Data-driven microscopy: placing high-fidelity data in a population-wide context

André, Oscar

2023

Document Version:
Publisher's PDF, also known as Version of record

Data-driven microscopy

placing high-fidelity data in a population-wide context

OSCAR ANDRÉ
DEPARTMENT OF CLINICAL SCIENCES, LUND | FACULTY OF MEDICINE | LUND UNIVERSITY

Data-driven microscopy

Data-driven microscopy

placing high-fidelity data in a population-wide context

by Oscar André

LUND
UNIVERSITY

Doctoral Dissertation

With due permission from the Faculty of Medicine at Lund University,
this doctoral thesis will be defended on Thursday, the 8th of June 2023 at 09:00,
in Belfragesalen, Biomedical Centre (BMC D15), Lund, Sweden

Faculty opponent

Associate Professor Guillaume Jacquemet

Åbo Akademi University

Turku, Finland

Organization
Division of Infection Medicine
Department of Clinical Sciences, Lund
Faculty of Medicine
Lund University
Lund, Sweden

Document name
DOCTORAL DISSERTATION

Date of disputation
2023-06-08

Sponsoring organization

Author(s)
Oscar André

Title and subtitle
Data-driven microscopy: placing high-fidelity data in a population-wide context

Abstract
Microscopy is a cornerstone in our ability to conduct scientific research. By allowing us to visualize the intricate nature at the microcosm, it has transformed the way we approach and think about biological research. As microscopy has advanced as a technique, it has seen significant advancements in both throughput and fidelity, allowing us to extract powerful spatio-temporal information at the single cell level. However, microscopy still faces challenges in achieving high-quality, population-wide data while maintaining high fidelity. This issue is further compounded by the reliance on human operators when acquiring data in the high-fidelity domain, which may risk the introduction of bias during acquisition. Moreover, microscopy encounters limitations in sample reusability, as the utilization of multiple microscopy techniques on the same cellular entities can be cumbersome and challenging. Addressing these limitations is crucial for a more comprehensive understanding of cellular heterogeneity and the diverse plethora of phenotypes present within cell populations.

My thesis has largely focused on solving these issues. The methods presented herein are designed to place high-fidelity data in its population-wide contex, and let the data, instead of the operator, be a dictating factor in what we acquire in the high-fidelity domain. Furthermore, we provide a novel mean for performing correlative imaging, where spatial patterns in the data distribution allow for an easy calibration of a previously acquired dataset on new microscopy systems. This mindset of prioritizing data extracted from the acquired images above visual interpretation was further contextualized in two collaborations, where I contributed to the characterization of isolated antibodies against SARS-CoV-2, as well as help explore the capacity of *Streptococcus pyogenes* to bind fibronectin in the presence of antibodies, potentially revealing a niche-specific interaction.

Key words
data-driven microscopy; event-driven microscopy; high-content screening; automation; image analysis; data analysis; cell heterogeneity; host-pathogen interactions; *Streptococcus pyogenes*; SARS-CoV-2

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title
1652-8220

ISBN
978-91-8021-418-6

Recipient's notes

Number of pages
165

Price

Security classification

Data-driven microscopy

placing high-fidelity data in a population-wide context

by Oscar André

LUND
UNIVERSITY

Cover illustration front: The image depicts a microscope surrounded by data (representing data-driven microscopy). Population-wide data (blue) is flowing through the microscope, dictating where it images next in high fidelty, resulting in an outstream of high fidelity datapoints (red). Because the high fidelity data is driven by the population-wide data, the high fidelity data can be placed in the same context as the population, thus populating the same cloud. (Credits: Stray Studios)

Cover illustration back: The illustration was extended with the use of DALL-E 2, a deep learning model developed by OpenAI. The resulting back cover was further retouched by my incredible skills in graphical design softwares.

Faculty of Medicine, Department of Clinical Sciences, Lund

ISBN: 978-91-8021-418-6
ISSN: 1652-8220

Printed in Sweden by Media-Tryck, Lund University, Lund 2023

Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

Printed matter 3041 0903

MADE IN SWEDEN

Dedicated to my family

Preface

As I approach the end of this journey, I find myself wanting to put some of the thoughts I've had writing this thesis into words.

About six years ago, during the last year of my Master's degree in Biomedicine, I participated in a course in microscopy at the Biology Department. I did not select this course from the point of view of 'This is what I want to do', but rather that it was the most interesting course out of the options I was presented. At the time, I was rather tired of what I was doing, I knew I had a profound fascination about biology and the molecular processes that makes up who we are, but the fatigue of studying for so long had finally caught up with me. Little did I know what I signed up for when I took this course, and where it would take me.

It was during this course that me and my supervisor, Pontus, met. He had just returned from his post-doc and was having a presentation on the work he had been conducting. How he and his co-authors had setup a system to measure the force that is generated when a cell binds to a surface, and all the clever engineering and problem-solving that took place in order to get it working. I was so fascinated by this, that I decided to ask him about running my final project for my degree in his lab.

I think it was during this time that he saw what a massive nerd I was. I had always been the kid who sat in front of the computer, 'wasting my hours away'. However, despite my interests, I never learned how to code. As with a lot of people, I thought you needed to proficient in math in order to code, something that was not my strength (also a motivating factor for studying biology). Thankfully, Pontus proved me otherwise, and we embarked on a journey together, with the aim of being able to capture and study bacterial infections using the microscope in an automated way.

Fast forward till today, six years later, with me sitting here wrapping up my thesis. Setting out I wanted my thesis to reflect this journey, and explore all the subjects that relate to my work. However, I cannot shake the feeling off of not having truly explored everything at depth, to give the subjects the respect they deserve. Nevertheless, I do hope it gives a glimpse of what these last six years have been like. Cheers!

Table of Contents

List of publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **Data-driven microscopy allows for context-specific acquisition of high-fidelity image data**
Oscar André, Johannes Kumra Ahnlide, Nils Norlin, Vinay Swaminathan, and Pontus Nordenfelt
Cell Reports Methods
Issue 3, volume 3

II **Correlative imaging using inherent spatial-geometric relationships of cell nuclei**
Oscar André, Johannes Kumra Ahnlide, Swathi Packirisamy, Nils Norlin, Vinay Swaminathan and Pontus Nordenfelt
Manuscript in preparation

III **Spike-Dependent Opsonization Indicates Both Dose-Dependent Inhibition of Phagocytosis and That Non-Neutralizing Antibodies Can Confer Protection to SARS-CoV-2**
Wael Bahnan, Sebastian Wrighton, Martin Sundwall, Aanna Bläckberg, Urban Höglund, Olivia Larsson, Hamed Khakzad, Magdalerna Godzwon, Maria Walle, Elisabeth Elder, Lotta Happonen, **Oscar André**, Johannes Kumra Ahnlide, Thomas Hellmark, Vidar Wendel-Hansen, Robert PA. Wallin, Johan Malmström, Lars Malmström, Mats Ohlin, Magnus Rasmussen, Pontus Nordenfelt
Frontiers in Immunology
Volume 12

IV **Group A streptococci induce stronger M protein-fibronectin interaction when specific human antibodies are bound**
Sebastian Wrighton, Vibha Kumra Ahnlide, **Oscar André**, Wael Bahnan, Pontus Nordenfelt
Frontiers in Microbiology
Volume 14

Publications not included in this thesis:

A human monoclonal antibody bivalently binding two different epitopes in streptococcal M protein mediates immune function
Wael Bahnan, Lotta Happonen, Hamed Khakzad, Vibha Kumra Ahnlide, Therese de Neergaard, Sebastian Wrighton, **Oscar André**, Eleni Bratanis, Di Tang, Thomas Hellmark, Lars Björck, Oonagh Shannon, Lars Malmström, Johan Malmström, Pontus Nordenfelt
EMBO Molecular Medicine
Issue 2, volume 15

Developmental cues license megakaryocyte priming in murine hematopoietic stem cells
Trine A Kristiansen, Qinyu Zhang, Stefano Vergani, Elena Boldrin, Niklas Krausse, **Oscar André**, Pontus Nordenfelt, Mikael Sigvardsson, David Bryder, Jonas Ungerbäck, Joan Yuan
Blood Advances
Issue 24 volume 6

Role of serotype and virulence determinants in internalization and persistence of *Streptococcus pyogenes* biofilm bacteria in epithelial cells in vitro
Feiruz Alamiri, **Oscar André**, Supradipta De, Pontus Nordenfelt, Anders P Håkansson
Frontiers in Cellular and Infection Microbiology
Accepted for publication

Acknowledgements

To my supervisor, Pontus Nordenfelt
I find myself at a loss of words trying to express the impact your guidance and mentorship has had on me. You're ability to listen and understand is unparalleled, and I am eternally grateful for you showing and teaching me so much of what I know, and for allowing me to truly explore my interests. And for having that magical ability to put me back on earth and ease my nerves. I am where I am today, largely thanks to you!

To my co-supervisors:
Anders P Håkansson: I am truly grateful having you as a co-supervisor. Your story has been an inspiration and your caring spirit has been a much appreciated presence. Thanks for taking the time to be my co-supervisor!

Anders Wittrup: I know we have not had much contact except for in the start of my journey. Regardless, I am thankful it being you there supporting me! Keep doing what you do, you are an inspiration to many of us.

To my close collaborators:
Vinay Swaminathan, it's hard to know where to start with you. For starters, I am truly grateful for everything you've taught me. But at the same time, the amount of times you've complained about my music... I'd like to see you make a better 80s list than my masterpiece!

Nils Norlin, for being an amazing colleague. You're enthusiasm and passion for your craft is an inspiration, and I look forward to future collaborations and discussions!

A thanks to the members of NordenfeltLab, past and present, in no particular order:

Martin, for being a great friend, colleague, and role model! After 5 years I can confidently say that I think everyone would be better of with a little bit of Martin in their lifes! Don't hesitate if you need help with a bandaid or have questions about the network (which you know I wont be able to answer), im your guy 8-).

Sebastian, you are truly amazing, and I am so grateful that we stumbled upon each other. My PhD studies have been significantly ($p \leq 0.0001$) more enjoyable having you around!

Therese, crazy to think that we started off together so many years ago, and to look where we are now! Your absence is deeply felt, as your presence has always been a source of joy and positivity.

Vibha, for transforming the office into a more delightful and fun place! You are an amazing friend, and I am eager to having you around more often!

Arman, for all our great discussions, after works and qualities as a friend! And for being such an easy target in Super Smash, its been a great joy having you as a victim when I've needed to let off some steam ;)

Wael, for being an amazing, modern and civilized cave man! You are an incredible teacher, but most importantly, friend, and I am truly grateful for our paths crossing!

Arsema, for being such a great student and, for the most part, colleague. Even as my student, you've taught me much, often making our roles seem reversed! Thank you!!

Berit, the mother (mutti) of the lab! You're cheerful laughter makes every fika so much better, and I am amazed at you're work ethic and expertise in the lab!

and a special thanks to **Johannes Kumra Ahnlide**, my brother in arms, for your unrivaled patience. I cant believe you've survived this long listening to my music. It is truly a feat of strength! In all seriousness, all that pales in comparison to your qualities as a friend, and I am truly happy its been you sitting there besides me.

A special thanks to two very important people:

Mattias, I know we have not had the pleasure of working together directly, but every time we greet and interact I am thrown back to my first semester at the Biomedicine program. The story you told us about your journey lives with me till this very day, and you have been a big inspiration during all these years! Thanks for being an amazing role model!

Anita, I don't understand how other people finish their PhD without having you in their lives. Thanks for all the help throughout the years, you are truly amazing!

To my friends outside of work:

Tobias, I've known you for a third of my life now! What an amazing time its been, so many memories and stories to tell that we've been on together! And **Olivia**, a much needed addition to the team! I knew Tobias would be in safe hands (see what I did there) under your wings ;).

Matthias, Patrik, Rasmus, Harald, Alle, Macke, world best alliance-only guild, lets go!!! Thanks for letting me escape work and enjoy listening to you guys crying for 3 hours 2 times a week! In all seriousness, I cherish every minute of our little online adventures!

Martin, lost but not forgotten, or what should I say? I miss your presence dearly!!

To my family...

Mamma and Pappa, I owe it all to you. Even though you might not have fully known what I've done the last 5 years, I know you're proud of me. And dad, I wish you were here. This is the first big milestone in my life that you're not around to see, and it pains me beyond words. I miss you and I hope I'm making you proud!

Lina and Sofie, you are my role models in more ways than you believe. The love and affection you shown your families are awe-inspiring. It will be hard for me to be as good of a mother as the two of you are, but I sure will try!

Billy och Jonte, you have been a part of my life for longer than I remember, and a big part at that.

Mormor, din förmåga att alltid vara så lätt till skratt och din obegränsade kärlek är svårt att beskriva. Bara tanken av dig får mig att le! Alla har något att lära av dig, för du är ovärderlig mormor!

Och självklart **Theo, Tim, Elliot och Hedvid**. Ni är småsyskonen jag aldrig fick själv. Jag blir galen bara av att tänka på era hyss. Ta världen med storm <3.

And of course, a thanks to my beloved **Swathi**, for being my being my sunshine and warmth. Thinking about my future with you puts a smile in my heart.

Populärvetenskaplig sammanfattning på svenska

Mikroskopi är idag ett fundamentalt verktyg inom forskning, där det tillåter oss att skåda in och utforska våra prover i hög detalj. Mycket utav utvecklingen av nya mikroskopimetoder har strävat efter att öka den detaljnivå vi kan uppnå. Samtidigt har utvecklingen inom hårdvara, med tillgång till bättre och mer kraftfulla instrument, lett till utveckligen av metoder där fokuset är att studera en hel population av celler. Till skillnad från när vi studerar ett fåtal celler i hög detalj, tillåter det oss att sätta perspektiv på det vi ser. Det ger oss en förmåga att säga vad det normala beteendet som man kan förvänta sig är, och vilka celler som sticker ut i en population. Med andra ord, vad som är intressant.

Samtidigt finns det ett stort intresse av att veta hur varje individuell cell beter sig. Varje cell är, precis som oss människor, unik. De har olika historia, olika ålder och befinner sig i olika tillstånd. Precis som våra celler i kroppen är unika, är även de cellerna som kan orsaka sjukdom unika. För att förstå varför vissa personer är mer känsliga mot sjukdom, och hur en infektion svarar på våra behandlingar behövs en förståelse och an förmåga att studera celler på individuell nivå, samtidigt som vi bibehåller ett perspektiv utifrån populations-nivå.

Denna brist på perspektiv har länge varit ett problem inom mikroskopi. Den vanliga lösningen på detta problem är att vi, som människor, kan tolka en bild och peka på vad det är som är intressant eller inte. Vi är, trots allt, extremt duktiga på att tolka visuell information. Men detta är inte en helt felfri lösning. Som människor kan vi vara relativt okonsekventa, vi tolkar oftast utifrån hur vi vill att datan ser ut. Med andra ord, vi saknar förmågan att vara objektiva i vår metodik för att samla in bilder i hög detalj.

Min avhandling har till stor del handlat om att utveckla ett verktyg som tillåter oss att sätta perspektiv på det vi studerar med mikroskopi. Detta har lett till **Arbete 1**, där vi presenterar en allmän strategi (data-styrd mikroskopi) för hur vi kan arbeta med mikroskopi för att samla in data på en hel population, samtidigt som vi kan samla in data med hög detalj på relevanta fynd i populationen. Vi presenterar även här en teknisk lösning, och utför metoden i tre olika scenarion: ett för att studera en population av celler mer allmänt, ett för att fånga det ögonblick som bakterier infekterar mänskliga celler, och ett där vi studerar och fångar in data på relevanta (från ett populations-kontext) cancerceller och följer dem över tid. Denna metod tillåter oss att samla in data i hög detalj på ett objektivt sätt, och att sätta perspektiv på det vi studerar.

I **Arbete 2** har vi vidare utvecklat på vår metod, där vi försöker lösa problemet att hitta en och samma cell i flera olika mikroskop. Eftersom vi, genom mikroskopi, jobbar på en så ofantligt liten skala, är det oftast väldigt svårt att orientera sig och hitta rätt inom ett prov. Det är lite som att spela På spåret och gissa vart man är, fast utan alla ledtrådar man får på varje nivå. Eftersom vi har tillgång till data på en hel population, så utgick vi från att det borde finnas samband mellan celler och deras grannar i ett prov som är unika för just

dem. Genom att använda sig av dessa unika samband kom vi fram med en lösning där vi snabbt kan kalibrera ett prov på ett nytt mikroskop. Det öppnar dörrarna för oss forskare att återanvända prov, att lättare justera provet med nya markörer (för det vi vill visualisera inom cellerna), och att kunna tolka ett prov med data insamlat från flera system.

COVID-19 pandemin var en stor omställning för samhället och vården. Likväl var det en stor omställning för många forskningslabb, där en kapplöpning startade för att så snabbt som möjligt förstå sig på hur viruset fungerar och hur vårt immunförsvar svarar på dess infektion. Det var i detta kontext som mitt **tredje arbete** utfördes. Genom den erfarenhet jag samlat på mig inom mikroskopi och att analysera bilder på stora dataset, bidrog jag med hjälp för att studera hur framtagna antikroppar kan förhindra bindningen av virus-lika partiklar till celler. Antikroppar är ett protein som immunförsvaret producerar i respons mot en patogen. En bättre förståelse kring hur antikroppar verkar, och vad skillnaden mellan en bra och en dålig antikropp är kan leda till framtagningen av bättre vaccin-program och behandlingar inom sjukvården.

I **Arbete 4** medverkade jag i ett arbete där bakterien *Streptococcus pyogenes* var i fokus. *S. pyogenes* enda värd är människor, och ansvarar för över 600 miljoner infektionsfall per år globalt. På bakteriens yta dominerar ett protein, M-proteinet, ett multi-funktionellt protein som bakterien (bland annat) använder sig för att binda till ytor och förhindra immunförsvarets förmåga att göra sig av med bakterien. I arbetet upptäckte vi att fibronektin binder till bakterien (specifikt M-proteinet) olika mycket beroende på mängden antikroppar som finns i miljön. Fibronektin är ett protein som vi människor producerar, och bidrar (bland annat) till att skapa den miljön som celler befinner sig i. Mängden fibronektin varierar beroende på var i kroppen man kollar. Till exempel, i saliv har du en relativt låg mängd fibronektin jämfört med i blodet. Detta ledde till hypotesen att bakterien är special-anpassad för olika miljöer i dess förmåga att undkomma immunförsvaret. En bättre förståelse kring hur bakterien är anpassad till våra olika miljöer och dess infektionsförlopp kan leda till bättre och mer anpassade behandlingar inom sjukvården.

Abbreviations

ADC analog-to-digital converter

APD avalanche photodiode

CCD charged-coupled device

CLEM correlative light-electron microscopy

CMOS complementary metal-oxide semiconductor

CNN convolutional neural networks

CTE charge transfer efficiency

DDA data-dependent acquisition

DDM data-driven microscopy

DIA data-independent acquisition

DIC differential interference contrast

EMCCD electron-multiplying CCD

FOV field of view

FP fluorescent protein

FWC full well capacity

GAS group A streptococcus

GFP green fluorescent protein

GUI graphical user interface

HCS high content screening

HTP high throughput

LED light-emitting diode

NA numerical aperture

NF-$\kappa\beta$ nuclear factor $\kappa\beta$

PALM photoactivated localization microscopy

PH phase contrast

PMT photomultiplier tube

PSF point spread function

QE quantum efficiency

ROS reactive oxygen species

SDCM spinning disk confocal microscopy

sGFP split GFP

SIM structured illumination microscopy

SnR signal-to-noise

STED stimulated emission depletion

STORM stochastic optical reconstruction microscopy

T3SS type 3 secretion system

Yops Yersinia outer proteins

Biology from a microscopist perspective

When God said 'Let there be light', he surely must have meant perfectly coherent light.
— Charles Townes

To understand the power and limitations of using microscopy when studying biology, we must have an understanding of the nature of light, and how we can bend and shape it for our purpose. We must understand how light can be measured, interpreted and evaluated, and how we may be our biggest enemies in uncurtaining of the truth. We must also understand that, like with most observations, the very act of observing has an impact on the outcome, and how we can work to minimize this impact when using microscopy. This hold especially true when studying the very fragile fabric of life, cells. Finally, we must understand the uniqueness of life itself, how our evaluations of cells at a population level may not propagate truthfully to the single-cell level, and the spontaneous rise of cellular heterogeneity in what is seemingly a homogeneous population.

