*Digital Comprehensive Summaries of Uppsala Dissertations*
*from the Faculty of Science and Technology 2324*

# Microfluidics and AI for single-cell microbiology

PRANEETH KAREMPUDI



ACTA UNIVERSITATIS
UPSALIENSIS
2023

ISSN 1651-6214
ISBN 978-91-513-1932-2
urn:nbn:se:uu:diva-514317

UPPSALA
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in B21, Biomedicinskt centrum (BMC), Husargatan 3, Uppsala, Friday, 1 December 2023 at 09:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Associate Professor Yoav Shechtman (Technion - Israel Institute of Technology, Israel).

**Abstract**

Karempudi, P. 2023. Microfluidics and AI for single-cell microbiology. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 2324. 57 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-1932-2.

Most of the biological sciences deal with understanding the relationships between phenotypes and the underlying molecular mechanisms of organisms. This thesis is an engineering, computational, and experimental exercise in expanding the scope and scale of phenotype-genotype mapping techniques in single-cell microbiology using microscopy, microfluidics, and image processing. To this end, we use mother-machine-based microfluidic devices together with recently developed techniques in deep learning and optics. We use optical microscopes to observe cells of different genotypes, physically move cells, and image molecules inside them.

We have designed a novel microfluidic device to expand the throughput of single-cell lineage tracing an order of magnitude compared to existing methods. We demonstrate the ability to isolate single cells from such a device using optical tweezers after phenotypic characterization in real time. We have developed analysis algorithms of various kinds with the prime intention of performing high-throughput real-time image processing in conjunction with experimental runs to identify interesting cells for further investigation.

We have also developed an experimental protocol for bacterial species identification using fluorescence-in-situ hybridization (FISH) in microfluidic chips to complement an existing phenotype-based antibiotic-susceptibility test (AST). We apply this method together with deep-learning-based cell segmentation and tracking algorithms, and image classification methods to perform species-ID of up to 10 species in 2-3 hrs.

Lastly, we have developed a 3D dot localization method to investigate how the chromosome structure changes during the *E. coli* cell cycle. Different loci on the *E. coli* chromosome were labeled using DNA-binding fluorescent proteins and imaged using an optical setup with an astigmatic point-spread-function. Mother-machine devices were used to constrain the movement of cells to the lateral plane during growth. A deep-learning-based single-molecule localization method was adapted for this application and used to map the chromosomal loci's physical position in 3D as a function of cell size during the *E. coli* cell cycle.

*Keywords:* Microfluidics, Artificial intelligence, Deep learning, Single-cell microbiology

*Praneeth Karempudi, Department of Cell and Molecular Biology, Molecular Systems Biology, Box 596, Uppsala University, SE-751 24 Uppsala, Sweden.*



ISSN 1651-6214
ISBN 978-91-513-1932-2
URN urn:nbn:se:uu:diva-514317 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-514317)

*To Nirmala, Manohar and Mounica*

# List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Vinodh Kandavalli*, **Praneeth Karempudi***, Jimmy Larsson, Johan Elf. Rapid antibiotic susceptibility testing and species identification for mixed samples. Nature Communications 13, 6215 (2022); doi: https://doi.org/10.1038/s41467-022-33659-1

II **Praneeth Karempudi**, Elias Amselem, Daniel Jones, Zahra Khaji, Maria Tenje, Johan Elf. Real-time pooled optical screening with single-cell isolation capability. bioRxiv 2023.09.21.558600; doi: https://doi.org/10.1101/2023.09.21.558600

III **Praneeth Karempudi***, Konrad Gras*, Elias Amselem, Spartak Zikrin, Johan Elf. Three-dimensional localization of fluorescent proteins in living *Escherichia coli*. bioRxiv 2023.10.20.563292; doi: https://doi.org/10.1101/2023.10.20.563292

Reprints were made with permission from the publishers. * represents equal contribution

# Contents

# 1. Introduction

This thesis is an engineering, computational, and experimental exercise in expanding the scope and scale of phenotype-genotype mapping techniques in single-cell microbiology using microscopy and image processing. In terms of engineering advances, we have designed microfluidic devices for single-cell biology, expanding the throughput of single-cell lineage tracing an order of magnitude compared to existing methods. Regarding experimental advances, we demonstrate the ability to isolate single cells from such a device after phenotypic characterization in real time. We have also designed an experimental protocol for bacterial species identification in microfluidic chips. To tie everything together, we have developed analysis algorithms of various kinds with the prime intention of performing high-throughput real-time image processing in conjunction with experimental runs. On a tangent, we have developed 3D localization methods to investigate the dynamic movement of bio-molecules inside cells over multiple cell-division cycles within a microfluidic device. Over the next few sections, I will briefly introduce the different components used to make these advances and describe them in detail in the following chapters.

## 1.1 Engineers studying microbes

Most of the biological sciences deal with understanding the relationships between a phenotype and the underlying molecular mechanisms. *Phenotype* refers to observable traits of an organism such as morphology, size, growth rate, motion, or macroscopic function. *Molecular mechanisms* describe the underlying interactions between molecules and polymers such as DNA, RNA, proteins, etc, that result in the phenotype. *Genotype* refers to the genetic makeup of an organism, describing molecules/polymers that drive an organism. Understanding molecular mechanisms in relation to macroscopic behavior is critical to developing cures for diseases, whether cancer or infectious diseases caused by agents such as viruses, bacteria, or fungi. Relating molecular mechanisms to phenotypes on different spatial and temporal scales under various environmental conditions is central to biological research. The environments in which organisms operate are quite complex and it is not feasible to observe these organisms' behavior in the wild. So, we engineer controlled environments in the lab to study model organisms, hoping to elucidate generalizable principles and/or specific mechanisms.

The organisms I study in this thesis are bacteria whose size varies from 0.5 $\mu$m to several $\mu$ms. Using techniques from micro-fabrication, we have designed fluidic devices to capture and study single cells using an optical microscope. We can follow the cells as they are growing and perform assays of various kinds with them. In the devices, we can quickly and precisely manipulate the growth medium of the cells and expose them to different molecular agents. In particular, we have used a specific model called mother-machine devices [1, 2], which provides unique capabilities of long-term cell-culturing and lineage-tracking.

In Chapter 2, I describe the design of these devices and the advances made by us to increase their throughput. We also present a unique method to extract single cells from any channel of choice from this device.

## 1.2 Seeing Microbes and molecules inside them

Microscopy has revolutionized science, and our ability to see small things has driven significant scientific progress over the past four centuries. Most biological objects, such as bacterial cells, are optically transparent and provide poor contrast in standard bright-field microscopy. We use phase-contrast microscopy to capture differences in optical density of bacterial cells to improve contrast and aid image processing.

As revolutionary as traditional light microscopy has been in science, applications of fluorescence microscopy have been equally important in biology, if not more. Fluorescence phenomena applied in biological contexts help us tag specific bio-molecules and cellular structures that are not identifiable in a standard optical microscope. Appropriate excitation and emission filters and light sources enable the imaging of multiple fluorescence channels concurrently. In this way, we can image sub-cellular components and measure dynamic properties of bio-molecules such as protein expression and molecule localization to name a few. I have made extensive use of fluorescence phenomena throughout the thesis work. Species-specific fluorescent probes were used to tag ribosomal-RNA for species identification in **Paper I**. Constitutively expressed fluorescence by cells was used to show single-cell isolation from fluidic devices in **Paper II**. Fluorescent proteins were used to label different chromosomal loci and the replisome complex in *E. coli* cells in **Paper III**.

The standard resolution of an optical microscope is limited by diffraction and is typically $\sim$250 nanometers. With clever optical engineering, statistics, and experimental methods to tag molecules, one can increase the resolution of the microscope to a few nanometers and generate super-resolved images. In Paper III, we engineer a standard epi-fluorescence microscope to encode 3D locations of fluorescently labeled molecules into 2D images and capture them on a scientific CMOS camera [3]. This imaging mode together with

phase-contrast microscopy and molecule localization algorithms enable studying molecular movements over the bacterial cell cycle in 3D with 10-30 nm localization precision in the lateral plane and 50-70 nm in the axial plane.

In Chapter 3, I describe the imaging setups used in the papers and introduce the imaging data and the required analysis methods.

## 1.3 Moving microbes

Single-beam gradient force traps, also known popularly as optical tweezers, are used to trap microscopic particles with focused laser beams [4]. Microscopic objects, such as cells, suspended in culture media can be trapped at the center of a focused laser beam and moved around using the microscope stage. The force applied on the object to keep it centered is typically in the order of picoNewtons. Hence, careful balancing of flow rates and pressures is needed to be able to move cells inside fluidic devices. In 2019, it was shown that one can use optical tweezers to move cells inside mother-machine devices to isolate interesting phenotypes after high-throughput screening [5]. We observe cells growing in the fluidic devices and calculate which cells to pick out of the chip. The chosen cells are moved to a clean region on the chip containing no cells and flushed out for further analysis or investigation.

In Chapter 2, I describe the fluidic chip design and method to perform single-cell isolation from a high-throughput mother-machine device containing 100,000 cell traps. In Chapter 3, I describe the optical setup used for this method, and in Chapter 4, I describe the real-time image processing software that helps us decide which cells are interesting for extraction.

## 1.4 Calculations on imaging data

Extracting insights from imaging data using algorithms is central to developing and testing quantitative hypotheses from microscopy data. Modern image processing techniques are used heavily to aid segmentation, tracking of objects, calculation of object sizes and growth rates of cells, localization of molecules inside cells, and other means of quantitative image characterization. Advances in machine learning, especially deep learning, have revolutionized computer vision and image processing techniques over the past decade, resulting in higher quality results on most of the common tasks we need for analyzing microscopy data.

In this thesis, I describe a suite of methods for analyzing experimental microscopy data that uses mother-machine devices. The central element in the analysis pipelines of each paper is typically a U-net, a neural network architecture that is useful for mapping an image to multiple images. The characteristics of the network and its use vary by application and are described in detail

in Chapter 4. We use the U-net architecture for variants of cell segmentation, localization of fluorescent proteins in 3D, mother-machine trap segmentation, etc. We have also developed and used cell-tracking algorithms as well as classifiers trained on imaging data when required. To aid real-time analysis of data for specific applications related to single-cell isolation after phenotyping, we have adapted the analysis algorithms to speed up the calculations. We have also built a GUI to run the experiment and monitor the results.

## 1.5 Outline

In the following chapters, I describe the core advances presented in this thesis in relation to the constraints of each specific paper. In Chapter 2, I describe the design and experimental capability of the new fluidic chips after a brief overview of the existing devices. In Chapter 3, I give a brief overview of the microscopy techniques that we have used and the imaging data that we have acquired in the various projects. In Chapter 4, I briefly describe the analysis algorithms implemented for analyzing images produced in different contexts described in Chapter 3. After these chapters, I summarize the results from each paper and finally, I provide an overview of all the advances in both English and Swedish.

# 2. Microfluidics for single-cell microbiology

Advances in semiconductor fabrication techniques have resulted in a computing revolution over the past 70 years. Semiconductor fabrication techniques primarily deal with patterning small features on wafers from Silicon-based materials to construct complex electronic devices. The core methods used in wafer fabrication are lithography, etching, and thin-film deposition. Using deep-uv lithography, it is possible to produce a feature size of 3 nm, but most academic clean rooms can only go down to around 1 $\mu$m features using normal UV lithography (i-line, g-line, h-line) and a few tens of nanometers using electron-beam lithography. These techniques are predominantly used in the fields of microfluidics to make structures with size ranges varying from tens of nanometers to hundreds of microns with very high precision.

The use of polymeric materials like PDMS [6], enables replica-molding of the structures created on Silicon wafers using semiconductor fabrication techniques. PDMS has nice properties such as optical transparency, high fidelity for smaller-feature replication, low background auto-fluorescence, biocompatibility, and gas permeability, making it suitable for biological assays with optical readout. Combining soft-lithographic materials with creative fabrication steps has led to the development of microfluidic devices for studying biological properties of single cells, colonies, and aggregates of cells at various spatial and temporal scales and throughputs.

## 2.1 Microfluidics in microbiology

A variety of microfluidic techniques have been successfully applied to study the biological properties of microbes. They fall broadly into two categories, one for studying microbial colonies and another for studying single cells [7]. Colonies of cells growing in nano-wells [8] or chambers contained with quake values [9] are useful for imaging colony properties, but offer low throughput and are complicated to seed. Nano-wells can be used to image colonies from single cells and reaction chambers offer the ability to add and remove reagents and perform reactions on isolated colonies. The throughput of nano-wells and chambers is typically limited to a few hundred cells.

Droplet microfluidics [10] offers very high throughput, typically tens of thousands to millions of cells or colonies, at the cost of imaging quality due to the oil-water interface. Droplets have been predominantly used for single-colony sequencing as well as cell sorting for screening applications. In PDMS-based single-cell devices, however, cells can be imaged at a much higher resolution than in nano-wells or droplets. They are based on trapping mono-layers

of cells in chambers where they can be imaged at high spatial and temporal resolutions over long experimental time frames. Among these devices, variants of the mother-machine design [1, 2] offer significant advantages such as isolation of single cells as well as lineages (Fig 2.1).

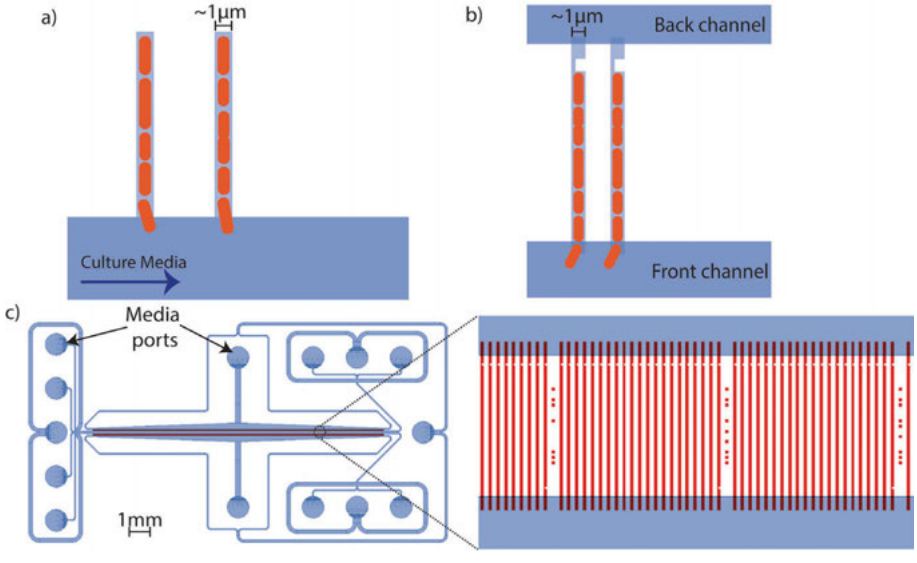



*Figure 2.1.* Mother-machine device and variants. a) Closed-end trap design showing the trapped cells. b) Modified trap design to allow fluid flow over cells inside the traps. The image also shows the larger media supply channels, i.e, the front channel and the back-channel, whose pressures are controlled to keep cells inside the trap. c) A complete chip design showing all the media ports and a small field-of-view of the chip containing growth channels.

