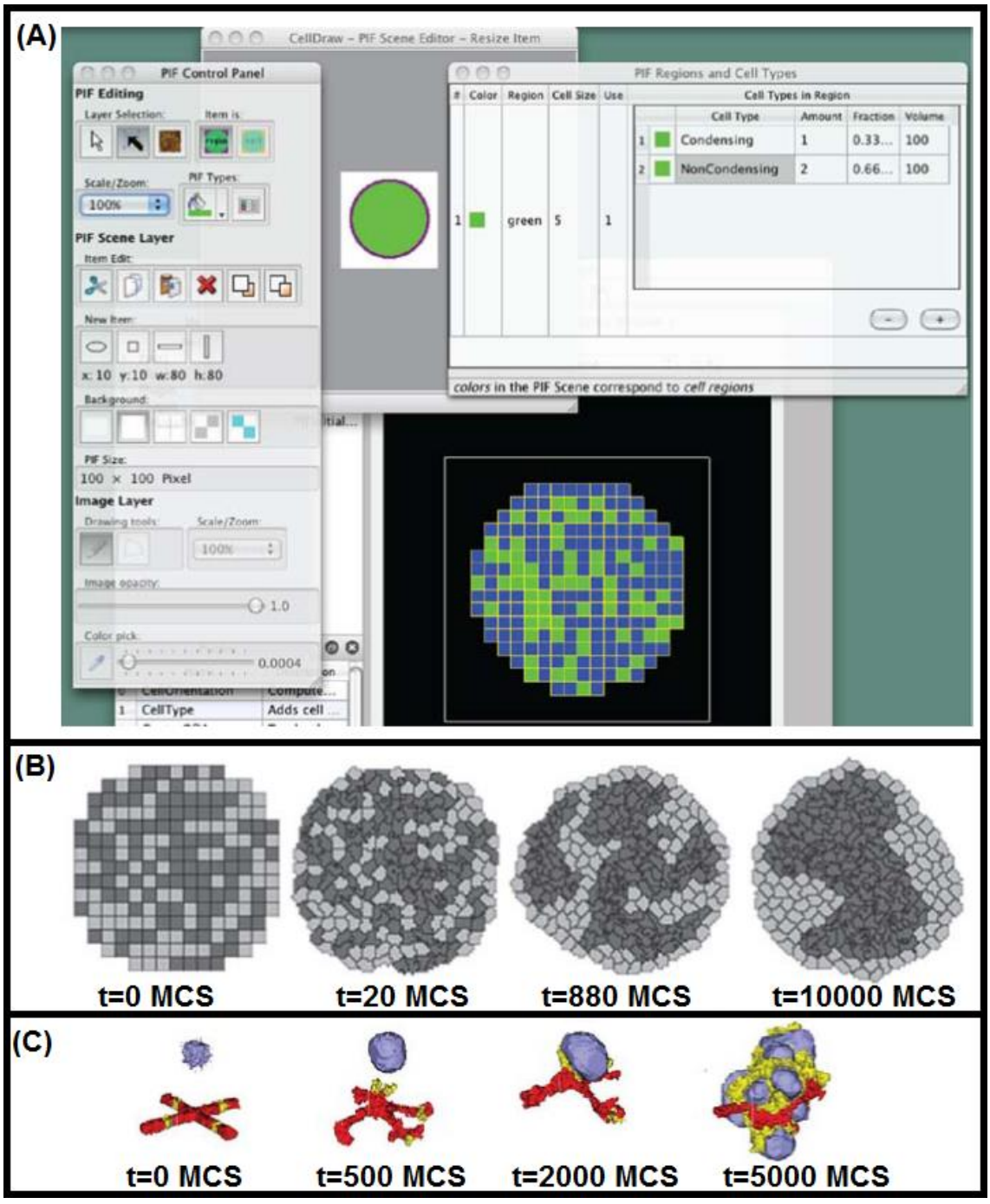

Figure 11 - (A) Graphic User Interface of 'CompuCell3D' showing the drawing graphics tool (B) Snapshots of the cell-sorting simulation from 'CompuCell3D'. MCS is one Monte Carlo Step (C) Snapshot of vascular tumor simulation taken at different steps. Taken from (Swat et al., 2012).

Instead of focusing on individual cell modelling, 'CompuCell3D' generates large cell populations by adopting the statistical large-Q Potts model to simulate the reorganization of uniformly distributed cell-like objects to assure that a natural cell shape is achieved (Swat et al., 2012).

Most of the publically available toolboxes for cell image generation, which are summarized in Table 2, and the reviewed standalone generators have used either mathematical models of cell shape, or in some cases used data mining techniques to acquire the models and its parameters directly from experimental images, providing realistic simulations of biological processes, working towards a 'Virtual Cell' model (Buck et al., 2012). The future work of the image generators will be focused on cell functionality, as cell morphology has been already studied extensively. The study of cell functional models (spatial and temporal cellular organization) can be used to implement mathematical models of cell migration, which is still being envisioned by scientists due to the complex nature of cell migration (Zhang, Wu, & Altschuler, 2013) or implement mathematical models of cell growth and division (Qu, Weiss, & MacLellan, 2004), allied with the usage of time-lapsed microscopy.

Table 2 - Review of the publically available toolboxes for cell microscopy image simulation

Toolbox	Availability	Reference
'SIMCEP'	http://www.cs.tut.fi/sgn/csb/simcep/tool.html	(Ruusuvuori et al., 2008)
'SimuCell'	http://www4.utsouthwestern.edu/altschulerwulab/simucell/	(Rajaram Satwik et al., 2012)
'CytoPacq'	http://cbia.fi.muni.cz/simulator/index.php	(Svoboda & Ulman, 2012)
'CellOrganizer'	http://cellorganizer.org/Downloads/	(Murphy, 2012)
'CompuCell3D'	http://www.compucell3d.org/	(Swat et al., 2012)

To tackle this problem, bacteria such as *E. coli* is the perfect organism to be used as a model, as several biological mechanisms at the molecular level are fundamentally conserved along various species, extending a considerable amount of our knowledge about those processes, trough experimental studies based on *E. coli* and its symbionts.

3.2. Cell Modelling – Spatial and Temporal Organization of Bacteria

E. coli is potentially the best studied organism, making it the basis for an impressive number of scientific breakthroughs, even in the medical field. *E. coli* is an organism that lives symbiotically

in the intestines of other organism, although some strains may cause gut diseases and sepsis (Hunter, 2008).

E. coli also has significant information in orthologous genes, which are present in various organisms such as humans, animals, plants and other bacteria. This suggests that this is an important model organism to be studied and will be kept being adopted in various experimental laboratories (Hunter, 2008). This means that along with the advances in microscopy and sequencing techniques, both the *E. coli* K-12 strain and B strain are the perfect candidates to study cellular structures and cellular processes, such as cell growth and division through computational and mathematical modelling of spatial and temporal bacterial organization (Kruse, 2012; Wanner, Finney, & Hucka, 2005).

Previous efforts to tackle the *E. coli* cell modelling problem have been extensively reviewed (Wanner et al., 2005) and include the creation of a common language to represent biological models, namely the Systems Biology Markup Language (Hucka et al., 2004), the development of numerous mathematical and empirical models found in the literature will have to be researched along with accessing specialized information stored in databases, such as the International E.coli Alliance Database Portal (Kröger, 2010) or the advances in the computational cell modelling.

In order to create realistic models of bacterial cell behaviour, it is necessary to understand the available information on bacterial spatial and temporal organization, namely the cell shape and size, kinetic models of cell motility, division and growth and models of location and functionality of subcellular structures at the molecule level (Kruse, 2012).

The next sub-chapters will concentrate on these topics, focusing on the *E. coli* species, but also making analogies with other bacterial species.

3.2.1. Bacterial Cell Morphology - Shape, Size and Spatial Arrangement

Bacterial cells can be classified by its shape and by their spatial organization. As was observed in Figure 6 and can be observed in Figure 12, *E. coli* has a rod-shape (bacillus), while other bacteria have shown a vast diversity of shapes, such as spherical (coccus), intermediate shapes (coccobacillus) or curved/corkscrew shapes (spirochete, spirillum and vibrio), each of them with its specific purpose (Cabeen & Jacobs-Wagner, 2005). Bacteria can also have a wide range of cell sizes (volumes that range from 0.02 to 400 μm^3), where even a vast variability can be observed within the same species (Koch, 1996; Zinder & Dworkin, 2006). These variations can be explained due to cell adaptation to external factors, such as lack of nutrients leading to starvation, situations of extreme temperatures (low and high) or of extreme dryness (Koch, 1996). It has been shown that the lower bound for cell size is maintained by the cellular mechanisms that cope and adapt to the environment, while the higher bound is normally limited by diffusion of nutrients along the cell. For example, studies using *E. coli* as a model organism have shown how temperatures between 22ºC and 42ºC affect two different *E. coli* strains in different growth media (Trueba, Spronsen, Traas, & Woldringh, 1982).

In terms of spatial arrangement, bacteria can be organized in single forms or be grouped in pairs (diplo prefix), in chains (strepto prefix), as can be seen in Figure 12. Cocci bacteria can also organize in groups of 4 (tetrad), 8, 16 or 32 (sarcinae) or in grape-like clusters (staphylo prefix). Bacilli bacteria can organize in palisade structures (side by side) or can be in unstructured spatial clusters. An example of the *E. coli* spatial arrangements is shown in Figure 12, visualized by phase-contrast microscopy obtained at the Laboratory of Biosystem Dynamics.