This chapter serves as an introduction to what it is like studying biology from a microscopists perspective.

The scattering of light

The use of microscopy involves the examination of tiny structures and details that are not visible to the naked eye. But what happens when light is focused on these structures? To answer this question, we must fist realize that, from the microscopes perspective, there is a perceivable infinite number of points in space that these structures can be located. In fluorescence microscopy, these points are continuously illuminating upon excitation, and

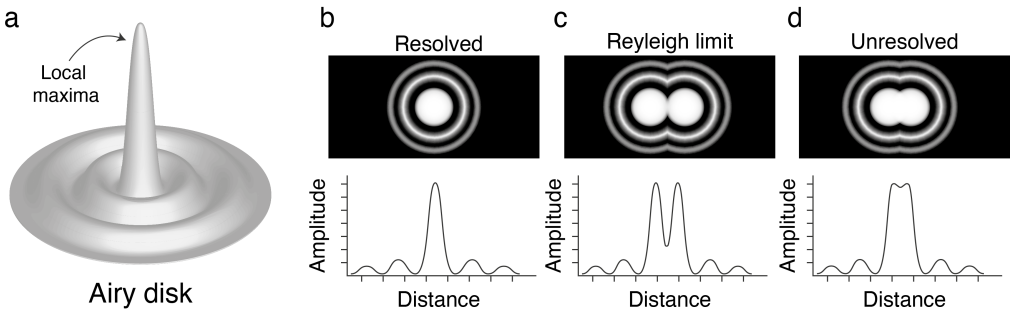

Figure 1: Resolution depends on the ability to resolve individual sources of light. **a**. Illustration of an airy disk, the theoretical illumination from a point source. **b** Points far apart can be resolved. **c** Points meeting the Rayleigh criteria can still be resolved. **d** Points closer than the Rayleigh criteria can not be resolved.

their collective measure results is the acquired image. If we imagine an isolated point, the light from this point in the microscope will diffract, creating a ripple like effect. This diffraction, or spreading, is also known as the point spread function (PSF) and can be visualized as Airy disks (after George Biddell Airy[1]; see Fig. 1a and b) in the microscope. The characteristics of the Airy disk fundamentally dictate our ability to resolve individual points.

The problem arises when there are two or more points in close proximity. Our ability to separate the two points as independent light sources depends on the distance and partial overlap between the local maxima of the two Airy disks. To determine the resolution of these point sources, a practical criterion is used which takes into account the distance at which the first Airy disk's local maxima coincides with the secondary disk's local maxima by:

$$r = \frac{1.22\lambda}{2NA_{obj}} \tag{1}$$

This criteria is known as the Rayleigh's criterion, where the numerical aperture (NA) of the objective and the wavelength λ of the light. The NA can be formulated as:

$$NA = n \sin\theta, \tag{2}$$

where n is the refraction index of the media and θ is the collection angle.

In bright-field images, that is, when there is a continuous illumination of the sample from all wavelengths of the visible spectra, the Rayleigh's criteria is given by:

$$r = \frac{1.22\lambda}{NA_{cond} + NA_{obj}} \tag{3}$$

In fluorescence microscopy, because we both excite and collect the light of the specimen using the objective, the objective acts as the condenser $(NA_{cond})^2$. Being one and the same component leads to the denominator in the equation 3 effectively becoming $2NA$.

We now know that our ability to resolve points as separate light sources (in other words: a high resolution) depends on our ability to achieve the smallest distance possible Rayleigh's limit r. But how do we go about optimizing for resolution? The most practical method for doing so is by creating an as large denominator as possible through the use of objectives with high NA, or achieving a small numerator by limiting the use of light to that of the shorter wavelengths. However practical it may be, the use of high NA comes with its own benefits and drawbacks that has to be taken into consideration when using microscopy.

When examining equation 2, we can see that a limiting factor for achieving a high NA is the refractive index n. Therefore, we must increase the refractive index between the objective and the specimen so that a higher NA can be obtained. To do this, mediums with a higher n than air (which has a n of 1), has to be used such as immersion oil (n differs depending on the oil) or water ($n = 1.33$). Furthermore, the refractive index of the sample has to be taken into consideration, and it is favorable to match the refractive index between the immersion media, sample and the solution in which the specimen is in. Failing to do so risk introducing spherical abbreviations and artifacts, decreasing resulting image quality.

The creation of an image

As light travels through the microscope and interacts with the sample, its intricate details are revealed. However, without the ability to capture and record this information, these details may go unnoticed or be lost entirely. Therefore, it is essential to have reliable methods for capturing and storing the images produced by the microscope. In the past, the eye acted as the primary detector, and the information was captured through hand-drawn depictions of the underlying observations. However, with the advent of modern technology, we now have the ability to digitally store this information using specialized detectors.

The process of capturing an image using an electronic camera can be viewed through three different stages; interaction with the photon, storing the interaction as an electrical charge, and the analog readout of the charges. This process takes place on the chip of the camera, and has to be repeated millions of times when capturing a single image. The reason for this is because each chip is composed of a large matrix of a photosensitive element, pixels (short-hand for 'picture element'), where each pixel corresponds to a single readout. Since each pixel corresponds to a single readout, its physical dimensions and composition determines a number of factors such as resolution and sensitivity.

Detectors for digital microscopy

The use of cameras for life-science application has been extensively reviewed before (see[3–5]). Cameras in the life-science industry can largely be divided into two categories: charged-coupled device (CCD) and complementary metal-oxide semiconductor (CMOS) cameras. These types of cameras work by converting incoming photons into an electrical charge through an array of lightsensitive photodiodes[6]. Each pixel in the array functions as a well, accumulating a charge proportional to the number of photons that interact with its photodiodes during the camera's exposure time. Each well has a upper limit in its capacity to maintain a charge, referred to as the full well capacity (FWC). Photodiodes reaching their FWC during their exposure time cease to increase in their accumulated charge, resulting in a saturated pixel and loss of information[7,8].

Comparison between CCD and CMOS

The major differences between these technologies lie in the method for reading the charge held in each pixel. For CCD cameras, the accumulated charges are shifted along transfer channels to a readout amplifier using a series of voltage steps. In other words, the read-out is performed one row at a time, and a single read-amplifier is used for the whole chip[7,9]. In CMOS cameras, each pixel is readout individually through its own amplifier, and each column of pixels has its own analog-to-digital converter (ADC)[8]. This means that, in a CMOS camera, each pixel's charge can be converted into a digital signal independently, allowing for a faster readout time, lower power consumption and a lower generation of heat.

Low light situations

In low-light situations, specialized cameras, such as scientific CMOS (sCMOS) and electron-multiplying CCD (EMCCD) cameras, are designed to provide improved performance.

sCMOS cameras are a type of CMOS camera that combines the advantages of both CCD and CMOS technologies. They offer high sensitivity, low readout noise, high frame rates, and a wide dynamic range, making them well-suited for low-light imaging applications[3]. These cameras have been used in a variety of applications, including single-molecule imaging, super-resolution microscopy, and live-cell imaging.

EMCCD cameras, on the other hand, are a specialized type of CCD camera specifically designed for extremely low-light imaging applications. EMCCD cameras utilize electron multiplication technology to amplify the signal before digitization, effectively reducing readout noise and improving the signal-to-noise ratio in low-light conditions[5]. The elec-

tron multiplication process occurs in a dedicated gain register, allowing for flexible control of the amount of amplification by adjusting the voltage applied to the register.

Detectors for confocal microscopy

In confocal microscopy, detectors play a crucial role in collecting emitted fluorescence signals from the sample. Photomultiplier tubes (PMTs) have been the traditional choice for confocal microscopy due to their high sensitivity and ability to detect very low light levels[10]. PMTs work by converting incident photons into electrons through the photoelectric effect. These electrons are then multiplied through a series of dynodes, generating a detectable current proportional to the number of incident photons. The high amplification provided by PMTs allows for the detection of weak signals in low-light conditions[11]. Alternatively, avalanche photodiodes (APDs) offer a more compact and efficient solution for detecting low-light signals, as they have lower noise levels and faster response times than PMTs. A key difference in their working principles is that APDs[12] utilize an avalanche process, where incident photons generate primary electrons which are accelerated by a strong electric field, leading to the ionization of more atoms and a cascade multiplication of electrons. When compared to cameras such as CCDs and CMOS, PMT and APD distinguishes themselves in that the incoming data is handled as a single stream that has to be recombined into an image post acquisition.

Detector parameters

When choosing a detector for digital microscopy, several key parameters should be taken into consideration to ensure accurate and high-quality imaging. These parameters include sensitivity to noise, readout speed, pixel size, sensor size, quantum efficiency, dynamic range, and bit depth.

Noise

When performing microscopy, we need the ability to distinguish between actual signal (detected photons originating from our sample) and signal stemming from the equipment itself. The ratio between true and false signal (from the perspective of trying to observe the specimen) is referred to as the signal-to-noise (SnR)[5,13,14]. Each step in the process of acquiring an image, from interacting with a photon to the subsequent conversion to an analog signal, is subject to noise, and can be categorized into four categories: Poisson noise (or photon shot noise), dark noise, readout noise and fixed-pattern noise[8].

Poisson noise

Poisson noise arises from the fact that the interaction we are trying to measure - the number of photons - is a stochastic process and follows Poisson counting statistics[14,15]. This means that there is a constant fluctuation in the number of photons originating from a single light source that interact with our photodiode. This noise cannot be eliminated but can be reduced by increasing the number of emitted photons from the light source (by, for instance, increasing the intensity of the light source) or increasing the exposure time (which allows for a higher statistical likelihood of 'true' signal due to a larger sample size).

However, it is important to note that increasing the intensity of the light source or exposure time also increases the likelihood of other sources of noise, such as dark noise or readout noise as we will see next.

Dark noise

Dark noise (a.k.a dark current or thermal noise) is the fluctuations in the amount of electronic charge accumulated in each pixel in the absence of photons. This buildup is a consequence of thermal excitations that liberate electrons, and is directly correlated with the temperature of the chip.

We can drastically decrease the impact of dark noise by cooling the camera during acquisition. Each 20°C decrease in temperature reduced the accumulated dark noise levels by an order of magnitude, and cooling the camera to 0°C can reduce the dark noise to negligible quantities for most microscopy applications. Additionally, in the case for CCDs, cooling the chip of the camera had additional benefits. Each time a charge is transferred along its channel, there is a possibility of leaving some charge behind. This results in a blurring and dimming of the regions farthest from the amplifier. Cooling the chip improves its charge transfer efficiency (CTE)), reducing some of these artifacts[7].

Readout noise

Readout noise stems from an imprecision in the measurement of each charge packet by the read amplifier. Read noise typically increases as a function of readout speed, where higher readout speeds decreases our ability to make a precise measurement. While read noise may vary with readout rate, it is independent of exposure time and the number of photons collected. Thus, as with poisson noise, the level of readout noise can be minimizes by collecting more photons.

Fixed-pattern noise

When a photon hits a photodiode, it generates an electrical charge that needs to be accumulated over the camera's exposure time. However, not all photodiodes have the same response to the same amount of light, resulting in a fixed-pattern noise. This noise can be caused by variations in the charge collection efficiency between photodiodes, as well as other factors like dust particles or uneven illumination. As a consequence, some pixels may be more sensitive than others, producing consistent differences in brightness across multiple exposures. In sCMOS cameras, variations in amplifier gain between pixel-to-pixel and column-to-column can also contribute to fixed-pattern noise[8]. However, improvements in the design of sCMOS have helped to reduce this effect[16,17].

Fortunately, flat-field correction methods can be applied after image acquisition to reduce fixed-pattern noise. These methods involve using a calibration image taken with uniform illumination to estimate and correct the variations in sensitivity across the pixels. Although fixed-pattern noise is more noticeable in CMOS cameras than in CCD cameras due to differences in their design, it can be effectively removed with proper correction techniques[1].

Pixel size

The size of each pixel on the chip of the camera plays a critical role in our ability to gather and transform photons into visual information. The smaller the pixel, the finer we can distinguish details that is present in the sample. However, this only works to a certain limit, considering our ability to distinguish individual objects between each other depend on their proximity and their PSF.

In order to achieve optimal spatial resolution, it is essential to carefully match the pixel size to the optical resolution of the microscope system. Here, a general rule of thumb is the Nyquist theorem[18], which states that the sampling rate of the imaging system must be at least twice the maximum spatial frequency present in the specimen to accurately represent the information. Smaller pixels enable higher spatial resolution, and thus allow for the detection of finer details in the specimen. To ensure adequate sampling for high-resolution imaging, an interval of 2.5 to 3 samples for the smallest resolvable feature is suggested[6,7]. However, as pixel size decreases, the number of photons that can be captured by each pixel decreases, which in turn can lead to a decrease in the SnR.

On the other hand, larger pixels allow for the capture of more photons, providing a higher SnR and in that context, improved image quality. However, the use of larger pixels, as we can imagine considering Nyquist criteria, instead lead to an under-sampling and a loss of spatial resolution. To find the appropriate pixel size for a given microscope system, one should consider the objective lens' NA and the wavelength of the light (λ). According to

the Abbe's diffraction limit formula[19] and the Nyquist theorem[18], the optimal pixel size can be calculated as:

$$Optimal\ pixel\ size = (0.5 * \lambda)/NA \tag{4}$$

Sensor size

In addition to pixel size, the size of the sensor of the camera is an important factor when acquiring images. The sensor size determines how large of an area that is captured of the projected image. Naturally, a camera with a large sensor size is capable of acquiring a larger area of the field of view. However, there may be trade-offs that one has to consider when picking a camera with a larger sensor[20,21].

- **Pixel density:** When picking a camera because of the size of its sensor, one must not neglect the size of the individual pixels. If the sensor size is increased but contains the same amount of pixels, the pixel size will be larger, potentially leading to a loss of spatial resolution. On the other hand, if the sensor size is increased while the pixel size remains the same, the number of pixels in the acquired image will increase, potentially leading to higher resolution but with an impact on other factors such as file size, oversampling, and signal-to-noise ratio[20,21].
- **File size:** If the sensor size is increased while the pixel size remains the same, the number of pixels in the acquired image will also increase. This results in larger image files due to the increased number of pixels. Larger files can lead to increased storage requirements and slower data transfer speeds, which may be a concern in high-throughput or time-lapse applications[20]. However, with the fast read- and write-speed that we see today with modern drives, this is rarely a limiting factor.

When selecting a camera for microscopy applications, it is important to consider the trade-offs between sensor size, pixel density, and file size. The optimal choice will depend on the specific requirements of the imaging system, such as the desired field of view, spatial resolution, and imaging speed. Additionally, the compatibility of the camera sensor with the microscope's optical system should be taken into account to ensure adequate sampling of the full resolution available with a particular lens.

Quantum efficiency

Quantum efficiency (QE) is a crucial factor in determining the sensitivity of a camera system. QE refers to the ability of the photodiodes to successfully register the incoming

photons. Unfortunately, we cannot register all the incoming photons; some photons will miss the photodiode entirely, some will be absorbed by surrounding material, and some simply fail to elicit a charge within the diode. QE is determined as the percentage of the photons hitting the chip that is successfully registered and measured[6,7,13].

Multiple factors influence QE of a camera system. These include the wavelength of the incoming photons and, to a lesser degree, the temperature of the chip. The QE of a camera is typically represented as a graph that plots the percentage of photons detected by the diode against the wavelength of light. This graph is commonly referred to as the QE curve. The peak of this curve varies depending on the material used in the diode's construction and represents the wavelength at which the photodiode has the highest efficiency in converting light into an electrical signal.

Dynamic range and bit depth

Once the photon has been registered in the photodoide and converted into a voltage by the read amplifier, the voltage must be digitized by the ADC. In the ADC, the voltage is transformed into a analog signal representing a digital gray value, which corresponds to the signal amplitude (i.e. proportional to the number of photons hitting the diode or pixel)[6]. The digital gray value ranges between 0 and 1, where 0 represents no signal and 1 represents the maximum charge capacity the pixel can store. However, as pointed out above, pixels will never be 0 because of the noises associated with capturing an image. The number of steps within this range is dictated by the bit depth of the camera, and is typically expressed as 2^n, where n is the dynamic range in bits. For example, a pixel in a 12-bit camera can display 2^{12}(4 096) different values between 0 (absolute black) and 1 (complete saturation; white). However, as demonstrated in Figure 2, there is a point at which a higher bit depth might provide excess information about the sample, depending on the structure being studied.

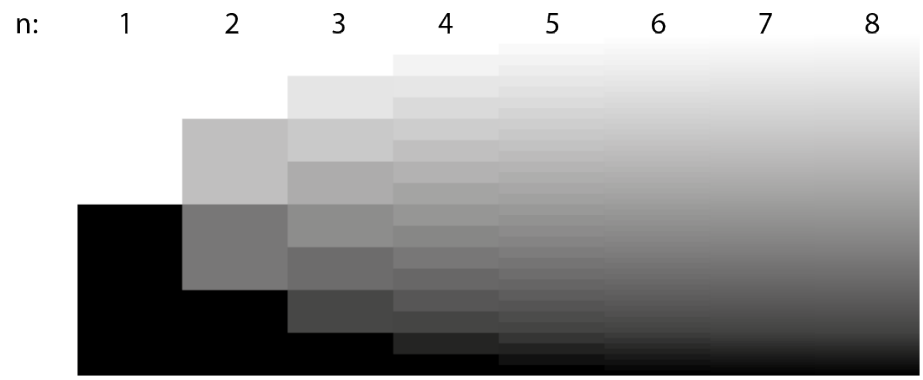

Figure 2: Visualization of the gradient of gray values as the number of bits increase. One bit allows for only two values (2^1, e i. black and white) whereas in the 8-bit, the gradient is smoother with 256 (2^8) gray values.

In practice, the dynamic range is often defined as the ratio of the maximum measurable signal to the minimum measureable signal. For CCD cameras, this is typically specified as the ratio of the FWC (maximum possible signal) to the read noise (minimum detectable signal considering the SnR).[5] However, the usable dynamic range of the camera is typically smaller than this maximum value due to nonlinear effects and the fact that the minimum quantifiable signal intensity in photoelectrons is actually well above the read noise level[7,8]. For instance, a camera with a 16,000 electron FWC and a readout noise of 10 electrons would have a dynamic range of 1600:1, or between 10- and 11-bit resolution (1021 and 2048 possible gray values respectively). Thus, a camera with a bit depth of 12 (4096 possible gray values) would be more than enough to sufficiently sample the intensity information. Storing the images in a higher bit depth than the camera can distinguish simply implies that we are oversampling the data, resulting in larger file sizes than what is necessary.

Widefield and Confocal microscopy

The microscope was a revolutionary tool that allowed us to see things we had never seen before, to witness the invisible and explore the microscopic world with newfound clarity. The pioneers of this new field of study include Robert Hooke, who first described the microscopic structure of cork cells (and also coined the term cell)[22,23], Marcello Malphigi, who studied the anatomy of plants and animals[23], and Antonie van Leeuwenhoek, who is credited with the discovery of bacteria and other microorganisms[22]. With the microscope, these scientists and many others gained a new perspective on the world and the creatures that inhabit it, and a seed of modern biology was sown.

As our understanding of the world around us grew, so did the demand for more advanced microscopy techniques. This led to a series of innovations, such as the development of compound microscope and the subsequent invention of electron microscopes, pushing the boundaries of magnification and resolution. The late 20th and early 21st centuries witnessed the integration of digital imaging and computer-assisted analysis, which transformed microscopy by enabling researchers to capture, manipulate, and analyze images with unprecedented precision and efficiency. Today, widefield and confocal microscopy stand as two of the most commonly used and relevant techniques.