## 2.2 Mother-machine devices and variants

The mother-machine [1] is a fluidic design where bacterial cells are trapped at the end of a long channel with access to fresh media allowing long-term single-cell imaging. The cell at the end of the trap is called the mother cell. Each trap has a length that typically fits a few bacteria and has the approximate height and width of a single *E. coli* cell. The mother cell grows and fills the rest of the trap with its progeny, allowing monitoring of properties over multiple generation cycles. Fresh media is supplied to the ends of the channels and required nutrients reach the mother cell via diffusion or flow.

In a previous study [2] from our lab, the original mother-machine device [1] was modified to overcome its main drawback, the difficulty in loading the traps with cells. The standard mother machine is loaded with a medium containing a high density of cells as it relies on diffusion for the bacteria to reach the ends of

the trap or an additional centrifugation step [11]. By modifying the ends of the trap to an open end with a wedge, one can load the channels to high occupancy in a few minutes using pressure-driven flow controllers. Fresh media can also flow over the mother cells to the back channels (Fig 2.1). This modification, however, requires the fabrication of 300 nm or sub-micron structures using e-beam lithography and etching of these sub-micron structures on the silicon mold.

The new design also allows increasing the length of the channel to fit 10-15 bacteria without compromising loading efficiency or cell viability [12]. Fig 2.1 illustrates the modification and the chip design with media ports and larger fluidic channels that supply media or reagents to the cells in the traps. Cells in the traps can be subjected to dynamic environmental conditions by changing the fluid that reaches the cells using pressure controllers on the fluid reservoirs connected to the media ports.

## 2.3 Applications

The quick capture that was made possible with the clever new design allowed us to load bacteria from a urine sample into the traps, leading to a diagnostic application for urinary tract infection (UTI)-related antibiotic-susceptibility tests (AST)[2]. These modified mother machine devices can also be used to capture bacteria from positive blood cultures for sepsis diagnosis [13]. In **Paper I**, we used this modified mother machine device to do mixed-species AST followed by species identification. With this method, we could AST and identify 95% of the most common sepsis-causing pathogens in 2-3 hours. In **Paper III**, we also used this device to constrain cells in traps, making it an effective tool to perform fluorescent protein localization approaches at higher throughput than previously possible.

Mother-machine devices have another unique feature that enables effective high-throughput mapping of phenotype to genotype in pooled libraries. After a few generations of cell growth, each individual trap is completely filled with the progeny of the mother cell. If a pooled library of cells where each cell has a unique recoverable barcode is loaded into the chip [14], a high-throughput mapping of phenotype to genotype is achievable. One could potentially map any optically measurable property and map it to its genotype. However, this method is quite limited in the sense that one would have to use complex cloning techniques to build libraries with unique barcodes and perform multiple rounds of genotyping protocols to recover them.

A single-cell isolation method where one can pick any cell after phenotypic observations would help to solve this problem.

## 2.4 Single-cell isolation

Mother-machine devices provide unique possibilities to screen for properties that vary from cell to cell using precise quantitative methods. However, the mother machine does not allow the collection of interesting phenotypes for further investigation. While isolating cells in droplets provides a convenient way to achieve high-throughput single-colony measurements, we cannot make detailed quantitative measurements before isolating interesting droplets using sorting techniques [15, 16]. FACS is another popular way to isolate cells of interest based on fluorescence, but this technique is also limited by the complexity of operation, sensitivity of fluorescence measurements, and lack of dynamic information retrieval from the samples [17].

Recovering cells of interest from fluidic devices after probing their properties is a challenging task. In the context of mother-machine devices, it has been recently shown [5] that one can isolate single cells of a particular lineage from the mother-machine traps cleanly and collect them for further processing. Some regions on the fluidic devices are maintained clean and are connected to the cell-growing area by pressurized valves [18, 19]. Interesting cells are caught with optical traps and the microscope stage is used to move the cell from the growth area to the clean area through a temporarily open valve. The clean channels are then flushed to collect single cells. This process can be repeated to collect as many cells as one desires. Fig 2.2 illustrates this scheme. The key challenge of using optical traps for moving cells is keeping the cells alive after exposure to focused laser beams by minimizing the photodamage. In [5, 20], the laser power and duration of trapping cells at a focal point were optimized to keep cells alive with growth rates similar to that of non-tweezed cells. Mutation rates were also calculated at exposures up to 75s and were found to be insignificant [5].

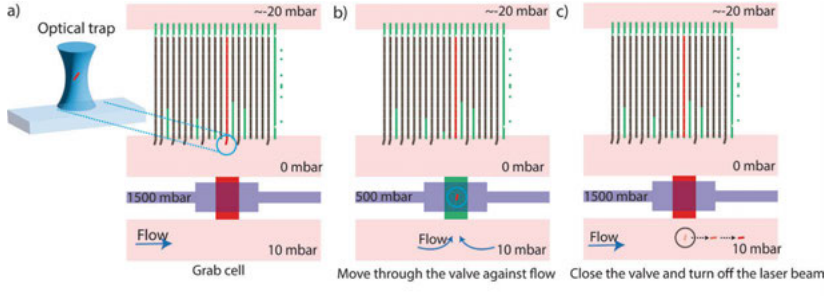



*Figure 2.2.* Single-cell isolation scheme using optical tweezers. a) A cell from the end of a selected mother-machine growth channel being trapped at the focal point of a laser beam. The cartoon also shows a clean channel and a pressurized valve in closed mode. b) Cell being moved from the growing region to the clean channel while the valve is open. c) Cell being released in the clean channel after the valve is closed. This figure is reproduced from Paper II.

In the work described in this thesis, we have incorporated the capability of isolating and collecting single cells of interesting phenotypes into our modified mother-machine designs. The major limitation of the isolation process is that we have to pick cells from the ends of the traps. The force provided by the laser in the operating range for cell viability is not sufficient to drag cells from inside the traps reliably.

## 2.5 Scaling device throughput

Mother-machine devices made so far are fundamentally limited in the number of traps to several thousand before one gets limited by the number of ports one needs to connect for fresh media, back channels, and loading ports. The open-ended design also requires the pressure at the back-channel side to be kept lower than the front-side pressure for the cells to stay in the traps. We address these challenges by introducing 3D channel connections where fluids flow in and out of the plane of imaging.

While 3D connections in microfluidics have been demonstrated well in materials such as paper using layer stacking techniques [21] and custom resins using 3D printing [22], making them in PDMS reliably at smaller dimensions has been relatively difficult to achieve until recently [23]. The technique described in [23] essentially boils down to making tall pillars on a wafer and spin-coating PDMS to a thickness smaller than the height of the pillars. These pillars leave holes in the PDMS when the membrane is lifted from the wafer. Applying a coat of $C_4F_8$ on the wafer makes the surface super-hydrophobic, preventing the PDMS from covering the pillars. We apply this technique to achieve 3D flows when we scale our device throughput. Fig 2.3 shows 3D through-holes created using this technique, which allows fluid to flow on multiple layers.

## 2.6 Our 100k chip

We designed a chip with 100,000 traps with the prerequisite for isolating cells from any particular trap of interest. The 100K chip has 25 rows of modified mother-machine channels, each containing 4000 traps. Each row also has a parallel clean channel that is used to collect cells of interest. Each row is connected to the clean channel by 28 valves placed evenly so that a cell extraction process can be completed in $\sim$ 40 seconds. The chip only has 8 media inlet-outlet ports, which is a significant improvement from the 22 required by [5] for 16,000 traps arranged in 4 isolated units. In **Paper II**, we loaded this device with fluorescent and non-fluorescent cells in a $\sim$ 1:10 ratio and picked out 5 fluorescent cells from different rows of the chip as a control experiment.

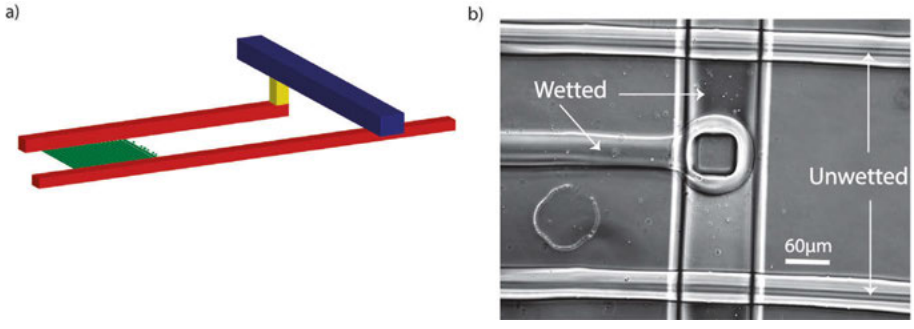



*Figure 2.3.* 3D connections and flows. a) A cartoon showing flow channels arranged in 3D. Channels in the lower layer are shown in red and channels in the top layer are shown in blue with interconnect shown in yellow. b) A realization of the 3D connection in the microfluidic chip, showing the fluid flow from a channel in the top layer to the bottom layer via the through-hole

The exact details of the control experiment are described in **Paper II**. Fig 2.4, illustrates the features of this chip.

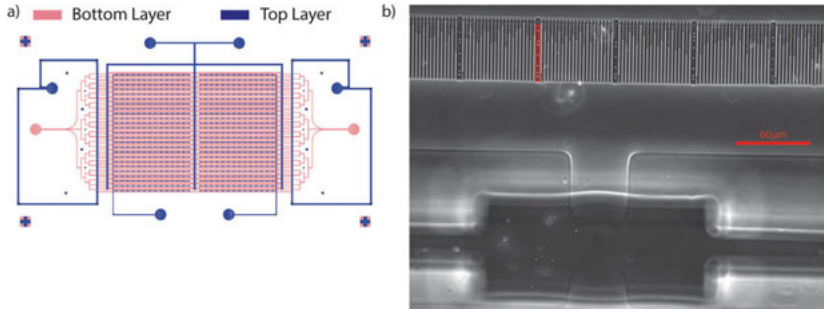



*Figure 2.4.* 100k chip. a) Chip design showing the overall arrangement of channels b) A phase-contrast image of a single field-of-view in 40x magnification cells in the growth channels and a closed valve. Red bounding box area indicates the unique bar-code region designed on either side of every block of 21 cell traps. Scale bar is 60 $\mu m$.

## 2.7 Limitations of large-scale fluidic devices

The construction of accurate lineages of cells growing in mother-machine traps requires imaging at high frequencies (typically a few minutes between frames). The current imaging time of the 100K chip in phase contrast using a high-NA 40x oil immersion objective and a standard CMOS camera is 10 minutes. This imaging frequency sets the limit on what phenomena one can study using large-scale devices. Imaging of fluorescence properties adds additional time to the process. If one is not interested in constructing lineages, these devices can

be used for large-scale phenotyping of the average properties of the cells in a single trap, followed by single-cell extractions.

When performing large-scale imaging-based phenotyping in this context, the data analysis to identify interesting cells is in itself a challenge. We address this problem in detail in Chapter 4.

## 2.8 Potential applications of the 100K chip

In general, the 100K chip can be used to identify and isolate cells with interesting phenotypes from any large strain library or other assembly of genetically diverse bacteria. The method is useful when the phenotype can be assessed by high-resolution imaging. Specific examples include screening libraries with variations in promoter or operator sequences for specific gene expression patterns and investigations of the influence of rare phenotypes on antibiotic tolerance [24]. Screening based on localization of dots (Chatper 4.9) inside cells can be used to study phenotypes, where chromosomal loci labels exhibit interesting distributions inside cells.

# 3. Probing biology using microscopy

Microscopy is one of the key elements in all the work done in this thesis. In this chapter, I describe the three different microscope configurations used in the papers and show images acquired using them.

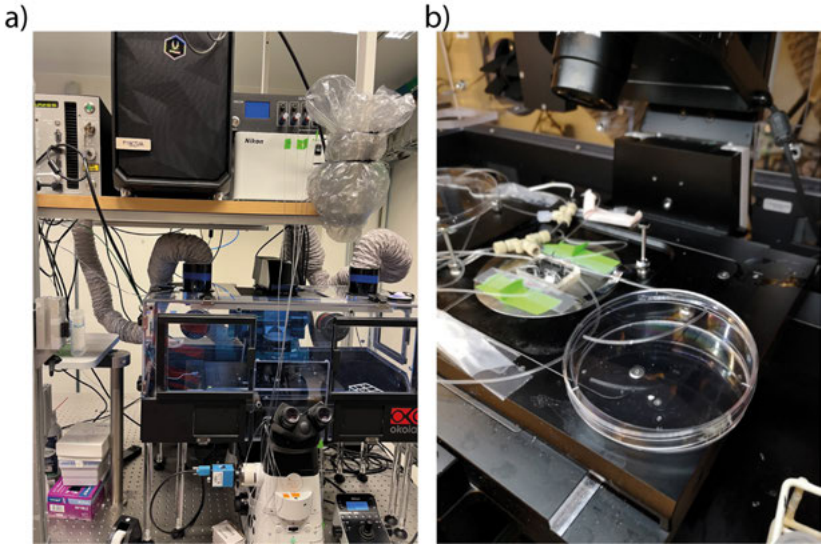



*Figure 3.1.* a) The microscope set-up. b) A chip mounted on the microscope.

## 3.1 Phase-contrast and Epi-fluorescence microscopy

The standard microscope we use for imaging is an inverted microscope from Nikon (Ti2-E, Fig 3.1) capable of both phase-contrast and epi-fluorescence microscopy. The microscope is equipped with both 100x and 40x oil immersion objectives with NA 1.45 and 1.3 respectively and an incubation hood (Oko lab, Cage incubator) for temperature control of the chip environment. We use an industrial CMOS camera (The Imaging source; DMK 38UX304) with 4096x3000 pixels capable of capturing 2 blocks in 100x or 5 blocks in 40x in one field-of-view (FOV). Representative phase-contrast images in both 100x and 40x are shown in Fig 3.2. The microscope is also equipped with light sources and excitation and emission filters to enable imaging in 5 different fluorescence channels (Alexa 488, Cy3, Cy5, Texas Red, and Venus). Fluorescence images, except the ones used for fluorescent protein localization, were captured on the same CMOS camera.

In **Paper I**, we used this setup to perform time-lapse phase-contrast microscopy followed by a genotyping protocol where cells were labeled with sequence-specific fluorescent probes. Phase-contrast images were used to construct lineages of cells growing in the traps. After the cells had grown for 60 mins, the genotyping protocol was performed and the fluorescence signals from the 4 imaging channels were used to assign species labels to cell lineages. Species-wise growth rates were calculated from these lineages to determine the susceptibility of each species to the antibiotic used. To get species-specific susceptibility profiles, we had to establish a reliable spatial mapping between the fluorescence images and the phase-contrast images, a problem that the mother-machine chip solves by design. Capturing phase-contrast and fluorescence images on the same camera let us avoid the image transformation steps usually required for mapping images from different cameras. Chapter 5 summarizes the results of this paper.

In **Paper II** where we used the 100k chip, we switched to a 40x objective as it provides a much larger FOV for faster imaging. The additional modifications that we made to the standard setup to enable single-cell tweezing are described in the next section.