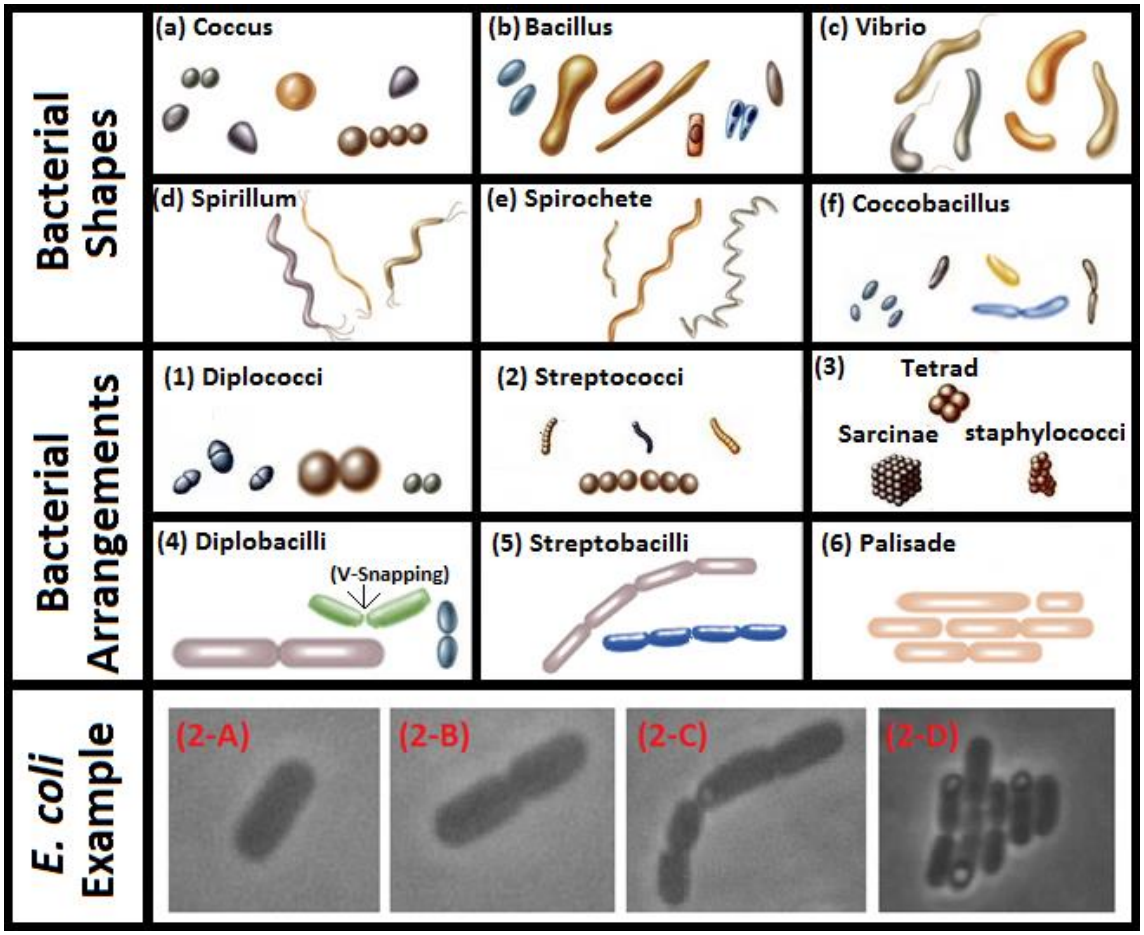

Figure 12 –Bacterial Shapes (a, b, c, d, e and f) and Bacterial Spatial Arrangements (1, 2, 3, 4, 5 and 6). *E. coli* examples: (2-A) Single bacillus; (2-B) diplobacilli; (2-C) streptobacilli; (2-D) palisade.

A typical bacterial cell envelop is mainly composed by a cytoplasmic membrane and peptidoglycan (also known as murein) cell wall. As can be seen in Figure 13, bacteria can also be divided in two groups regarding a fundamental difference in the cell envelope: Gram-negative and Gram-positive bacteria. In the first group (which is the case of *E. coli*) a bacterial outer membrane is also present (with intercalating pore-forming proteins, called porins), with lipopolysaccharides connected to the exterior of that outer wall. The interior of the outer wall is then connected to a very thin murein wall by a lipoprotein (Höltje, 2004). On the other hand, in the second group (which is the case of human pathogenic bacterium Streptococcus pneumonia), the cell envelope consists of a very thick murein wall (sometimes more than 10 times thicker than the first group) with teichoic acids spread across the murein. The Gram-positive bacteria also have a cytoplasmic membrane as the Gram-negative (Höltje, 2004). The structural differences and similarities between the cell envelope of Gram-negative and Gram-

positive bacteria is represented in Figure 13. The shape is maintained and determined by the way murein is incorporated during cellular elongation, especially in rod-shape organisms, such as *E. coli* (U Henning, Rehn, & Hoehn, 1973) and B. subtilis (Carballido-López & Formstone, 2007), as the murein is the main cell wall structure that supports the stress from the outside (Höltje, 1998), as computational physical models have been develop to study how defects in the murein can affect *E. coli* shape (and the shape robustness to murein damage) and how different murein defect patterns can build bacterial shape patterns such as curved rods and spirochaetes, as shown in Figure 13-3 (Casey Huang, Mukhopadhyay, Wen, Gitai, & Wingreen, 2008). Along with the cell wall, other cytoskeleton proteins are associated with bacterial shape, such as FtsZ (tubulin homologue), MreB (actin homologue) and crescentin (Cabeen & Jacobs-Wagner, 2005; Carballido-López & Formstone, 2007). These proteins influence how the cell wall is created and hydrolysed during cell growth and division, respectively, influencing their sizes, shapes and spatial and temporal organization (Carballido-López & Formstone, 2007). The role of cell size in bacterial growth has been discussed alongside a model for cell growth control in different nutrients, as seen in Figure 13-4 (Chien, Hill, & Levin, 2013).

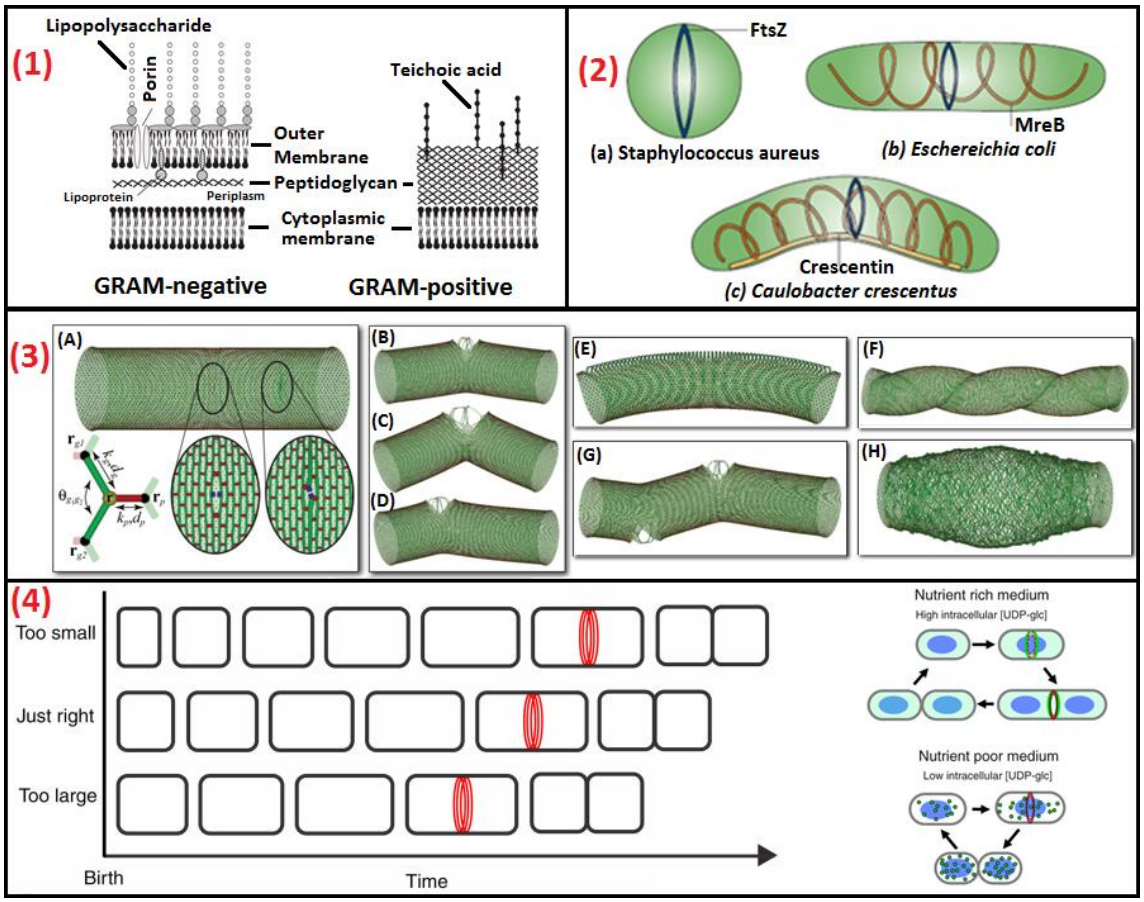

Figure 13 – (1) Structure of the cell envelope of Gram-negative (left) and Gram-positive (right) bacteria and their differences. Taken from (Höltje, 2004). (2) Bacteria cell shapes and cytoskeletal elements (Cabeen & Jacobs-Wagner, 2005). (3) Murein deformations (Casey Huang et al., 2008). (4) Cell size control by growth and division processes and how they can cope with a poor nutrient medium (Chien et al., 2013)

Bacterial morphology is closely related to important mechanisms to the bacterial cell activity, such as cell growth/elongation and cell division, so it is also important to understand how

these mechanisms are regulated in the time and space (Cabeen & Jacobs-Wagner, 2005; Carballido-López & Formstone, 2007; Ulf Henning, 1975).

3.2.2. Bacterial Cell Growth and Division

Bacterial cell cycle is normally divided in three stages, specifically the period between its "birth" and the initiation of DNA replication, the replication period where the cell increases its mass and size (Cell Growth) and finally the binary fission process into two new daughter cells (Cell Division), which will be repeated over the next generations (J. D. Wang & Levin, 2009), as shown in Figure 14-1.

The creation of new murein polymer can lead to cell growth through cell elongation, as murein is inserted in the sidewalls at the middle of the cell (see Figure 14-2-b) or at the poles (see Figure 14-2-c).

The creation of the division septum at the middle of the cell then leads to a division event (this is also the main process for cell growth in spherical cells, where cell elongation does not occur, as seen in Figure 14-2-a), where two daughter cells are created (Cabeen & Jacobs-Wagner, 2005).

Each of those processes have their own protein and enzymatic apparatus, which work in specific places of the cell wall (Cabeen & Jacobs-Wagner, 2005; Carballido-López & Formstone, 2007), as can be observed in Figure 14-3.

The FtsZ cytoskeleton protein along with various other proteins create the division septum at the middle of the cell (as two proteins MinC and SlmA that are present in the rest of the cell, inhibit the assembly of the FtsZ ring required for division (K. D. Young, 2010). The MreB cytoskeleton protein is mainly associated with the elongation apparatus, as mentioned in the previous section (K. D. Young, 2010).

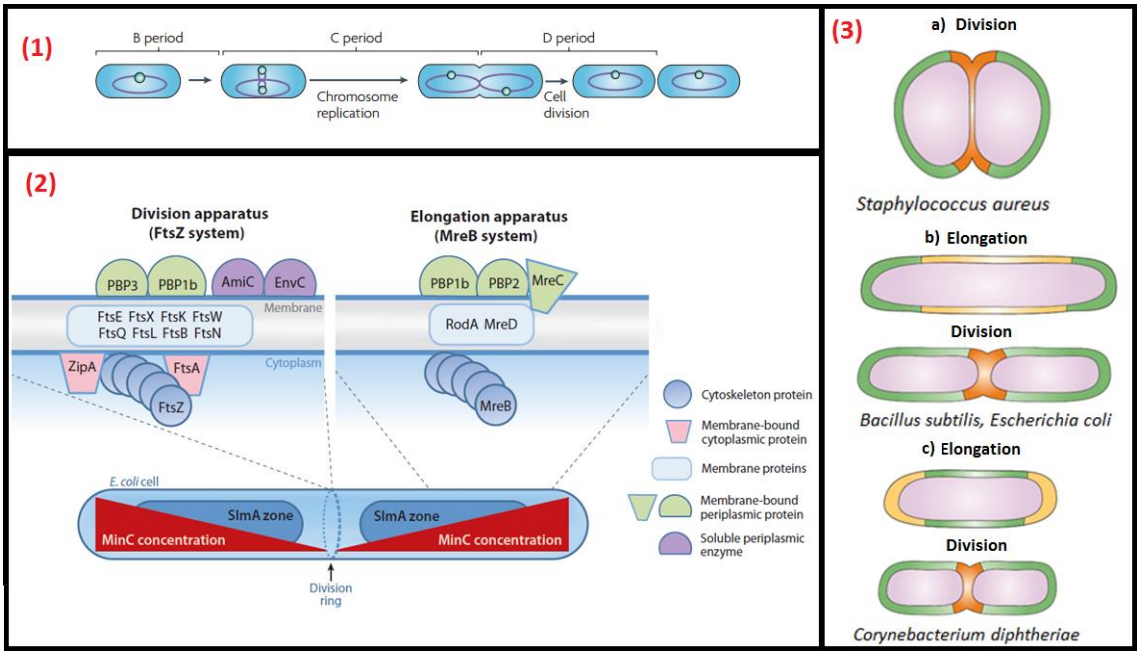

Figure 14 – (1) Bacterial Cell Cycle. Taken from (J. D. Wang & Levin, 2009). (2) Division and Elongation Protein apparatus. Taken From (K. D. Young, 2010). (3) Division and Elongation processes in different bacterial organism. Taken from (Cabeen & Jacobs-Wagner, 2005).