Widefield

Widefield microscopy might be seen as the most straighforward of the microscopy techniques. As the name may suggests, in widefield microscopy, a large section of the specimen is illuminated and captured at a time. To understand how we can generate an image in widefield, lets take a look at two conventional methods of light microscopy and follow the

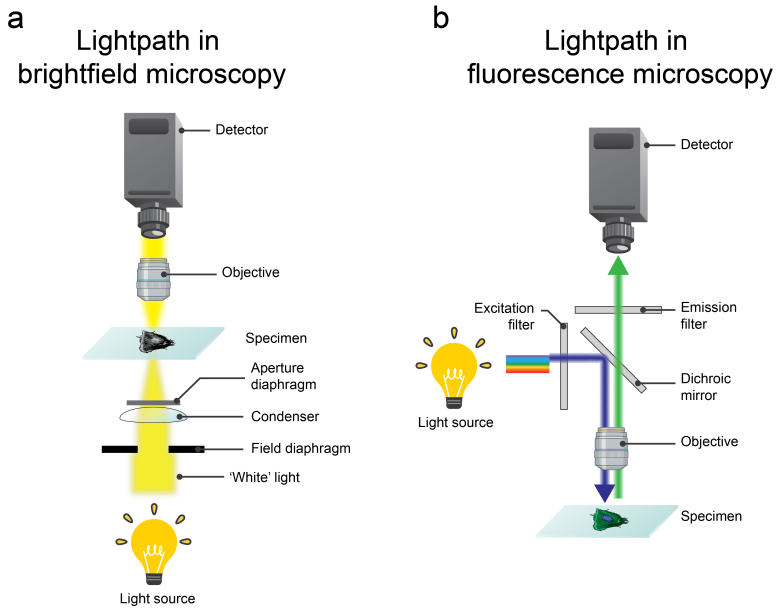

Figure 3: The lightpath as it travels through the microscope. **a** In brightfield, a lamp (typically a tungsten-halogen) illuminates the sample by first passing through the field diaphragm and the condenser. The condenser focuses the light into a beam. A aperture diaphragm is typically associated with the condenser for controlling contrast. The light passing through the specimen is subsequently collected with an objective and registered on a detector. **b** In fluorescence microscopy, the light (a lamp or LEDs) is first passed through an excitation filter to narrow down the wavelength to an appropriate excitation range for the fluorophore (in this example: UV). A diachroic mirror allows the passing of some wavelengths whereas others are reflected. The light is focused by the objective onto the specimen, where the excitation light excites any fluorophores situated in the specimen. Emitted light (green) are collected with the objective, and as the light travels back down, it is reflected towards the detector by the dichroic mirror. The lightpath in microscopes can vary greatly depending on setup, but the principles remain the same.

light as it travels through the microscope.

Brightfield

Brightfield microscopy is a common method in widefield microscopy that involves the unfiltered transmission of white light through the specimen. This light is typically generated using a lamp or light-emitting diodes (LEDs), and travels through a series of diaphragms and the condenser before it interacts with the specimen.

The diaphragms, such as the field and aperture diaphragms, help control the amount and the angle of the incoming light, while the condenser focuses the light onto the specimen (see Figure 3). The condenser has to be aligned with the focal plane of the specimen to ensure that the incoming light is uniform and focused correctly on the entire field of view. Köhler illumination [24] is often used in this context to optimize the distribution of light and enhance image contrast, and consists of a series of steps to ensure the condenser is properly aligned with the focal plane of the specimen and the objective [2].

However, when working with transparent or low contrast specimens, such as living cells or thin tissue sections, the contrast in the brightfield images may be insufficient for any valuable information to be extracted. To overcome this, various contrast-enhancing techniques have been developed. The two most common ones are:

- **Phase contrast:** Phase contrast (PH), invented by the Nobel Prize winner Frits Zernike[25], is a technique that exploits the differences in the refractive index of different parts of the specimen to generate contrast[26]. In PH microscopy, an annulus and a phase plate is introduced in the light-path, shifting the phase of the light passing through the specimen relative to the light passing through the surrounding medium. This phase shift results in variations in brightness and contrast in the image, and thereby helps in our ability to resolve low-contrast specimens (see Figure 4a).
- **Differential inteference contrast:** Compared to phase contrast, differential interference contrast (DIC) utilize polarized light and optical interference to generate contrast based on the structure of the specimen as well as gradients in the specimen's refractive index. In DIC, the light is first polarized and split into two orthogonally polarized beams that travel through the specimen along slightly different paths. When these beams are combined, they interfere with each other, generating an image with enhanced contrast that reveal fine detail and structures within the sample[27]. See Figure 4b for an illustration of the light-path in DIC.

When comparing the two, phase contrast is generally easier and cheaper to setup than DIC. However, phase contrast produce halo effects around the specimen, which appear as bright or dark rings around the edges. These artifacts can sometime obscure fine details and make it difficult to interpret the image accurately. On the other hand, DIC offers improves contrast and resolution when compared to phase contrast, and allow for the visualization and perception of fine structure of the specimen. However, DIC requires more complex and sometimes expensive polarizers and prisms.

Fluorescence

When working with fluorescence widefield microscopy, the path the light travels differs when compared to brightfield. Instead of using white light, fluorescence microscopy involves the excitation of fluorophores within the specimen using specific wavelengths of light. As this light interacts with the fluorophores, the light is absorbed and the fluorophore enters an excited state. Eventually the fluorophore exits its excited state, releasing the absorbed energy in the form of a photon. However, during this process, some energy is inadvertently lost, causing the emitted light to be in a longer wavelength compared to the

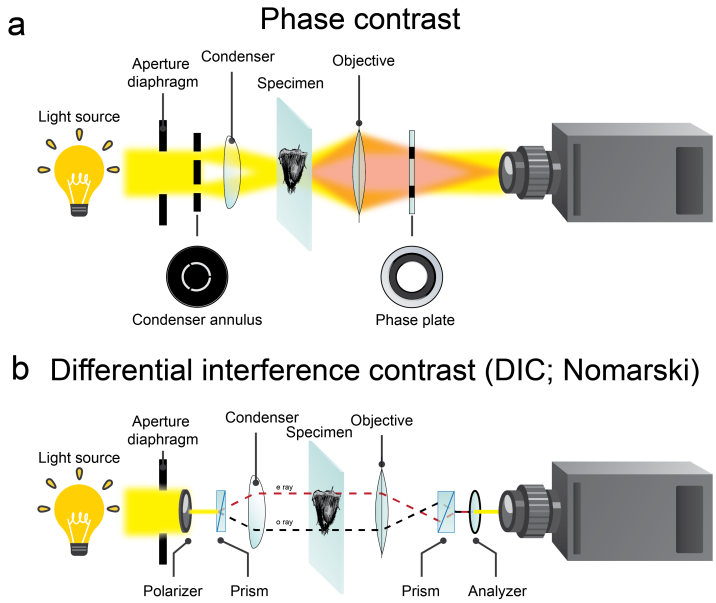

Figure 4: The principle behind phase contrast and differential interference microscopy (DIC). *a* In phase contrast, a condenser annulus is introduced in front of the condensor, causing parallel light wavefronts to emanate from the holes in the annulus. These wavefronts interact differently with the specimen depending on the reflactive index and length of travel. The phase-shift caused by the specimen is finally filtered by a phase ring. This process results in a difference in signal amplitude depending on the phase. Unlike DIC, the circular nature of the condenser annulus and the phase plate prevents orientation dependant artifacts. **b** In DIC techniques, contrast is achieved by utilizing the polarization of light. In the case of Nomarski DIC, polarization is achieved by passing the light through a polarizer. A Nomarski prism subsequently splits the linear polarized light into two orghogonal components (denoted e ray and o ray in the schematic, where e=extraordinary and o=ordinary). As the light interacts with the sample, the e ray and o ray travel are interfered by the sample, resulting in a phase difference. Once the light is collected using the objective, another Nomarski prism (angled perpendicular to the first prism) combines them into a single beam of light. Finally, the light emerging from the second Nomarski prism is passed through an analyzer to form an image.

excitation light. This phenomena is known as Stoke's shift, and it is the foundation in our ability to perform fluorescence microscopy[1,28].

While mercury or xenon lamps have traditionally been used as a high-intensity light source capable of illuminating in specific wavelengths, in recent years, light-emitting diodes (LEDs) have gained popularity due to their lower heat production, longer lifespans, and more precise control of the excitation wavelengths. As the excitation light enters the microscope, it is first passed through a filter which selects a narrow range of wavelengths optimal for exciting the fluorophore. The light then passes through a dichroic mirror, which allow the light to travel through towards the specimen through the objective.

Upon reaching the specimen and the specimen fluoresce, the emitted light is captured by the objective and travels back down through the same path as the incoming excitation light. However, when interacting with the dichroic mirror, the longer wavelength light is reflected and focused towards the detector. In widefield microscopy, the detector is often a CCD camera or a CMOS sensor, which captures the incoming light and produce the final image.

Because of the fact that the light is unhindered in widefield microscopy, the speed of acquisition is considerably faster when compared to techniques such as confocal microscopy. However, the presence of out-of-focus light contributes to the final image, resulting in a decrease in contrast and the ability to distinguish finer details, and is particularly a problem when working with thicker specimens or samples with high autofluorescence. In terms of resolution, widefield microscopy is limited by the diffraction of the light, which restricts the resolution to around $200\,\mathrm{nm}$ laterally (x-y) and $500\,\mathrm{nm}$ to $700\,\mathrm{nm}$ axially (z), depending on the NA of the objective and the wavelength of light used (see Chapter 1: The scattering of light).

Confocal

Confocal microscopy, in contrast to widefield, introduces a new element to the light-path that significantly alters the way in which images are generated. The key difference in confocal is the use of a pinhole places in conjugate focal plane of the objective. This pinhole effectively blocks out-of-focus light from reaching the detector, which greatly improves the contrast and axial resolution of the resulting image[10].

In confocal, the light source is often a laser, which provides high-intensity, monochromatic light to the microscope. The incoming light is directed through a set of scanning mirrors, which allow the beam to scan across the specimen in a raster pattern. This enables the microscope to collect data from one point of the time, creating a point-by-point image of the specimen.

As the specimen fluoresce, the emitted light is collected by the objective and directed back throught he dichroic mirror. The emitted light then passes through the pinhole, which only allows in-focus light to reach the detector (a PMT or APD). This results in a significant reduction of the background signal, and an increase in image contrast and axial resolution ($300\,\mathrm{nm}$ to $500\,\mathrm{nm}$) compared to widefield.

Comparison between widefield and confocal

When comparing widefield and confocal, it is evident that the two techniques come with their unique advantages and drawbacks. While both techniques are relatively equal in xy-resolution, they differ in other aspects that may impact their suitability for particular applications.

Image contrast and axial resolution

Confocal microscopy offers improved image contrast and axial resolution compared to widefield microscopy, due to the use of a pinhole to block out-of-focus light. This makes confocal microscopy particularly advantageous when imaging thicker specimens or samples with high autofluorescence, where widefield microscopy may struggle with high background signals and reduced contrast. However, that being said, deconvolution algorithms can be applied in which the out-of-focus light in an image is compensated for.

The process of deconvolution involves using mathematical algorithms to reverse or compensate for these effects, thereby enhancing the image quality, resolution and contrast in the final image. Deconvolution can be described mathematically[29]:

$$image(x, y, z) = object(x, y, z) * PFS(x, y, z) \tag{5}$$

where * denotes the convolution operations, object(x,y,z) represents the original, undisturbed image, and PFS(x,y,z) is the point spread function of the system.

In other words, one must know the PFS of the system in order to restore the 'original', undisturbed image. Inaccuracies or incomplete information about the PFS can introduce artifacts or errors in the deconvolved image, potentially compromising data reliability. Furthermore, as we can see in equation X, one must capture the three dimensional data (z-stacks) in addition to the focal plane of interest in order to correctly perform deconvolution[29]. This can significantly slow down acquisition and lead to additional photobleaching and phototoxicity, particularly when used with higher magnifications.

Speed of acquisition

Widefield microscopy offers a faster image acquisition speed compared to confocal microscopy, as it captures a large section of the specimen simultaneously. In contrast, confocal microscopy scans the specimen point-by-point, which takes more time to generate an image. For applications that require rapid imaging or real-time observations, widefield microscopy may be a more suitable choice.

Photobleaching and phototoxicity

Confocal microscopy, due to its point-by-point scanning and the use of high-intensity laser light, may cause increased photobleaching and phototoxicity in the specimen compared to widefield microscopy. This can be particularly problematic when imaging live cells or

other sensitive samples. Widefield microscopy, with its lower light exposure, may be a more suitable choice for such specimens.

In conclusion, the choice between widefield and confocal microscopy depends on the specific requirements of the application, the sensitivity of the specimen, and the available resources. Each technique offers unique benefits and limitations, and researchers must carefully consider these factors when selecting the most appropriate method for their imaging needs.

Relevance and considerations in the presented work

Naturally, the choice of imaging system and equipment were of great importance in the my work. Considering my work primarily revolves around acquiring large, live-cell population-wide datasets of the specimen in question, high throughput was essential when picking the microscopy technique.

The throughput of a widefield microscope is without a doubt hard to question, providing an easy means for fast acquisition with a relatively low photon-dosage, thereby minimizing phototoxicity and bleaching of the cells. Here, the model of the microscope made a big difference. When I began working we had just purchased and installed a new microscope; specifically the Nikon Ti-E2 model. What makes this model stand out is its $25\,\mathrm{mm}$ field of view (FOV), making it acquire up to $50\,\%$ more data in a single image compared to the more common $18\,\mathrm{mm}$ FOV. This drastically increases the throughput of the system, and reduces the redundancy of image-stitching for large samples. Furthermore, the choice of camera was an important consideration for our applications. In order to utilize the full FOV that the microscope was capable of, without sacrificing resolution (in the form of pixel size) or sensitivity (CMOS and QE), there was really only one camera on the market at the time capable of meeting these needs: the Nikon Qi2. While there are other cameras that outperforms the Nikon Qi2 in terms of QE and pixel size (the Qi2 has a peak QE of $77\,\%$ and $7.3\,\mu\mathrm{m}^2$ pixels), these parameters are still relatively good, especially considering the large sensor (36.0×23.9mm). While there may be other cameras capable of reaching a higher theoretical resolution, this was not our top priority.

Breaking the diffraction limit

In conventional microscopy techniques such as widefield and confocal, we are limited in our ability to resolve and distinguish points based of the diffraction limit of our configuration. Considerable work has gone into the challange of breaking this diffraction limit and to open up new avenues in terms of image resolution.

These techniques are referred to as super-resolution, and they have had tremendous impact on our understanding of biological processes at the nanoscale.

STED microscopy

Stimulated emission depletion STED microscopy is a super-resolution technique that was first proposed by Nobel Prize winner Stefan Hell in 1994[30,31]. STED microscopy works by selectively depleting the fluorescence surrounding molecules, thereby sharpening the effective PSF and improves the resolution beyond the diffraction limit.

The process of (STED) relies on two laser beams; one for excitation of the fluorophores and another, called the STED beam, for de-excitation. The STED beam has a doughnut-shaped intensity profile, which selectively depletes the fluorescence from molecules at the periphery of the excited region, leaving only a smaller central area to emit fluorescence. By scanning the sample point-by-point, STED microscopy can generate images with resolutions down to $20\,\mathrm{nm}$ to $50\,\mathrm{nm}$, significantly surpassing the diffraction limit.

PALM and STORM microscopy

Stefan Hell was not alone in receiving the Nobel Prize in Chemistry 2014[31]. Among him, Erik Betzig and Willam E. Moerner shared the prize for their development of photoactivated localization microscopy (PALM)[32] and stochastic optical reconstruction microscopy (STORM)[33], respectively. PALM[32] and STORM are localization-based super-resolution techniques that rely on the stochastic activation and localization of individual fluorophores within a sample.

By controlling the activation and subsequent imaging of a sparse subset of fluorophores at a time, the position of each fluorophore can be determined with a higher, nanometer precision compared to conventional microscopy. By repeating this process thousands of times, and accumulating the positions of all fluorophores, a high-resolution image can be reconstructed, typically achieving resolutions down to the $10\,\mathrm{nm}$ to $30\,\mathrm{nm}$ range.

Structured Illumination Microscopy

Structured illumination microscopy (SIM)[34] is a super-resolution imaging technique that improves the resolution of an image by utilizing patterned illumination to extract higher spatial frequency information from the sample. The excitation light is modulated with periodic patterns, typically achieved by passing it through a patterned grid or using a spatial light modulator, and then projected onto the specimen.

When the sample is illuminated with these structured patterns, it creates interference patterns that contain high-frequency spatial information about the sample, which is normally inaccessible in conventional widefield or confocal microscopy due to the diffraction limit. By acquiring multiple images with different phase shifts and orientations of the illumination pattern, SIM allows the reconstruction of a high-resolution image, effectively doubling the resolution compared to the diffraction limit.

The process of acquiring an image in SIM involves three main steps[34]:

1. **Acquiring raw images:** Multiple images (typically 9 to 15) are captured with varying phase shifts and orientations of the structured illumination pattern. This ensures that high-frequency information from different spatial directions is collected.
2. **Demodulation:** The raw images are processed to extract the high-frequency information contained within the interference patterns. This is achieved through a series of mathematical operations, including Fourier transformations and filtering, which separate the low- and high-frequency components of the images.
3. **Reconstruction:** The demodulated information is combined to generate the final super-resolved image. This is done by merging the high-frequency components from the demodulated images with the low-frequency components from the conventional image.

SIM is particularly advantageous for live-cell imaging, as it is less phototoxic and faster than other super-resolution techniques like STED or PALM/STORM[35]. Additionally, SIM can be used with a wide range of fluorophores and can be adapted for multicolor imaging, allowing the simultaneous visualization of multiple structures within a sample. However, it is worth nothing that SIM typically provides a more modest resolution improvement compared to other super-resolution techniques, achieving a resolution of approximately $100\,\mathrm{nm}$ to $130\,\mathrm{nm}$ in practice[35,36]. Furthermore, the image reconstruction process in SIM can be sensitive to noise and sample drift, which may introduce artifacts or errors in the final image if not properly accounted for[37].

On the combination of multiple techniques

When performing experiments utilizing microscope, the choice of method depends on the restrictions of the specimen and the requirements of the experiment in question. These limitations can be in the form of specimen thickness and resolution requirements that narrow down the potential choice to a handful of techniques. However, it is sometimes necessary to not only acquire data from one system, but many. This acquisition of data, where

information from one technique is supplemented and correlated with data from another technique, is refereed to as correlative microscopy.

One of the most common applications where data from two modalities are acquired is in correlative light-electron microscopy (CLEM) [38–40]. CLEM combines the high-resolution structural data of electron microscopy with the molecular specificity of fluorescence microscopy, providing a more comprehensive understanding of the specimen.

There are also applications where information is gathered on multiple modalities using solely light microscopy (for a review, see [41]). Light microscopy offers various advantages, including the ability to study living specimens, less complex sample preparation, and a wide range of contrast mechanisms. The complexity and challenges with performing correlative microscopy lies in the correlation of the data and images. Moving the sample from one microscope to the next means that the images you take are on different coordinate systems, and finding the same region of interest between the two is hard even for an experienced microscopist. Furthermore, problems quickly arise when working in the high-fidelity space, where nano-scale offsets in calibration causes the data to be misaligned.

Most applications of correlative light microscopy (that I am aware of) has been performed where the two imaging techniques are combined into one system [42–48]. This removes the problem of having to solve alignment of the coordinate space, leaving only the problem of aligning the images taken themselves. While I have stumbled upon solutions for solving the alignment problem when using multiple microscopes, or when the experiment calls for sample preparation and thus the removal of the sample off the stage in-between acquisition [49–52], there does not seem to be a general solution for quickly calibrating a sample and perform correlative light microscopy on multiple systems.

How life created light

Using microscopy, we possess the ability to dive deep down into the cell, observing its complex machinery and intricate processes in real-time. However, without fluorescent proteins (FPs), like the green fluorescent protein (GFP) and its variants, investigating the inner workings of life would be substantially harder. Moreover, because they can be genetically encoded, FPs have given us a way to directly link and see individual molecules in real time, revealing their intrinsic behaviour and interactions with unprecedented detail. Without this bio-molecular tool, we would be limited to indirect methods that provide only a fraction of the information that FPs can offer.

The GFP was discovered in the 1960s when Osamu Shimomura et al. [53], who was interested in the fluorescent capabilities of *Aequorea victoria*, isolated the protein as a by-product while studying its associated counterpart aequorin (a calcium sensitive protein capable of emitting

light of the blue spectrum)[54]. Despite its discovery, the true potential of this protein was not realized until Douglas Prasher successfully cloned the gene for GFP[55], and Martin Chalfie successfully expressed it in *Escherichia coli* and *Caenorhabditis elegans*[56]. However, the wild-type variant of GFP was poorly fluorescent, and had in fact a stronger peak in the ultraviolet spectra rather than the green[57]. Additionally, the protein, being evolved to fold in the colder climate of the *Aequorea victoria*, was in its wild-type unstable for mammalian expression. To understand the transformation that GFP has underwent to where it is today, we must first explore the structure of the protein and its fluorescent capability.