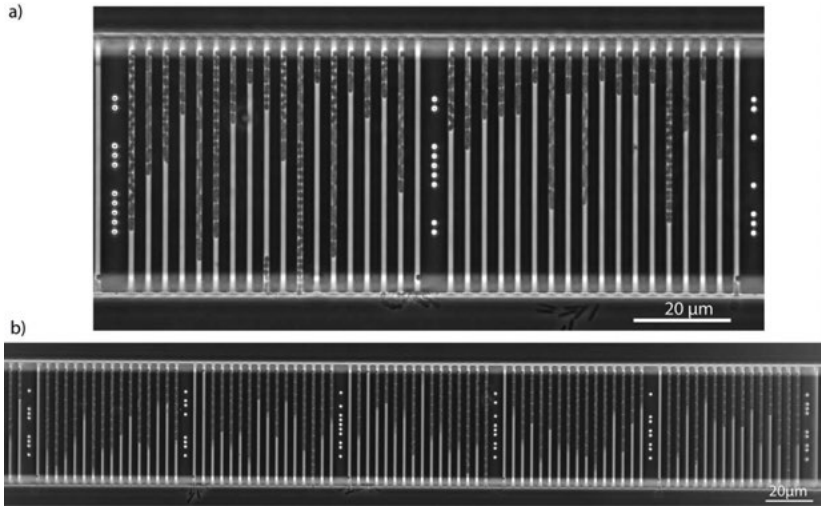



*Figure 3.2.* Phase contrast imaging of the mother-machine in 100x and 40x a) 100x image showing two blocks of the mother-machine device b) 40x image showing 5 blocks in one field-of-view. A block is 15-21 channels separated by a barcode

## 3.2 Optical tweezing

For tweezing cells out of the chip, we have chosen a laser of wavelength 1030 nm and a maximum power of 1W. We motivate this choice by conclusions

from previous studies characterizing photo-damage of *E. coli* in optical traps [5, 20]. The microscope configuration is modified with a home-built module containing a Spatial-light-modulator (SLM), beam expanders, and dichroic mirrors. Using this module, a focused laser beam spot is created at the focal plane (Fig 3.3). The beam can be visualized on the camera overlayed on the phase-contrast images. To enable seeing the cells while we are tweezing without harming the camera, we insert a band-pass filter in the camera path (Fig 3.3). We achieve reliable grabbing of cells when the power delivered to the cells is $\sim$ 50 mW and the fluid flows are controlled appropriately. This power is measured at the objective without a fluidic chip mounted. The time a cell is physically hit by the beam is controlled using an epi-fluorescence shutter to minimize exposure.

The detailed procedure for controlling the flows and valves and executing single-cell extraction is described in **Paper II**. To demonstrate the effectiveness of this method, we used a mixture of wild-type *E.coli* and *E. coli* constitutively expressing the mVenus fluorescent protein. The standard epi-fluorescence mode was used to check the fluorescence before single-cell extractions. Control experiments are also described in detail in **Paper II** and summarized in Chapter 6.

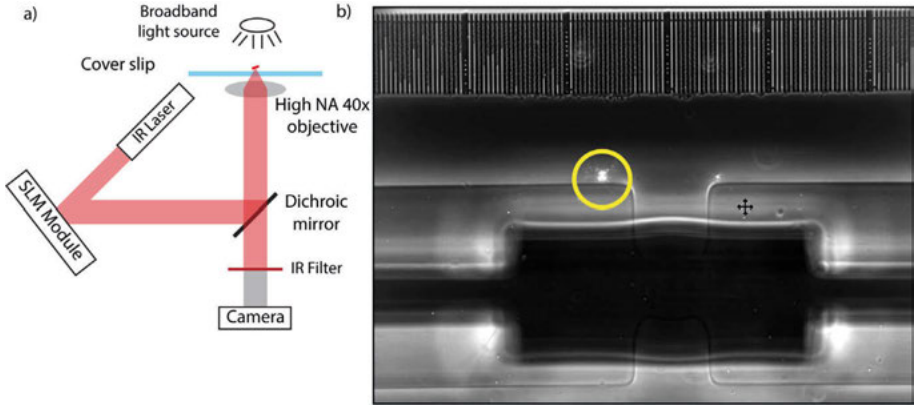



*Figure 3.3.* Optical trapping setup a) A simplified schematic diagram of the optical tweezing setup b) Focused laser beam spot visualized on the camera without the IR filter (center of the yellow circle).

## 3.3 Imaging setup fluorescent protein localization and its calibration

The base microscope used for localizing fluorescent proteins is a standard phase-contrast and epi-fluorescence setup described in Section 3.1. We image multiple fluorescent proteins binding in a diffraction limited spot (called a

dot from here on). The microscope is modified to enable imaging dots in 3D. Phase-contrast and fluorescence imaging have different optical paths. Phase-contrast images are captured on an industrial CMOS camera as before, but fluorescence images are captured on a scientific CMOS camera (Photometrics Kinetix 22) due to its superior noise and gain characteristics and larger field-of-view. A laser of wavelength 515 nm is used as a light source for the single molecule imaging of mVenus-labeled proteins. The signal produced by a single mVenus protein is too weak to be detected against the cell background. Hence, several mVenus proteins are bound in a diffraction-limited spot for localization proposes. Dot localization in 3D is achieved by encoding the z-axis information of the fluorescent protein's location into its image using astigmatism [3]. The optical path of the microscope between the emission filter and the scientific CMOS camera is modified using a cylindrical lens to achieve astigmatism.

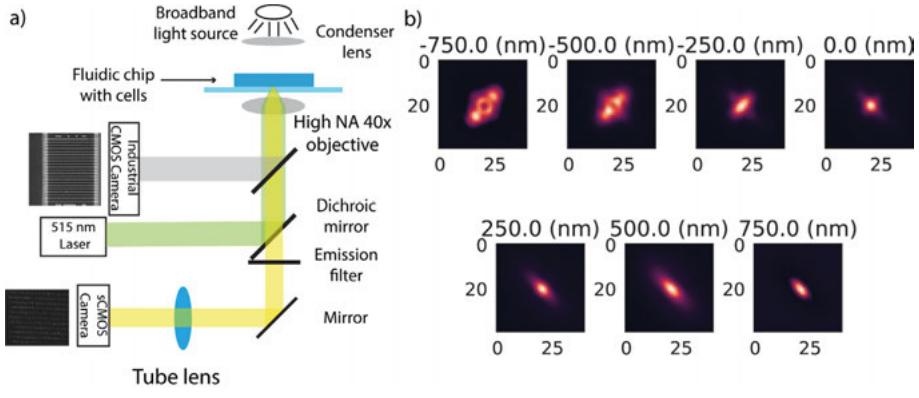



*Figure 3.4.* Dot localization imaging setup and images. a) A simplified schematic diagram of the optical setup. Phase contrast images and fluorescent images of cells growing in a mother-machine fluidic device are captured on two different cameras. b) Point spread functions sampled from the spline model at different z-axis locations.

The point-spread-function (PSF) of the microscope was approximated using cubic splines fitted on images of bright fluorescent 100 nm beads imaged on the astigmatic optical path. The SMAP tool [25] was used for fitting cubic splines. The pixel-wise noise and gain parameters of the camera were calibrated before imaging and are described in **Paper III**. The spline model of the PSF and camera calibration parameters were used to train neural network models for dot localization in 3D (Chapter 4.9). Fig 3.4 shows a schematic diagram of the optical setup and the fitted PSF sampled along the z-axis to illustrate astigmatism.

We used this setup together with the mother-machine device described in the previous chapter to perform dot localization in *E. coli* over multiple division cycles. Different loci on the *E. coli* chromosome as well as the *E. coli* replisome

complex were labelled with fluorescent proteins and imaged on this setup. An example image of cells with emitters is shown in Fig 3.5a.

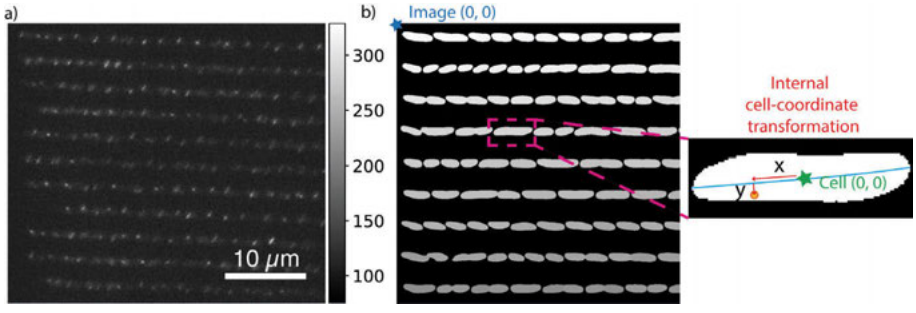



*Figure 3.5.* a) Sample fluorescence image of *E. coli* cells growing in a mother-machine device, where chromosomal loci are labeled with mVenus-conjugated DNA-binding proteins. b) Mapping of dot localization (orange dot) from image coordinates to the cells' internal coordinates using the cell backbone (blue line). The image origin is shown with a blue star and the cell origin (green star) is the mid point of the cell backbone line.

Phase-contrast images were acquired in conjunction with fluorescence images and used to map localizations from global image coordinates to coordinates with respect to the cell's long axis and short axis using a cell segmentation mask (Fig 3.5b). Accurate transformation between images captured on the two different cameras was necessary to realize this mapping. 100 nm fluorescent beads were localized on both cameras and an appropriate image transformation was fit for this purpose.

We briefly describe the emitter localization process in chapter 4.9 and summarize the results of **Paper III** in chapter 7.

# 4. Image analysis methods

Modern microscopic techniques facilitate studying biological processes at various spatial and temporal scales. We are able to acquire large microscopy datasets at unprecedented rates [26, 27, 28] due to advances in hardware (cameras, stages, electronics, optical components, data storage, etc) and software. Analyzing microscopy images is a central problem faced by most biologists. In the context of image analysis, the things of primary interest to us are tasks such as cell segmentation, cell tracking, fluorescence measurements of various kinds, and single-molecule localization. In the following sections, I describe the advances made by us in relation to the requirements of the projects I have worked on during this thesis.

## 4.1 Image processing using neural networks

Classical algorithms in image processing rely on a deterministic sequence of operations on pixels of images using different convolutional filters, pixel aggregation methods, noise reduction techniques, thresholding methods, etc. Handcrafting operations on images was standard procedure until a decade ago and required deep domain experience with image manipulation. However, over the past decade, image processing has been revolutionized by advances in training large neural networks reliably, the use of commercial graphic processing units, and the availability of large datasets [29, 30, 31, 32, 33, 34, 35].

Modern deep-learning models consist of a series of layers containing parameters/weights that define operations on data. These parameters are learned during the training of the network. New ways of stacking layers, training deep layers stably, great software stack in Python, and fast computations on GPUs have driven progress in deep learning. Deep-learning models also have the advantage of learning and combining higher-order features that can't be handcrafted. Modern computer vision is dominated by deep-learning models where many problems are formulated as transformations from inputs to outputs, with a neural network doing the transformations using learned parameters. In supervised learning tasks, such as the tasks we are primarily interested in in this work, outputs/labels are ground truths obtained with human annotations or realistic simulations.

## 4.2 Loss functions and training schemes

Loss functions map discrepancies between the values predicted by a model and the ground truth onto a real-valued number. In most image processing tasks we are interested in in this thesis, loss functions vary depending on the task. They are described in detail in the supplementary materials of each paper.

Minimization of the value of the loss function is done using iterative algorithms such as SGD and its variants [36, 37]. Training schemes describe the process of adjusting the parameters during the optimization process until convergence of the loss function to a minimum is achieved. Learning rate, batch size, and learning-rate schedule are amongst the most common training hyperparameters updated during the training processes in most of the models described in this thesis.

## 4.3 Central working horse aka U-net

Many image processing tasks can be modeled as problems of finding a mapping from an input image (HxW) to one or more images of the same size. Among recent developments in convolutional neural network methods, the U-net architecture [38] is a robust way to map images to images. The U-net has a down-sampling path and an up-sampling path, with connections between these paths where the spatial resolutions of the image volumes match (Fig 4.1). Each down-sampling stage consists of convolution-normalization-non-linearity blocks, followed by a reduction in spatial resolution. Each up-sampling stage consists of increasing spatial resolution and additions from the down-sampling path. After the final up-sampling stage, the required number of output image channels are generated using 1x1 convolutions. The number of stages, number of convolution layers, and size of convolutions are hyperparameters. The run-time performance of the network can also be controlled by the number of layers and stages in the network. The universal architecture is capable of integrating features over different scales and learning mapping from one image domain to another.

I used the U-net architecture (Fig 4.1) extensively for the segmentation and dot localization tasks described in this thesis. Although U-net-based architectures have also been used for cell-tracking [39] and fluorescence classification, we looked for more lightweight algorithms for these tasks. In the following sections, I describe each image-processing task briefly and contrast the methods I developed for the thesis against standard techniques.

## 4.4 Cell segmentation

Segmentation is the process of partitioning the pixels of an image into groups that represent a common object, or internal regions of an object, that have use-

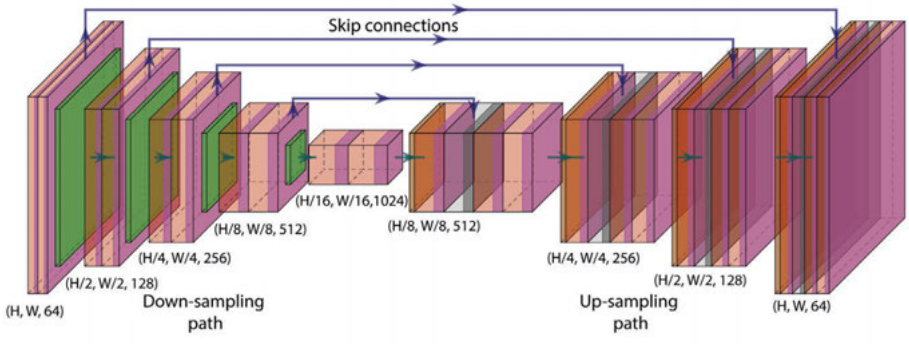



*Figure 4.1.* U-net architecture illustrating down-sampling, up-sampling, skip connection paths, and sizes of the image/tensor volumes.

ful semantics associated with them. When we refer to cell segmentation, we are mainly interested in grouping pixels that belong to a particular bacterial cell, thereby generating a unique label for each cell in the image. If the objects in the image are separable, i.e., they don't share a common boundary, the problem simplifies to a binary classification of each pixel into two classes, background or object. Connected-component labeling methods can then be used on these binary images to generate unique labels for each object. If the cells share a boundary, methods that label instances of different objects uniquely are preferred. Essentially, segmentation can be re-framed as mapping of images from domain $\mathrm{R}^{H\times W} \in \{0,1\}$ into $\mathrm{N}^{H\times W} \in \{1,2,3...N_{max}\}$, where $N_{max}$ is the largest number of objects in the image (Fig 4.2).