Mathematical models of the temporal and spatial organization of the bacterial cell cycle (Anderson, Mendelson, & Watkins, 2000; Fan, Tuncay, & Ortoleva, 2007; Lan, Wolgemuth, & Sun, 2007) are required to model cell elongation and cell division. These mathematical models arise from numerous experimental studies at the single cell level, especially using *E. coli* as a model organism (Kruse, 2012).

A mathematical model showed how the FtsZ ring acts as force generator in order to predict the contraction speed and force and how the cell shape arises, as shown in Figure 15-1 (Lan et al., 2007). Another model showed how the chromosome can be segregated during cell elongation and cell division, implying an influence of the MreB cytoskeleton protein that is involved in the bacterial elongation apparatus. This process modifies the membrane pressure and influences the DNA segregation, as shown in Figure 15-2-A (Fan et al., 2007). It is important to note that they use cell shape assumptions where the width of the cell doesn't change over time, and that cells do not deform, as shown in Figure 15-2-A and Figure 15-2-B. As was previously seen in Figure 13-3, a computational model of peptidoglycan insertion along the cell wall showed how the cells can grow and maintain their shapes (Casey Huang et al., 2008). Another model of cell growth, based on *Bacillus subtilis,* showed that the rate of cell division can be dependent not only on cell size but also on the its age (Anderson et al., 2000).

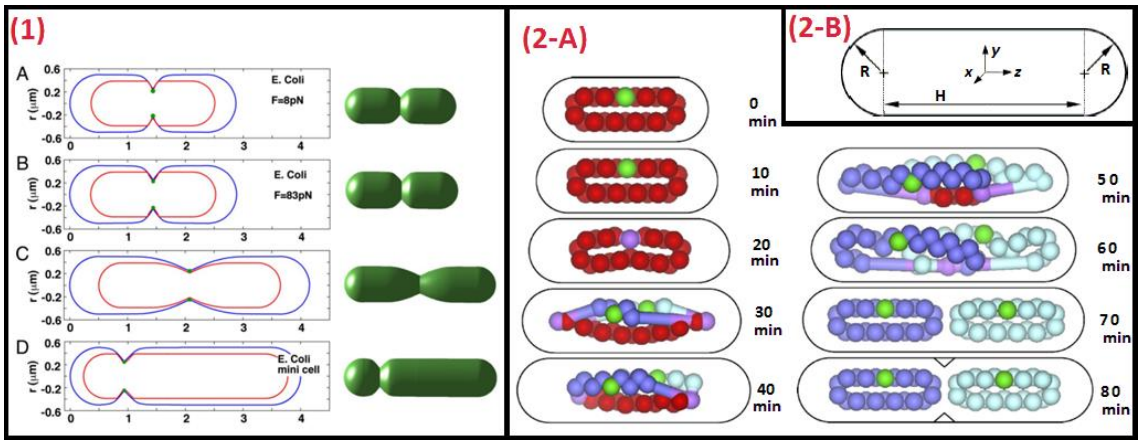

Figure 15 – (1) Taken from (Lan et al., 2007). (2) (A) Models of DNA Segregation. (B) Computational model of the *E. coli* rod shape. Taken from (Fan et al., 2007).

Bacterial growth as a colony can also be dependent on the capability to move in the direction of more favourable conditions, which at its basic form is normally associated with Brownian random movement or active movement towards a specific gradient, e.g. chemicals (chemotaxis) and temperature (thermotaxis) (Lauffenburger, 1982).

3.2.3. Bacterial Cell Motility

As stated in the previous chapter, each individual bacteria cell performs a random walk in the absence of external factors, but when these factors are present, active bacterial cell movement is activated, e.g., towards nutrient sources or by moving away from certain toxins

or stress conditions (Lauffenburger, 1982). The random walk is characterized by a smooth swim/run in a determined direction for a few seconds followed by a tumble (a change in direction, which last tenths of a second (Berg, 2004)

Bacteria such as *E. coli* have developed external structures for motility purposes (the flagella) that help them to propel themselves by acting as cellular motors (Berg, 2004; Tindall, Maini, Porter, & Armitage, 2008). The active motility mechanism is controlled by a distinct biochemical network (see in Figure 16-1) that transmits information from the extracellular environment, gathered by the membrane receptors to the flagella (Wadhams & Armitage, 2004).

While each cell behave independently, bacteria populations also display collective behaviour, as was previously observed in Figure 12, where bacteria was shown to be spatially arranged in various configurations. Bacteria can even organize in large clusters, due to its high rate of division, which can have specific macroscopic motility properties. Specifically they have been studied using individual *E. coli* cells were tracked inside those clusters, employed under the fluorescence microscopy (Mittal, Budrene, Brenner, & Van Oudenaarden, 2003).

The boundaries of the cluster is maintained by supressing the direction change (tumble) of individual bacteria in the centre of the cluster, as seen in Figure 16-2, while it is restored for cells at the edge of the cluster (Mittal et al., 2003). These experimental findings were confirmed by using a computational simulation, which confirmed that the tumble rate and the cluster morphology are determined by the sensory memory of cells (Mittal et al., 2003).

As previously stated in this Chapter introduction, to develop computational models of these mechanisms, it is essential to have mathematical models or empirical models based on empirical observations that are capable of describing these mechanisms.

In the case of cell motility, such mathematical models have been extensively studied and reviewed (Tindall et al., 2008). These models incorporate how an individual bacterial cell behaviour affect the population by interacting with other cell and they also incorporate how bacteria interact with the environment (Tindall et al., 2008).

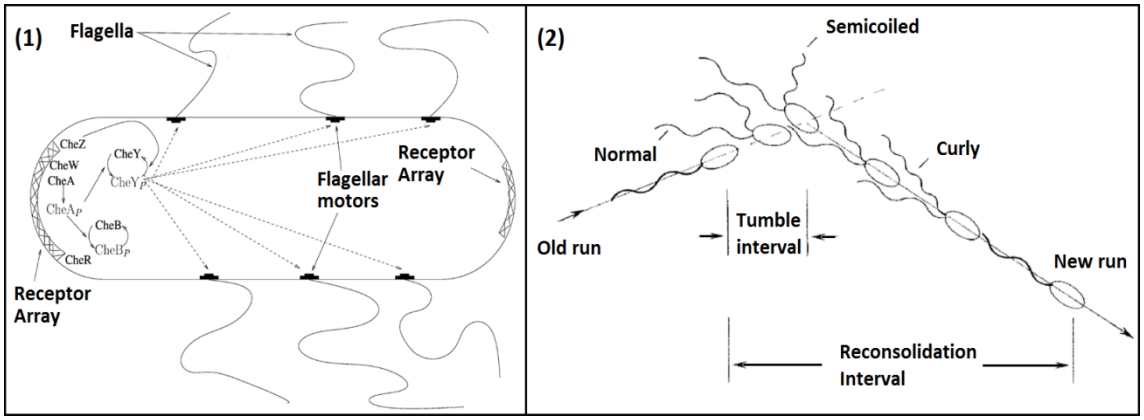

Figure 16 – (1) Motility biochemical network and anatomy of the flagellar systems. (Tindall et al., 2008). (2) Run and Tumble (Berg, 2004).

Unlike bacteria motility, the mechanism by which eukaryotic cells can migrate is still a problem that isn't totally solved, especially as the mathematical models of cell migration are still being envisioned (Zhang et al., 2013), which are essential for the simulation tools of temporal cellular activity of eukaryotic cells.

The previously described morphological processes are associated with functional processes that occur inside the bacteria, which can also be observed based on functional microscopy by examining the synthesis of Ribonucleic acids (RNAs) and proteins by labelling these molecules with fluorescent probes.

3.2.4. Bacterial Internal Functions and Cellular Structures

Gene expression is the process of synthesizing a functional gene product, by using the specific gene information. This process starts with Transcription, where RNA transcript is synthesized by using the information contained in a determined region of the DNA, which is executed by the RNA polymerase (RNAp).

In prokaryotes, as opposed to eukaryotes, the produced RNA doesn't need to be processed or transported to other areas of the cell, which means that the RNA is ready for Translation, the process in which ribosomes present in the cell produces a specific protein from the messenger RNA that was synthesized in the previous mechanism (Transcription).

Various microscopy and biochemical techniques are used to observe the spatial organization of bacterial DNA (which normally is condensed in the Nucleoid), RNA and proteins, as can be observed in Figure 17-1, especially at the single cell and single molecule level (Weng & Xiao, 2014). One of such techniques was based on the fusion of RNA, the bacteriophage MS2 coat protein and GFP proteins, creating a RNA-MS2-GFP complex (Golding et al., 2005) or using the λ bacteriophage to create such complexes (Baron-Benhamou, Gehring, Kulozik, & Hentze, 2004).

Other reviewed techniques include the use of single molecule FISH to detect RNA's, photo-activated localization microscopy (PALM) to detect proteins, fluorescent repressor-operator system (FROS) to detect DNA, stochastic optical reconstruction microscopy and single-molecule-tracking to detect proteins and nucleic acid targets (Weng & Xiao, 2014).

The visualization of the *E. coli* nucleoid in 4 dimensions (spatial and time) was done by HupA-mCherry staining and using phase contrast and wide-field epifluorescence microscopy, while the FROS technique was used to detect the origin of replication (oriC), as seen in Figure 17-2. PALM has been used to detect the protein Kaede, at the single copy level, and study its spatial distribution and diffusion, alongside the detection of ribosomes and chromosomal DNA, using wide-field microscopy, as seen in Figure 17-3 (Bakshi, Bratton, & Weisshaar, 2011).

Fluctuations of the RNA and protein numbers have been detected (McAdams & Arkin, 1997; Ozbudak, Thattai, Kurtser, Grossman, & van Oudenaarden, 2002), showing that these

processes have a stochastic nature and can have an important effect in the cell behaviour (Michael B Elowitz, Levine, Siggia, & Swain, 2002).