Fluorescent proteins

Green fluorescent proteins

The GFP is a 28-kDA protein capable of emitting green fluorescence upon the stimulation of light from the blue to ultraviolet spectrum. The protein folds into a barrel like structure, where its 11 β-sheets surrounds the fluorescent chromophore forming core in a helical lattice (see Figure 5). This chromophore consists of the three amino acids, Serine 65, Tyrosine 66 and Glycine 67 that are covalently bound to each other through series of post-trancriptional modifications[55,58]. Notably, these modifications are dependant on oxygen, and as a by-product produces hydrogen peroxide which can be toxic to the organisms producing the protein. Roger Y. Tsien et al.[57,59] pioneered the work of improving the photostability of the protein, as well as increasing its capacity to emit light in the green spectra, by replacing the Serine 65 with Threonine (S65T).

Yet, these modifications did not solve all issues that hindered the GFP from widespread adoption. The wild-type variant was unstable at the warmer temperature needed for mammalian expression[60], largely attributed to the fact that the protein was evolved to fold in the colder habitat of *Aequorea victoria*. Furthermore, the wild-type variant was prone to dimerization because of a hydrophobic patch in one of its β-sheet, and . The amino acids contributing to this hydrophobic patch was, most notably, Ala 206, Lys 221, and Phe 223[61]. This was troublesome, as anything tagged with a GFP would now be prone to dimerization as well, even though the two tagged proteins might not find itself in that conformation naturally. This was eventually solved by replacing one or several of these amino acids to positively charge residues, such as A206K (Alanine 206 replaced with Lysine)[62,63]. Any interaction between monomers would now be suppressed due to the electrostatic repulsion caused by the positively charged amino acids.

Naturally, it is difficult to cover all the work that has been performed on the wild-type variant of GFP to the versions that we have today. However, an important exception is the development in shifting the excitation and emission wavelengths of the GFP towards the longer wavelengths. Without the additional proteins capable of emitting light in other

wavelengths, we would be limited to studying a single kind of molecule at a time. Work by Roger Y. Tsien et al.[58] showed that it was possible to manipulate and alter the fluorescent spectra of the GFP by substituting amino acids in or around the chromophore. For instance, if the Tyrosine 66 is replaced with a Histamine or Tryptophan, the protein instead emits light in the blue or cyan spectra respectively. However, the absence of red FPs, at the time, limited the range of biological systems that could be studied simultaneously.

Red fluorescent proteins

In parallel to the improvements on GFP, there was considerable interest in unlocking the use of FPs in the longer wavelengths of the visible spectra. Lukyanov et al.[65] hypothesized and showed that the fluorescence seen in reef corals, some among them in red, were not due to unidentified pigments, but rather of other GFP-like proteins. This work lead to the identification of 6 homologous proteins and share the same β-sheet structure as GFP. Most notably was their isolation of drFP583, a variant with a peak excitation of 558nm and with an emission peak of 583nm. This variant, later referred to as DsRed (stemming from the organism it was isolated from, *Discosoma*, and its red color profile), was the wild-type variant of many of the red FPs we have today. However, three major issues resided with this variant; from crystallization it was found to have a strong affinity to form a tetrameric structure[66–68], slow maturation[69–71], and its residual green fluorescent component[65,71]. Further mutagenesis eventually solved these issues[72,73], growing the plethora of fluorescent proteins with different spectral profiles (e.g. mOrange, mStrawberry, mCherry[74] and mScarlet[75]) that we have today.

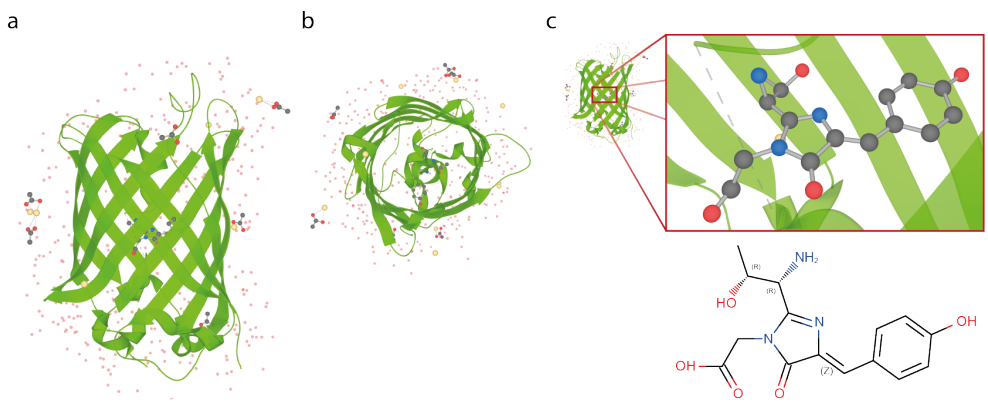

Figure 5: The three dimensional structure of the superfolder GFP resembles a barrel-like structure with 11 β-sheets surrounding the fluorescence-capable chromophore. **a** Rendering of the 3D structure of superfolder GFP. Water molecules and ligands can be seen surrounding the structure. **b** Top view of the 3D structure. **c** The containing chromophore and its chemical structure. The amino-acids within and around the chromophore dictate the characteristics of the fluorescence. Images are from the RCSB PDB (`csb.org`) of PDB ID 2B3P[64].

Split GFP

One creative development of GFP came with the introduction of split GFP (sGFP). In this technique, first demonstrated by Cabantous et al.[76], the GFP molecule is divided into two fragments: a larger N-terminal fragment (residue 1-214) and a smaller C-terminal fragment (residue 215-238). The N-terminal fragment, known as GFP1-10, encompasses the first 10 β-strands of the GFP and adopts a structure resembling the entire GFP, with the exception of the missing 11th β-strand. On the other hand, the C-terminal fragment, GFP11, comprises this absent 11th β-strand. In their isolation, these fragments are non-fluorescent. However, upon association, GFP1-10 and GFP11 complete the β-barrel structure and the containing chromophore, restoring fluorescence in the process. The unique properties of sGFP have provided opportunities for various applications, including, but not limited to, the investigation of protein-protein interactions[77–79] and cellular localization[78,80–83], thereby broadening the scope of GFP-based technologies in molecular and cellular biology.

During my studies, we had a project running where utilizing the sGFP variant was key to our research question. Our primary interest was to further characterize the delivery of effector proteins through the type 3 secretion system (T3SS) employed by the human pathogen *Yersinia pseudotuberculosis*. The delivered proteins, known as Yersinia outer proteins (Yops), are effector proteins that are translocated to the host cell upon contact, allowing the bacterium to evade the immune system. For more detail on the T3SS and Yops in *Y. pseudotuberculosis*, I refer to the review by Gloria I. Viboud and James B. Bliska[84]. In short, the T3SS is a sophisticated protein complex used by numerous gram-negative bacteria to secrete effector proteins into the milieu or into the membrane or cytosol of host eukaryotic cells. The major defining characteristics of the T3SS are the presence of a needle-like structure extending from the basal body, activation by host cell contact, and the delivery of effector proteins. This needle-like structure of the T3SS, through which the effector proteins pass, is remarkably narrow, with a diameter of a mere 28 Å[85,86]. To overcome this size-restriction, the delivery of effector proteins is orchestrated by an unfolding of proteins prior to translocation. The specificity of which proteins are translocated is dictated by the presence of a secretion-signal on either the amino-acid level or the mRNA[87].

Attempts to study the translocation of Yops by tagging them with GFP have been unfruitful. The GFP is too large (approximately 28 Ångstroms in diameter) to fit through the needle of the T3SS and lacks the necessary secretion-signal to be unfolded prior to translocation. However, other fluorescent tools, such as the β-lactamase TEM-1 system, have been utilized to follow the delivery of effector proteins in *Escherichia coli* to the host cell[88–90]. As the fluorescence measured is a product of a chemical reaction, it would be troublesome to quantify the amount of effector proteins delivered or correlate the amount of delivered effector proteins to the cellular outcome. We hypothesized that, instead of using the com-

plete GFP, the 11th β sheet (GFP11) of sGFP would be sufficiently small to fit through the narrow corridor of the T3SS, circumventing the need for unfolding and the presence of a secretion-signal.

After successfully cloning GFP11 onto YopE, we observed an increase in fluorescence upon bacteria-cell contact. However, the positive interactions were rare and weak in signal, and the background fluorescence in the negative control was relatively high, making it difficult to assess the results. Additional development would have to be made in order to increase the signal-to-noise ratio of the delivered Yops. One means of increasing the fluorescence per translocated Yop was to tag each Yop with multiple GFP11 subunits in tandem [81]. However, doing so meant answering the following questions: could the tandem-complex translocate at all? And how would additional GFP11 subunits affect the speed of which the effector protein is translocated? Ultimately, due to time limitations, we decided to pause the project.

Live cell imaging

The desire to study and observe living specimens has been around since the dawn of microscopy. This process of studying living specimens refers to as live cell imaging, and allow us to in real-time investigate dynamic cellular processes in their native state. Excellent review articles exists on the subject of live-cell imaging using cell culture (see [91,92]). Without risking of repeating what they already have so excellently put into words, I do want to mention some of the key considerations one has to take when performing live-cell imaging, especially in the context of fluorescence microscopy.

Phototoxicity

All living things are fragile. When we peer into the microscopic world of living single cells, we have to make sure our tools and methods are as unobtrusive as possible. That can be difficult in the context of fluorescence microscopy, considering the very basic element we build our observation with, light, can be highly damaging to cells in virtually any amount [93–95]. At the intensity-levels used in fluorescence microscopy, light not only have the strength to physically break cellular structures (like the break of double-stranded DNA [96]), but also to change the environment of the cells. Fluorophores, as they undergo excitement, sometimes fail to emit light as they fall back to their rested state. Instead, the high energy electrons generated can react with the dissolved oxygen in the media. The product of this reaction, reactive oxygen species (ROS), is highly damaging to cells left unaccounted for. An increase from the homeostatic levels of ROS, also referred to as oxidative stress, can negatively affect several cellular structures and processes, such as membranes, lipids, proteins, lipoproteins and DNA [97–102]. Albeit cellular systems having been evolved in order to handle oxidative

stress[103] produced by cellular processes themselves, an excessive amount caused by fluorescence imaging will nevertheless alter cell state and impact the experimental outcome. For this purpose, it is important to consider the monitoring of cellular processes affected by phototoxicity, like the timing of mitotic stages[104], or the monitoring of ROS directly through a detection kit. Other, more direct measures like decreaseing the intensity of the excitation light or reducing the exposure time may help in terms of sample health.

Cell homeostasis

Another consideration that might seem obvious is the maintenance of happy cells throughout the imaging session. This means keeping the cells in homeostatic temperatures and CO_2 concentrations suitable for that particular cell line. Today, we have excellent incubator champers that we can equip onto our microscopes, providing the specimen a 37°C environment. An important note here however is to make sure the microscope body and objectives are kept at similar temperatures as the stage environment surrounding the cells. The microscope body acts as a massive heat sink, and if the environment has not been been calibrated for at least 24 hours, heat will participate from the sample through the objective into the microscope body. This heat-transfer is even worse during the use of immersion media such as oil or water, causing there to be a direct channel where heat can effectively travel down from the sample and where we image to the microscope body. Failing to maintain a homeostatic temperature can have major implications on the cells' states[92].

In terms of providing an ample CO_2 rich environment, this is most effectively done by providing a CO_2/air mixture to the atmosphere. The usual target here is to have a $5\,\%$ to $10\,\%$ CO_2 atmosphere (depending on media) in order to buffer and maintain the pH of the media. However, it is also possible to grow cells in a unaltered atmosphere, but does require the introduction of an additional buffer, like HEPES, that can act over longer periods of time. In my experience, because of the usual small environments that the specimens are in, it can be difficult to maintain the correct atmosphere and pH levels over time due to the atmospheric leakage and difficulties in calibrating the enclosure. Thus, I have found that the additional HEPES buffer may help keeping the cell culture going for longer ($> 24\,\mathrm{h}$) even in the presence of CO_2. However, careful consideration should be taken when using any additional complement, and adequate controls should be in order to make sure the effect on the biological process in question is not affected.

Use of fluorophores

First, one has to consider the means of staining. The most common approach, where DNA expressing the protein of interest in fusion with a fluorescent protein (FP) is introduced to the cells through transfection, results in the additional production of the protein of interest

than what is otherwise produced. In other words, this may lead to the over-expression of the protein of interest in the cells[105]. This might be undesirable since you leave the territory of physiological relevance. Furthermore, this means of introducing fluorescence is hampered by the lack of control in terms of how much DNA each cell 'absorbs'. Each cell in the culture will be transfected in varying amounts, leading to a heterogeneous level of expression in the population. Alternatively, one could perform gene-editing on the protein of interest through CRISPR/Cas9[106], where a fluorescent protein would be introduced in conjunction with the protein of interest. However, performing gene editing is a more complex and time-consuming process compared to simple transfection, and requires additional controls to be in order to ensure insertion was performed correctly.

Other methods, such as introducing antibody fragments in conjunction with a fluorescent tag targeting the protein of interest, may be advantageous compared to direct FP tagging[105]. Considering the rather bulky GFP, its fusion to a protein of interest may impede proper folding and/or its function through steric hindrance. On the other hand, direct FP fusion allow for a 1 to 1 relationship between the protein of interest and the FP, making it easier to quantitatively assess the protein of interest. However, such a relationship might make S/N unsatisfactory, considering the lower amount of FPs overall, compared to where you have multiple fluorescent probes per protein.

The imaging technique

The choice of imaging techniques can have a profound effect on the health of the specimen. Different techniques simply requires the use of different dosages of light in order to achieve satisfactory levels of SnR. For example, considering the same dosage of light, a widefield will yield an image that is on average 10 to 15 times more intense than that of a point scanning confocal microscope[107]. This is largely due to the nature of the confocal technique, where the pinholes reject a considerable amount of light passing through. Nevertheless, confocal microscopy remains a viable option for live-cell imaging, particularly when employing spinning disk confocal microscopy (SDCM).

SDCM is a widely-used technique where two disks, each containing thousands of pinholes arranged in a spiral pattern, are introduced to the light-path. The Yokogawa scanner enhances the system's light efficiency by adding microlenses to the second disk[108]. The combination of the high scanning speed (up to 360 frames per second) and a high QE CCD camera, SDCM significantly reduces the amount of photobleaching and toxicity (10 to 15 times) when compared to point scanning[109]. While SDCM does not achieve as low a photon dose as widefield microscopy, it's significantly lower photon dose compared to point scanning makes it well suited for high magnification live-cell imaging, especially in 4D[92].

Image interpretation and analysis

What we see depends mainly on what we look for
— John Lubbock

As microscopy has undergone a digital transformation, unlocking its full potential requires not only a comprehensive understanding of the optics and hardware, but also an understanding of how we can interpret and analyse the generated data. Without the ability to computationally analyze the data, we are left to our own imagination to guide us when using microscopy. In many ways, it is unfortunate that our brain is so incredibly sophisticated at this task of visually interpreting an image. When we observe an image, we can easily identify and distinguish between objects in the field of view, we can place what we see in the context of the sample and we (at least believe) that we can make some sort of temporal interpretation of what has and will happen to the objects outside of our immediate observation. Additionally, the human brain suffers from inconsistency, where objects of the same kind might be interpreted differently. It is clear that despite its remarkable capabilities, the human brain will inadvertently introduce bias and errors, clouding any attempts at an objective image interpretation.

In this chapter, we will delve into the various aspects of computational image analysis in the context of digital microscopy. We will untangle how an image is represented digitally, how we can extract information about objects and regions, and how this process is affected by different techniques. We will also explore some of the most popular software solutions for computational image analysis available for researchers without expertise in coding.

A picture is worth a thousand words

The basics of an image may seem trivial to us. We are surrounded by visual information that we continuously have to collect (through the eye and its millions of photoreceptors)

and interpret (through the visual cortex). It is remarkable how our brain can sift through the noise, filter and focus in on the relevant information and connect that information to our memory and imagination. In many ways, the saying that 'a picture is worth a thousand words' holds true. The same can be said from the perspective of computational image interpretation and digital microscopy. Here, instead of eyes we have a camera that can almost seamlessly collect information, and a central processing unit (CPU) and graphics processing unit (GPU) that acts as a visual cortex and brain, performing the task of data extraction and classification. However, in contrast to the human process of image interpretation and analysis, a computational system will always 'see' the same given the same set of instructions, no matter the time of day or glucose levels.

In reality, the digital representation of an image is nothing more than a sequence (typically referred to as a vector or array) of numbers that, when separated into blocks of columns or rows, take the form of an image. Each number is a measurement of the amount of light corresponding from a point in the sample over the course of the exposure time. For instance, a camera with a chip size of 1024×1024 pixels will generate a sequence of $1\,048\,576$ (1024×1024) numbers. This point of view makes it evident that, for a computer, the saying 'a picture is worth a thousand words' is a gross underestimation (see Figure 6).

Coordinate conventions and indexation

As we've alluded to, the result of partitioning of an image into rows and columns is a 2-dimensional matrix of real numbers. Here, each pixel is indexed according to its row- and column-position, giving it a spatial representation in the context of the image. For historical reasons (that I will not dive into here), two paradigms exists for indexing the data contained in a data structure like a vector or array: (a) zero-based indexing and (b) one-

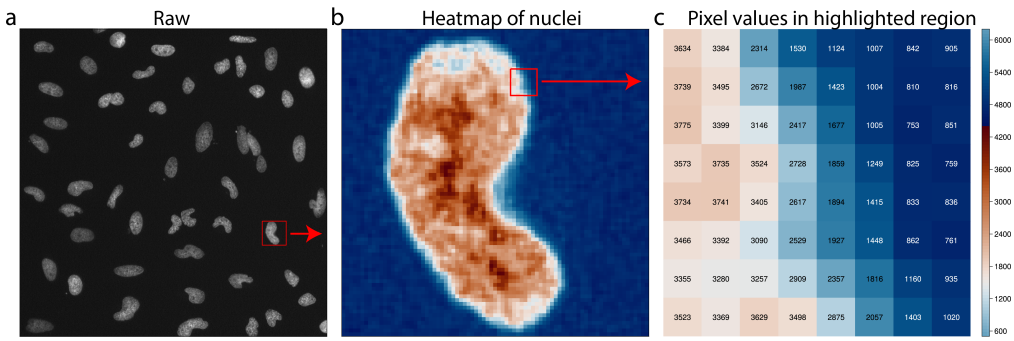

Figure 6: An image captured on the microscope is made up of individual pixels. **a** illustrates the raw image captured as seen in the microscope. **b** A zoomed in region containing a single nuclei. Here, we can begin to see the individual pixels. Image is color-coded according to the vales contained in each pixel. **c** A zoomed in region of panel b plotted in a heatmap. Here, we can see the individual values of each pixel that makes up the image. This small region compromises approximately 0.06‰ of the whole image.

based indexing. Zero-based indexing is used in some of the most influential programming languages today (like Java, Javascript, C, C++, C# and Python), and simply determines that the first element in a data-representation like a list is at the 0th position. In one-based indexing, the indexing starts from 1 and is more popular in scientific programming languages like Fortran, R, Matlab and Julia. See Figure 7 for an illustration of the two types of indexing.

In the context of digital image representation and processing, the sequences (lists or arrays) used to store pixel values can contain various types of data, not just numbers. While numbers are the basic building blocks in programming and are extensively used in image processing, other data types can also be employed to represent information in specific applications. For example, sequences can hold boolean values, complex numbers, or even custom data structures, depending on the requirements of the image analysis tasks at hand. This flexibility allows for diverse approaches to image processing, catering to a wide range of applications and methodologies.

Number representation and types

In programming languages, there are multiple ways to represent numerical values. The most common representations are integers and floating-point numbers. Integers are, like you would assume, numbers without any decimal component, while floating-point numbers have a decimal component and can thus represent numbers with fractional parts. The precision of the number that these types can represent is limited by the number of bits allocated to each type.