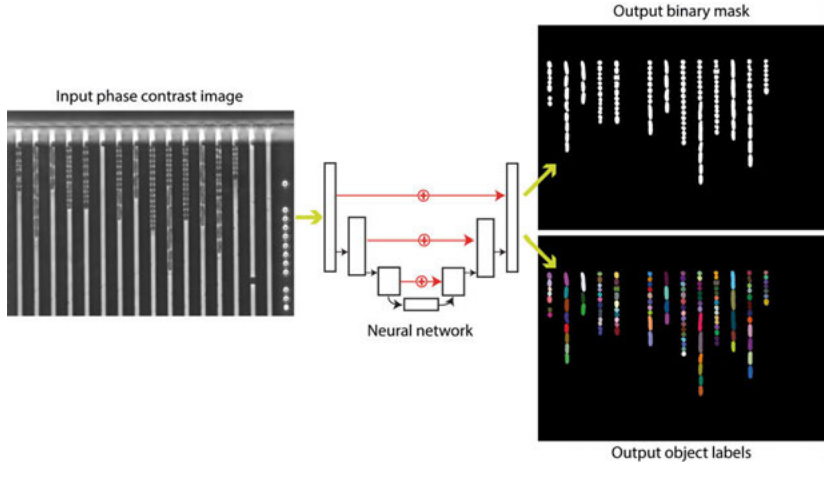



*Figure 4.2.* Expected input and output from the cell segmentation process for a) binary classification and b) object labelling

.

Algorithms can also be classified along different axes such as computational complexity, parallelization on computing hardware, and generalization with respect to imaging conditions and cell identity. Per-object-ellipse fit (POE) [40] and SuperSegger [41] are methods that rely on pixel intensities in phase-contrast images and properties of the object being segmented, typically elliptical/rod-shaped bacteria grown on agarose pads. These classical algorithms usually rely on a few parameters set by the user depending on the data acquired. U-net-based methods have recently been shown to perform segmentation better than most classical algorithms. Unlike classical algorithms, neural networks have a large number of parameters that need to be learned from data.

U-nets have been successful in generating binary masks of cells imaged in different modalities [42], and 3D extensions of this kind of architecture have been used to segment image volumes [43]. U-net-like dense architectures have also been used to predict intermediate representations of objects, which can later be used to generate labels for each object. For example, in StarDist [44], the radial distance of the pixel to the center of the object to which it belongs is predicted for each pixel. In the case of Cellpose [45] and Omnipose [46] algorithms, radial fields based on heat diffusion and eikonal equations are predicted respectively. These fields are then traced to generate object labels. Fig 4.3 summarizes the transformations from input phase-contrast image to output labels using the Omnipose algorithm.

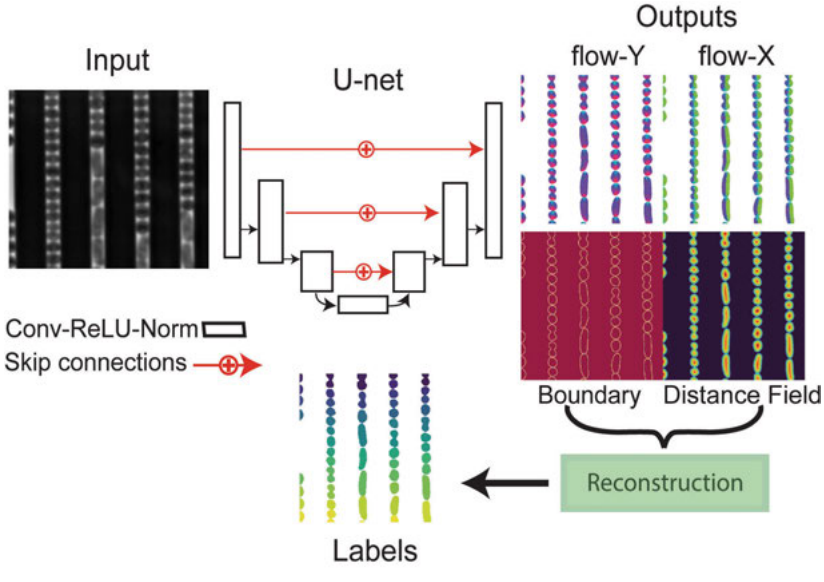



*Figure 4.3.* Omnipose prediction pipeline showing the inputs and outputs of the U-net architecture. The outputs from the network are reconstructed to produce unique labels for each cell. The figure is reproduced from Paper I

The training of neural networks is done on datasets acquired using both hand-labeling and semi-automatic labeling strategies, such as using fluorescent markers to label cells and curating the data to obtain final labels. In the work described in **Paper I**, we used the Omnipose method to segment cells of different shapes and sizes. The outputs of the network 4.3 were used to assign each pixel a number such that different cells received a unique ID. However, this assignment was noisy and required a spatial clustering step, which is the slowest step in this method. The clustering of noisy labels was expedited $\sim 2$x using a parallelized clustering algorithm [47]. We also showed that the U-net for binary classification of *E.coli* performed as well as Omnipose, but Omnipose generalizes to cells of different shapes and sizes and offers the advantages of separating touching cells (Fig 4.4). In **Paper II**, where we developed a real-time image processing system for concurrent image acquisition and analysis, we used the U-net for binary classification of cells and mother-machine traps. By tuning the number of channels in each layer of the U-net, the run-time performance of this operation was reduced to make real-time segmentation on large images (4096 x 1024 pixels) feasible.

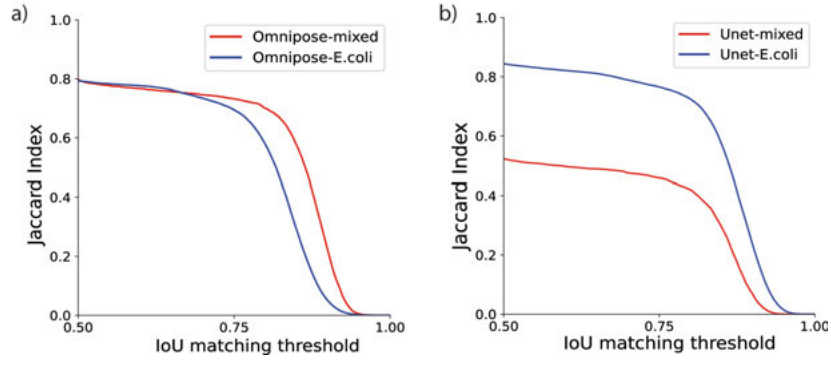



*Figure 4.4.* Performance improvements in Omnipose and U-net for binary classification. Metrics showing that Omnipose does well on cells with mixed morphologies compared to a binary classification using U-net.

## 4.5 Cell tracking

Cells growing in microfluidic traps can be tracked over time to monitor changes in their state such as growth, division, death, or leaving the trap. The goal of cell-tracking algorithms is to construct cell lineages by linking cells from one frame to the next. Tracking algorithms involve formulating multiple hypotheses of possible lineages and pruning these hypotheses based on criteria specific to the problem. This pruning process is performed by formulating a cost function for each possible state in the hypothesis and minimizing the cost function to find the optimal lineage [48, 49]. Due to the exponential nature of formu-

lating multiple hypotheses, the number of hypotheses is typically mitigated by using only a few frames or by limiting the set of possible links using spatial and/or temporal constraints. In addition, cell-tracking algorithms need to consider possible mergings and splittings of objects in the image due to segmentation errors.

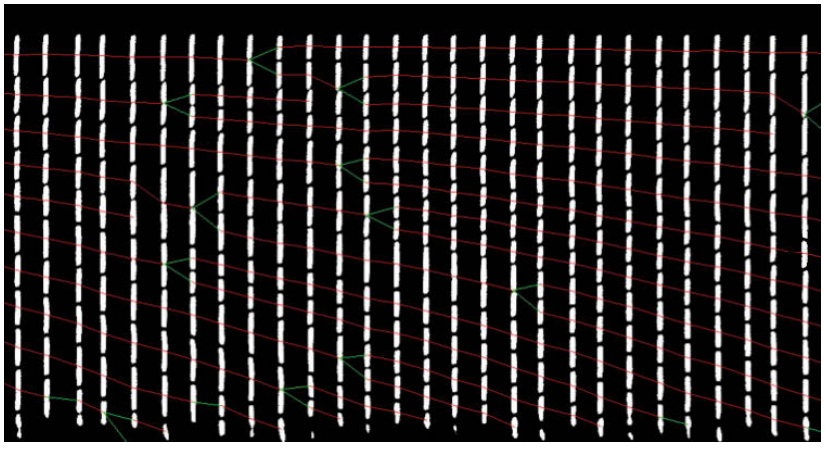


*Figure 4.5.* Lineage tracing of cells in one mother-machine trap.

In cell-tracking algorithms for mother-machine devices, the primary way to constrain the number of multiple hypotheses is by using overlapping measures such as Jaccard indices to construct a set of possible links between frames. Cost functions are usually defined for each type of event, such as division, merging, growth/movement, or leaving the channels, and tuned for every new dataset before minimization and identification of the optimal set of links. The optimal set of links can be found considering all available frames or frame-by-frame. These links are used for constructing cell lineages. Methods using overlapping measures fail to perform well when the input data has high variability in growth rates, shapes, and size. Overlapping metrics are sensitive to drifts between frames and require registration of successive frames. Recently, algorithms for predicting mother-daughter relationships using U-net [39] have been shown to give very high performance, but are computationally expensive because the network inference is done based on full images for each individual cell. Evaluating the performance of different tracking algorithms for mother-machine datasets is difficult due to the lack of standard collectively agreed-upon ground-truth datasets.

The Cell Tracking Challenge [50] bench-marked various segmentation and tracking algorithms over the past decade on a variety of biological datasets. Although deep learning-based algorithms have made significant progress on segmentation tasks, no statistically significant differences were observed in the performance of tracking tasks between machine learning and non-machine learning methods. For the work in this thesis, we switch between different

tracking approaches depending on the type of datasets. We developed and used three different tracking algorithms. In **Paper I**, we have mixed-species cultures growing in the traps, which require algorithms robust against cells having vastly different overlapping measures and growth rates. Using the recent developments in Siamese-networks [51] and graph formulations [52] for object-tracking, we developed a network-based approach for constructing links between cells of adjacent time points. We only modeled growth/movement and division events for this application and used cell properties such as centroid location, area, ellipticity, etc. Other events, such as cells leaving or merging due to segmentation errors, are assigned in the post-processing. Training data for this tracking algorithm was obtained using simulations of cells of different shapes and sizes growing in traps, similar to CellSium [53]. Fig 4.6 shows the input and output of the tracking network, example training data, and performance measures of the network evaluated on 100 simulated time-series stacks.

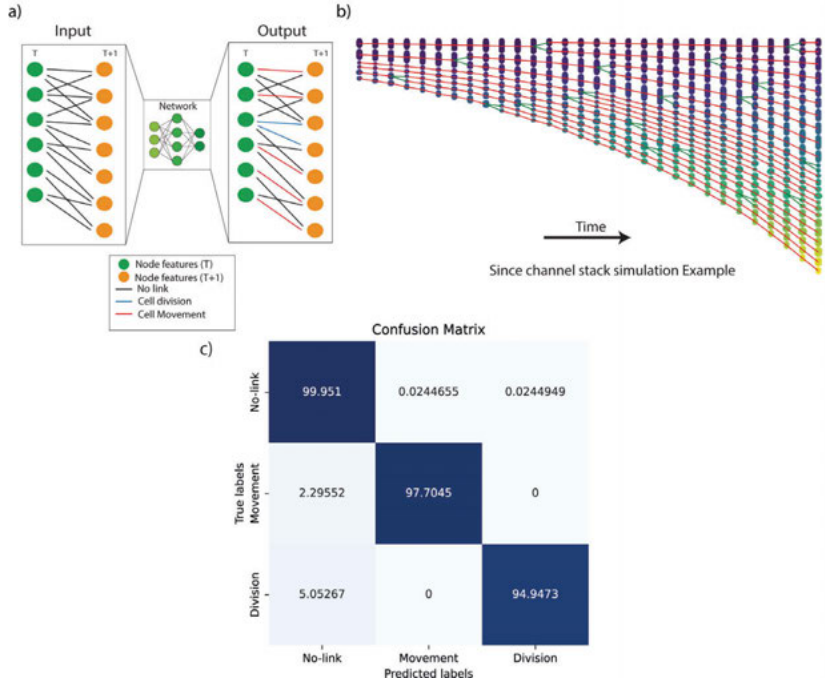



*Figure 4.6.* Cell tracking a) Illustration showing the input and output to the cell tracking network for classifications of links between cells from one frame to the next. b) Simulation of cells growing in a single mother-machine trap used for training the cell tracking network. c) Confusion matrix showing the link classification performance.

In **Paper II**, where real-time performance is required, we used a tracking scheme from [54] modified to assign priorities based on a one-dimensional Jaccard index and controlled for errors in post-processing. This frame-by-frame approach is suitable for real-time track construction. In **Paper III**, where glob-

ally optimal tracks should be constructed over all available image frames, we used a previously developed tracking algorithm [48].

## 4.6 Barcode detection

Mapping a mother-machine trap from timepoint $t$ to timepoint $t-1$ is a critical step in lineage construction. When using a 100x objective with the fluidic chip containing only two rows of mother-machine traps (Fig 2.1c), we can array field-of-views such that the trap ID barcodes appear in the same region in every image acquired. During post-experiment image analysis, trap-to-trap mapping in time is done using barcode image registration and intensity profiles on phase-contrast images or by counting traps on a trap segmentation mask. However, on the larger chip with 100K traps, the variations in fabrication of the two layers result in non-uniformities making trap-to-trap mapping difficult.

For single-cell isolation tasks, we need to localize the traps in our chips from which we would like to grab cells with interesting phenotypes and extract them for further analysis. The chip design includes barcodes every 21 traps (block), that can uniquely decode the row and channel block numbers on the 100K chip. We use a YOLOv3 [55] network with modifications to its anchor box sizes to detect and keep track of both the left and right barcode for every trap. This indexing also helps us verify that we are picking cells from the right block on the chip and solves the problem with varying barcode locations on different field-of-views. Fig 4.7 shows bounding boxes of barcodes detected on a 40x phase-contrast image.

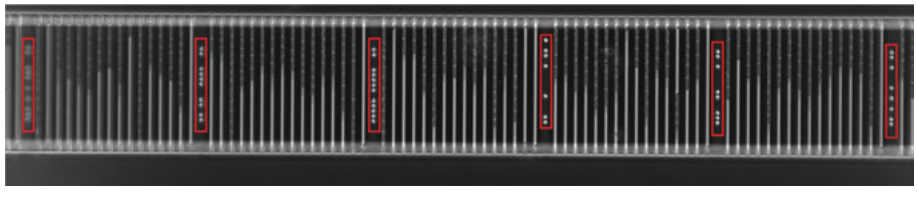


*Figure 4.7.* A phase-contrast image in 40x with detected barcodes shown in red bounding boxes.