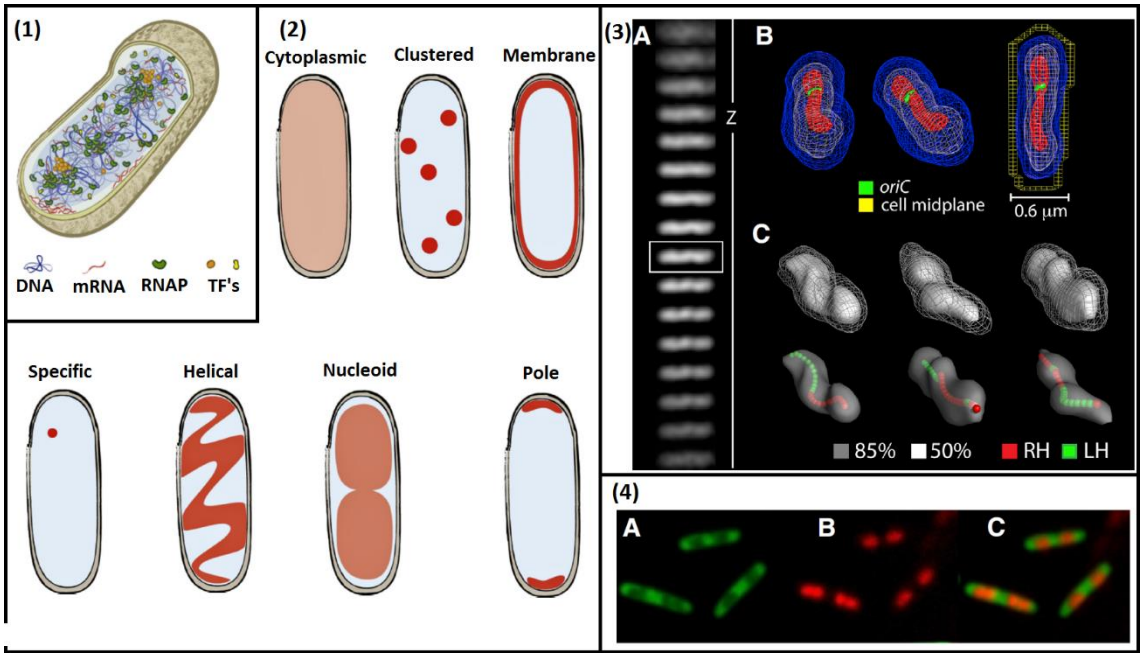

Figure 17 – Visualization of internal structures. (1) Distribution of transcription molecules (Weng & Xiao, 2014). (3) Nucleoid Visualization in 3D (Fisher et al., 2013). (4) Ribosome distribution (Bakshi et al., 2011).

The observation of the spatial distribution of these molecules at the single cell level, and studying the kinetics of segregation to the cell pole and partitioning during division can establish a correlation between those events and cellular aging, i.e., loss of reproductive vitality, which is one of the main objectives of the project in which this PhD work is integrated, which is referenced in Chapter 4.5.

3.3. Cellular Aging

Prokaryotic organisms, such as bacteria, appear to be functionally immortal, when in suitable environments, as each cell perpetuates itself by dividing into two daughter cells with the same genotype as the mother cell.

For that, the stockpiling of unwanted substances or degradation of internal components needs to be managed effectively with mechanisms such as the deliberate asymmetry in the partitioning of intracellular material during division (Stewart, Madden, Paul, & Taddei, 2005). This was first encountered in unicellular organisms that exhibit highly asymmetric divisions, such as yeast and *Caulobacter crescentus* (Ackermann, Stearns, & Jenal, 2003; Jazwinski, 2002).

Other unicellular organisms, such as *Escherichia coli*, have apparently a morphologically symmetrical division, making the aging process less straightforward to comprehend. Studies of individual cell lineages have shown that two morphologically identical sisters can exhibit functional asymmetries (Stewart et al., 2005). The detected asymmetry can be indicative of aging, as unwanted protein aggregates tend to concentrate at the older pole of the mother cell

and that accumulation can cause a slower division rate of the daughter cells (Stewart et al., 2005).

Another study using *E. coli* also investigated how a different complex construct (Tsr-GFP fusion protein) accumulated at the old poles, which can be used to identify the old and the new poles, when no information about the cell ancestry is available (Ping, Weiner, & Klecknerr, 2009). One related study also concluded that the asymmetric deposition of aggregates can increase the bacterial population fitness and also resulting in higher rates of growth in the daughter cells that have less damaged aggregates, due to the misfolding of proteins (Winkler et al., 2010).

The link between cellular aging and many age-related diseases such as Alzheimer's disease, spongiform encephalopathies (Dobson, 2001), Huntington's disease, Alexander's disease, cataracts (Harding, 2002), Motor Neuron disease and Frontotemporal Dementia (Wolozin, 2012) further enhances the importance of studying the aging process at a simpler level, as it will then shed light how these mechanisms work in complex organisms.

To make accurate conclusions about these processes, even at a simple organism level such as the *E. coli*, we need to use reliable tools and methods for image processing, capable of detecting and tracking individual molecules at the single-cell level. To validate such image processing tools, we need to apply the developed image generator to realistic spatial and temporal models of bacterial organization.

Chapter 4

Research Methodology

This section presents the aimed contribution to answer the proposed main research question, a detailed work plan together with a schedule for the four years that compose the doctoral program (referencing the scientific method), the proposed validation methodology, a tentative dissemination plan, including article publication in indexed journals and participation in conferences and courses and the integration with other research projects.

4.1. Aimed Contribution

To answer the proposed main research question, the proposed hypothesis and methodology approached aims to contribute by creating bacterial morphological and functional activity models and produce a set of "gold-standard" time-lapsed microscopy images. We will use Monte Carlo methods make numerical simulations of the time evolution concerning the entire system. For this we will need to specify a reaction for each of the temporal processes that we want to simulate and the rate constants associated with these processes.

We will develop empirical models of realistic cell shapes, based on the skeletons of bacterial images. These skeletons, which represent the major axis, are then dilated with a disk form with the radius of the minor axis to create the rod shape of *E. coli* cells. This is similar to previously developed parametric shape model where the centre line is just a straight line (Lehmussola et al., 2011). Another way to create the centre line is to create a random walk between the two points of the major axis. Other approaches include the use of Machine Learning techniques to simulate bacterial shapes by using non parametric shape simulation (Lehmussola et al., 2011). That approach might need to include the spatial arrangement of bacteria, as the shape of neighbouring bacteria can influence the shape of the others.

Mathematical modelling of the rod shape of *E. coli* cells can be done by defining a rectangle (black line in Figure 18) with the length of the major axis (horizontal green line in Figure 18) and the height of the minor axis (vertical green line in Figure 18) and taking the convex hull of two equal semi-circles with the radius of half of the minor axis and placing their centres at the major axis line, by a distance of half of the minor axis from the border (see Figure 18).

Figure 18 – Mathematical modelling of the rod shape of *E. coli* cells (red colour). Minor and Major Axis in Green. The semi circles have a radius defined as half of the minor axis.

Cell Growth can be defined as a linear temporal process until it reaches a division event (when it reaches two times the initial size) as the kinetic constant of *E. coli* cell doubling time has been reported to be around 3600s in favourable conditions (Lloyd-Price, Tran, & Ribeiro, 2014; Yu et al., 2006).

We also have to simulate the spatial modelling of cell growth (which for *E. coli* cells is along the major axis and has to be done by forming new pixels on the tips), as seen in Figure 19-A. When bacteria are in clusters, the cell growth might have to be stopped or will have to be done by pushing the other touching cell or by bending and growing in different directions.

During Cell Division we will need to "cut" the parent cell in two daughter cells, as observed in Figure 19-B. If these daughter cells are organized in specific arrangement (as seen in Chapter 3.2.1 Cell Modelling – Spatial and Temporal Organization of Bacteria) we need to create a cluster to maintain these cells in that specific arrangement.

The distribution of bacterial cell sizes can be obtained from the literature review of experimental studies. It should be noted that these distributions can be dependable on the applied external conditions.

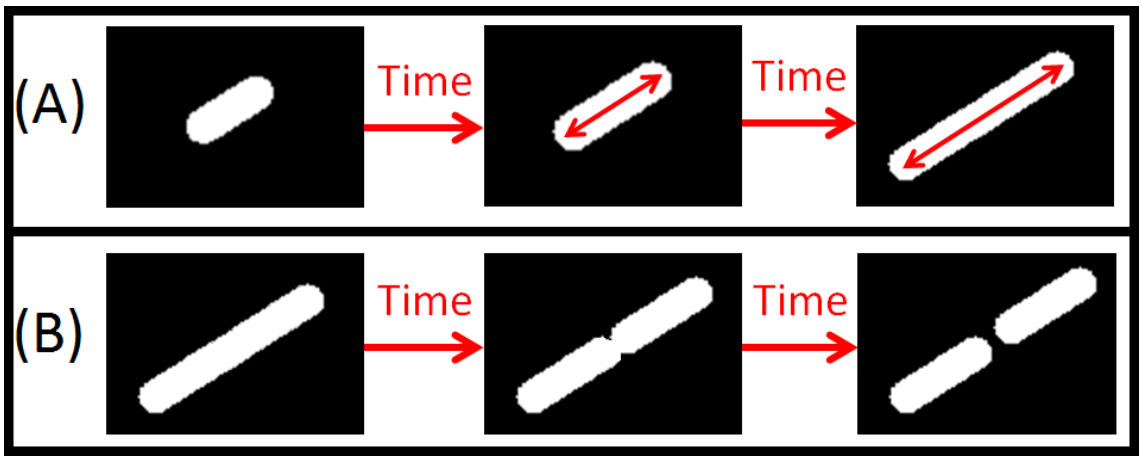

Figure 19 – (A) Spatial simulation of cellular growth along the major axis of the cell. (B) Spatial simulation of cell division at the centre of the major axis.

Cell Movement will be modelled by making a geometric translation of the centre of the bacteria (moving all the other pixels in that direction, as seen in Figure 20-C) or by rotation of the axis angle of the bacteria (changing the orientation angle as seen in Figure 20-B) without moving the centre of the bacteria or by mixing both types of movement.

As in the modelling of bacteria growth, sometimes the movement will be blocked by the presence of other bacteria, where in that case we can also stop the movement, push the other cell or bend the moving bacteria.

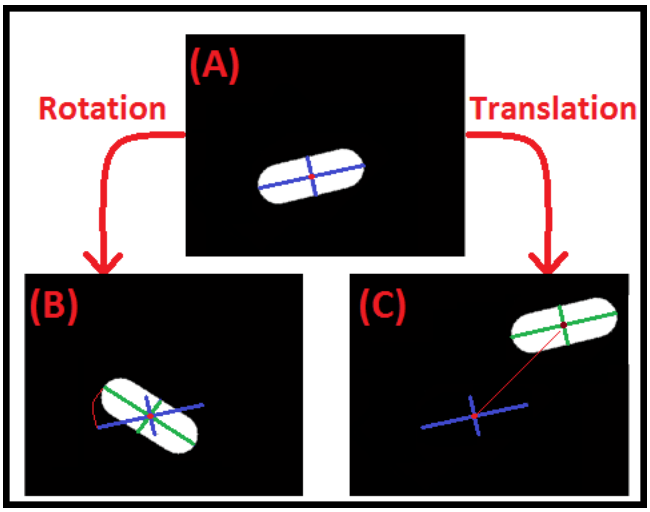

Figure 20 – (A) Initial state of the bacteria before movement. (B) Rotation movement, the centre remains in the same place and the axis move their orientation angle. (C) Translation movement, the centre moves, but the axis remains with the same angle. Note: In both (B) and (C) Blue line represents the initial state of the Axis and the green line the new state.