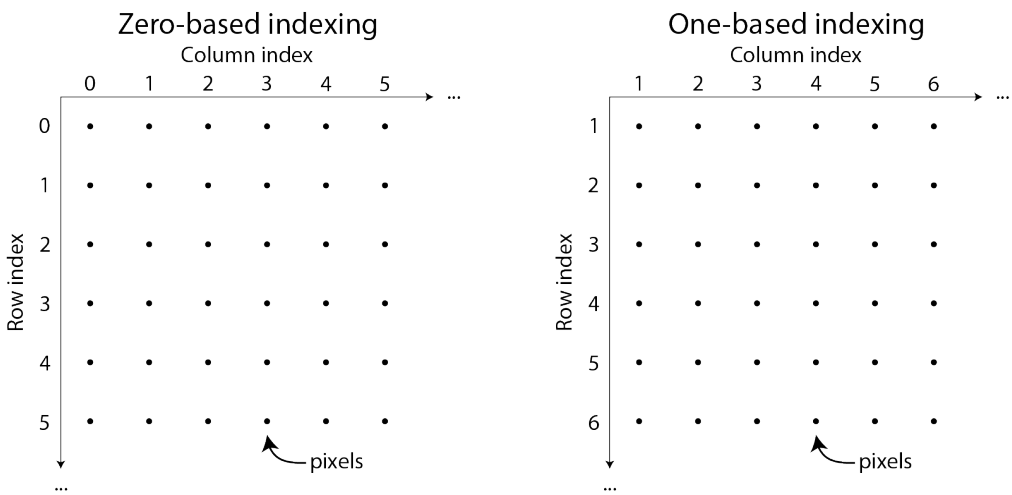

Figure 7: Indexation of data-collections can either be zero-based (left panel) or one-based (right-panel).

Integers are, in addition to the number of bits, limited to the number it can represent by whether it is signed or unsigned. Signed integers can represent both positive and negative numbers, while unsigned integers can only represent positive numbers and zero. For example, an 8-bit unsigned integer can represent values from 0 to 255 (2^8 possible values in total) whereas a signed 8-bit integer can represent values from -128 to 127.

Floating-points numbers, on the other hand, provide a way to represent a much larger range of real numbers, including those with fractional components. They are typically represented using the IEEE 754 standard[110], which defines both single-precision (32-bit) and double-precision (64-bit) floating point numbers. The standard specifies a format that includes a sign bit, an exponent, and a fraction. This representation allows for numbers with a wide range of magnitudes and precision, but due to the finite number of bits used (again, an 8-bit float can only represent 256 possible values), floating-point numbers can only provide an approximation of real numbers.

Programming languages (like Julia[111]) keep track of what a chunk of bits represents using a type system. This type system is used by the compiler (and ultimately the processor) to figure out how operations on the data should be evaluated. For instance, in Julia, **'UInt32'** is an un-signed 32-bit integer whereas **'Int32'** is a signed 32-bit integer. When Julia evaluates addition of these two numbers, it understands that to represent the results, both numbers have to be converted into a 64-bit signed integer (**'Int64'**) in order to perform the operation and contain the results.

Color spaces and representations

In the context of digital microscopy and images in general, various color spaces and representations can be used to describe pixel values. Some of the most common color spaces include RGB (Red, Green, Blue), HSV (Hue, Saturation, Value), and grayscale. Each color space has its unique characteristics and advantages for specific applications. For example, RGB is a widely used color space for display systems, while HSV is useful for color-based segmentation tasks. Grayscale, on the other hand, only contains intensity values. Depending on the color space and representation used, different data types and number representations may be employed to store pixel values.

Collections and data-structures

When working with data such as numerical values and strings (i.e. text; "hello world") or characters, a good way would be to organize them into collections that are suited for the upcoming operations. There are many ways to organize data, for instance:

- **Vectors:** a one dimensional series of values,
- **Matrices:** two dimensional data structure,
- **Arrays:** N-dimensional collection.

Moreover, there are other data-structures that may be more suitable for organizing ones data, like:

- **Dictionaries:** a glossary of key-value pairs,
- **Sets:** a data structure containing unique values,
- **Tuple:** a fixed-length immutable (cannot be changed once instantiated) data structure,
- **NamedTuple:** similar to a tuple, but each value is associated with a name.

These data-structures, not mentioning ones you can access by external libraries, allow for procedures that are both powerful and versatile.

Example: Implementing shading and flat-field correction

When working with images and image analysis, it is important to be aware of the type of image one is dealing with. To highlight this, we will go through an example of where we correct for uneven illumination (through shading correction) as well as the dark current (flat-field correction), a common problem when performing any kind of light microscopy. Because I find the syntax of Julia to be the easiest on the eyes, all code shown from here on out will be expressed as if it was to be executed using Julia.

The first thing we need to to is to load the images into memory. For this, we'll first declare their path and load them using the Images package.

```
using Images

path_to_image = "images/nuclei.tiff"
path_to_flat_field = "images/flat_field.tiff"
path_to_shading = "images/shading.tiff"

image = load(path_to_image)
flat_field = load(path_to_flat_field)
shading = load(path_to_shading)
```

Code block 1: Declaring the path to the images and loading them into memory.

Running the code above (see Code block 1) in a terminal will output the data-structure of the last line of code (loading the shading image; see below). Here we can see that we now have a 367×514 large Array, with the trailing curly brackets declaring the **type** - `Gray{N0f16}`, and the **size** - (2) that the Array can contain. This means that the Array in its current state can only contain values of the `Gray{N0f16}`. The trailing information 'with eltype `Gray{N0f16}`' simply states the elements' type.

```
367×514 Array{Gray{N0f16},2} with eltype Gray{N0f16}:
 Gray{N0f16}(0.0)      Gray{N0f16}(0.0)      Gray{N0f16}(0.0)      ...
 Gray{N0f16}(0.0)      Gray{N0f16}(0.0)      Gray{N0f16}(0.0)      ...
 Gray{N0f16}(0.0)      Gray{N0f16}(0.0)      Gray{N0f16}(0.0)      ...
 ...                   ...                   ...                   ...
```

The next step would be to define a function that corrects for the shading as well as the flat-field image. Functions are a contained sequence of code that has a local variable scope (meaning its variables cannot be seen outside its scope), accepts arguments and returns the result. In fact, in Julia, the result of a function can be a data-structure (like a corrected image) but also another function. This means that for our dataset, we can first create a function which takes two images, the flat-field and the shading image, calculates how the data should be transformed and return a new function where input is a single image (the

one we need to correct). This way, we can declare how shading and flat-field correction is performed early on in the code and reuse the procedure (given that we may have more data in our dataset). This would look like:

```
function generate_correction(shading::Array{T, 2}, flat::Array{T,2}) where T
    shading_norm = maximum(shading) ./ shading

    image -> (max.(image, flat) .- flat) .* shading_norm
end
```

Code block 2: Defining a function in Julia in order to perform flatfield and shading correction. Returns a new function that we can apply to our images.

A few noteworthy things happen here:

- **The `::Array{T, 2}` syntax:**

  This syntax declares which types of argument the function accepts. In this case, T can be any type (declared through `where T` in the end), but it allows us to specify that the two arguments should contain the same type of element.

- **The `.-`, `.+` and `./` syntax:**

  This syntax means that the basic operation (subtraction, addition and division) should be performed per-element (i.e. broadcasting). Normally these operations expect two inputs. Using broadcasting, we can perform the operation on elements (pixels) in the image in one line. An alternative expression that could be implemented here is: `image -> @. (max(image, flat) - flat) * shading_norm)` The `@.` states that all operations should be performed element-wise.

- **The `image -> ...` expression:**

  Here, we declare an anonymous function. This function takes one argument (for the sake of clarity, I've named it image) and performs the right-hand side expression using this argument. Since this is the last expression of the function, and I have not declared a return-statement elsewhere, this anonymous function is returned.

Moving on, we can now generate our anonymous function and apply it to our image (see Code block 3).

```
using Images

function generate_correction(shading::Array{T, 2}, flat::Array{T,2}) where T

    shading_norm = maximum(shading) ./ shading
    image -> (max.(image, flat) .- flat) .* shading_norm
end

path_to_image = "images/nuclei.tiff"
path_to_flat_field = "images/flat_field.tiff"
path_to_shading = "images/shading.tiff"

image = load(path_to_image)
flat_field = load(path_to_flat_field)
shading = load(path_to_shading)

corr = generate_correction(shading, flat_field)

corr(image)
```

Code block 3: Generating and applying our correction to the image.

In our example, we had a simple case of reading images in a directory into memory using the Images library, we declared a function for generating the correction-procedure as a new function, and we applied this to our image of interest. We could as well just have written the sequence using variables for the different images in the first place. The reason we did not do that is to improve re-usability promote readability. For instance, we most likely have multiple images that we wish to perform our correction on. This can be done by:

```
...
file_paths = readdir(path_to_image_seq)

# readdir only gives us the filename, we need the full path like:
file_paths = joindir.(path_to_image_seq, file_paths)

image_seq = map(file_paths) do img_path
    load(img_path)
end

corr = generate_correction(shading, flat_field)

# image sequence is a ::Vector{Matrix{T}}
map(corr, image_seq)
```

Code block 4: In a scenario where we have a series of images that need to be corrected, we can utilize the **map** function.

Here (see Code block 4), we utilize the map-function in two locations. Map is a procedure which takes a function and a collection, and executes the function for all elements in the collection. In our implementation, we expand the map with the ('do') syntax, which enables us to declare an anonymous function. This allows us to have a temporary local variable (img_path; each file in our directory), that we iteratively call `load` on. The neat thing here is that the resulting data-structure will be the same as the input (a vector) when using map. Finally, when we have access to our `corr` function, we can map over the list of images, applying the correction on all images in one line (see Code block 4).

```
...
shading_sequence = map(shading_file_paths) do img_path
    load(img_path)
end

shading_stack = cat(shading_sequence, dims=3)

shading = mean(shading_stack, dims=3)
```

Code block 5: Calculation of the shading image. In our case, we have multiple images taken for estimating the shading in our system. These can be loaded in using the **map** function, creating a List of images in the memory. To concatenate the list into something more manageable, we apply the **cat** function with the keyword-argument **dim** of 3. This results in the a 3 dimensional Array. To calculate the mean shading per pixel and reduce the 3 dimensional array to a 2D image, we can again utilize the dims keyword-argument.

Moreover, several presumptions were made during this example, such as the estimation of the shading and flat-field images in our system. This would typically be calculated by taking the mean value for each pixel from a series of images taken in the same system. This could be done in a similar manner as when we loaded the images, with the exception of first creating a 3-dimensional array where all images are stacked on top of each other, and utilize the `dims` optional argument of the Julia's mean function (see Code block 5).

Image segmentation and classification

In the context of digital microscopy, image segmentation and classification are crucial steps in order to detect and analyse the content of the image. Segmentation refers to the process of partitioning an image into distinct regions, while classification assigns labels or categories to these segmented regions based on specific features or characteristics. The task of extracting and subsequently measuring the content of an image varies depending on the specimen and modality used during acquisition. Acquisition on the microscope should be optimized for high S/N while, at the same time, making sure pixels are not saturated (i.e. pixels on the camera-chip reach their FWC, which leads to the loss of information).

When we see an image, our brain immediately picks out content that is rich in S/N and texture. Consider the raw image of nuclei in Figure 8, how much signal compared to background would you estimate this image contains? In fact, in this rather rich image (~ 50 nuclei), 89% of the data is made out of background (estimated using Otsu's theshold[112]). For this image, the task would be to: (a) identify individual nuclei, (b) separate and label the different nuclei and (c) measure the features of the nuclei. This different steps have to be adjusted depending on the specimen and signal characteristics of the image, with the goal of finding an algorithm able to handle this specific content in different circumstances (e. g. different S/N or, in our case, density of the nuclei). In our example, (a) could be achieved by thesholding on the background, (b) would simply be done by labeling all objects not touching and (c) would depend on the research question.

Image segmentation techniques

There are several methods for image segmentation, ranging from simple thresholding techniques to more complex machine learning-based methods. Some of the most commonly used techniques include:

- **Thresholding:** This is a simple yet effective method for segmenting an image based on pixel intensity values. Typically, a threshold value is calculated based on estimating the background intensity, and pixels with intensities above this threshold can be assigned to different regions or objects. Popular algorithms here are: Otsu's[112], Adaptive[113] and Sigma-clipped[114].
- **Edge detection:** This method identifies edges or boundaries within an image by detecting areas with rapid changes in intensity values. Common edge detection algorithms include Sobel[115], Canny[116], and Laplacian of Gaussian (LoG)[117].

- **Region growing:** This method starts with a seed point or region and iteratively expands the region by including neighboring pixels that meet specific criteria, such as intensity or texture similarity[118,119].

- **Watershed:** This technique treats the image as a topographic surface and segments the image by finding watershed lines that separate regions with different intensity values[120].

- **Clustering:** This method groups pixels based on their similarity in intensity or other features, such as color, texture, or spatial proximity. Examples of clustering algorithms include K-means and hierarchical clustering[121,122].

- **Machine learning-based methods:** These approaches utilize machine learning algorithms, such as random forests, support vector machines, or deep learning models like convolutional neural networkss (CNNs), to segment and classify image regions based on learned features or patterns[123].

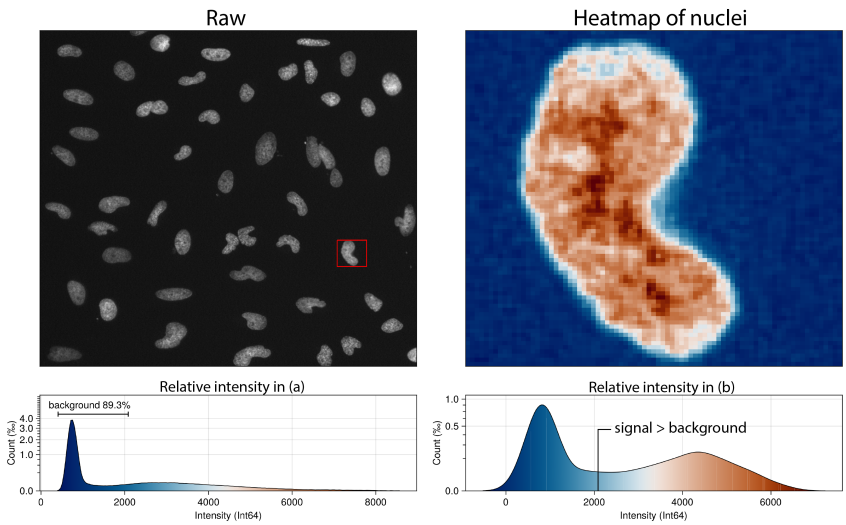

Figure 8: Illustration of the signal distribution of a fluorescence image. The raw image depicts cell nuclei imaged at 20X magnification (NA=0.75; top left). The majority of the content in the image is made out of background (89 %; lower left panel). A heatmap of a nuclei in the image (red square in the raw image) reveals a clear cutoff between background and signal (white color; top right image). The threshold between background in the complete image using Otsu's algorithm is for this nuclei a conservative estimate (bottom right panel).

Example: Image segmentation example

Because theory is best accompanied by real-world examples, I have put together an example of an image-segmentation pipeline using the nuclei example from before. The goal of this short exercise is to compare two segmentation algorithms (**Otsu's** and **sigma-clipped**) as well as to see the full process of loading an image to visualizing data. From an image-analysis perspective, this image is rather simple to segment.

Loading packages and defining functions

To begin, we need to load the necessary packages and define functions that we will use later on. Like mentioned previously, the reason we do this at the start of the code is for readability and consistency.

In Code block 6, we start of by defining three functions, two functions for each thresholding technique and one function to filter out objects depending on their area (number of pixels). In more detail, these functions do:

- **otsu_segmentation:** This function segments the image using Otsu's method, which estimates the optimal threshold value for separating the image into foreground and background classes based on their intensity values. The algorithm works by minimizing the within-class variance or maximizing the between-class variance, making it effective when the image histogram has distinct peaks corresponding to the background and the object.
- **sigma_segment:** This function segments the image using a sigma-clipping approach, which iteratively estimates the background intensity by removing outliers based on the mean and standard deviation of the intensity values in the image. Pixels with intensities that are more than a specified number of standard deviations away from the mean are considered outliers and are excluded from the background estimation process. The function returns the median and standard deviation of the background, which we use to segment the image with (by providing the n_sigma argument).
- **filter_on_size:** This function filters objects in an image based on size. Using countmap from StatsBase, it creates a dictionary with the count of unique pixel values, representing object sizes. An empty output array (**out**) of the same size and type as the input image (**lb_img**) is initialized. The function iterates over non-zero input pixels, providing the value (**v**) and index (**i** and **j**), and looks up the count of that value in the dictionary (**object_sizes**). If the size associated with this value falls within the allowed range (between **minsize** and **maxsize**), it assigns the pixel value to the output array. Finally, the filtered output array is returned.

```
using Images
using SparseArrays
using StatsBase
using SegmentationUtils
using RegionProps

# Functions for segmenting and extracting data
function sigma_segment(image, n_sigma)
    # median and std of background
    m, s = sigma_clipped_stats(image)

    return image .> (m + s * n_sigma)
end

function otsu_segment(image)
    t = otsu_theshold(image)
    return image .> t
end

function filter_on_size(lb_img, minsize, maxsize=Inf)
    # Returns a Dict with the labels as keys, and the
    # values being the count of each key
    object_sizes = countmap(nonzeros(lb_img))

    out = zeros(eltype(lb_img), size(lb_img))

    iterator = Iterators.zip(nonzeros(lb_img), findnz(lb_img)...)
    for (v, i, j) in iterator
        # if the size is outside the provided limits
        if maxsize > object_sizes[v] > minsize
            out[i,j] = v
        end
    end

    return out
end
```

Code block 6: Declaration of the segmentation functions. Here, we declare **sigma_segment** and **sigma_segment**, two functions that segment the image based on different approaches for estimating the background. Lastly, a function that takes a labeled image and filter the objects based on size (the number of pixels).

Applying our segmentation algorithms to the image

Once we have defined our algorithms, the next step is to apply them to our image. As we've seen, the **otsu_segment** algorithm simply takes the image and applies the estimated threshold to the image (see Code block 7). For the **sigma_segment**, we need to provide the number of standard deviations that we estimate our content lies above. In our case, using an **n_sigma** of three, after empirical testing, segments the image well. In our case, both methods are followed by a closing operation, a morphological operation that smoothens object contours and closes small holes or gaps within segmented objects. See Figure 9 for

a comparison between the resulting segmentations using **Otsu's** and **Sigma-clipped**. For a more comprehensive comparison of segmentation algorithms in the use of life-science application, I refer the reader to the following reviews by Anne L. Plant et al.[124], Jisha John et al.[125] and Anne E. Carpenter et al.[126].

```
path_to_image = "images/nuclei.tiff"
img = load(path_to_image)

# Otsu segmentation
seg = otsu_segment(image) |> opening

# Segmentation using histogram clipping
seg = sigma_segment(img, 3) |> opening
```

Code block 7: Loading of our raw image into memory and performing our segmentation algorithms. To remove small objects (noise) and smooth edges, we can perform an opening operation on the segmented image.

Measure object features and data-analysis

Once we have our segmented masks, it is time to analyze the properties of the content they represent. However, before we do anything, we need to be able to distinguish between the objects in the image. Since the segmentation algorithms only return a BitArray (an array containing true or false), we need to convert this into a format where all objects are distinguishable. For this, we can use the **label_components** from the Images library to assign a unique value to each connected component. Once each object has been labeled, a good approach is to already now filter out objects that we know we do not need to measure (like small objects). Here, we can utilize the function we wrote previously to keep objects that are of a certain size (in our case, > 100 pixels).

The benefit of performing image analysis in scientific languages with a big community is that everything does not have to be written from the ground up. We do not need to write handlers and interpreters for how images are loaded, nor write functions for common methods such as estimating Otsu's threshold nor label connected components in an image. The same holds true for measuring object features in an image. Inspired by the regionprops function in Matlab, my colleague Johannes Kumra Ahnlide wrote an equivalent function in Julia, that takes a raw image, a segmented image and a series of values that corresponds to the labels we wish to measure (provided by the **selected** keyword argument. Regionprops extracts intensity-based information (e.g. **mean**, **median**, **minimum** and **maximum** intensity) about each object in the raw image. Additionally, it performs measurements of some basic morphometrical features (e.g. circularity, area, the perimeter) each object and returns a generator that, when collected, results in a series of NamedTuples. This collection can then immediately be used to, for instance, save to a **.csv** file or loaded into a DataFrame

(data-structure for tabular data; provided by the DataFrames library) for further analysis. See Code block 8 for an implementation of this procedure.

Interpreting the data

Interpreting image data can be challenging without understanding the underlying properties and characteristics of the objects being analyzed. In our case, we are working with nuclei that are relatively homogeneous in size, shape, and texture. However, segmentation algorithms may not always be perfect for all scenarios, making it crucial to evaluate and adjust the methods used for obtaining the data. Plotting the data in association with the image data can be a powerful tool for detecting potential issues in the segmentation procedure and enhancing our understanding of the data. In Julia, this would be relatively simple considering we already have access to the **DataFrames** library. By complementing it with the **Colors** and a suitable plotting library of our choice, we can effectively visualize and render the data in a comprehensive manner.