## 4.7 Real-time system

To facilitate the identification of interesting phenotypes in **Paper II**, a real-time image processing sequence was designed by combining the tools described in the previous sections (Fig 4.8). Microscope event acquisition was controlled using Pycromanager [56]. Monitoring experimental results during the run of the experiment was made possible by a graphical user interface (GUI). The GUI can be used to monitor cell lineages and growth rates of individual traps

and flag interesting channels for single-cell isolation. The pipeline is divided into multiple stages designed to work in parallel on both GPUs and CPUs depending on the task involved. For this work, we only used cell segmentation, tracking, barcode detection, and growth measurements in the pipeline, which means every operation could be done in parallel in less than 500ms on images of size 4096x1000 pixels containing $\sim$ 100 traps each. Additional tasks can be added to the pipeline as long as they run in < 500ms. The 500ms upper limit is set by the microscope stage movement and focus stabilization. Any operation taking longer than the upper limit results in the accumulation of data over time and measures are required to avoid memory blowout.

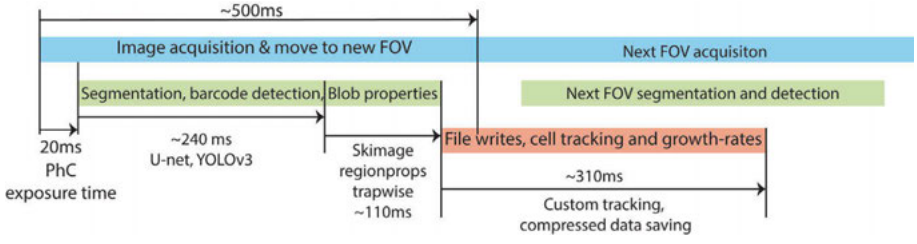



*Figure 4.8.* A schematic figure showing the real-time system for cell segmentation, tracking, and growth rate calculations

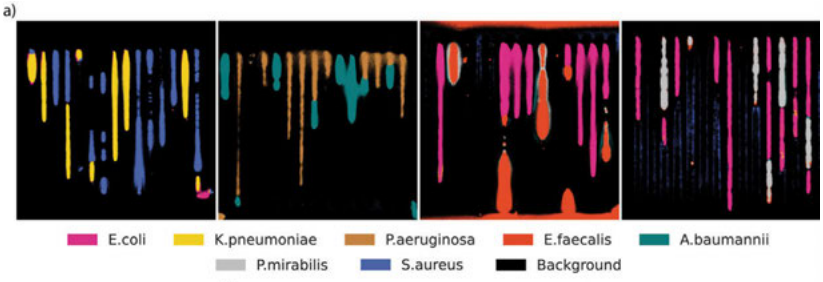
## 4.8 Classification tasks on fluorescence data

In assays that recover the genotype information of each trap in the mother-machine device, a fluorescence signal of some sort is generally used. Methods to convert the fluorescence signal to the corresponding genotype are critical in mapping phenotype-genotype relationships. Generally, barcoding schemes [14, 57, 58] based on fluorescence *in situ* hybridization (FISH) [59, 60, 61] are used for this purpose. In **Paper I**, where we needed to recover species information, abundant ribosomal RNA was tagged to generate a signal. We built a species-wise classifier for 8 classes (7 for different bacterial species and one for background) using random forests on the signal from 4 imaging channels. This species information was later used for species-wise AST. Fig 4.9 shows FISH classification results on the two-color-combination assay used to decode species information.

## 4.9 Dot localization

Single-molecule localization methods are used to localize fluorescently labeled molecules inside cells with a precision of a few tens of nanometers, far below the diffraction limit. The localization precision of a diffraction-limited spot

a)




b)

| Species | Precision | Recall | f1-score |
|---|---|---|---|
| E.coli | 0.9910 | 0.9977 | 0.9943 |
| K.pneumoniae | 0.9962 | 0.9980 | 0.9971 |
| P.aeruginosa | 0.9934 | 0.9974 | 0.9954 |
| E.faecalis | 0.9912 | 0.9705 | 0.9807 |
| A.baumannii | 0.9924 | 0.9921 | 0.9922 |
| P.mirabilis | 0.9705 | 0.9966 | 0.9834 |
| S.aureus | 0.9659 | 0.9800 | 0.9729 |
| Background | 0.9910 | 0.9291 | 0.9446 |

*Figure 4.9.* FISH classification a) Four example images showing pixel-wise classifications of the seven species classifier. b) Metrics showing the classifier performance.

(called dot from here on) depends on several factors such as the number of photons emitted by the fluorophores, camera noise, background photons emitted by the sample, exposure times, and information content in the point-spread-function. To perform dot localization in 3D, we engineered the microscope as described in chapter 3.3 to produce images that encode 3D locations of fluorescent emitters.

Until recently, most successful methods for dot localization of general point-spread functions were based on spline models and maximum-likelihood estimation techniques [62, 63]. Neural networks have now captured this domain and replaced maximum-likelihood-estimation (MLE) methods as state-of-the-art [64, 65, 66]. As there are no good labeling experiments to generate ground-truth training data for localization algorithms, we rely on accurate simulations of the point-spread function, fluorophore blinking dynamics, camera physics, and background noise models to generate ground-truth data.

Cubic spline models of the point-spread function are built from z-stack images of fluorescence beads immobilized on coverslips. Spline models are used to approximate the point-spread function and model fluorescent emitter image generation as a function of the real values in x, y, and z (Fig 3.4c). Emitter-localization algorithms for STORM/PALM-based imaging [67, 68, 69] assume a fixed-background pattern sampled from a uniform distribution or structured diffuse background sampled from perlin-noise models at varying spatial scales [70]. However, these assumptions are not valid for our use case (Fig 3.5a), hence requiring a custom cell background model. We used the cubic spline model together with models of the background photons and camera noise to

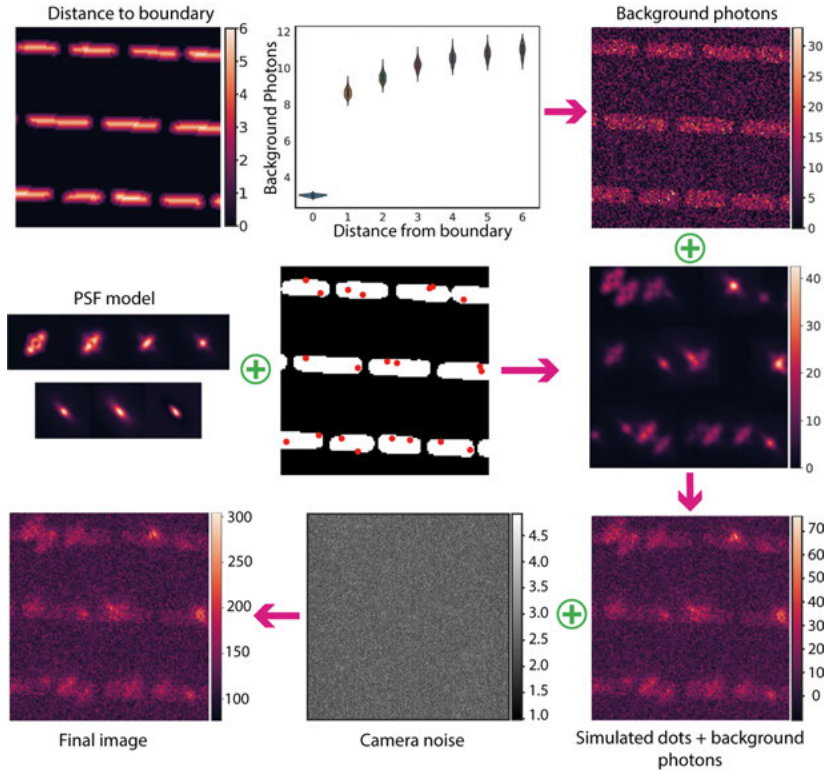



*Figure 4.10.* Sampling emitters inside cells with known emitter locations, cell background model, PSF model, and camera noise.

generate realistic images, where the locations of emitters are known with sub-pixel resolution (Fig 4.10).

In **Paper III** we used the mother-machine device (Fig 2.1) to image *E. coli* with fluorescently labeled chromosomal loci. Models for the background photons from the cells were estimated from experimental data. Fig 4.10 shows cell background as a function of the distance field inside cells and sampled simulated images of emitters and background photons. We set the simulation parameters, such as ranges of photon counts and density of emitters inside cells, using values observed in experiments. It can be observed Fig 4.10 that photon count to background falls in the low SNR regime observed in traditional STORM/PALM datasets.

The network architecture that we used to fit localizations is also based on a U-net. We used a neural network-based dot localization method [64] with a field-dependent variation[66] that adds a pixel-location-dependent field to the data so that the model learns pixel location-dependent data features like sCMOS camera noise in our case. The inputs of the networks are images from cells with dots, and xy location on the camera chip. There are 10 output channels, one for the probability of having an emitter in a pixel, 3 for offsets of x,y,z

from the center of each pixel, 3 for $\sigma_x, \sigma_y, \sigma_z$, and 1 for the PSF image of predicted emitters. The outputs were used to construct a list of localizations and their uncertainties. Fig 4.11 shows the inputs and outputs of the localization network.

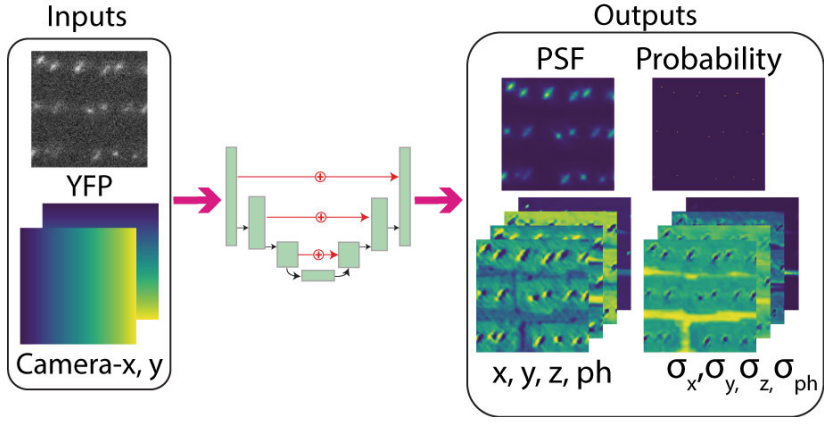



*Figure 4.11.* Dot localization network illustration showing input and output images

The loss functions used to train this model were combinations of the localization loss for constraining localization precision, count loss for the number of emitters on an image, cross-entropy for the probability of emitters, and mean-square error for the noise-free PSF image. Each of these losses was weighted to be in the same order of magnitude. The loss functions are described in detail in the supplementary material of **Paper III**.

The localizations obtained from the network are lists of xyz coordinates and corresponding uncertainties relative to the origin of the image. We convert these localizations into the cell's internal coordinates (Fig 3.5) using the image cell segmentation masks and image transformations described in chapter 3.3.

In chapter 7, I summarize the results of applying this dot localization approach to study several loci of the *E. coli* chromosome as well as the replisome complex.

# 5. Paper I

Antibiotic susceptibility testing (AST) is a critical part of reducing the problem of antibiotic resistance observed worldwide [71, 72]. Identifying the pathogen species as well as its susceptibility to antibiotics is needed at point-of-care [73], especially in life-threatening conditions such as sepsis [74]. The current gold standard for species identification, MALDI-TOF, requires pure cultures of single bacterial species which can take several days to acquire from a patient sample [75]. This procedure is particularly challenging for mixed species infections observed in sepsis, catheter-associated UTIs [76], and community-acquired pneumonia, among others [77]. In this paper, we build on the previously published phenotype-based susceptibility testing [2] to perform both species identification and susceptibility testing on the same sample. We primarily target sepsis-related pathogens and cover 95 % of the most common bacterial species causing sepsis.

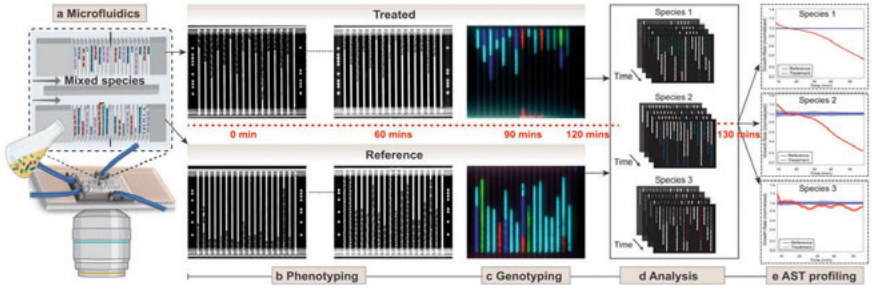



*Figure 5.1.* a) A cartoon of the microfluidics setup with the mixed species sample loaded on the chip. b) Time-lapse phase-contrast images of the cells in the traps when grown in media with (top) and without (bottom) antibiotics. c) Fluorescence images of the bacteria with ssDNA probes targeting the ribosomal RNA of specific bacteria for species identification. d) Analysis of time-lapse stacks and species ID using deep learning for segmenting and tracking cells. e) Detection of AST profiles for individual pathogens at a given antibiotic concentration. Part of Fig. 1a created using www.biorender.com. Figure reproduced from Paper I.

For all the assays described in this paper, we used the mother machine device shown in Fig 2.1c with two parallel separate rows of $\sim$ 3000-4000 mother machine traps. One row is supplied with media containing antibiotics and the other with plain media. The traps were loaded with a mixed culture of different bacteria. After 60 minutes of phenotyping in 100x using phase-contrast microscopy, we performed a genotyping protocol to label each species with

fluorescent probes. The genotyping protocol is based on the fluorescence *in situ* hybridization (FISH) method. After the genotyping protocol, we imaged the cells in fluorescence channels and mapped the signal back to the species. We used the species information to label the cell tracks and calculate species-wise growth rates with and without antibiotic exposure. Figure 5.1 shows the overall workflow of the assay.

The number of species we can distinguish by labeling them with specific fluorescence probes is limited by the number of fluorescence channels we can image on the microscope. With four channels, we can image four species at a time using direct probing where one fluorescence probe corresponds to one species. By using combinations of two or more fluorophores to distinguish individual species, we can increase the number of detected species in one round of the genotyping protocol. In this paper, we used up to two specific fluorescence probes to tag a species and increased the number of species detected in one round of the genotyping protocol to ten (Fig 5.2). This was done using species-specific adapter probes with multiple arms that could be tagged with different fluorophore combinations. Fig 5.3 (left column) illustrates the FISH signal observed when seven species were genotyped in the chip.

The phase-contrast images were segmented using the Omnipose neural network and the cells were tracked using the custom algorithms described in Chapter 4. Cell tracks were corrected for errors and single-cell growth rates were calculated by fitting exponential curves on a rolling window of five frames, corresponding to 10 min. Growth rates of each species were calculated for the cells receiving antibiotics and the non-treated control at the single-cell level. Example results of the normalized growth rate curves where two species were loaded in the chip with multi-color probing used for genotyping are shown in Fig 5.3 (right column).

To conclude, we show an experimental protocol and image processing algorithms that can perform rapid antibiotic susceptibility testing with species identification in a $\sim$ 2-hour time frame for seven species.