The modelling of internal structures will be focused on simulating the nucleoid, which does not have a parametric shape, so the best approach is to use training shapes from real microscopy images that show the nucleoid (DAPI nucleic staining or mCherry staining). Bacterial internal temporal processes such as the production of fluorescent RNA complexes and proteins is being studied in the Laboratory of Biosystem Dynamics (LBD) from Tampere University of Technology, so we can use that information to make a spatial and temporal model of these processes to produce fluorescence images.

In terms of external environmental conditions, we can consider the bacterial response to external factors such as temperature (heat-shock, cold-shock), pH stress, oxidative stress, nutritional stress or even exposure to antibiotics and how they influence the processes and distributions that we want to simulate. This can be done by using the information acquired in the Literature or by collaborating with the LBD group to perform experimental studies to gather such statistics.

The simulation of the image acquisition processes has already been done extensively in previous works, so we will need to adapt these methodologies depending on the illumination, contrast and florescence methods that we want to simulate.

We aim to contribute to the validation of image processing tools but also to help as a simulation tool in studies with *E. coli* cells, namely those related to functional aging and how it is related to the disposal of molecular aggregates during cell division, as using a simulation toolbox we can generate almost an unlimited number of images to be processed, where in a real laboratory setting this cannot happen, as doing experiments have costs and sometimes can be time-consuming.

4.2. Detailed Work Plan and Scheduling

The detailed planned work was developed accordingly to the classical research method as it is presented in Figure 21.

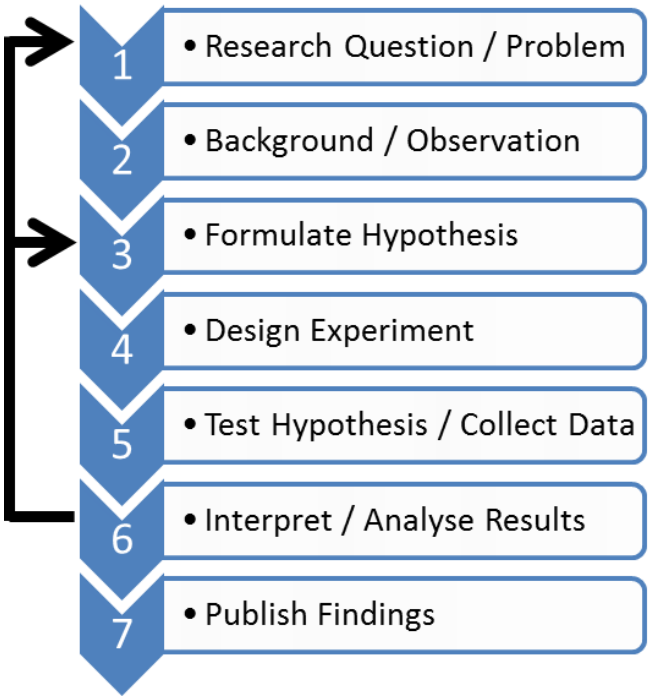

Figure 21 - Classical research method. Adapted from the hand-outs of the Scientific Research Methodologies and Techniques course of the PhD Program in Electrical and Computer Engineering, by Professor Luis Camarinha-Matos.

The first year will be focused on completing the courses of the PhD Program in Electrical and Computer Engineering and the first three steps in the classical research method.

The first step started the identification of the open research problems and the definition of the main research question.

The second step includes the literature review of previous research works, generating a critical analysis on topics related to microscopy imaging, such as multimodal and multidimensional imaging, microscopy imaging processing methods and its validation by generating simulated microscopy images. The state of the art also includes the modelling of bacteria cellular organization, both spatial and temporal.

The third step includes the formulation of the hypothesis, related to the main research question and according to the current point in the state of the art.

The second year will be focused on the fourth step of designing the experiments, which include the development of bacterial morphological models, namely the cell shape, size and spatial arrangement and bacterial temporal biological functions such as cell division, cell growth and motility. To test the hypothesis and collect data on this first part of the model, an artificial microscopy image generator will be developed, producing morphological images, such as phase contrast and differential interference contrast, which will have to be interpreted and

analysed (sixth step in the classical research method). These work packages will continue through the first half of the third year, as shown in Figure 22.

During the third year, the modelling of internal structures of bacteria will be the main focus, which includes the production of RNA and proteins, but also the modelling of internal structures such as the nucleoid. This will prompt the development of the second stage of the artificial microscopy image generator, which will produce fluorescence images, by simulating the production of fluorescent proteins or using different staining techniques (e.g. DAPI nucleic staining), highlighting the bacterial internal structures. It is possible to integrate morphological and functional information by fusing both types of images, which will then be used to validate the image fusion techniques that are being developed in Project SADAC.

The first part of the fourth year will be concentrated in final model validations and refinements of the developed tool and its resulting images. The final part of the year will be directed to the writing of the thesis, which is also part of the seventh step, which is described in more details in Sub-Section 4.4 - Dissemination Plan, regarding also the plan to publish the findings in indexed journals and the participation in international conferences.

A tentative scheduling for the four years (from 2014 to 2017) is presented in Figure 22, representing the above information of the work plan.

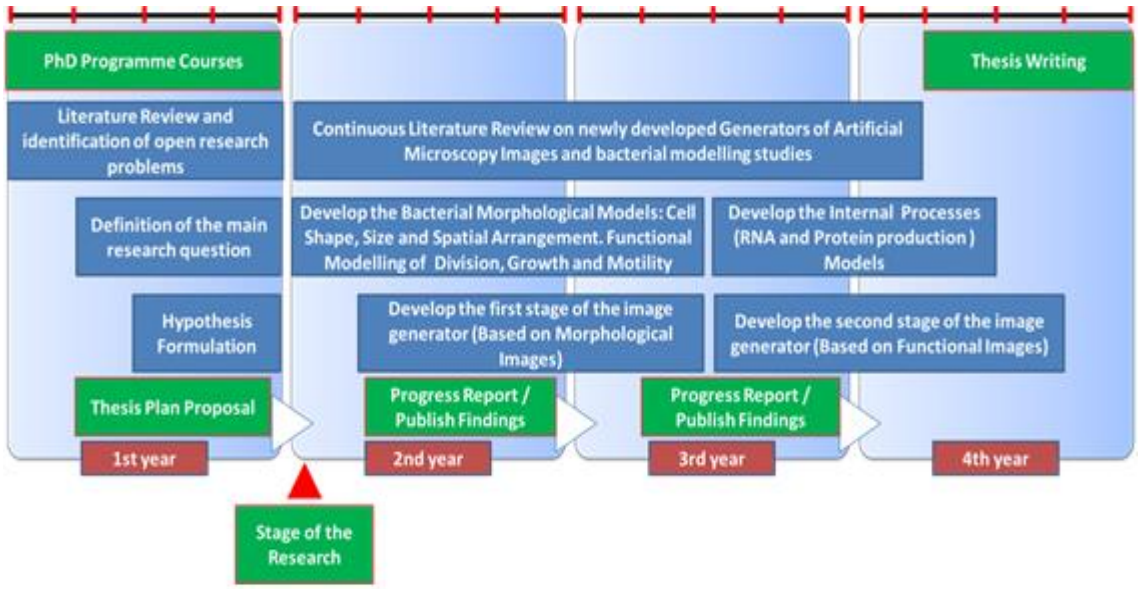

Figure 22 – Scheduling for the four years of the PhD programme and the research work.

4.3. Validation Methodology

Confirmation of a research hypothesis and its validation is the final objective in a research work, so it is very important to plan ahead how to answer the research question, validating the expected results.

This validation process can be done by peer validation, by submitting the developed work to peer-reviewed journals indexed to the Web of Science, to renowned international conferences (with peer-reviewing process and with relevant technical programmes). The next Sub-section

(Dissemination Plan) presents a group of identified dissemination channels relevant to this research work.

Qualitative validation can be done by surveying groups of microscopy experts. These experts will be acquainted and contacted via the participation in relevant international conferences and workshops, such as EMBO Practical Course “Microscopy, Modelling and Biophysical Methods” and the International Workshop on Computational Systems Biology. This survey can evaluate the quality of the produced artificial images and compare them with real microscopy images.

It is also expected to develop a website for the publication of image benchmark in various environmental conditions, which can be accessed to validate novel image processing algorithms. This website could also serve as a contact point to the microscopy experts that will make the qualitative validation, which could use the site to directly evaluate the images.

The integration of this research work on a research project funded by FCT - Fundação para a Ciência e a Tecnologia will assist in the validation of this work, as the artificial microscopy images will be compared with the real microscopy images that are being produced during Project SADAC. This project also requires the development of image processing tools such as novel algorithms of single cell segmentation and spot identification which will be validated using the image generation tool. Also regarding the scope of the project, artificial image generator can also be used to simulate solutions of a null-model that can be compared with measurements done in laboratory (e.g. to study the effects of the nucleoid by removing it from the cells).

Using the real *E. coli* images collected in various conditions in the Laboratory of Biosystem Dynamics, we can compare the obtained artificial images, validating the artificial image generation tool. For this, we need to identify correct model indicators such as cell sizes and shapes (distributions of bacterial spatial organization), motility velocity, division and growth rates (distributions of bacterial temporal organization), and also the production and localization of subcellular structures (fluorescent proteins, nucleoid, etc).

4.4. Dissemination Plan

This research work dissemination will rely on submitting/publishing in at least in three International Journals indexed in the Web of Science and International Conferences relevant to the research work. In this context, Table 3 illustrates a tentative dissemination plan, detailing the name of the conferences, journals and courses and the expected participation dates.

Table 3 – Dissemination Plan, with a selection of courses, conferences and journals, relevant to the research area.