Consider the data from our previous example. In Figure 10, I have plotted a scatter plot of the mean intensity versus the area of each nucleus, color-coding a third feature (circularity) in both the scatter plot and the segmented image. As we can see in the image, two

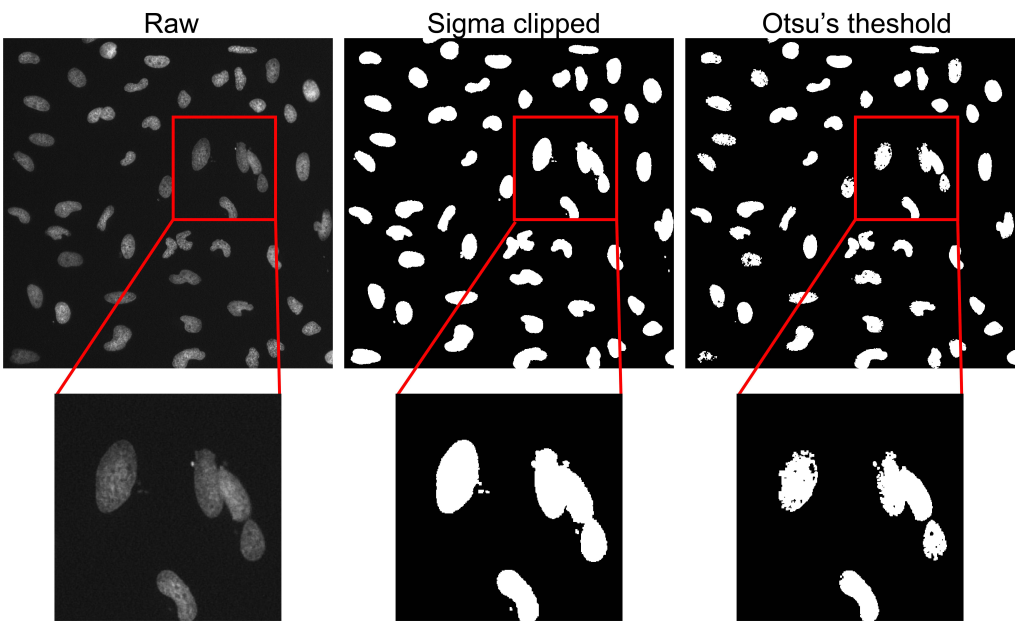

Figure 9: Different techniques for segmenting a image of nuclei. The raw image (left column) when segmented using sigma-clipped (middle column) and otsu's (right column) algorithm for estimating the threshold between background and signal. Sigma-clipped segmentation using three standard deviations performs better overall (but picks up small unwanted fragments) compared to Otsu's, in this case, more aggressive threshold (nuclei are not kept intact).

```
...
minsize = 100
filt_labels = filter_on_size(
    sparse(label_components(img)),
    minsize
)
labels = unique(nonzeros(sparse(filt_labels)))
properties = regionprops(img, filt_labels, selected=labels)

# if we wish to export it
using CSVFiles
save("props.csv", properties)

# or if we wish to continue data-analysis in Julia
using DataFrames
df = DataFrame(properties)

...
```

Code block 8: In order to extract data from the segmented image, we can utilize the RegionProps package (inspired by the regionprops from MATLAB; written by Johannes Kumra Ahnlide). First, we filter out objects that are smaller than what we estimate cell nuclei are. Next, we look up all the unique values in our labeled image. These are just integers referring to the label of each object in the image. Using the regionprops function, we measure properties (e.g. intensity- and shape-based features) of the objects in the raw image. Finally, we can either return a .csv file to export the data, or use a table-like datas tructure (DataFrames) for further analysis in Julia.

objects stand out in terms of circularity. However, the scatter plot did not sufficiently distinguish these improperly associated objects (what appears to be multiple nuclei in close proximity). Only one object appears to be an outlier on the area-axis. Depending on the questions asked and the analysis objectives, these objects could potentially introduce errors and wrongful interpretations down the line. A further investigation (data not shown) revealed that these objects could be distinguished using the perimeter, area, and circularity as a means for separating them from the otherwise homogeneous population of single nuclei. By visualizing the data in conjunction with the original image, we can better identify the sources of potential discrepancies, adjust our segmentation analysis accordingly, and make more informed decisions.

Available software for image analysis

In the context of microscopy and life-sciences, it is unreasonable to expect that every user will possess the necessary coding skills to write the analysis for the images they obtain. Indeed, not everyone shares an enthusiasm for learning programming solely for the purpose of image analysis. Nevertheless, extracting information in an unbiased and robust manner necessitates the implementation of image analysis methodologies. To address the challenge

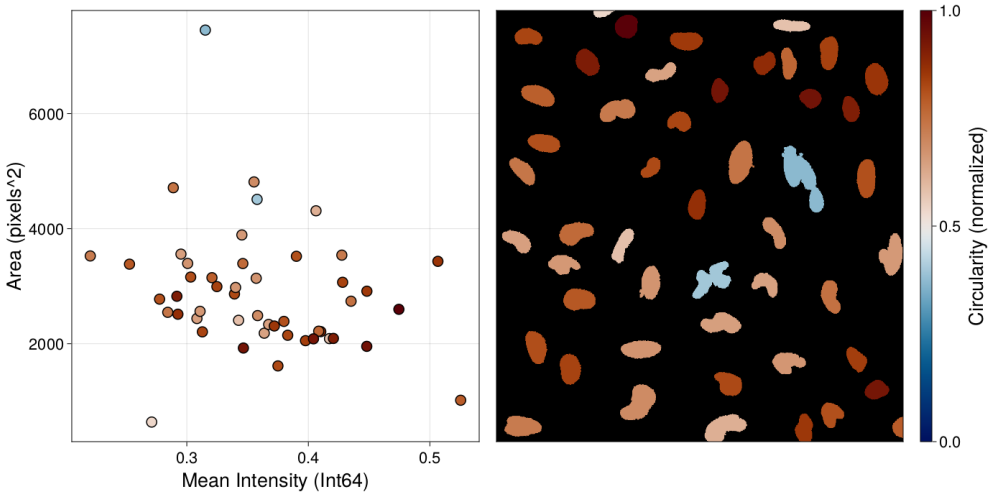

Figure 10: Illustration on two forms of visualizing the data of our nuclei. Left panel displays a scatterplot of the area and the mean intensity of the nuclei in the image. Here, we can see one datapoint that stands out from the rest. Right panel displays the segmented objects in an image. Both plots are color coded according to the circularity.

of limited coding experience, several free, open-source software options with user-friendly graphical user interfaces (GUIs) have been developed. These platforms provide accessible solutions for individuals with varying levels of programming experience.

- **Fiji/ImageJ:** Fiji[127] (standing for Fiji is just ImageJ) is an open-source, Java-based image processing platform that builds upon the core functionality of ImageJ[128]. Developed by the National Institutes of Health (NIH), ImageJ is a versatile image processing and analysis software with applications in life sciences, including microscopy and bioimaging. Both Fiji and ImageJ offer extensive libraries of plugins and macros that can be easily customized to suit specific requirements, enabling tasks such as image filtering, segmentation, and quantification with minimal programming knowledge[129].

- **CellProfiler:** Developed by the Broad Institute, CellProfiler[130,131] is an open-source image analysis software specifically designed for high-throughput, large-scale biological image analysis. With a user-friendly interface, CellProfiler allows biologists and other researchers to create custom pipelines for processing and analyzing large datasets without requiring extensive programming skills. This software is useful for quantifying phenotypic characteristics, such as cell morphology, protein localization, and gene expression patterns.

In addition to Fiji/ImageJ and CellProfiler, other open-source alternatives include:

- **Napari:** Napari is a fast, interactive, and multi-dimensional image viewer for Python, which can handle a wide range of imaging data types. Its plugin-based architecture allows users to extend its functionality, making it a versatile tool for various image analysis tasks[132].
- **Icy:** A flexible, open-source software platform focused on bioimage informatics, providing both easy-to-use tools and advanced image processing algorithms[133].
- **QuPath:** An open-source digital pathology and whole-slide image analysis software, particularly useful for analyzing tissue sections and tissue microarray slides[134].

While the software options mentioned offer valuable tools such as access to powerful pre-built segmentation algorithms and a graphical user interface (GUI), some potential drawbacks and limitations should be considered:

- **Limited customization:** While the available software options offer a wide range of tools and features, there might be specific tasks that require additional customization or scripting. Users with limited programming skills may, even though they revert to a software with user friendly GUI, need to seek help from the software's user community or external resources.
- **Scalability and performance:** Some software platforms may not be optimized for handling very large datasets or high-throughput analysis, which could result in performance bottlenecks or slower processing times. It is essential to ensure that the chosen software can handle the volume of data you intend to process. Additionally, you cannot ignore the foundation that the software was built upon. For instance, consider Fiji and ImageJ. Being built on Java and at a time where the datasets were smaller, Fiji does not stand up in terms of performance in high-throughput situations compared to languages high-performance languages (like Julia) or languages utilizing C for image analysis (like Python).
- **Updating and maintenance:** Open-source software relies on the user community and developers for updates and maintenance. However, this 'drawback' holds true for custom-written code in languages like Python and Julia as well, as most of the image-analysis functionality comes in the form of community built libraries and packages.
- **Lack of specialized tools:** Some research fields may require highly specialized image analysis tools that are not available in the general-purpose software options. In these cases, users may need to explore niche or commercial software solutions, or develop custom tools to meet their specific needs.

Despite these potential drawbacks, the open-source software options available for image analysis can provide powerful solutions for many users in microscopy and life sciences. It is important to note that the challenges faced when using these software platforms are often similar to those encountered when learning to code in languages like Python[135,136], Julia[137], and MATLAB[138] on your own.

The key is to carefully evaluate the software's capabilities and limitations in relation to the specific requirements of your project and to invest time in learning and mastering the chosen platform. With proper evaluation and commitment, researchers can harness the power of both open-source software and custom-written code in languages like Python, Julia, and MATLAB to address their image analysis needs.

In conclusion, the choice between open-source software and custom-written code in programming languages depends on individual preferences, project requirements, and the level of expertise. If the prospect of learning how to code is not exciting, one should steer towards the open-source softwares, but make sure to learn the basics of image analysis. On the other hand, if someone need high customizability and full control, try pick up a language and code! It is essential to understand the benefits and limitations of each option and choose the one that best fits the needs of the researcher and the project.

My story and lessons learned

When I embarked on my PhD-studies, I had little programming knowledge. My first encounter with a programming language was just six months prior when, during my master's degree, I had began coding in Matlab. I was given the opportunity to develop a programmatic solution to finding bacteria and cells interacting with each other, and subsequently image them using a higher magnification and NA objective. Valiantly trying to understand and learn from my supervisors code, the notions of pointers and compilers was still all new to me. After some time, I managed to create an executable file that we could run on the microscopy-computer in order to detect events between bacteria and host-cells. However, it was rather static in the way that no other analysis could be performed, it was slow and I don't remember if the coordinate space for the events was fully transformed to match that of the microscopes.

It wasn't until my colleague Johannes, who's an expert coder, joined the lab that we switched to a, at the time, rather new programming language called Julia. I remember I scoffed at the idea, why would you need another programming language other than Matlab, considering I had a working solution? Nevertheless, he pointed out some problems and I gave in to what felt like starting from scratch again. The syntax was new and foreign, and I remember being mad that indexation was done using square brackets rather than parentheses, something I

find rather silly today considering the very stripped down syntax of Julia. I also remember me having problems and him showing me his solutions in what appeared to be written in another language other than Julia. At the time, I had no idea about concept of syntactic sugar (syntax which make expressions 'sweeter' to the eyes and easier to understand) and how the compiler interprets code, and that knowledge in this domain could produce code that was (for me at the time) best described as wizardry.

Since the beginning I had a somewhat maximalist view on the topic of object-oriented versus functional programming. To me, it simply made sense organizing code into a functional approach where we focus on the transformation of the input to output, rather than trying to model the world around us using objects. This point of view has changed to the point where I on occasion borrow ideas from the object-oriented style, like the considerations in regards to encapsulation and polymorphism during the design of custom structures and types, and how they interface with other functions. However, I think my point of view, coming from the side of functional programming, has helped me focus on what matters most to me: the code describing the transformation of data between states.

As we continued developing new analyses and solving image-analysis problems in Julia, the aim of our project had evolved into the development of a framework where we could load different analysis-pipelines. We focused on abstraction and responsibility, isolating processes which were common for all pipelines (for instance, the conversion of pixel position into the stage-coordinates of the microscope), so that we had a solid infrastructure. This resulted in us hosting services using Julia, transferring the images over HTTP, and the ability for the user to load pre-defined image-analysis pipelines at the start of their experiments. During this journey, I had also began looking into the domain of web-development, experimenting with languages like ELM and Elixir. However, I had little time for this in the scope of our project, and had to postpone that phase of my journey to after my PhD.

Going from a complete novice to where I am today opened my eyes to the realm of computer science, and it sparked an interest beyond my imagination. I learned to be resilient, to not be daunted by the ugly red text that appears in ones face as the code vomits errors. I learned the art of problem-solving and, perhaps most importantly, how to learn. I discovered that no matter how far you've come in your coding journey, around the corner is a intimidating boss (monads) waiting at the end of the dungeon, ready to slap you in the face. This taught me to be humble, to always keep learning, and the embrace new challenges. As I look forward to the future with anticipation, I am eager to see the innovative software projects I will have the opportunity to contribute to, and the diverse scientific fields they will encompass.

Heterogeneity in cell biology

It is not the strongest of the species that survive, but the one most responsive to change.
— Charles Darwin

In this final chapter, we will embark on a journey to uncover the intricacies of cellular heterogeneity, a phenomenon that permeates every aspect of life sciences. With the foundations of microscopy and image analysis behind us, we will now focus on the biological implications of these variations among individual cells. Cellular heterogeneity not only challenges our understanding of the biology at play, but it also has profound implications for disease diagnosis and treatment, as well as the development of targeted therapeutics.

To fully appreciate the scope and impact of cellular heterogeneity, we will begin by examining the various factors that contribute to its existence. From external factors in the cellular environment to the inherent stochasticity of biological processes, we will explore the complex interplay of forces that shape each cell's unique characteristics. Building on this understanding, we will delve into the microscopy techniques that has the potential to revolutionize our ability to study heterogeneity at unprecedented scales and resolutions. Moreover, we will explore the topic of heterogeneity in diseases, how variations at the cellular level impacts our understanding of disease progression, and how this perspective ties into the work of my thesis.

Sources of heterogeneity

The very idea of heterogeneity is not foreign to us: in fact, we all perceive ourselves as unique individuals, that no walking soul on this planet is an exact copy of one another. This holds true for identical twins, where the notion of the individual holds true even in the presence of genetic similarity. Then why, when we shift our focus towards the smallest living organisms, cells, do we anticipate homogeneity in their properties?

During experiments, it can be difficult to assess the source of the variability one observes. It could just as well stem from inaccuracies in the tools we use in order to perform the observation in the first place. To compensate for this, it is common to reconcile towards the appeal of statistics of large numbers to correct for any potential inaccuracies. However, in doing so, the importance of each individual molecule is diminished. Moreover, cellular function often involves a very small number of molecules, like DNA that, in a single or very few copies, gives organisms their unique genetic identity. So, can we truly disregard the significance of individual molecules and the distinctiveness of cellular properties?

In reality, discrete and inherently random biochemical reactions involved in gene expression give rise to significant variability even among genetically identical cells (for review, see[139–143]). Indeed, it has long been known that random fluctuations in the expression of individual genes play a crucial role in shaping cellular heterogeneity, even in cells grown under homogeneous conditions[144–147]. The presence of DNA and its encoded genes in very low numbers[148] means that these fluctuations do not simply average out. Instead, they can lead to discernible differences between otherwise identical cells, highlighting the importance of considering both deterministic and stochastic processes when studying cellular behavior. As we recognize the importance of heterogeneity in cellular behaviour, we must ask ourselves: which factors influences this heterogeneity?

The World Outside: External Factors

External factors can impact cellular heterogeneity through subtle environmental differences. In fact, no matter how much we try and optimize for consistency across experiments, some external factors are out of our control. Variations at the microscopic and nanoscopic scale in terms of nutrient density or the composition of diffusable molecules used for cell-cell signalling are in part stochastic processes[141,149–151]. Additionally, the cellular response upon these changes may vary as well, considering the biochemical networks used by cells to sense the molecular landscape may fluctuate in composition and localization[152–154].

In the context of fluorescence microscopy, cells are immediately under pressure from the very nature of our observations. Light, in virtually any amount, will be damaging to the cells[93–95]. This is a result of high energy electrons failing to be emitted as photons during fluorescent excitation, and instead react to the dissolved oxygen in the environment[92]. The product of this reaction, reactive oxygen species (ROS), is highly damaging to cells left unaccounted for. An increase from the homeostatic levels of ROS, also referred to as oxidative stress, can negatively affect several cellular structures and processes, such as membranes, lipids, proteins, lipoproteins and DNA[97–102]. Albeit systems have been evolved in order to handle oxidative stress[103], an excessive amount caused by fluorescence imaging will nevertheless alter cell state and impact the experimental outcome.

Furthermore, one has to consider the variability in experimental procedures and data acquisition that is introduced by the investigator/lab that performs the experiment. This variability can significantly impact the reproducibility and comparability of microscopy data, as demonstrated in a recent study by Staffan Strömblad et al.[155] that assessed the sources of variability in high-content imaging data of migrating cancer cells across three different laboratories. The authors found that the highest technical variability occurred between laboratories, and to a lesser extent, between individuals conducting the experiments. This variability can hinder the ability to perform high-quality analysis, and our ability to assess cellular heterogeneity overall. However, it is important to note that certain data processing approaches, such as batch effect removal, can help overcome this variability and enable a more reliable analysis of image-based datasets that originate from different laboratories.

The Dance of Chance: Inherent Factors

When discussing the sources of cellular heterogeneity, it's crucial to acknowledge that nearly all adjustments made to achieve a controlled setup are imperfect, despite our best efforts. Some factors, like the external ones, cannot be completely excluded for since we cannot (yet) control the fluctuations and gradients of environmental parameters that affect gene expression at that level. Despite this, a substantial body of work, both from mathematical models and experimental studies, supports the notion that inherent factors, such as stochastic gene expression, influence cellular heterogeneity[139–143,156–158].

Another noteworthy example of processes that influence cellular heterogeneity are those of periodic nature. These periods, or pulses, are observed throughout the animal kingdom, from bacteria[146,159–162] to fungi[163,164] and mammals[165]. This is an important mechanism for temporally controlling cell function. For instance, in mammals, the pulsatile nature of the stress response pathway mediated by p53, which regulates the DNA damage response[166–168], and nuclear factor $\kappa\beta$ (NF-$\kappa\beta$), involved in immune responses[169–171], highlights the significance of periodic mechanisms in temporally controlling cell function. Of course, this in in addition to the more well studied and beautifully orchestrated systems like the cell cycle, calcium dynamics and multicellular phenomena based on coordinated pulsing[163]. The specific state of each cell in terms of these rhythms contributes to the overall phenotype at both the single-cell and population-wide level.

Additionally, cellular age is an influencing factor that contributes to the phenotypic heterogeneity[172]. Cellular age arises from asymmetrical cell division, and has been demonstrated in budding yeast[173,174] as well as in bacterial systems, such as in *Methylobacterium extorquens*[175]. This age-dependent variation can impact cellular behavior, including response to environmental changes, stress tolerance, and survival under various conditions.

The many faces of disease: Heterogeneity in Pathology

As the time has come for me to write this section, I have found my self in a rather peculiar situation. Waking up the other night I could feel something was not alright. And what at first started off as nothing more than a tingling sensation in my throat was met with itching and a runny nose two days later. I am getting sick! Funny is the irony that life plays you. From what I am getting sick, I do not know. However, I do find my self imagining all the hurdles and boundaries that, whatever it is, has to break through in order to cause the symptoms I'm experiencing.