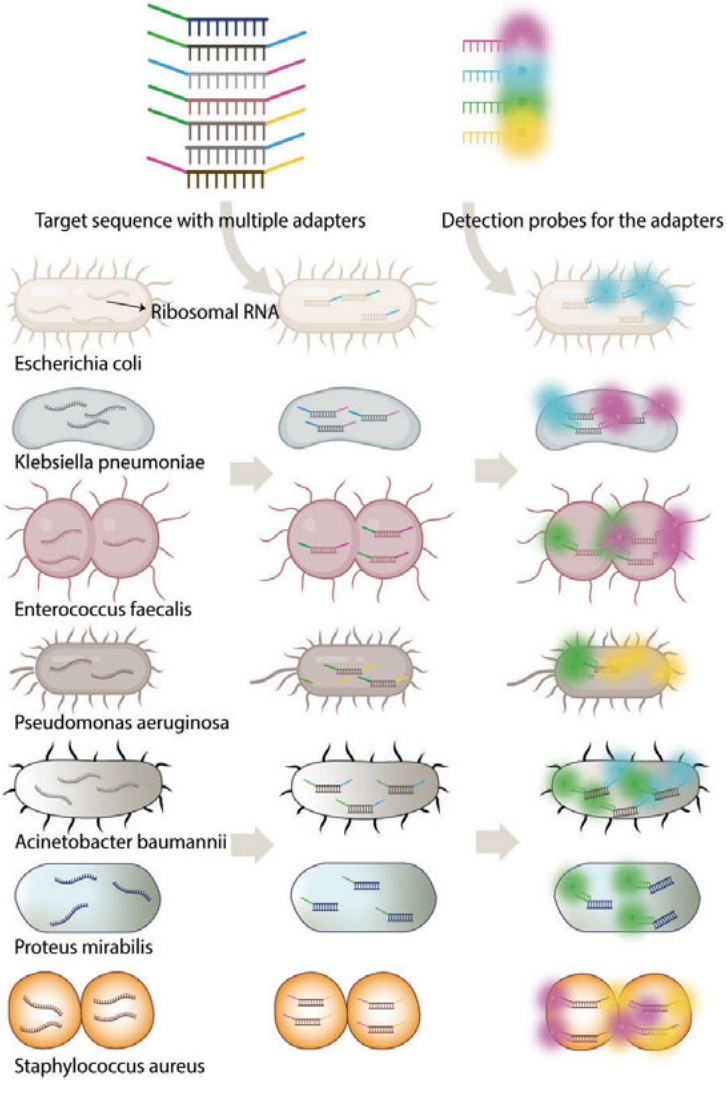



*Figure 5.2.* Overview of the combinatorial FISH probing for multi-species identification. A cartoon illustrating the different bacterial species with their ribosomal RNA (left). Illustration of the specific sequences with multiple adapters targeting the ribosomal RNA of individual bacteria and its hybridization to the target rRNA (middle). Detection probes with different fluorophores. Hybridization of detection probes to the adapter sequences along with unique sequences that are targeted to the species specific rRNA (Right). Figure reproduced from Paper I

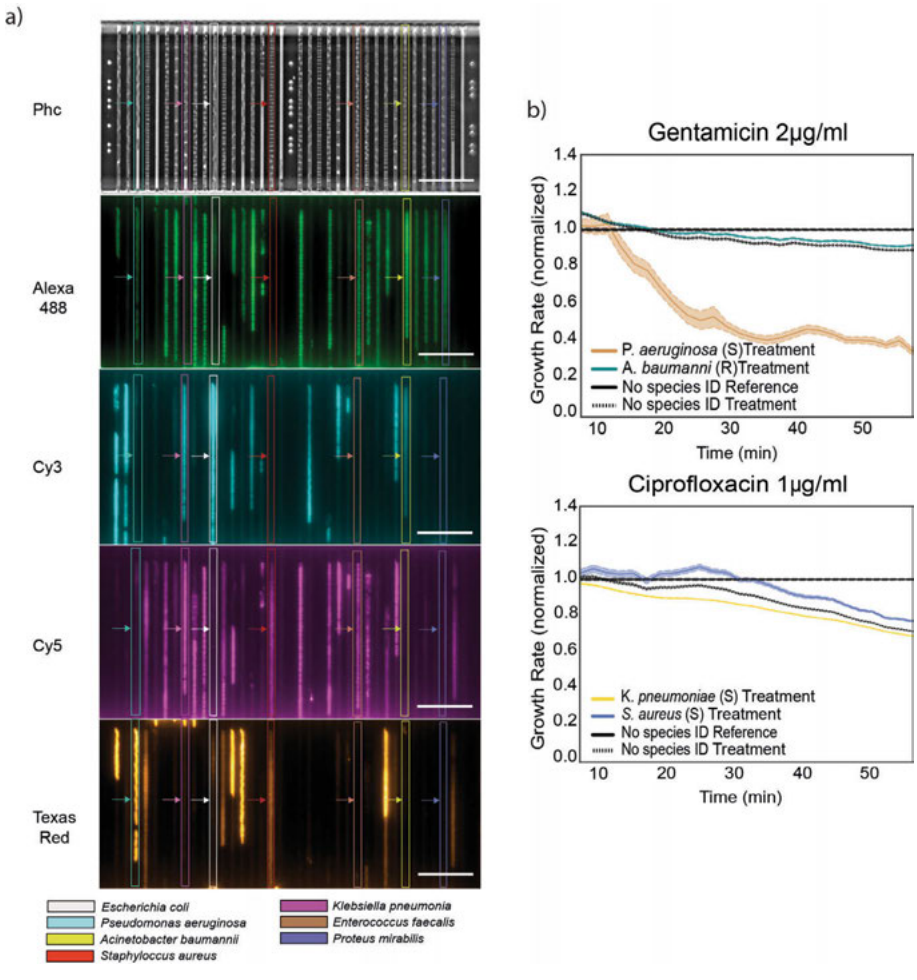



*Figure 5.3.* a) Example images (Scale bar 20$\mu$m) of mixed species loaded in the microfluidic chip and probed using combinatorial FISH for species identification. After the hybridization step, cells were imaged in different channels (PhC, Alexa 488, Cy3, Cy5, and Texas Red). The bacterial species are marked in white (Escherichia coli), magenta (*Klebsiella pneumoniae*), cyan (*Pseudomonas aeruginosa*), brown (*Enterococcus faecalis*), yellow (*Acinetobacter baumannii*), navy blue (*Proteus mirabilis*), and red (*Staphylococcus aureus*). b) Normalized growth curves from two different AST + Species-ID experiments. Figure reproduced from Paper I

# 6. Paper II

The goal of this paper was to show a single-cell extraction method on a large-scale mother-machine device. We used the 100K chip described in Chapter 2 and performed control experiments to demonstrate the ability to extract single cells from the chip. The control experiment was performed using a 1:10 mixture of *E.coli* cells in the chip where the rarer fraction expressed fluorescence. By repeatedly extracting fluorescent cells from different regions of the chip, we showed that we could reliably get single cells of our choice out of the chip. The method to extract single cells is described in Chapter 2.4 and the optical trapping setup used for this single-cell extraction process is described in Chapter 3.2. In Figure 6.1, we show snapshots of the extraction process for one cell.

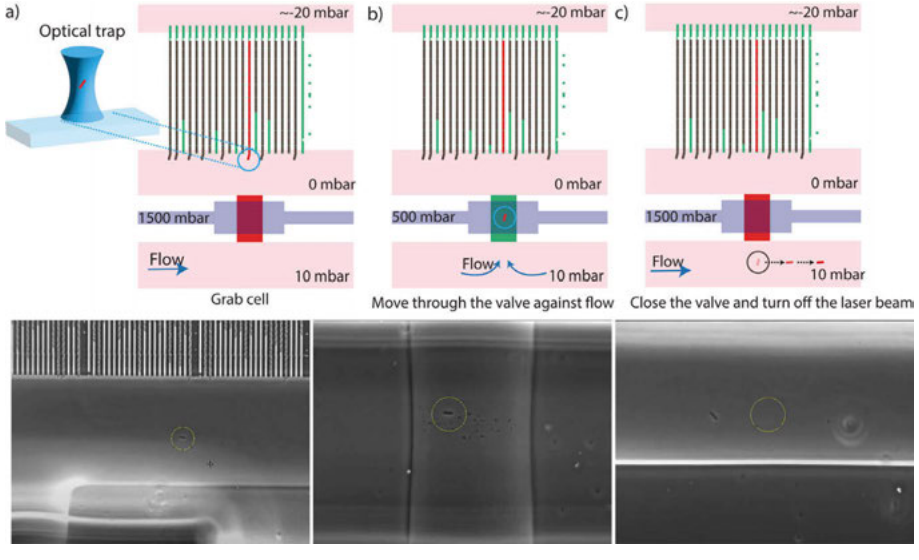



*Figure 6.1.* Tweezing process a) Top: A cell from the end of a selected mother-machine growth channel being trapped at the focal point of a laser beam. The cartoon also shows a clean channel and a pressurized valve in closed mode. Bottom: Phase contrast image with a cell trapped in an optical trap (yellow circle) corresponding to the top image b) Top: Cell being moved from the growth region to the clean channel while the valve is open. Bottom: Phase contrast image with cell being tweezed through the open valve. c) Top: Cell being released in the clean channel after the valve has been closed. Bottom: Image showing cell leaving the optical trap after the laser has been turned off. Part of this figure is reproduced from Paper II

We repeatedly picked five fluorescent cells out of the chip in three independent experiments to ensure the repeatability of the process. We showed that we could reliably extract 13 out of 15 cells, while 2 got stuck at the exit port filters. Fig 6.2 shows two agar plate images imaged on a fluorescence plate reader, where the left plate contains a dilution of media used for loading the chip and the right plate has the cell picked out of the chip. From this we can conclude that we can achieve single-cell extraction using this chip.

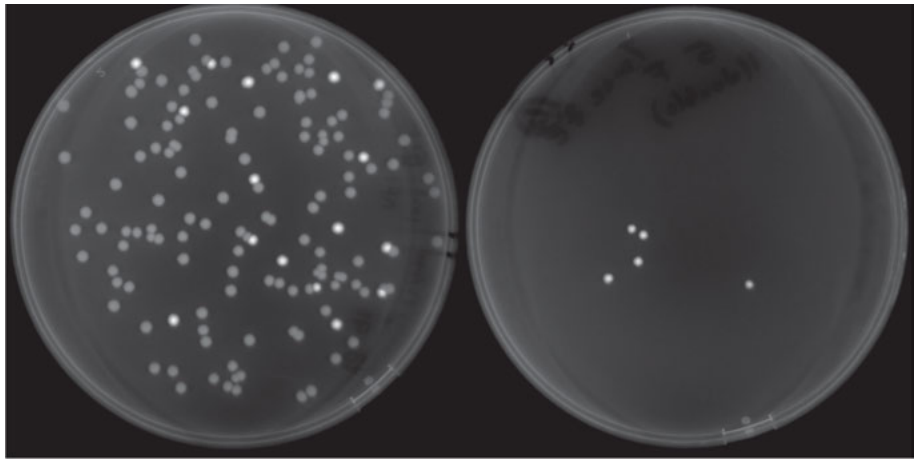


*Figure 6.2.* Control plates imaged on a fluorescence plate reader in the Venus channel. Left: Colonies from the cell mixture loaded into the 100K chip. Right: Colonies from the cells tweezed out of the chip.

The second important part of this paper is the companion software that we used to identify interesting traps from which to extract cells, in real-time. The real-time pipeline includes cell segmentation, cell tracking, single-cell growth rate calculations, and image acquisition sequence programming. Most of these components are described in detail in Chapter 4 and in Paper II.

# 7. Paper III

In this paper, we were interested in studying the 3D distributions of different loci of the *E. coli* chromosome as well as the replisome complex. Three strains of *E. coli* were constructed. In these strains, binding sites for a DNA-binding protein fused to a fluorescent protein (YFP) were introduced on the chromosome near the origin, terminus, and midway between the origin and terminus, respectively. The replisome complex was also tagged with YFP.

We used the mother-machine device (Fig 2.1c) to constrain the cells in a mono-layer and grow them with a fresh supply of media. The optical setup described in Chapter 3.3 was used to acquire images in both fluorescence and phase-contrast mode at each time point. A neural network-based dot localization method (Chapter 4.9) was used together with image transformations between phase contrast and fluorescence (Chapter 3.3) to generate distributions of loci inside the cells as a function of cell size.

The loci detected in the fluorescence images were mapped onto the cell segmentation masks obtained from segmenting phase-contrast images (Fig 3.5b). The position of the loci relative to the cell backbone was computed for the x and y positions of the dots. The boundaries of a cell in the z-axis were obtained from rotating the y-axis boundaries. The z-axis positions relative to the cell could not be determined with this method as we cannot determine the origin of this axis from one image alone. The z-axis positions obtained from the neural network are relative to the origin of the point-spread function (Fig 3.4b). Hence, the origin on the z-axis was moved with a fixed offset after correcting for variations in z across the field-of-view (FOV) such that the dots fell inside the boundaries of the cells for illustration purposes.

We pooled all localizations from tracked cells of a single FOV on the mother-machine device over 300 frames and binned them by cell size. Fig 7.1 shows the distributions of loci as a function of cell size for three different regions on the chromosome. The first three rows show the distributions of the loci as a function of x, y, and z-axes of the cell. The last row represents the yz distribution, which is the main result of this paper. We showed that the midway locus avoids the nucleoid of the cell. The terminus locus had a much tighter distribution localized in the central volume of the cell, while the origin was more spread out. We also showed that these distributions were symmetric in y and z as expected from the rotational symmetry feature of the cells.

The *E. coli* replisome was also studied in a similar fashion. Fig 7.2 shows the corresponding distribution plots on the x, y, z, and yz axes.