Dissemination Channel	Name	Date
Courses	EMBO Practical Course, Heidelberg, Germany Microscopy, Modelling and Biophysical Methods	7-20 September 2014
Conferences	11th International Workshop on Computational Systems Biology (11th WCSB), Lisbon, Portugal • Participation as Local Organizer • Abstract of preliminary work was approved for Oral Presentation and published in Conference Proceedings	15/16 May 2014
	2015 IEEE 4th Portuguese Meeting on Bioengineering (ENBENG) • Full Paper of preliminary resuls was approved for Oral Presentation and published in Conference Proceedings	26-28 February 2015
	DoCEIS 2015 - Doctoral Conference on Computing, Electrical and Industrial Systems (Poster presentation) • Participation as Local Organizer in Financial Committee • Presentation of Poster of Thesis Plan	13-15 April 2015
	International Conference on Bioinformatics and Biomedicine (BIBM), IEEE	2016 (expected)
	DoCEIS 2016 - Doctoral Conference on Computing, Electrical and Industrial Systems (Oral presentation)	2016 (expected)
	International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics - CIBB	2017 (expected)
Journals	Cytometry. Part A : the journal of the International Society for Analytical Cytology	2015 (expected)
	IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)	2016 (expected)
	BMC bioinformatics	2017 (expected)

4.5. Integration with other Research Activities

The current work is being developed in the CA3 group (CA3 - Computational Intelligence Research Group) of CTS, UNINOVA in cooperation with Laboratory of Biosystem Dynamics (LBD) from Tampere University of Technology. This work is integrated in Project SADAC – Study of the kinetics of asymmetric disposal of aggregates in cell division and its correlation to functional aging from in vivo measurements, one event at a time – with the reference PTDC/BBB-MET/1084/2012, funded by FCT - Fundação para a Ciência e a Tecnologia.

This work is also supported by the Portuguese Foundation for Science and Technology (FCT/MCTES) through funded PhD Scholarship (ref. SFRH/BD/88987/2012).

References

Ackermann, M., Stearns, S. C., & Jenal, U. (2003). Senescence in a bacterium with asymmetric division. *Science*, *300*(5627), 1920. doi:10.1126/science.1083532

Al-Kofahi, Y., Lassoued, W., Lee, W., & Roysam, B. (2010). Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Transactions on Biomedical Engineering*, *57*(4), 841–52. doi:10.1109/TBME.2009.2035102

Ambühl, M. E., Brepsant, C., Meister, J.-J., Verkhovsky, a B., & Sbalzarini, I. F. (2012). High-resolution cell outline segmentation and tracking from phase-contrast microscopy images. *Journal of Microscopy*, *245*(2), 161–70. doi:10.1111/j.1365-2818.2011.03558.x

Anderson, K. R., Mendelson, N. H., & Watkins, J. C. (2000). A new mathematical approach predicts individual cell growth behavior using bacterial population information. *Journal of Theoretical Biology*, *202*(1), 87–94. doi:10.1006/jtbi.1999.1051

Bahlmann, F. H., Degroot, K., Duckert, T., Niemczyk, E., Bahlmann, E., Boehm, S. M., … Fliser, D. (2003). Endothelial progenitor cell proliferation and differentiation is regulated by erythropoietin Rapid Communication. *Kidney International*, *64*, 1648–1652.

Bakshi, S., Bratton, B. P., & Weisshaar, J. C. (2011). Subdiffraction-limit study of Kaede diffusion and spatial distribution in live Escherichia coli. *Biophysical Journal*, *101*(10), 2535–44. doi:10.1016/j.bpj.2011.10.013

Baron-Benhamou, J., Gehring, N. H., Kulozik, A. E., & Hentze, M. W. (2004). Using the λN peptide to tether proteins to RNAs. *Methods in Molecular Biology*, *257*, 135–54. doi:10.1385/1-59259-750-5:135

Berg, H. C. (2004). *E. coli in Motion*. (E. Greenbaum, Ed.) (pp. 1–34). Springer-Verlag Berlin Heidelberg.

Bjornsson, C. S., Lin, G., Al-kofahi, Y., Smith, K. L., Shain, W., & Roysam, B. (2009). Associative image analysis: a method for automated quantification of 3D multi-parameter images of brain tissue. *Journal of Neuroscience Methods*, *170*(1), 165–178. doi:10.1016/j.jneumeth.2007.12.024.Associative

Bolte, S., & Cordelières, F. P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy*, *224*(3), 213–232. doi:10.1111/j.1365-2818.2006.01706.x

Bonnet, N. (2000). Artificial intelligence and pattern recognition techniques in microscope image processing and analysis. *Advances in Imaging and Electron Physics*, *514*(114), 1–77.

Bonnet, N. (2004). Some trends in microscope image processing. *Micron (Oxford, England : 1993)*, *35*(8), 635–653. doi:10.1016/j.micron.2004.04.006

Brown, S., Yarden, O., Gollop, N., Chen, S., Zveibil, A., Belausov, E., & Freeman, S. (2008). Differential protein expression in Colletotrichum acutatum : changes associated with

reactive oxygen species and nitrogen starvation implicated in pathogenicity on strawberry. *Molecular Plant Pathology*, *9*(2), 171–190. doi:10.1111/J.1364-3703.2007.00454.X

Buck, T. E., Li, J., Rohde, G. K., & Murphy, R. F. (2012). Toward the virtual cell: automated approaches to building models of subcellular organization “learned” from microscopy images. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *34*(9), 791–9. doi:10.1002/bies.201200032

Cabeen, M. T., & Jacobs-Wagner, C. (2005). Bacterial cell shape. *Nature Reviews. Microbiology*, *3*(8), 601–10. doi:10.1038/nrmicro1205

Carballido-López, R., & Formstone, A. (2007). Shape determination in Bacillus subtilis. *Current Opinion in Microbiology*, *10*(6), 611–6. doi:10.1016/j.mib.2007.09.008

Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., … Sabatini, D. M. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, *7*(10), R100. doi:10.1186/gb-2006-7-10-r100

Casey Huang, K., Mukhopadhyay, R., Wen, B., Gitai, Z., & Wingreen, N. S. (2008). Cell shape and cell-wall organization in Gram-negative bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19282–19287.

Chapnick, J. V., Noz, M. E., Maguire, G.Q., J., Kramer, E. L., Sanger, J. J., Birnbaum, B. A., & Megibow, A. J. (1993). Techniques for multimodality image registration. In *Bioengineering Conference, 1993., Proceedings of the 1993 IEEE Nineteenth Annual Northeast* (pp. 221–222).

Chien, A., Hill, N. S., & Levin, P. A. (2013). Cell Size Control in Bacteria. *Current Biology*, *22*(9), 1–23. doi:10.1016/j.cub.2012.02.032.Cell

Chowdhury, S., Kandhavelu, M., Yli-Harja, O., & Ribeiro, A. S. (2013). Cell segmentation by multi-resolution analysis and maximum likelihood estimation (MAMLE). *BMC Bioinformatics*, *14 Suppl 1*(Suppl 10), S8. Retrieved from http://www.biomedcentral.com/content/pdf/1471-2105-14-S10-S8.pdf

Coelho, L. P., Shariff, A., & Murphy, R. F. (2009). Nuclear Segmentation In Microscope Cell Images A Hand-Segmented Dataset And Comparison Of Algorithms. In *Proc IEEE Int Symp Biomed Imaging* (pp. 518–521). doi:10.1109/ISBI.2009.5193098.NUCLEAR

Coutu, D. L., & Schroeder, T. (2013). Probing cellular processes by long-term live imaging--historic problems and current solutions. *Journal of Cell Science*, *126*(Pt 17), 3805–15. doi:10.1242/jcs.118349

Danuser, G. (2011). Computer vision in cell biology. *Cell*, *147*(5), 973–8. doi:10.1016/j.cell.2011.11.001

Deshmukh, M., & Bhosle, U. (2011). A survey of image registration. *International Journal of Image Processing*, *5*(3), 245–269.

Dobson, C. M. (2001). The structural basis of protein folding and its links with human disease. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *356*(1406), 133–145. doi:10.1098/rstb.2000.0758

Du, X., & Dua, S. (2010). Segmentation of fluorescence microscopy cell images using unsupervised mining. *The Open Medical Informatics Journal*, *4*, 41–9. doi:10.2174/1874431101004020041

Eils, R., & Athale, C. (2003). Computational imaging in cell biology. *The Journal of Cell Biology*, *161*(3), 477–81. doi:10.1083/jcb.200302097

Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, *403*(6767), 335–8. doi:10.1038/35002125

Elowitz, M. B., Levine, A. J., Siggia, E. D., & Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science*, *297*(5584), 1183–6. doi:10.1126/science.1070919

Fan, J., Tuncay, K., & Ortoleva, P. J. (2007). Chromosome segregation in Escherichia coli division: a free energy-driven string model. *Computational Biology and Chemistry*, *31*(4), 257–64. doi:10.1016/j.compbiolchem.2007.05.003

Fisher, J. K., Bourniquel, A., Witz, G., Weiner, B., Prentiss, M., & Kleckner, N. (2013). Four-dimensional imaging of E. coli nucleoid organization and dynamics in living cells. *Cell*, *153*(4), 882–95. doi:10.1016/j.cell.2013.04.006

Fusco, D., Accornero, N., Lavoie, B., Shenoy, S. M., Blanchard, J.-M., Singer, R. H., & Bertrand, E. (2003). Single mRNA Molecules Demonstrate Probabilistic Movement in Living Mammalian Cells. *Current Biology*, *13*(2), 161–167. doi:10.1016/S0960-9822(02)01436-7

Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature*, *403*(6767), 339–42. doi:10.1038/35002131

Glasbey, C. A., Buildings, K., Eh, E., & Martin, N. J. (1996). Multimodal microscopy by digital image processing. *Journal of Microscopy*, *181*(3), 225–237.

Golding, I., Paulsson, J., Zawilski, S. M., & Cox, E. C. (2005). Real-time kinetics of gene activity in individual bacteria. *Cell*, *123*(6), 1025–1036. doi:10.1016/j.cell.2005.09.031

Gordon, A., Colman-lerner, A., Chin, T. E., Benjamin, K. R., Yu, R. C., & Brent, R. (2007). Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nature Methods*, *4*(2), 175–181. doi:10.1038/NMETH1008

Gotelli, N. J., & Bldg, S. B. (1996). Null versus neutral models: what's the difference? *Ecography*, *29*(5), 793–800.

Graner, F., & Glazier, J. (1992). Simulation of biological cell sorting using a two-dimensional extended Potts model. *Physical Review Letters*, *69*(13), 2013–2016. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10046374

Grigoryan, M., Hostetter, G., Kallioniemi, O., & Dougherty, E. (2002). Simulation Toolbox for 3D-FISH Spot-Counting Algorithms. *Real-Time Imaging*, *8*(3), 203–212. doi:10.1006/rtim.2001.0280

Häkkinen, A., Muthukrishnan, A.-B., Mora, A., Fonseca, J. M., & Ribeiro, A. S. (2013). CellAging: a tool to study segregation and partitioning in division in cell lineages of Escherichia coli. *Bioinformatics (Oxford, England)*, *29*(13), 1708–1709. doi:10.1093/bioinformatics/btt194

Hand, A. J., Sun, T., Barber, D. C., Hose, D. R., & Macneil, S. (2009). Automated tracking of migrating cells in phase-contrast video. *Journal of Microscopy*, *234*(1), 62–79.

Harding, J. J. (2002). Viewing molecular mechanisms of ageing through a lens. *Ageing Research Reviews*, *1*(3), 465–479. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12067598

Hartwell, K., Miller, P. G., Mukherjee, S., Kahn, A. R., Stewart, A. L., Logan, D. J., … Golub, T. (2013). Niche-based screening identifies small-molecule inhibitors of leukemia stem cells. *Nature Chemical Biology*, *9*(12), 840–8. doi:10.1038/nchembio.1367

Henning, U. (1975). Determination of cell shape in bacteria. *Annual Review of Microbiology*.