Every surface of our body, except a few exceptions like the urethral tract and our eyes, are covered with bacteria and other organisms. Most of the time, our co-inhabitants are not harmful. In fact, without them, we would not survive. For instance, take the pathogen that is causing a ruckus in my body as of now, triggering the immune-system and the symptoms I'm feeling. Let us presume it is an bacteria, albeit it being unlikely. The first hurdle it had to overcome is to establish a foothold. It has to be able to grow, and the microbiome inhabiting my upper respiratory tract and nasal cavities are not making it easier. For whatever it is infecting me, it has to fight for real-estate and food in order to even have a chance of infecting me.

It is in these environments that cellular heterogeneity reveals its significance. The diverse characteristics of individual cells within a population enable them to adapt and thrive, even in highly competitive environments. This inherent heterogeneity has profound implications for understanding the development and progression of various diseases[176]. Thus, we cannot close our eyes to the characteristics of each individual in the quest to understand and treat diseases more effectively. By appreciating the complexity and adaptability of these cellular populations, we can begin to explore new avenues for treatment and prevention of infections and other health challenges.

In the following sections, we will delve into how cellular heterogeneity influences the manifestation and progression of different types of diseases, with an emphasis on the areas related to my thesis and the works I have been involved in. Specifically, we will examine the roles of heterogeneity in bacterial infections, viral diseases, and cancer.

Bacterial infections and the ability to cope with dynamic environments

As with every organism, the goal is to survive. This means that, in addition to the ability to adapt over time, organisms must also be capable of adapting to sudden and drastic changes in the environment if they are to stand a chance. This ability manifests itself in several ways in bacterial population, and may be the reason why some bacterial infections are easier to clear than other.

This resistance to environmental changes is best exemplified by the survival of *Escherichia coli* in the presence of antibiotics. At first, when a genetically identical population of bacteria are exposed to the antibiotics, the majority of the population is killed. Over time however, the rate of killing decreases, until a small population of cells are left[177]. This survival cannot be hedged on the evolution of antibiotic resistance; the exposure is simply too quick for any such mechanism to take place. Instead, the survival is hedged on a small population of slow-growing bacteria, known as persistors[178,179], and was first reported for almost 80 years ago in the case of staphylococcal infections treated with penicillin[180]. This slow growth grants tolerance to the antibiotics and can result from various external or internal factors, such as nutrient shifts[181] or the expression of certain virulence genes[182]. For instance, in *E. coli*, the formation of persisters has been linked to fluctuations in the expression of an intracellular toxin called HipA. When the level of HipA surpasses a certain threshold, cells become dormant and tolerant to antibiotics[183].

Phenotypic heterogeneity also plays a role in the evasion of host immune systems by certain bacterial species. For example, *Salmonella* can form a subpopulation of bacteria that do not express flagella, allowing them to avoid detection and elimination by eukaryotic defense pathways[184]. This strategy highlights the adaptive benefits of maintaining a diverse population in response to environmental changes and stresses.

Heterogeneity in *Streptococcus pyogenes* and the M-protein

Streptococcus pyogenes, or group A streptococcus (GAS), is a gram-positive human pathogen responsible for > 600 million infections each year. Most commonly, these infections result in relatively mild disease development (e.g. strep throat or impetigo). However, the infection can also progress into more severe, life threatening manifestations, such as necrotizing fascitis, sepsis and streptococcal toxic shock syndrome, with a mortality rate of 15 to 25 %[185,186]. This wide range of outcomes, especially when the high prevalence of asymptomatic carriers (1 % to 5 %)[187,188] is taken into account, underscores the importance of understanding the cellular heterogeneity when addressing these infections.

The diverse disease outcomes associated with *S. pyogenes* infections can be largely attributed to the M-protein, a dominant feature on the bacterium's surface. This protein is highly variable, so much in fact that the GAS strain classification relies primarily on the M-protein makeup. The M-protein is responsible for a wide range of functions important for the survival of the pathogen, including resistance to phagocytosis and antibacterial activity of histones, as well as adherance and intracellular invasion[189–191]. Historically, the different GAS strains were typed using serotype-specific anti-serum against the M-protein[186], but is now more commonly typed based on sequencing in hypervariable region in the *emm* gene encoding for the protein[185,192], with over 250 emm types having been identified at the time of writing[186,193,194].

In many ways, the tight co-evolution of the bacterium and the host and its immune system can be seen as a clash between two, highly complex, heterogeneous system. On one hand, you have the immune-system, its different cell types and highly variable antibodies trying to recognize and opsonize the pathogen. On the other side, the functional importance of the M-protein has forced it to not only have hyper-variable regions, but to also adapt other strategies in order to evade the immune system. For instance, some M-protein has the capacity to bind to the Fc-region of antibodies, causing the antigen-recognizing Fab-domains to be pointing away from the bacterium[195,196]. In addition to this camouflage, it has also been shown to display molecular mimicry towards the human myosin protein. This pushes the immune-system to produce antibodies that are potentially cross-reactive, and may be contributing factor to disease progression in rheumatic heart diseases[186,197].

In conclusion, the cellular heterogeneity in *S. pyogenes* infections, driven largely by the highly variable M-protein (considering the now more predominant *emm* typing), plays a crucial role in the diverse disease outcomes observed. A better understanding of this heterogeneity and the underlying mechanisms is essential for developing improved diagnostics, treatments, and preventive measures to combat the wide spectrum of diseases caused by this pathogen.

At the interface of where two heterogeneous systems collide

The adaptive immune system is highly specialized at combating pathogens and foreign particles that may cause harm to our health. A central part of adaptive immune system is the production of antibodies, or immunoglobulins, large Y-shaped proteins made up of four polypeptide chains (two identical heavy chains and two identical light chains; for a review, see[198,199]). Antibodies are responsible for a wide-range of functions, including opsonization, neutralization and complement activation by binding to antigens on foreign proteins[199]. Considering the enormous range of possibilities when it comes to the available antigens that can be out there, it is remarkable how such a heterogeneous system have emerged in order to protect us. On top of this, during antigen presentation and the maturation of B-cells, the derived antibodies should not only bind to the antigen sufficiently enough to elicit an immune response, but also be incapable of binding to any host-protein in the process.

I was fortunate to be part of a rather exciting project (**Paper 5**; not included in this thesis) with the intent of extracting and producing antibodies from isolated B-cells reactive towards the M-protein of *Streptococcus pyogenes*. This project[200] was spear-headed by Wael Bahnan (a senior scientist at the time) and Pontus Nordenfelt (my supervisor), who made a rather remarkable finding. Among the antibodies they extracted, one antibody stood out functionally, both eliciting an increased activation of NF-$\kappa\beta$, but also in its ability to promote a higher bacterial association in THP-1 cells (a phagocytic cell line) compared to the

other antibodies extracted. Because the antibody in question (referred to as Ab25 in the paper) also showed protective function in mice models, structural epitope characterization was warranted in order to investigate the mode of binding and what set this antibody apart from the rest. Remarkably, what was found was that Ab25 was capable of binding to two distinctly different epitopes on the M-protein, specifically a region that has previously only been associated with Fc-mediated binding[201,196]. This is emphasized by the inability of another antibody, referred to as Ab49, which bound to one of the same epitopes as Ab25, but was unable to elicit an immune response. This lead to the presentation of a new antibody-binding mechanism, where the antibody, albeit having structurally identical Fab (fragment antigen-binding) regions, can bind to the same protein through two structurally different epitopes and only then trigger immune activation.

Learning from the COVID-19 Pandemic

A few months ago, I told my co-worker and friend Johannes that I would like to witness a miracle. Something so crazy that my beliefs would be questioned, the kind of stuff that gets written in stone tablets for surviving generations to remember. Like witnessing a comet hitting the moon or an event of similar magnitude (without the loss of human lives that is, I do want everyone to live happily ever after). It was when I reiterated this story to my supervisor that he said "Did you forget about COVID-19?"

He was right, the COVID pandemic was of that magnitude, causing the world to pause and our priorities to be re-evaluated. The COVID-19 pandemic has been an unprecedented event in modern history, affecting millions of lives, economies and healthcare systems around the globe. While it has brought immense challenges and suffering, it has also provided us with valuable lessons in understanding the importance of cellular heterogeneity, both in terms of the virus' ability to spread, but also in how our body responds and fights off the infection.

The virus responsible for the COVID-19 pandemic, SARS-CoV-2, exploits host cellular functions for its own replication by specifically targeting the angiotensin-converting enzyme 2 (ACE2) receptor on human cells. This interaction facilitates the virus's entry into the cell, where it commandeers the host's cellular machinery to reproduce and generate new viral particles. Early attempts in treating patient suggested the use of convalescent plasma or monoclonal antibodies, in part because of the partial success seen in such strategies in treatments of Respiratory Syncytial Virus[202] and Ebola[203].

The heterogeneity of the virus itself, through genetic mutations and the emergence of different variants, has further complicated the pandemic[204–206]. Some variants have shown increased transmissibility[207], resistance to neutralization by antibodies, or partial escape from vaccine-induced immunity[208–210], which has led to alterations in public health strategies

and vaccine development.

Moreover, the COVID-19 pandemic has emphasized the significance of individual variability in disease progression and outcomes. Patients infected with the virus have exhibited a wide range of symptoms and severity, from asymptomatic carriers to those suffering severe complications or death. This variability can be attributed to factors such as genetic predisposition[211,212], pre-existing health conditions, and differences in individual immune responses[213]. Cellular heterogeneity, both in the host and the virus itself, plays a crucial role in determining the course and outcome of the infection.

One critical aspect of understanding the immune response to the virus is recognizing the variability in individual responses. An individual's immune response can differ significantly[214–216], with some mounting a robust response capable of clearing the virus, while others may experience a more subdued or delayed response that may lead to severe disease[215]. This has implications not only for treatment but also for vaccine development, as vaccines need to elicit a strong and effective immune response to provide protection against the virus.

These lessons from the COVID-19 pandemic underscore the need for continued research on cellular heterogeneity in diseases, as well as the development of tailored treatment strategies based on individual patient characteristics. As we move forward, it is crucial to integrate these insights into our understanding of other diseases and their management, ultimately improving patient outcomes and their quality of life.

Assessing heterogeneity using microscopy

For much of its history, microscopy has not been a suitable means for the study of heterogeneous populations due to the large amount of data necessary to obtain meaningful and statistically relevant results. However, advancements in imaging technology, automation and computation analysis have improved the method for the last 20 years. Today, microscopes can be equipped with larger, faster and more sensitive detectors, allowing us to capture more of our sample in a single image. We have improved upon the hardware, in terms automated stage controllers, filter wheels and imaging modalities. Furthermore, in the domain of computation, the continued development of cheaper and powerful components allow us to perform increasingly sophisticated image and data analysis. This digital revolution of the microscope now allow us to truly tap into its power of spatio-temporal resolution at the scale necessary to study cellular heterogeneity.

Today, this use of microscopy is referred to as high throughput (HTP) microscopy, or high content screening (HCS) when in combination with automated image analysis of multiple parameters. HTP microscopy and HCS allow us to automatically capture, store

and analyze terabytes of data in a single experiment, quantities that would otherwise overwhelm any human operator[150,217–224]. This process of acquiring data provides us with rich phenotypic data about the spatio-temporal properties at the single-cell level, while simultaneously maintaining the scale of the whole population and a large number of observations[150,218–220,225–227]. Furthermore, HTP microscopy and HCS solves some of the problems traditionally associated with microscopy, such as it being a labor-intensive method, hard to reproduce results (owing to the low number of observations) and biased interpretation caused by the human factor[228–232].

The Omics Revolution: Systems Microscopy

The omics revolution has significantly advanced out understanding of biological systems by providing comprehensive and holistic views of molecular components and their interactions[233–235]. Incorporating omics technologies, such as genomics[236–238], transcriptomics[239–241], proteomics[242–246], and metabolomics[247–249], into the field of systems biology has led to a deeper understanding of complex molecular networks governing cellular behavior[250–254]. The omics fields have produced some groundbreaking achievements, such as the development of AlphaFold[244] in proteomics and the completion of the Human Genome Project[238] in genomics. However, while the power of these techniques are immeasurable, especially in combination with one another, they lack the spatial context needed to resolve inter-cellular variation and intracellular spatial variations over time[255–258].

This need for spatial, temporal and inter-cellular information has prompted the integration off advanced microscopy techniques with systems biology approaches, a concept referred to as 'systems microscopy'[256]. Systems microscopy combines HTP microscopy with quantitative multiparametric data acquisition and population data derived from numerous levels of resolution. As the concept of systems microscopy emerged, several challanges were identified, including the standardization of data quality, repeatability, throughput in image analysis, public databases and the bioinformatic infrastructure needed to parse, organize, interrogate and share image-derived quantitative data as well as metadata[256,259,260]. While some of these challenges have been solved since its emergence[130,261–266], much work and cooperation is still needed before we can truly call microscopy an 'omics' technology and harness its power for studying cellular heterogeneity.

Data-driven microscopy

While HCS has become a central method for assessing large quantities of data at the population-wide context, it is still hindered by throughput in high-resolution applications. This had lead to the development of techniques where a data filtering and selection step is integrated into the image analysis pipeline, and the coordinates of points of interests are sent

back to the microscope for imaging in the higher resolutions[223,267–270]. This process, most commonly refereed to as feedback microscopy or smart microscopy, where the microscope can be programmed to respond to image content in real-time.

One of the earliest implementation of a pipeline where segmentation and event detection was incoroporated into the acquisition steps were that of Ellenberg et al[227]. To test their setup, they aimed at targeting and image cells in the two most transient mitotic stages, prophase and anaphase. To do this, they coupled the image acquisition pipeline with automated segmentation, feature extraction and a trained classifier able to detect cells in the correct phases, and subsequently switched to a, what they refer to as, 'complex imaging' mode. In addition to showing the capability of feedback microscopy, the released software accompanying the paper, Micropilot, was shown on multiple microscopy manufacturers hardware, enabling for more vendor-agnostic approach to microscopy.

The outcome of these techniques is the ability to steer the microscope and acquire high-resolution images of cells of interest based on predefined segmentation algorithms or detection algorithms optimized for recognizing events in a population of cells. However, the means for finding the cells of interest are in that of a more static nature, where the segmentation and data filtration algorithms are pre-defined. What this entails is that the population-context is at large disregarded, and the methods are more akin to targeted event detection based on pre-defined criteria. While successful at identifying cells of interest in the context of the experiment, the methodology of feedback microscopy or event-driven microscopy lacks the necessary focus on the population-wide context needed to study cellular heterogeneity.

In contrast to the terms feedback or event-driven microscopy, I am more inclined (although the risk of being biased) to use the term data-driven microscopy (DDM). DDM aims at being a more general approach, where data describing the population is either used directly or taken into consideration when steering the microscope. In many ways, feedback microscopy and event-driven microscopy are steps in the process of DDM, but do not suffice on their own for the study of cellular heterogeneity due to the lack of population-wide context.

The present investigation

Objectives

The objective of this thesis has been to develop a method for studying single cell interactions using microscopy at both the population-wide level as well as high resolution. What started off as a project with the goal of being able to target bacteria interacting with human cells for high magnification, developed into something much broader and generally applicable (**Paper I**), and an exciting spin-off project that I am looking forward to finishing (**Paper II**). In addition to my own work, the experience and expertise I've acquired throughout my thesis has enabled me to help and collaborate with other researchers in their work. These collaborations have taken many forms, from training of people on equipment and experimental design using microscopy (**Paper III**), to the acquisition and analysis of data (**Paper IV**).

Results and comments

Paper I

Data-driven microscopy allows for context-specific acquisition of high-fidelity image data

Background

Techniques such as high throughput (HTP) microscopy and high content screening (HCS) allow researchers to acquire and analyse large datasets at a population-wide level using microscopy[228,230,229,226]. This has several benefits compared to other HTP techniques, such as the spatial information and the ability to resolve information of processes over time. However, most high-resolution applications are typically acquired from manually selected points, resulting in data that lack the population-context potential bias.

Several solutions such as feedback microscopy or intelligent microscopy have been proposed, where automated detection of cells of interested are in-cooperated into the targeting-selection for high-resolution[267,268]. However, there is no generally accepted solution for how feedback-based microscopy could effectively integrate image analysis with hardware control in a way that allow for information from the image analysis to inform and drive hardware control. Additionally, no solutions have been presented that allow for targeted image acquisition based on population parameters, and where individual cells are located in the context of the overall population.

Key results

- We present a general framework, data-driven microscopy (DDM), that allows for high-fidelity single cell data to be acquired based off the population-wide context.
- DDM consists of two imaging strategies, data-independent acquisition (DIA) aiming at population-wide characterization at the single cell level, and data-dependent acquisition (DDA) which aims at the targeted high-fidelity acquisiting of cells of interest within the population.
- DDM is a modular framework where the image analysis pipeline can be developed and loaded separately into the framework, allowing for different applications to be performed.
- As proof of principle, DDM was applied to three diverse experimental settings.
- A plugin in the subject of characterizing transfection efficiency showcase the synergistic relationship of combining a population-wide dataset and high-fidelity dataset.

Using the high-fidelity data, the population-wide estimation could be further enhanced.

- An infection plugin was developed for the identification of interactions between fluorescent bacteria and host-cells. This showcased the capacity of DDM to perform population-wide characterization and high-fidelity imaging of rare events in a live-cell setting.
- A cell migration plugin was developed for the purpose of characterizing live-cell behaviour over time. For this, a cancer cell line (H1299) was characterized, grouped depending on their migration speed, and targeted for high fidelity imaging. This experiment further showcased the ability of DDM to acquire, assess and place high-fidelity live single cell data in a population-wide context.

Discussion

In **Paper I**, we present a general framework, DDM, capable of acquiring and placing high-fidelity single cell data in the context of its population. This has for long been a problem, even in the presence of methods such as HCS, where a few, typically manually, selected points dictate what data we acquire in the high-fidelity domain. With DDM, we have implemented a data-centric approach to image acquisition, where population-wide data describing the population directly dictate and steer the acquisition in higher-fidelity applications. Using DDM, the user is removed from the process of image interpretation. Instead, by defining image segmentation plugins, the user can describe more methodologically how the data extraction should be performed, and interact with the extracted data in order to find cells of interest for automated high-fidelity imaging. This way, the evaluation of the data is performed in a data-centric fashion, and not influenced by visual phenomena that mesmerize and cloud our judgement in the image data.

In contrast to previous work[225], DDM distinguishes itself by providing a general solution for acquiring high-fidelity data based on population-characteristics. Furthermore, the majority of DDM is written in Julia, a programming language that is getting increasingly more attention in the field of bioimaging[137]. This choice differs from other related work, which generally relies on the more popular programming language, Python. Although the decision to use Julia might initially seem like a limitation given the widespread adoption of Python and R in the research community, it is essential to note that Julia is interoperable with both languages. We have ensured that our framework includes an simple interface for incorporating image analysis written in Python. The intent of the paper, however, is not the technical means for achieving data-driven microscopy, but rather to present the idea, and showcase our implementation of it. Hopefully, DDM inspires the bioimaging community to adopt and adapt data-centric methodologies in their pursuit of understanding complex biological systems.

Paper II

Correlative imaging using inherent spatial-geometric relationships of cell nuclei

Background

This paper is a spin-off from **Paper I** where we initially wanted to complement data that we acquired from one microscopy-system with additional data of the same cells of interest but from a second microscopy systems equipped with additional modalities. Typically, imaging the same region in multiple systems, referred to as correlative imaging, relies on the identification of key-markers that one can use to calibrate in-between the different systems. These can either be derived from the introduction of artificial markers (i.e. beads). However, this approach may have an impact on the biology and puts a strain on the experiment setup as it may render some microscopes incompatible with the fluorescent profile of the markers. On the other hand, several algorithms exists (e.g. BRISK[271], SIFT[272], SURF[273] and ASIFT[274]) that derive the key-points from the features within the image-data, such as the texture and intensity profiling. While powerful, these algorithms depend on an overlap in the image data (e.g. the same fluorescenct channel or biological information) or a similar-enough image quality for identifying the key-points, and thus also put a strain on the experiment design.