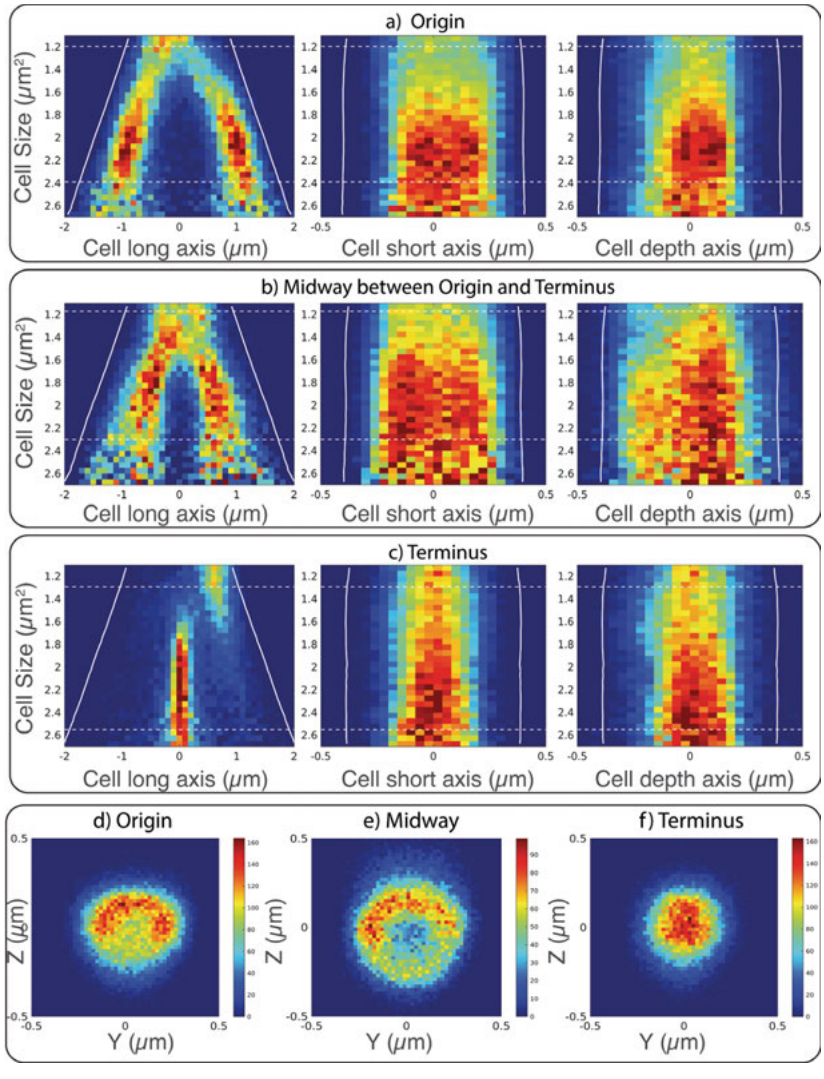



*Figure 7.1.* Localization of different loci on the *E. coli* chromosome. Labeled loci: a) Origin b) Midway between origin and terminus c) Terminus. a-c) The x-axis in each figure represents the cell long axis or x-coordinate (left), cell short axis or y-coordinate (center), and cell depth axis or z-coordinate (right). The y-axis represents cell size in $\mu m^2$. The positions of each loci are pooled from 20 mother-machine traps with growing cells imaged every 2 minutes for 600 minutes. White lines indicate the average position of the cell poles, for the cell long axis figures, or the cell widths, for the cell short and depth axes figures. d-f) Y-Z distributions of the different loci. Bin size for this plot is 20 nm x 20 nm. Figure reproduced from Paper III.

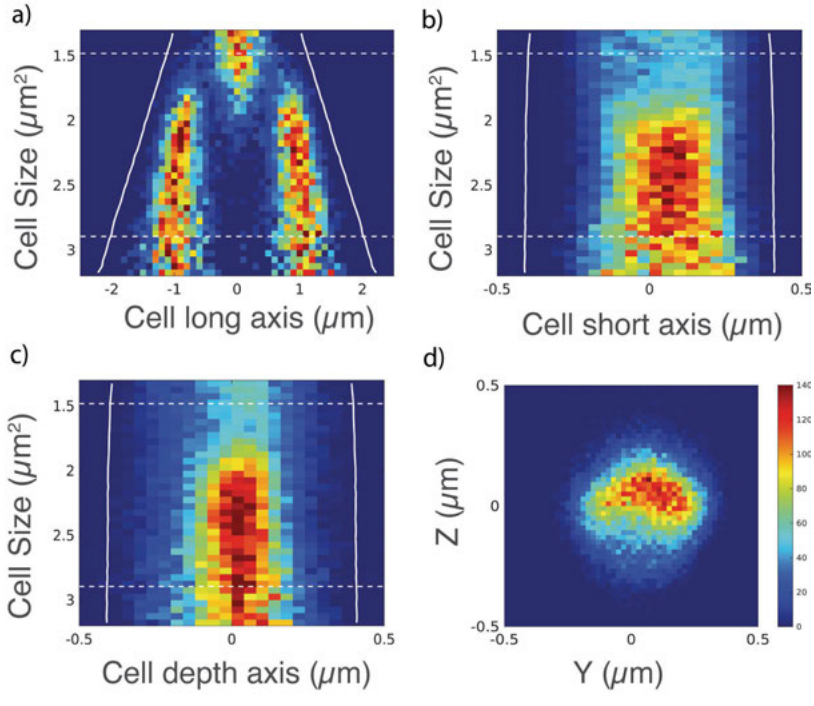



*Figure 7.2.* Histograms of emitter locations as a function of cell size for replisome-labeled cells. a) Cell long axis (x-coordinate) vs Cell size in $\mu m^2$ b) Cell short axis (y-coordinate) vs Cell size in $\mu m^2$ c) Cell depth axis (z-coordinate) vs Cell size in $\mu m^2$ d) YZ histogram of all positions. Bin size for this plot is 20 nm x 20 nm. The color bar represents number of localizations per bin in this plot. Figure reproduced from Paper III.

# 8. Popular Summary in English

Mapping observable traits of micro-organisms such as growth, size, morphology, and movement to molecular mechanisms is the central subject of modern microbiology. Advances in biochemistry since the middle of the 20$^{th}$ century have expanded our knowledge about the molecules that are fundamental to all living systems. With the help of modern microscopes, we can observe cells growing at single-cell resolution, and combined with fluorescence labeling, we can follow the molecules inside them. Under special circumstances, it is possible to physically move cells using a focused laser beam, without significant physiological effects to the cell being moved. In this work, we use microscopy and imaging-based approaches to increase the scale and scope of the methods used to study fundamental processes in bacteria.

A typical bacterial cell such as the model organism *E. coli* grows and divides every 20-60 minutes. Imaging traits of bacterial cells and their progeny requires engineering approaches to constrain them in space over long time periods while they grow and divide over several generations. Fluidic devices that contain small trenches to trap monolayers of bacteria have been developed using semi-conductor fabrication techniques and polymeric materials such as Polydimethylsiloxane (PDMS). In these devices, one can supply fresh growth media or any reagents to the cells to study how they react to changing environmental conditions, induce protein production, or simply perform chemical reactions of various kinds. Typically, one such device has the capacity to study a few thousand single-cell lineages at a time, with no capability to screen for cells with interesting properties and perform further experiments on them. During the work of this thesis, we demonstrated a new fluidic device with an order of magnitude more capacity in terms of lineages studied. At any given time during the experiment, it is possible to extract a single viable cell from any of these lineages.

High-throughput quantification of attributes of objects of interest, in our case bacterial cells and molecules inside them, requires computer algorithms. Over the past decade, developments in the area of deep learning have revolutionized image processing. Image processing results for tasks such as cell segmentation, tracking, and single-molecule localization tasks have radically improved with these methods. Using deep learning methods together with classical approaches to image processing, we developed pipelines with capabilities to segment cells of any morphology, localize single-dot fluorescence emitters, perform single-cell lineage construction, and classify fluorescence signals using multiple images.

We used a combination of fluidic devices, microscopy techniques, and AI software for three distinct projects. First, we developed a diagnostic assay for rapid antibiotic susceptibility testing and species identification of multi-species samples in $\sim$ 2 hrs. Second, we developed a large-scale fluidic device with 100,000 cell traps to study individual cell lineages in real time. We also demonstrated single-cell extraction capability in this device. Third, we developed a method to study 3D localization of fluorescent emitters in bacterial cells, in the context of studying distributions of different loci of the *E. coli* chromosome and the replisome complex.

To conclude, we have developed a powerful set of tools to aid high-throughput single-cell bacteriology.

# 9. Sammanfattning på Svenska

Inom den moderna mikrobiologin är det viktigt att kunna koppla mikroorganismernas observerbara egenskaper, exempelvis tillväxt, storlek, och morfologi, till de bakomliggande molekylära mekanismerna. Sedan mitten av 1900-talet har vår kunskap om livets byggstenar ökat markant tack vare tekniska framsteg inom biokemin och med hjälp av moderna mikroskop kan man observera tillväxten hos enstaka celler. I kombination med fluorescensmikroskopi är det möjligt att observera enstaka molekyler inuti cellerna. Under speciella omständigheter kan man också fysiskt flytta celler med hjälp av en fokuserad laserstråle, utan att skada cellen som flyttas. I detta arbete använder vi mikrofluidik, mikroskopi och olika redskap för bildanalys för att göra experimenten ovan i stor skala.

En typisk bakteriecell, exempelvis modellorganismen *E. coli*, växer och delar sig var 20:e-60:e minut. Dynamiska studier av bakteriecellerna och deras avkomma kräver tekniska metoder för att begränsa dem i rummet under långa tidsperioder medan de växer och delar sig. Mikrofluidiska odlingskammare med små kanaler där enstaka bakterier kan fångas upp och växa under flera generationer kan tillverkas med hjälp av halvledartekniker och polymera material som polydimetylsiloxan (PDMS). Genom att tillföra olika tillväxtmedier eller reagens till cellerna som är fångade i dessa vätskekanaler kan man studera hur de reagerar på förändrade miljöförhållanden, inducera proteinproduktion eller utföra olika kemiska reaktioner. Vanligtvis har en sådan mikrofluidisk odlingsenhet kapacitet att studera några tusen bakteriekolonier åt gången, och det är inte möjligt att isolera intressanta stammar vid experimentets slut för identifiering eller vidare studier. I denna avhandling demonstrerar vi en ny fluidisk odlingsmetod med en storleksordning större kapacitet i termer av hur många stammar vi kan studera i samma experiment. Från dessa kammare är det också möjligt att extrahera livskraftiga enstaka celler från kulturen när som helst under experimentet gång.

Att storskaligt kvantifiera attribut hos föremål av intresse, i vårt fall bakterieceller och molekyler inuti dem, kräver oftast användning av datoralgoritmer. Under det senaste decenniet har utvecklingen inom lärande system revolutionerat vad vi kan åstadkomma med hjälp av bildbehandling. Till exempel har metoderna för cellsegmentering, spårning och lokalisering av fluorescenta foci förbättrats radikalt. Genom att använda metoder för djupinlärning tillsammans med klassiska tillvägagångssätt för bildbehandling har vi utvecklat pipelines med kapacitet att segmentera celler med godtycklig morfologi, lokalisera fluoroforer, konstruera stamträd för olika cellinjer och klassificera fluorescenssignaler.

Kombinationen av fluidiska odlingsenheter, mikroskopitekniker och AI-mjukvara har vi använt för tre olika applikationer. Först utvecklade vi en analysmetod för artidentifiering och snabb (<2 h) diagnostisering av antibiotikaresistens för prover med flera olika patogena bakterier. För det andra utvecklade vi en storskalig odlingskammare med 100 000 cellfällor för studier av individuella cellinjer i realtid. Från denna enhet är det också möjligt att extrahera enstaka levande celler. För det tredje utvecklade vi en metod för 3D-lokalisering av fluororescenta foci i bakterieceller, för att kunna studera distributionen av replisomkomplexet samt olika kromosomala loci i växande *E. coli*.

Sammanfattningsvis utvecklar vi kraftfulla verktyg för att underlätta storskaliga studier av grundläggande processer i enstaka levande bakterier.

# Acknowledgements

I thank Prof. Johan Elf setting major directions in all the projects.

I would like to thank Vinodh for running all the experiments in the AST project related to Paper I. I would like to thank Jimmy for taking care of everything in the lab and being my friend. Elias for all the technical conversations and other related systems-level thinking type conversations. Konrad for running the 3D localization experiments and Spartak for hooking me up with the right code at the right time and debugging results. Daniel Jones and Beer Sen for cloning strains. Prof. Maria Tenje, Atena, Örjan in Angstrom for hosting me during my time in the clean room and helping me with wafer fabrication in the first half of my PhD. Finally, Irmeli for running everything behind the scenes.

# References


[1] Ping Wang, Lydia Robert, James Pelletier, Wei Lien Dang, Francois Taddei, Andrew Wright, and Suckjoon Jun. Robust growth of escherichia coli. *Current Biology*, 20(12):1099–1103, June 2010.

[2] Özden Baltekin, Alexis Boucharin, Eva Tano, Dan I. Andersson, and Johan Elf. Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proceedings of the National Academy of Sciences*, 114(34):9170–9175, August 2017.

[3] Bo Huang, Wenqin Wang, Mark Bates, and Xiaowei Zhuang. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319(5864):810–813, February 2008.

[4] A. Ashkin and J. M. Dziedzic. Optical trapping and manipulation of viruses and bacteria. *Science*, 235(4795):1517–1520, March 1987.

[5] Scott Luro, Laurent Potvin-Trottier, Burak Okumus, and Johan Paulsson. Isolating live cells after high-throughput, long-term, time-lapse microscopy. *Nature Methods*, 17(1):93–100, November 2019.

[6] David C. Duffy, J. Cooper McDonald, Olivier J. A. Schueller, and George M. Whitesides. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). *Analytical Chemistry*, 70(23):4974–4984, October 1998.

[7] Ott Scheler, Witold Postek, and Piotr Garstecki. Recent developments of microfluidics as a tool for biotechnology and microbiology. *Current Opinion in Biotechnology*, 55:60–67, February 2019.

[8] Sara Lindström and Helene Andersson-Svahn. Miniaturization of biological assays — overview on microwell devices for single-cell analyses. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1810(3):308–316, March 2011.

[9] Soohong Kim, Joachim De Jonghe, Anthony B. Kulesa, David Feldman, Tommi Vatanen, Roby P. Bhattacharyya, Brittany Berdy, James Gomez, Jill Nolan, Slava Epstein, and Paul C. Blainey. High-throughput automated microfluidic sample preparation for accurate microbial genomics. *Nature Communications*, 8(1), January 2017.

[10] Tomasz S. Kaminski, Ott Scheler, and Piotr Garstecki. Droplet microfluidics for microbiology: techniques, applications and challenges. *Lab on a Chip*, 16(12):2168–2187, 2016.

[11] Ryan Thiermann, Michael Sandler, Gursharan Ahir, John T. Sauls, Jeremy W. Schroeder, Steven D. Brown, Guillaume Le Treut, Fangwei Si, Dongyang Li, Jue D. Wang, and Suckjoon Jun. Tools and methods for high-throughput single-cell imaging with the mother machine. *eLife*, July 2023.

[12] Da Yang, Anna D. Jennings, Evalynn Borrego, Scott T. Retterer, and Jaan Männik. Analysis of factors limiting bacterial growth in PDMS mother machine devices. *Frontiers in Microbiology*, 9, May 2018.