Henning, U., Rehn, K., & Hoehn, B. (1973). Cell envelope and shape of Escherichia coli K12. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(7), 2033–6. Retrieved from /pmc/articles/PMC433659/?report=abstract

Höltje, J.-V. (1998). Growth of the Stress-Bearing and Shape-Maintaining Murein Sacculus of Escherichia coli. *Microbiology and Molecular Biology Reviews*, *62*(1), 181–203.

Höltje, J.-V. (2004). Cell Walls, Bacterial. In *The Desk Encyclopedia of Microbiology* (pp. 239–250).

Hucka, M., Finney, A., Doyle, J. C., Kitano, H., Kovitz, B. L., Shapiro, B. E., … Schilstra, M. J. (2004). Evolving a lingua franca and associated software infrastructure for computational systems biology: the Systems Biology Markup Language (SBML) project. *Systems Biology, IEE Proceedings*, *1*(1), 41–53. doi:10.1049/sb:20045008

Hunter, P. (2008). The paradox of model organisms. *EMBO Reports*, *9*(8), 717–720.

Iwanowski, M., & Korzynska, A. (2010). Segmentation of moving cells in bright field and epi-fluorescent microscopic image sequences . In *Computer Vision and Graphics, International Conference, ICCVG 2010, Warsaw, Poland, September 20-22, 2010, Proceedings, Part I* (pp. 401–410).

Jazwinski, S. M. (2002). Growing old: metabolic control and yeast aging. *Annual Review of Microbiology*, *56*, 769–92. doi:10.1146/annurev.micro.56.012302.160830

Jung, C., & Scharcanski, J. (2005). Robust watershed segmentation using wavelets. *Image and Vision Computing*, *23*(7), 661–669. doi:10.1016/j.imavis.2005.03.001

Kalisky, T., Blainey, P., & Quake, S. R. (2011). Genomic analysis at the single-cell level. *Annual Review of Genetics*, *45*, 431–445. doi:10.1146/annurev-genet-102209-163607

Kamentsky, L., Jones, T. R., Fraser, A., Bray, M.-A., Logan, D. J., Madden, K. L., … Carpenter, A. E. (2011). Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics (Oxford, England)*, *27*(8), 1179–80. doi:10.1093/bioinformatics/btr095

Kandhavelu, M., Mannerström, H., Gupta, A., Häkkinen, A., Lloyd-Price, J., Yli-Harja, O., & Ribeiro, A. S. (2011). In vivo kinetics of transcription initiation of the lar promoter in Escherichia coli. Evidence for a sequential mechanism with two rate-limiting steps. *BMC Systems Biology*, *5*, 149. doi:10.1186/1752-0509-5-149

Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes : Active Contour Models. *International Journal of Computer Vision*, *1*(4), 321–331.

Koch, a L. (1996). What size should a bacterium be? A question of scale. *Annual Review of Microbiology*, *50*, 317–48. doi:10.1146/annurev.micro.50.1.317

Korzynska, A., & Iwanowski, M. (2012). Artifical Images for Evaluation of Segmentation Results : Bright Field Images of Living Cells. In *Third International Conference, Information Technologies in Biomedicine (ITIB) 2012, Gliwice, Poland, June 11-13, 2012. Proceedings* (pp. 445–455).

Kröger, M. (2010). IECA - International E.coli Alliance - E.coli Database Portal. Retrieved December 23, 2014, from http://www.uni-giessen.de/ecoli/IECA/index.php

Kruse, K. (2012). Bacterial Organization in Space and Time. In *Comprehensive Biophysics* (Vol. 7, pp. 208–221). doi:10.1016/B978-0-12-374920-8.00717-7

Lamprecht, M., Sabatini, D., & Carpenter, A. (2007). CellProfiler™: free, versatile software for automated biological image analysis. *BioTechniques*, *42*(1), 71–75. doi:10.2144/000112257

Lan, G., Wolgemuth, C. W., & Sun, S. X. (2007). Z-ring force and cell shape during division in rod-like bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16110–5. doi:10.1073/pnas.0702925104

Lauffenburger, D. (1982). EFFECTS OF CELL MOTILITY AND CHEMOTAXIS ON MICROBIAL POPULATION GROWTH. *Biophysical Journal*, *40*(December), 209–219.

Lehmussola, A., Ruusuvuori, P., Selinummi, J., Huttunen, H., & Yli-Harja, O. (2007). Computational framework for simulating fluorescence microscope images with cell populations. *IEEE Transactions on Medical Imaging*, *26*(7), 1010–6. doi:10.1109/TMI.2007.896925

Lehmussola, A., Ruusuvuori, P., Selinummi, J., Rajala, T., & Yli-harja, O. (2011). Synthetic Images of High-Throughput Microscopy for Validation of Image Analysis Methods. *Proceedings of the IEEE*, *96*(8), 1348 – 1360.

Lehmussola, A., Selinummi, J., Ruusuvuori, P., Niemisto, A., & Yli-Harja, O. (2005). Simulating fluorescent microscope images of cell populations. In *Conference proceedings : … Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE*

Engineering in Medicine and Biology Society. Conference (Vol. 3, pp. 3153–6). doi:10.1109/IEMBS.2005.1617144

Lloyd-Price, J., Häkkinen, A., Kandhavelu, M., Marques, I. J., Chowdhury, S., Lihavainen, E., … Ribeiro, A. S. (2012). Asymmetric disposal of individual protein aggregates in Escherichia coli, one aggregate at a time. *Journal of Bacteriology*, *194*(7), 1747–1752. doi:10.1128/JB.06500-11

Lloyd-Price, J., Lehtivaara, M., Kandhavelu, M., Chowdhury, S., Muthukrishnan, A.-B., Yli-Harja, O., & Ribeiro, A. S. (2012). Probabilistic RNA partitioning generates transient increases in the normalized variance of RNA numbers in synchronized populations of Escherichia coli. *Molecular bioSystems*, *8*(2), 565–71. doi:10.1039/c1mb05100h

Lloyd-Price, J., Tran, H., & Ribeiro, A. S. (2014). Dynamics of small genetic circuits subject to stochastic partitioning in cell division. *Journal of Theoretical Biology*, *356*, 11–9. doi:10.1016/j.jtbi.2014.04.018

Mäkelä, J., Kandhavelu, M., Oliveira, S. M. D., Chandraseelan, J. G., Lloyd-Price, J., Peltonen, J., … Ribeiro, A. S. (2013). In vivo single-molecule kinetics of activation and subsequent activity of the arabinose promoter. *Nucleic Acids Research*, *41*(13), 6544–6552. doi:10.1093/nar/gkt350

Malpica, N., & Solorzano, C. O. de. (2002). Automated Nuclear Segmentation in Fluorescence Microscopy. In A. Méndez-Vilas (Ed.), *Science, Technology and Education of Microscopy: an Overview. Volume 2 of Microscopy Book Series.* (pp. 614–621). Formatex.

Manders, E. M. M., Verbeek, F. J., & A, A. J. (1993). Measurement of co-localization of objects in dual-colour confocal images. *Journal of Microscopy*, *169*(3), 375–382.

Manders, E. M., Stap, J., Brakenhoff, G. J., van Driel, R., & Aten, J. A. (1992). Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy. *Journal of Cell Science*, *103*(3), 857–862. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1478975

Maška, M., Ulman, V., Svoboda, D., Matula, P., Matula, P., Ederra, C., … Ortiz-de-Solorzano, C. (2014). A benchmark for comparison of cell tracking algorithms. *Bioinformatics (Oxford, England)*, *30*(11), 1609–17. doi:10.1093/bioinformatics/btu080

McAdams, H. H., & Arkin, A. (1997). Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(February), 814–819.

Meijering, E. (2012). Cell Segmentation: 50 Years Down the Road. *IEEE Signal Processing Magazine*, *29*(5), 140–145.

Meyer, P., & Dworkin, J. (2007). Applications of fluorescence microscopy to single bacterial cells. *Research in Microbiology*, *158*(3), 187–94. doi:10.1016/j.resmic.2006.12.008

Mittal, N., Budrene, E. O., Brenner, M. P., & Van Oudenaarden, A. (2003). Motility of Escherichia coli cells in clusters formed by chemotactic aggregation. *Proceedings of the*

National Academy of Sciences of the United States of America, *100*(23), 13259–63. doi:10.1073/pnas.2233626100

Murphy, R. F. (2012). CellOrganizer: Image-derived Models of Subcellular Organization and Protein Distribution. *Methods in Cell Biology*, *110*, 179–193. doi:10.1016/B978-0-12-388403-9.00007-2.CellOrganizer

Niki, H., & Hiraga, S. (1997). Subcellular distribution of actively partitioning F plasmid during the cell division cycle in E. coli. *Cell*, *90*(5), 951–957. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9298906

Niki, H., Yamaichi, Y., & Hiraga, S. (2000). Dynamic organization of chromosomal DNA in Escherichia coli. *Genes & Development*, *14*, 212–223. doi:10.1101/gad.14.2.212

Otsu, N. (1979). A Tlreshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66.

Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., & van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. *Nature Genetics*, *31*(1), 69–73. doi:10.1038/ng869

Ozbudak, E. M., Thattai, M., Lim, H. N., Shraiman, B. I., & Van Oudenaarden, A. (2004). Multistability in the lactose utilization network of Escherichia coli. *Nature*, *427*(6976), 737–740. doi:10.1038/nature02298

Peng, T., & Murphy, R. F. (2011). Image-derived, three-dimensional generative models of cellular organization. *Cytometry. Part A*, *79*(5), 383–91. doi:10.1002/cyto.a.21066

Pertsinidis, A., Zhang, Y., & Chu, S. (2010). Subnanometre single-molecule localization, registration and distance measurements. *Nature*, *466*(7306), 647–51. doi:10.1038/nature09163

Ping, L., Weiner, B., & Klecknerr, N. (2009). Tsr-GFP accumulates linearly with time at cell poles, and can be used to differentiate “old” versus “new” poles, in E.coli. *Molecular Microbiology*, *69*(6), 1427–1438. doi:10.1111/j.1365-2958.2008.06372.x.Tsr-GFP

Powers, D. M. W. (2011). Evaluation : From Precision , Recall And F-Measure To Roc , Informedness , Markedness & Correlation. *Journal of Machine Learning Technologies*, *2*(1), 37–63.

Ptacin, J. L., Lee, S. F., Garner, E. C., Toro, E., Eckart, M., Comolli, L. R., … Shapiro, L. (2010). A spindle-like apparatus guides bacterial chromosome segregation. *Nature Cell Biology*, *12*(8), 791–8. doi:10.1038/ncb2083

Qu, Z., Weiss, J. N., & MacLellan, W. R. (2004). Coordination of cell growth and cell division: a mathematical modeling study. *Journal of Cell Science*, *117*(Pt 18), 4199–207. doi:10.1242/jcs.01294

Queimadelas, C. (2012). *Automated segmentation, tracking and evaluation of bacteria in microscopy images*. Faculdade de Ciências e Tecnologia - Universidade Nova de Lisboa. Retrieved from http://hdl.handle.net/10362/8435

Ruusuvuori, P., Aijö, T., Chowdhury, S., Garmendia-Torres, C., Selinummi, J., Birbaumer, M., … Yli-Harja, O. (2010). Evaluation of methods for detection of fluorescence labeled subcellular objects in microscope images. *BMC Bioinformatics*, *11*, 248. doi:10.1186/1471-2105-11-248

Ruusuvuori, P., Lehmussola, A., Selinummi, J., Rajala, T., Huttunen, H., & Yli-Harja, O. (2008). Benchmark Set Of Synthetic Images For Validating Cell Image Analysis Algorithms. In *Proceedings of the 16th European Signal Processing Conference, EUSIPCO*.