[13] Özden Baltekin, Alexander T. A. Johnsson, Alicia Y. W. Wong, Kajsa Nilsson, Berivan Mert, Lovisa Söderberg, Erik Wistrand-Yuen, and Volkan Özenci. Evaluation of an ultra-rapid antibiotic susceptibility testing method on positive blood cultures with *Escherichia coli*. *bioRxiv*, December 2021.
[14] Michael J Lawson, Daniel Camsund, Jimmy Larsson, Özden Baltekin, David Fange, and Johan Elf. In situ genotyping of a pooled strain library after characterizing complex phenotypes. *Molecular Systems Biology*, 13(10), October 2017.
[15] Heng-Dong Xi, Hao Zheng, Wei Guo, Alfonso M. Gañán-Calvo, Ye Ai, Chia-Wen Tsao, Jun Zhou, Weihua Li, Yanyi Huang, Nam-Trung Nguyen, and Say Hwa Tan. Active droplet sorting in microfluidics: a review. *Lab on a Chip*, 17(5):751–771, 2017.
[16] Guoyun Sun, Lisha Qu, Fidelis Azi, Yanfeng Liu, Jianghua Li, Xueqin Lv, Guocheng Du, Jian Chen, Chia-Hung Chen, and Long Liu. Recent progress in high-throughput droplet screening and sorting for bioanalysis. *Biosensors and Bioelectronics*, 225:115107, April 2023.
[17] Ming Li, Hangrui Liu, Siyuan Zhuang, and Keisuke Goda. Droplet flow cytometry for single-cell analysis. *RSC Advances*, 11(34):20944–20960, 2021.
[18] Marc A. Unger, Hou-Pu Chou, Todd Thorsen, Axel Scherer, and Stephen R. Quake. Monolithic microfabricated valves and pumps by multilayer soft lithography. *Science*, 288(5463):113–116, April 2000.
[19] Jessica Melin and Stephen R. Quake. Microfluidic large-scale integration: The evolution of design rules for biological automation. *Annual Review of Biophysics and Biomolecular Structure*, 36(1):213–231, June 2007.
[20] Keir C. Neuman, Edmund H. Chadd, Grace F. Liou, Keren Bergman, and Steven M. Block. Characterization of photodamage to escherichia coli in optical traps. *Biophysical Journal*, 77(5):2856–2863, November 1999.
[21] Martin M. Thuo, Ramses V. Martinez, Wen-Jie Lan, Xinyu Liu, Jabulani Barber, Manza B. J. Atkinson, Dineth Bandarage, Jean-Francis Bloch, and George M. Whitesides. Fabrication of low-cost paper-based microfluidic devices by embossing or cut-and-stack methods. *Chemistry of Materials*, 26(14):4230–4237, July 2014.
[22] Jose L. Sanchez Noriega, Nicholas A. Chartrand, Jonard Corpuz Valdoz, Collin G. Cribbs, Dallin A. Jacobs, Daniel Poulson, Matthew S. Viglione, Adam T. Woolley, Pam M. Van Ry, Kenneth A. Christensen, and Gregory P. Nordin. Spatially and optically tailored 3d printing for highly miniaturized and integrated microfluidics. *Nature Communications*, 12(1), September 2021.
[23] Tomas Silva Santisteban, Roland Zengerle, and Matthias Meier. Through-holes, cavities and perforations in polydimethylsiloxane (PDMS) chips. *RSC Adv.*, 4(89):48012–48016, 2014.
[24] Hervé Nicoloff, Karin Hjort, Bruce R. Levin, and Dan I. Andersson. The high prevalence of antibiotic heteroresistance in pathogenic bacteria is mainly caused by gene amplification. *Nature Microbiology*, 4(3):504–514, February 2019.
[25] Jonas Ries. SMAP: a modular super-resolution microscopy analysis platform for SMLM data. *Nature Methods*, 17(9):870–872, August 2020.
[26] Roy Wollman and Nico Stuurman. High throughput microscopy: from raw images to discoveries. *Journal of Cell Science*, 120(21):3715–3722, November

2007.
[27] Rainer Pepperkok and Jan Ellenberg. High-throughput fluorescence microscopy for systems biology. *Nature Reviews Molecular Cell Biology*, 7(9):690–696, July 2006.
[28] Andrew E. S. Barentine, Yu Lin, Edward M. Courvan, Phylicia Kidd, Miao Liu, Leonhard Balduf, Timy Phan, Felix Rivera-Molina, Michael R. Grace, Zach Marin, Mark Lessard, Juliana Rios Chen, Siyuan Wang, Karla M. Neugebauer, Joerg Bewersdorf, and David Baddeley. An integrated platform for high-throughput nanoscopy. *Nature Biotechnology*, March 2023.
[29] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015.
[30] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, January 2015.
[31] Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, and Laith Farhan. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1), March 2021.
[32] Thomas Serre. Deep learning: The good, the bad, and the ugly. *Annual Review of Vision Science*, 5(1):399–426, September 2019.
[33] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu. Object detection with deep learning: A review, 2018.
[34] Xiongwei Wu, Doyen Sahoo, and Steven C. H. Hoi. Recent advances in deep learning for object detection, 2019.
[35] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.
[36] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.
[37] Sebastian Ruder. An overview of gradient descent optimization algorithms, 2016.
[38] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science*, pages 234–241. Springer International Publishing, 2015.
[39] Owen M. O'Connor, Razan N. Alnahhas, Jean-Baptiste Lugagne, and Mary J. Dunlop. DeLTA 2.0: A deep learning pipeline for quantifying single-cell spatial and temporal dynamics. *PLOS Computational Biology*, 18(1):e1009797, January 2022.
[40] Petter Ranefall, Sajith Kecheril Sadanandan, and Carolina Wahlby. Fast adaptive local thresholding based on ellipse fit. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, April 2016.
[41] Stella Stylianidou, Connor Brennan, Silas B. Nissen, Nathan J. Kuwada, and Paul A. Wiggins. Supersegger: robust image segmentation, analysis and lineage tracking of bacterial cells. *Molecular Microbiology*, 102(4):690–700, September 2016.
[42] Thorsten Falk, Dominic Mai, Robert Bensch, Özgün Çiçek, Ahmed Abdulkadir, Yassine Marrakchi, Anton Böhm, Jan Deubner, Zoe Jäckel, Katharina Seiwald, Alexander Dovzhenko, Olaf Tietz, Cristina Dal Bosco, Sean Walsh, Deniz

Saltukoglu, Tuan Leng Tay, Marco Prinz, Klaus Palme, Matias Simons, Ilka Diester, Thomas Brox, and Olaf Ronneberger. U-net: deep learning for cell counting, detection, and morphometry. *Nature Methods*, 16(1):67–70, December 2018.

[43] Özgün undefinediçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation, 2016.

[44] Uwe Schmidt, Martin Weigert, Coleman Broaddus, and Gene Myers. Cell detection with star-convex polygons. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, pages 265–273. Springer International Publishing, 2018.

[45] Carsen Stringer, Tim Wang, Michalis Michaelos, and Marius Pachitariu. Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*, 18(1):100–106, December 2020.

[46] Kevin J. Cutler, Carsen Stringer, Teresa W. Lo, Luca Rappez, Nicholas Stroustrup, S. Brook Peterson, Paul A. Wiggins, and Joseph D. Mougous. Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation. *Nature Methods*, 19(11):1438–1448, October 2022.

[47] Yiqiu Wang, Yan Gu, and Julian Shun. Theoretically-efficient and practical parallel DBSCAN. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. ACM, May 2020.

[48] Klas E. G. Magnusson, Joakim Jalden, Penney M. Gilbert, and Helen M. Blau. Global linking of cell tracks using the viterbi algorithm. *IEEE Transactions on Medical Imaging*, 34(4):911–929, April 2015.

[49] Karl Rohr, William J. Godinez, Nathalie Harder, Stefan Wörz, Julian Mattes, Wolfgang Tvaruskó, and Roland Eils. Tracking and quantitative analysis of dynamic movements of cells and particles. *Cold Spring Harbor Protocols*, 2010(6):pdb.top80, June 2010.

[50] Martin Maška, Vladimír Ulman, Pablo Delgado-Rodriguez, Estibaliz Gómez de Mariscal, Tereza Nečasová, Fidel A. Guerrero Peña, Tsang Ing Ren, Elliot M. Meyerowitz, Tim Scherr, Katharina Löffler, Ralf Mikut, Tianqi Guo, Yin Wang, Jan P. Allebach, Rina Bao, Noor M. Al-Shakarji, Gani Rahmon, Imad Eddine Toubal, Kannappan Palaniappan, Filip Lux, Petr Matula, Ko Sugawara, Klas E. G. Magnusson, Layton Aho, Andrew R. Cohen, Assaf Arbelle, Tal Ben-Haim, Tammy Riklin Raviv, Fabian Isensee, Paul F. Jäger, Klaus H. Maier-Hein, Yanming Zhu, Cristina Ederra, Ainhoa Urbiola, Erik Meijering, Alexandre Cunha, Arrate Muñoz-Barrutia, Michal Kozubek, and Carlos Ortiz de Solórzano. The cell tracking challenge: 10 years of objective benchmarking. *Nature Methods*, 20(7):1010–1020, May 2023.

[51] Luca Bertinetto, Jack Valmadre, João F. Henriques, Andrea Vedaldi, and Philip H. S. Torr. Fully-convolutional siamese networks for object tracking. In *Lecture Notes in Computer Science*, pages 850–865. Springer International Publishing, 2016.

[52] Xinshuo Weng, Yongxin Wang, Yunze Man, and Kris M. Kitani. Gnn3dmot: Graph neural network for 3d multi-object tracking with 2d-3d multi-feature learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[53] Christian Carsten Sachs, Karina Ruzaeva, Johannes Seiffarth, Wolfgang Wiechert, Benjamin Berkels, and Katharina Nöh. CellSium: versatile cell simulator for microcolony ground truth generation. *Bioinformatics Advances*, 2(1), January 2022.

[54] Karina Ruzaeva, Jan-Christopher Cohrs, Keitaro Kasahara, Dietrich Kohlheyer, Katharina Nöh, and Benjamin Berkels. Cell tracking for live-cell microscopy using an activity-prioritized assignment strategy. In *2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS)*. IEEE, December 2022.

[55] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement, 2018.

[56] Henry Pinkard, Nico Stuurman, Ivan E. Ivanov, Nicholas M. Anthony, Wei Ouyang, Bin Li, Bin Yang, Mark A. Tsuchida, Bryant Chhun, Grace Zhang, Ryan Mei, Michael Anderson, Douglas P. Shepherd, Ian Hunt-Isaak, Raymond L. Dunn, Wiebke Jahr, Saul Kato, Loïc A. Royer, Jay R. Thiagarajah, Kevin W. Eliceiri, Emma Lundberg, Shalin B. Mehta, and Laura Waller. Pycro-manager: open-source software for customized and reproducible microscope control. *Nature Methods*, 18(3):226–228, March 2021.

[57] Daniel Camsund, Michael J. Lawson, Jimmy Larsson, Daniel Jones, Spartak Zikrin, David Fange, and Johan Elf. Time-resolved imaging-based CRISPRi screening. *Nature Methods*, 17(1):86–92, November 2019.

[58] George Emanuel, Jeffrey R Moffitt, and Xiaowei Zhuang. High-throughput, image-based screening of pooled genetic-variant libraries. *Nature Methods*, 14(12):1159–1162, October 2017.

[59] R I Amann, L Krumholz, and D A Stahl. Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *Journal of Bacteriology*, 172(2):762–770, February 1990.

[60] Rudolf Amann and Bernhard M. Fuchs. Single-cell identification in microbial communities by improved fluorescence in situ hybridization techniques. *Nature Reviews Microbiology*, 6(5):339–348, May 2008.

[61] Edward F. DeLong, Gene S. Wickham, and Norman R. Pace. Phylogenetic stains: Ribosomal RNA-based probes for the identification of single cells. *Science*, 243(4896):1360–1363, March 1989.

[62] Hazen P. Babcock and Xiaowei Zhuang. Analyzing single molecule localization microscopy data using cubic splines. *Scientific Reports*, 7(1), April 2017.

[63] Yiming Li, Markus Mund, Philipp Hoess, Joran Deschamps, Ulf Matti, Bianca Nijmeijer, Vilma Jimenez Sabinina, Jan Ellenberg, Ingmar Schoen, and Jonas Ries. Real-time 3d single-molecule localization using experimental point spread functions. *Nature Methods*, 15(5):367–369, April 2018.

[64] Artur Speiser, Lucas-Raphael Müller, Philipp Hoess, Ulf Matti, Christopher J. Obara, Wesley R. Legant, Anna Kreshuk, Jakob H. Macke, Jonas Ries, and Srinivas C. Turaga. Deep learning enables fast and dense single-molecule localization with high accuracy. *Nature Methods*, 18(9):1082–1090, September 2021.

[65] Elias Nehme, Daniel Freedman, Racheli Gordon, Boris Ferdman, Lucien E. Weiss, Onit Alalouf, Tal Naor, Reut Orange, Tomer Michaeli, and Yoav Shechtman. DeepSTORM3d: dense 3d localization microscopy and PSF design by deep learning. *Nature Methods*, 17(7):734–740, June 2020.

[66] Shuang Fu, Wei Shi, Tingdan Luo, Yingchuan He, Lulu Zhou, Jie Yang, Zhichao Yang, Jiadong Liu, Xiaotian Liu, Zhiyong Guo, Chengyu Yang, Chao Liu, Zhen li Huang, Jonas Ries, Mingjie Zhang, Peng Xi, Dayong Jin, and Yiming Li. Field-dependent deep learning enables high-throughput whole-cell 3d super-resolution imaging. *Nature Methods*, 20(3):459–468, February 2023.

[67] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, September 2006.

[68] Michael J Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–796, August 2006.

[69] Sebastian van de Linde, Anna Löschberger, Teresa Klein, Meike Heidbreder, Steve Wolter, Mike Heilemann, and Markus Sauer. Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nature Protocols*, 6(7):991–1009, June 2011.

[70] Leonhard Möckl, Anish R. Roy, Petar N. Petrov, and W. E. Moerner. Accurate and rapid background estimation in single-molecule localization microscopy using the deep neural network BGnet. *Proceedings of the National Academy of Sciences*, 117(1):60–67, December 2019.

[71] Alex van Belkum, Carey-Ann D. Burnham, John W. A. Rossen, Frederic Mallard, Olivier Rochas, and William Michael Dunne. Innovative and rapid antimicrobial susceptibility testing systems. *Nature Reviews Microbiology*, 18(5):299–311, February 2020.

[72] C. L. Ventola. The antibiotic resistance crisis: part 1: causes and threats. *P T*, 40(4):277–283, Apr 2015.

[73] A. M. Caliendo, D. N. Gilbert, C. C. Ginocchio, K. E. Hanson, L. May, T. C. Quinn, F. C. Tenover, D. Alland, A. J. Blaschke, R. A. Bonomo, K. C. Carroll, M. J. Ferraro, L. R. Hirschhorn, W. P. Joseph, T. Karchmer, A. T. MacIntyre, L. B. Reller, and A. F. Jackson and. Better tests, better care: Improved diagnostics for infectious diseases. *Clinical Infectious Diseases*, 57(suppl 3):S139–S170, November 2013.

[74] Anand Kumar, Paul Ellis, Yaseen Arabi, Dan Roberts, Bruce Light, Joseph E. Parrillo, Peter Dodek, Gordon Wood, Aseem Kumar, David Simon, Cheryl Peters, Muhammad Ahsan, and Dan Chateau. Initiation of inappropriate antimicrobial therapy results in a fivefold reduction of survival in human septic shock. *Chest*, 136(5):1237–1248, November 2009.

[75] Jenna Rychert. Benefits and limitations of MALDI-TOF mass spectrometry for the identification of microorganisms. *Journal of Infectiology*, 2(4):1–5, July 2019.

[76] Laura Ferreira, Fernando Sánchez-Juanes, Magdalena González-Ávila, David Cembrero-Fuciños, Ana Herrero-Hernández, José Manuel González-Buitrago, and Juan Luis Muñoz-Bellido. Direct identification of urinary tract pathogens from urine samples by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of Clinical Microbiology*, 48(6):2110–2115, June 2010.

[77] A. de Roux. Mixed community-acquired pneumonia in hospitalised patients. *European Respiratory Journal*, 27(4):795–800, February 2006.

Acta Universitatis Upsaliensis

*Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 2324

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-514317

ACTA UNIVERSITATIS
UPSALIENSIS
2023