Santinha, J., Martins, L., Häkkinen, A., Lloyd-price, J., Oliveira, S., Gupta, A., … Fonseca, J. (2014). iCellFusion – a Tool for Fusion and Analysis of Live-cell Images from Time-lapse Multimodal Microscopy. In *(in Revision)*. doi:10.3390/s140x0000x

Satwik, R., Benjamin, P., Nicholas, H., Steven, A., & Lani, W. (2012). SimuCell : a flexible framework for creating synthetic microscopy images a PhenoRipper : software for rapidly profiling microscopy images, *9*(7), 634–636.

Satwik, R., Benjamin, P., Nicholas, H., Steven, A., & Lani, W. (2012). SimuCell : a flexible framework for creating synthetic microscopy images a PhenoRipper : software for rapidly profiling microscopy images. *Nature Methods*, *9*(7), 634–636.

Selinummi, J., Seppälä, J., Yli-Harja, O., & Puhakka, J. (2005). Software for quantification of labeled bacteria from digital microscope images by automated image analysis. *BioTechniques*, *39*(6), 859–863. doi:10.2144/000112018

Shen, H., Nelson, G., Nelson, D. E., Kennedy, S., Spiller, D. G., Griffiths, T., … Kell, D. B. (2006). Automated tracking of gene expression in individual cells and cell compartments. *Journal of the Royal Society, Interface / the Royal Society*, *3*(11), 787–94. doi:10.1098/rsif.2006.0137

Shivanandan, A., Deschout, H., Scarselli, M., & Radenovic, A. (2014). Challenges in quantitative single molecule localization microscopy. *FEBS Letters*, (June). doi:10.1016/j.febslet.2014.06.014

Sliusarenko, O., & Heinritz, J. (2011). High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Molecular Microbiology*, *80*(3), 612–627. doi:10.1111/j.1365-2958.2011.07579.x.High-throughput

Stephens, D. J., & Allan, V. J. (2003). Light microscopy techniques for live cell imaging. *Science (New York, N.Y.)*, *300*(5616), 82–86. doi:10.1126/science.1082160

Stewart, E. J., Madden, R., Paul, G., & Taddei, F. (2005). Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biology*, *3*(2), e45. doi:10.1371/journal.pbio.0030045

Sung, M.-H., & McNally, J. G. (2011). Live cell imaging and systems biology. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine*, *3*(2), 167–82. doi:10.1002/wsbm.108

Svoboda, D., Kasik, M., Maska, M., & Hubeny, J. (2007). On simulating 3D Fluorescent Microscope Images. In *Computer Analysis of Images and Patterns -12th International*

Conference, CAIP 2007, Vienna, Austria, August 27-29, 2007. Proceedings (pp. 309–316). doi:10.1007/978-3-540-74272-2_39

Svoboda, D., Kozubek, M., & Stejskal, S. (2009). Generation of digital phantoms of cell nuclei and simulation of image formation in 3D image cytometry. *Cytometry. Part A : The Journal of the International Society for Analytical Cytology*, *75*(6), 494–509. doi:10.1002/cyto.a.20714

Svoboda, D., & Ulman, V. (2012). Generation of synthetic image datasets for time-lapse fluorescence microscopy. In A. Campilho & M. Kamel (Eds.), *ICIAR'12 Proceedings of the 9th international conference on Image Analysis and Recognition - Volume Part II* (Vol. 7325, pp. 473–482). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-31298-4

Swat, M. H., Thomas, G. L., Belmonte, J. M., Shirinifard, A., Hmeljak, D., & Glazier, J. a. (2012). Chapter 13 - Multi-scale modeling of tissues using CompuCell3D. In *Methods in cell biology* (Vol. 110, pp. 325–366). Elsevier Inc. doi:10.1016/B978-0-12-388403-9.00013-8

Taniguchi, Y., Choi, P. J., Li, G.-W., Chen, H., Babu, M., Hearn, J., … Xie, X. S. (2010). Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science (New York, N.Y.)*, *329*(5991), 533–538. doi:10.1126/science.1188308

Tindall, M. J., Maini, P. K., Porter, S. L., & Armitage, J. P. (2008). Overview of mathematical approaches used to model bacterial chemotaxis II: bacterial populations. *Bulletin of Mathematical Biology*, *70*(6), 1570–607. doi:10.1007/s11538-008-9322-5

Torr, P. H. S., & Zisserman, A. (1999). Feature Based Methods for Structure and Motion Estimation. In *Vision Algorithms: Theory and Practice, International Workshop on Vision Algorithms, ICCV '99, Corfu, Greece, September 21-22 1999* (pp. 278–294). doi:http://dx.doi.org/10.1007/3-540-44480-7_19

Trueba, F. J., Spronsen, E. A. Van, Traas, J., & Woldringh, C. L. (1982). Effects of Temperature on the Size and Shape of Escherichia coli Cells. *Archives of Microbiology*, *131*(3), 235–240.

Ulman, V., & Hubeny, J. (2007). On Generating Ground-Truth Time-Lapse Image Sequences And Flow Fields. In *In International Conference on Informatics in Control, Automation and Robotics - ICINCO 2007, RA-1, Angers: INSTICC* (pp. 234–239).

Vallotton, P., Mililli, L., Turnbull, L., & Whitchurch, C. (2010). Segmentation of Dense 2D Bacilli Populations. In *2010 International Conference on Digital Image Computing: Techniques and Applications* (pp. 82–86). IEEE. doi:10.1109/DICTA.2010.23

Van de Kraats, E. B., Penney, G. P., Tomazevic, D., van Walsum, T., & Niessen, W. J. (2005). Standardized evaluation methodology for 2-D-3-D registration. *IEEE Transactions on Medical Imaging*, *24*(9), 1177–89. doi:10.1109/TMI.2005.853240

Wadhams, G. H., & Armitage, J. P. (2004). Making sense of it all: bacterial chemotaxis. *Nature Reviews. Molecular Cell Biology*, *5*(12), 1024–37. doi:10.1038/nrm1524

Wählby, C., Kamentsky, L., Liu, Z. H., Riklin-raviv, T., Annie, L., Rourke, E. J. O., … Anne, E. (2013). An image analysis toolbox for high-throughput C. elegans assays. *Nature Methods*, *9*(7), 714–716. doi:10.1038/nmeth.1984.An

Wang, J. D., & Levin, P. a. (2009). Metabolism, cell growth and the bacterial cell cycle. *Nature Reviews. Microbiology*, *7*(11), 822–7. doi:10.1038/nrmicro2202

Wang, Q., Niemi, J., Tan, C.-M., You, L., & West, M. (2010). Image segmentation and dynamic lineage analysis in single-cell fluorescence microscopy. *Cytometry A*, *77*(1), 101–110. doi:10.1002/cyto.a.20812

Wang, Q., You, L., & West, M. (2008). *CellTracer : Software for automated image segmentation and line- age mapping for single-cell studies (Discussion Paper)* (pp. 1–2).

Wanner, B., Finney, A., & Hucka, M. (2005). Modeling the E. coli cell: The need for computing, cooperation, and consortia. In H. V. W. L. Alberghina (Ed.), *Systems Biology, Definitions and Perspectives. Topics in Current Genetics, Vol. 13* (Vol. 13, pp. 163–189). Springer-Verlag Berlin Heidelberg. doi:10.1007/b138743

Weng, X., & Xiao, J. (2014). Spatial organization of transcription in bacterial cells. *Trends in Genetics*, *30*(7), 287–297.

Wiesmann, V., Sauer, T., Held, C., Palmisano, R., & Wittenberg, T. (2013). Cell Simulation for Validation of Cell Micrograph Evaluation Algorithms. *Biomedizinische Technik (Berlim)*, *58*(Suppl. 1). Retrieved from http://www.degruyter.com/view/j/bmte.2013.58.issue-s1-L/bmt-2013-4272/bmt-2013-4272.xml

Winkler, J., Seybert, A., König, L., Pruggnaller, S., Haselmann, U., Sourjik, V., … Bukau, B. (2010). Quantitative and spatio-temporal features of protein aggregation in Escherichia coli and consequences on protein quality control and cellular ageing. *The EMBO Journal*, *29*(5), 910–23. doi:10.1038/emboj.2009.412

Wolozin, B. (2012). Regulated protein aggregation: stress granules and neurodegeneration. *Molecular Neurodegeneration*, *7*(56), 1–12. doi:10.1186/1750-1326-7-56

Wyawahare, M. V, Patil, P. M., & Abhyankar, H. K. (2009). Image Registration Techniques : An overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, *2*(3), 11–28.

Xiong, W., Wang, Y., Ong, S. H., Lim, J. H., & Jiang, L. (2010). LEARNING CELL GEOMETRY MODELS FOR CELL IMAGE SIMULATION : AN UNBIASED APPROACH. In *Proceedings of 2010 IEEE 17th International Conference on Image Processing* (pp. 1897–1900).

Young, J. W., Locke, J. C. W., Altinok, A., Rosenfeld, N., Bacarian, T., Swain, P. S., … Elowitz, M. B. (2012). Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy. *Nature Protocols*, *7*(1), 80–8. doi:10.1038/nprot.2011.432

Young, K. D. (2010). Bacterial shape: two-dimensional questions and possibilities. *Annual Review of Microbiology*, *64*, 223–40. doi:10.1146/annurev.micro.112408.134102

Yu, J., Xiao, J., Ren, X., Lao, K., & Xie, X. S. (2006). Probing gene expression in live cells, one protein molecule at a time. *Science*, *311*(5767), 1600–3. doi:10.1126/science.1119623

Zhang, E. R., Wu, L. F., & Altschuler, S. J. (2013). Envisioning migration: mathematics in both experimental analysis and modeling of cell behavior. *Current Opinion in Cell Biology*, *25*(5), 538–42. doi:10.1016/j.ceb.2013.04.004

Zhao, T., & Murphy, R. F. (2007). Automated learning of generative models for subcellular location: building blocks for systems biology. *Cytometry. Part A : The Journal of the International Society for Analytical Cytology*, *71*(12), 978–90. doi:10.1002/cyto.a.20487

Zinder, S. H., & Dworkin, M. (2006). Chapter 1.7 - Morphological and Physiological Diversity. In M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, & E. Stackebrandt (Eds.), *Prokaryotes* (pp. 185–220). New York, NY: Springer New York. doi:10.1007/0-387-30741-9

Zitová, B., & Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, *21*(11), 977–1000. doi:10.1016/S0262-8856(03)00137-9