Key results

- We describe an alternative means for performing correlative imaging by utilizing inherent spatial relationships between cells.
- Simulation data aimed at simulating real-live scenarious such as loss of objects (cells detaching or photobleached) and temporal dynamics (e.g. cell migration) show a robustness in correlative success in up to 3 µm and 10 min time differential between the datasets.
- Validation of the correlative method was performed by live-cell characterization and focal-adhesion characterization on three separate imaging systems (low magnification widefield epifluorescence, SIM and TIRF microscopy)

Discussion

In **Paper II**, we present a means for performing correlative imaging on multiple microscopy systems using inherent information about the specimen. This paper came as a by-product of when we were performing the experiments for **Paper I**. In our pursuit of studying cell migration and perform further investigation into the subcellular realm, we realized the need for other imaging techniques. This lead us to the aim of performing live-cell imaging on one

system, fix and stain the sample for paxillin (a proxy for visualizing focal adhesions), and subsequently image cells of interest systematically on secondary systems (SIM and TIRF). Because the steps of fixation and staining may alter the overall image quality, particularly in differential interference contrast (DIC), we were hesitant on using the traditional tools available for image registration. Additionally, because our purpose was to target cells of interest in our population using DDM, we had access to the spatial information of each cell. By utilizing this information, we came up with a relatively simple solution where the spatial position of each cell and their spatial relationship with their neighboring cells was used to calculate a scale- and rotation-invariant representation. Using this representation, we could calculate the transformation between the two imaging systems, thus allowing for cells captured on the primary system to also be accessible on the secondary SIM and TIRF systems.

Our means for performing correlative microscopy distinguishes itself in the form of not being reliant on the quality of the images captured in the different systems. Furthermore, our simulations, particularly when evaluation the allowance for perturbations in the estimated cell position, hint at the possibility of utilizing spatial information stemming from other fluorescent channels and markers. For instance, instead of using the nuclei for estimating the cell centroid in both systems, if the perturbation-tolerance is high enough, this could allow for the use of another channel or marker, such as the actin or staining of cytosol, to be used instead. This would allow for further flexibility, as some microscope-setups might not be compatible with each other in the sense of sharing the same configuration.

Paper III

Spike-Dependent Opsonization Indicates Both Dose-Dependent Inhibition of Phagocytosis and That Non-Neutralizing Antibodies Can Confer Protection to SARS-CoV-2

Background

No one could have anticipated the arrival of the COVID19 pandemic and its broad societal impact. Considerable efforts were made early on in order to generate neutralizing anti-Spike antibodies towards the SARS-CoV-2 virus in order to hinder viral entry into host cells. However, little effort had been made in characterizing the opsonic capability of neutralizing antibodies against SARS-CoV-2. On the other hand, non-neutralizing antibodies, compromising of the majority of the antibody repertoire generated by B cells, have other immunological functions such as mediating phagocytosis and complement-dependant immune activation. In this context, we were interested in whether or not anti-Spike antibodies might mediate phagocytosis as had previously been seen with influenza[275–277].

Key results

- Out work shows evidence that convalescent patient plasma and monoclocal anti-Spike antibodies induce phagocytosis but with diminishing returns when the antibody concentrations become too high.
- We demonstrate that the activation and inhibition of phagocytosis are independent of neutralization potential.
- We present data form an experimenal animal infection model that show that non-neutralizing antibodies can protect animals from SARS-CoV-2 infection.

Discussion

In **Paper III**, we were interested in exploring the role of non-neutralizing Spike-specific antibodies in the context of COVID-19 immunity. While much of the research efforts have focused on antibodies that neutralize the ACE2-Spike interaction, the potential impact of non-neutralizing antibodies remained largely unexplored. Our research revealed that Spike-specific antibodies can enhance or reduce Spike-bead phagocytosis by monocytes, depending on their concentration. This concentration-dependent modulation phenomenon could help explain the unclear clinical benefits observed with high concentration monoclonal antibody treatment for COVID-19. Furthermore, our findings highlight the potential role of non-neutralizing antibodies in conferring protection against SARS-CoV-2 infection by mediating phagocytosis.

To better understand the early events following Spike-monocyte contact, we used shorter incubation times (30 min) and performed dose-response analysis across varying plasma concentrations. The results were congruent with our in vitro experiments and in vivo animal infection data, suggesting a role for non-neutralizing antibodies in infection management.

My contribution to the paper came in the form of providing help with the imaging and image analysis of the bead-based neutralization assay. This was setup to acquire automatically akin to a HCS procedure. To compensate for intra-well variation, we utilized the median and standard deviation of the background on the aggregate of all images during the segmentation of the cells and beads. If the segmentation would have been done on a per image basis, some images would return falsely labeled background as objects. Considering the segmentation algorithms we use attempt at seperating foreground from background, images lacking cells or beads (thus only having background) would have returned segmented objects. These can to some degree be filtered out depending on the object size, but it is an unnecessary risk.

Our study has important implications for antibody therapy in COVID-19 patients. The concentration-dependent modulation of phagocytosis by anti-Spike antibodies could partially explain the varying clinical benefits seen with monoclonal antibody treatment. Moreover, our results indicate that non-neutralizing antibodies generated by vaccines or natural immunity might still offer protection against mutated variants of the virus, highlighting the importance of considering non-neutralizing antibodies in the development of therapeutic strategies for COVID-19.

Paper IV

Group A streptococci induce stronger M protein-fibronectin interaction when specific human antibodies are bound

Background

Streptococcus pyogenes, or group A streptococcus (GAS), is a widespread human pathogen causing a range of illnesses, some with high mortality rates. Fibronectin (Fn), a glycoprotein of the extracellular matrix (ECM), plays a crucial role in cell interactions with their environment. GAS binds Fn to enhance adhesion to and invasion of host cells through various Fn binding proteins, suggesting an evolutionarily driven process considering the abundance in proteins capable of Fn binding[278]. The GAS M protein, which dominates the bacterium's surface, have shown to posses the ability to bind Fn and thus increase the adhesion or trigger internalization[279,280]. In addition, some M proteins can bind the Fc region of antibodies, allowing them to reverse the orientation of the antibodies and reduce immune activation[196]. In this context, we were interested in the effect the presence of Fn has on the ability of antibodies to opsonize the M protein and trigger phagocytosis.

Key results

- We show that convalescent plasma, as well as certain monoclonal antibodies against the M-protein increase the Fn-binding affinity to the M1 GAS in an M protein-dependant manner.
- The enhanced Fn binding capability in the presence of monoclonal anti-M antibodies reduced phagocytosis in a monocyte cell line.
- The increased Fn binding of the M protein is dependant on antibody flexibility as well as intact Fc domains.
- We show that this phenomena is present in multiple different M types.

Discussion

In **Paper IV**, we investigated the interaction between group A streptococcus (GAS), a highly adapted, human-specific pathogen, and the immune system. GAS is known to manipulate the immune system through various mechanisms, and its M protein is a primary target due to its spatial configuration and dominance on the bacterial surface. Our study focused on the ability of GAS to bind to fibronectin (Fn), a high molecular weight glycoprotein of the extracellular matrix, which the pathogen can utilize to bind to host cells and evade the immune system.

We found that human antibodies (Abs) can induce increased Fn-binding affinity in M proteins, likely by enhancing the weak A-B domain binding. This enhanced Fn binding leads to a reduction in Ab-mediated phagocytosis, indicating that this constitutes a GAS immune escape mechanism. Our results showed that the Fc domain of Abs is necessary to trigger this phenomenon, and that Ab flexibility may also play a key role. Furthermore, we observed that our Abs could enhance Fn binding in 3 out of 5 emm type strains tested, belonging to different clades, suggesting that this may be a more generalizable phenomenon.

My contribution for the paper came in the form of training for and support for the acquisition of data using the SIM microscope. In addition, during the review-process of the paper, there had been an observation that Fn and the antibodies were co-localized on the surface of *S. pyogenes* upon inspection of the data acquired using the SIM microscope. This lead to me testing that hypothesis. The resulting data showed that, when using the complete dataset that was available to us, we could not with certainty prove that hypothesis. Based on this, we decided to remove the SIM analysis from the paper

Overall, our findings highlight a novel synergistic interplay between GAS and host proteins, which ultimately benefits the bacterium, and underscores the complex strategies employed by GAS to evade the immune system.

Future perspective

I'd like to dedicate this last section to explore what the future might hold. Now I cannot speak for how accurate this will be, for I am sure there will be surprises and unforeseen developments along the way. Instead, see this as an attempt at tying all the different technologies together and envisioning how they might synergistically contribute to our understanding of biological systems in the years to come.

As I look back on the incredible work that has been done before me (and enabled me to do what I've done), I cannot escape the feeling of being inspired. Never before have we been able to see deeper, far into the realm of the nanoscopic, allowing us to visualize processes that were once beyond the reach of our imagination. We have seen strides in automation and image analysis where we now are at a point of being capable of collecting and analyzing multi-parametric data of hundreds of thousands of cells simultaneously. Looking forward, I envision a future where these two, what have been rather parallel tracks of development, ultimately merge to where the power of high-throughput is combined with high resolution imaging. This is not without challenges however. Systems need to be in place in order to maximize for the acquisition of data rich in information. We need better tools and software for visualizing and interpreting the data. Here, I see a big space for incorporated machine learning tools, where the user and the machine learning algorithms work in synergy to

extract rich and novel insights from the multi-parametric space. Possible, this could even be guided by the software, where the machine learning algorithms see where information is lacking, and predicts and suggests (based off the metadata and what it has learned) what experiments that should be performed in order to test the newly generated hypothesis.

We have already seen extensive work in the field of machine learning and artificial intelligence (AI). Now, I am rather reluctant to the use of the word 'intelligent', as I do not find these tools particularly clever at this stage, and I am sceptical of the potential of generating true intelligent systems. However, that does not mean that these systems cannot mimic intelligence, and I do see the potential future of where these tools can act as a highly-specialized colleague that can work along side you. Just over twenty years ago, a computer had just beaten a world champion in chess, but the idea of a computer beating a human at Go seemed like a far-fetched idea given the substantial increase in complexity (10^{40}to10^{50} estimated number of legal moves in chess[281] compared to $\sim 10^{170}$ in go[282]). Back then, the idea that a machine could predict, with high accuracy, the folding of proteins seemed like a distant future. This reality was shattered recently with the introduction of AlphaGo[283] and AlphaFold[244]. It appears like the future is here sooner rather than later.

A key aspect that will greatly benefit from the integration of these technologies is our understanding of cellular heterogeneity. The convergence of high-throughput data acquisition, high-resolution imaging and the omics technologies will enable us to examine the plethora of phenotypes that exists within cell populations, revealing the underlying molecular signatures and spatial organization that gives rise to their diversity. Here, machine learning algorithms have a the potential to aid us in deciphering the complex interactions and relationships between various cell types, providing valuable insights into the role that cellular heterogeneity plays in development, tissue function, and disease progression. As we continue to refine and optimize these technologies, our capacity to investigate and characterize the nuances of cellular heterogeneity will grow, informing the development of novel diagnostic and therapeutic strategies tailored to the unique molecular profiles of individual cells and their microenvironments.

However, we cannot dismiss the risks that comes with using these tools either. These are all prediction models after all, which is why it is up to us to use them responsibly and interpret their results in the correct context. They might be a great companion for suggesting what is going to happen, but it is up to us to ask the questions and test the hypothesis. If we are going to artificially predict which molecule is able to bind to and alter the function of protein X, or generate an antibody with good opsonization and phagocytic properties against protein Y (which I do believe lies in the future), we have to do so responsibly. For instance, it is with reason that the body has so many checks and balances in the prevention of auto-reactive antibodies, and even this system fails at times. For this purpose, extensive validation has to be in order for in silico predicted antibodies to be used clinically. Furthermore, one could easily utilize the same tool used for predicting potential advant-

ageous molecules to develop harmful agents, a nightmarish scenario that have already been witnessed by researchers in the pharmaceutical field. In a comment published in Nature Machine Intelligence[284], researchers discussed a case where their model, which normally strives for low toxicity, were switched to optimize for toxicity. This resulted in the generation of over 40 000 molecules that were potentially harmful, among them known chemical warfare agents. If we do not have guidelines and restrictions in place for the use of these tools, their incredible power might be detrimental for our society.

Another important aspect that I believe will have a profound impact on our ability to study complex cellular systems is the proper annotation of the data. Normally when we analyze and interpret microscopy-based data, it is up to us researchers to evaluate the data in its correct context (e.g. the cell line, cell media, which protein is labeled for fluorescence and the like). If this context is, in addition to the metadata that is typically acquired by the microscope), transferred to the database, the richness of the data is further elevated. This would enable better aggregation of the data acquired in different experiments, microscopes and users, and let us get deeper insights from the data. Furthermore, cells do not live in isolation. Instead, we are built by a complex and intricate system of multiple different cell types. If we are to understand how these systems work, we need to enable for research to be done where the complexity of these systems are reflected in the experiments. This calls for an increased ability to study cellular heterogeneity, and I believe microscopy is the perfect tool for the job given the right environment and proper precautions.

As we embrace the convergence of omics technologies, microscopy, image analysis, and machine learning, we stand on the precipice of a new era in the study of biological systems. By overcoming the challenges that lie ahead, we can unlock the full potential of these technologies, leading to breakthroughs that will not only deepen our understanding of life, but also transform the way we approach healthcare, drug development, and personalized medicine.

References

1. David E. Wolf. The Optics of Microscope Image Formation. *Methods in Cell Biology*, 81:11–42, 2007.
2. H. Ernst Keller. Proper Alignment of the Microscope. *Methods in Cell Biology*, 81:43–53, 2007.
3. Saumya Saurabh, Suvrajit Maji, and Marcel P. Bruchez. Evaluation of sCMOS cameras for detection and localization of single Cy5 molecules. *Optics Express*, 20(7):7338–7349, 2012.
4. Monya Baker. Faster frames, clearer pictures. *Nature Methods*, 8(12):1005–1009, 2011.
5. Butch Moomaw. Camera technologies for low light imaging: overview and relative advantages. *Methods in cell biology*, 114:243–83, 2013.
6. James Pawley. More than you ever really wanted to know about charge-coupled devices. 01 2006.
7. Kenneth R Spring. Cameras for digital microscopy. *Methods in cell biology*, 114:163–78, 2013.
8. Talley J Lambert and Jennifer C Waters. Assessing camera performance for quantitative microscopy. *Methods in cell biology*, 123:35–53, 1 2014.
9. R.S. Aikens, D.A. Agard, and J.W. Sedat. Chapter 16 Solid-State Imagers for Microscopy. *Methods in Cell Biology*, 29:291–313, 1988.
10. James B Pawley. *Handbook of Biological Confocal Microscopy*. Springer Science & Business Media, New York, NY, 3rd edition, 2006.
11. R.W. Engstrom and RCA Corporation. *Photomultiplier Handbook*. RCA Corporation, 1980.
12. R.J. McIntyre. Multiplication noise in uniform avalanche diodes. *IEEE Transactions on Electron Devices*, ED-13(1):164–168, 1966.
13. Ivan Rasnik, Todd French, Ken Jacobson, and Keith Berland. Electronic cameras for low-light microscopy. *Methods in cell biology*, 114:211–41, 2013.
14. C. J. R. Sheppard, Min Gu, and Maitreyee Roy. Signal‐to‐noise ratio in confocal microscope systems. *Journal of Microscopy*, 168(3):209–218, 1992.
15. W. Schottky. Über spontane Stromschwankungen in verschiedenen Elektrizitätsleitern. *Annalen der Physik*, 362(23):541–567, 1918.
16. Abbas El Gamal, Boyd A. Fowler, Hao Min, and Xinqiao Liu. Modeling and estimation of FPN components in CMOS image sensors. *Solid State Sensor Arrays: Development and Applications II*, pages 168–177, 1998.
17. Martijn F. Snoeij, Albert J. P. Theuwissen, Kofi A. A. Makinwa, and Johan H. Huijsing. A CMOS Imager with Column-Level ADC Using Dynamic Column Fixed-Pattern Noise Reduction. *IEEE Journal of Solid-State Circuits*, 41(12):3007–3015, 2006.
18. C.E. Shannon. Communication in the Presence of Noise. *Proceedings of the IRE*, 37(1):10–21, 1949.
19. E. Abbe. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9(1):413–468, 1873.
20. Jan Hinsch. Mating Cameras to Microscopes. *Methods in Cell Biology*, 72:57–63, 2003.
21. W. Gray (Jay) Jerome. Practical Guide to Choosing a Microscope Camera. *Microscopy Today*, 25(5):24–29, 2017.
22. H. Gest. The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society. *Notes and Records of the Royal Society of London*, 58(2):187–201, 2004.
23. Patrizia Fughelli, Andrea Stella, and Antonio V. Sterpetti. Marcello Malpighi (1628–1694). *Circulation Research*, 124(10):1430–1432, 2019.
24. August Köhler. Ein neues beleuchtungsverfahren für mikrophotographische zwecke. *Zeitschrift für wissenschaftliche Mikroskopie und für Mikroskopische Technik*, 10(4):433–440, 1893.
25. Nobel Prize Outreach AB. Frits zernike – biographical, 2023. Accessed 24 Apr. 2023.
26. F. Zernike. Phase contrast, a new method for the microscopic observation of transparent objects. *Physica*, 9(7):686–698, 1942.
27. E.D. Salmon and Phong Tran. High‐Resolution Video‐Enhanced Differential Interference Contrast Light Microscopy. *Methods in Cell Biology*, 81(J. Cell Biol.1011985):335–364, 2007.
28. David E. Wolf. Fundamentals of Fluorescence and Fluorescence Microscopy. *Methods in Cell Biology*, 81:63–91, 2007.
29. Yu‐li Wang. Computational Restoration of Fluorescence Images: Noise Reduction, Deconvolution, and Pattern Recognition. *Methods in Cell Biology*, 81:435–445, 12 2007.
30. S W Hell and J Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters*, 19(11):780–2, 1994.
31. Nobel Prize Outreach AB. Press release, 2023. Accessed 24 Apr. 2023.
32. 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, 2006.
33. Michael J Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–796, 2006.
34. M. G. L. Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy*, 198(2):82–87, 2000.
35. Lothar Schermelleh, Alexia Ferrand, Thomas Huser, Christian Eggeling, Markus Sauer, Oliver Biehlmaier, and Gregor P. C. Drummen. Super-resolution microscopy demystified. *Nature Cell Biology*, 21(1):72–84, 2019.
36. Yicong Wu and Hari Shroff. Faster, sharper, and deeper: structured illumination microscopy for biological imaging. *Nature Methods*, 15(12):1011–1019, 2018.
37. Rainer Heintzmann and Thomas Huser. Super-Resolution Structured Illumination Microscopy. *Chemical Reviews*, 117(23):13890–13908, 2017.
38. Miriam S. Lucas, Maja Günthert, Philippe Gasser, Falk Lucas, and Roger Wepf. Bridging Microscopes 3D Correlative Light and Scanning Electron Microscopy of Complex Biological Structures. *Methods in Cell Biology*, 111:325–356, 2012.
39. Minoo Razi and Sharon A. Tooze. Correlative Light and Electron Microscopy. *Methods in Enzymology*, 452:261–275, 2009.
40. F. J. Timmermans and C. Otto. Contributed Review: Review of integrated correlative light and electron microscopy. *Review of Scientific Instruments*, 86(1), 01 2015. 011501.
41. Chad M. Hobson and Jesse S. Aaron. Combining multiple fluorescence imaging techniques in biology: when one microscope is not enough. *Molecular Biology of the Cell*, 33(6):tp1, 2022.
42. Jeroen Vangindertael, Isabel Beets, Susana Rocha, Peter Dedecker, Liliane Schoofs, Karen Vanhoorelbeke, Karen Vanhoorelbeeke, Johan Hofkens, and Hideaki Mizuno. Super-resolution mapping of glutamate receptors in C. elegans by confocal correlated PALM. *Scientific Reports*, 5(1):13532, 09 2015. Same system, Confocal microscopy and PALM.
43. László Barna, Barna Dudok, Vivien Miczán, András Horváth, Zsófia I László, and István Katona. Correlated confocal and super-resolution imaging by VividSTORM. *Nature Protocols*, 11(1):163–183, 01 2016. Same system, Confocal and STORM. Refer to using fiducial markers for

seperate systems for image registration.
44. Barna Dudok, László Barna, Marco Ledri, Szilárd I Szabó, Eszter Szabadits, Balázs Pintér, Stephen G Woodhams, Christopher M Henstridge, Gyula Y Balla, Rita Nyilas, Csaba Varga, Sang-Hun Lee, Máté Matolcsi, Judit Cervenak, Imre Kacskovics, Masahiko Watanabe, Claudia Sagheddu, Miriam Melis, Marco Pistis, Ivan Soltesz, and István Katona. Cell-specific STORM super-resolution imaging reveals nanoscale organization of cannabinoid signaling. *Nature Neuroscience*, 18(1):75–86, 01 2015. Same system, Confocal and STORM.
45. Štefan Bálint, Ione Verdeny Vilanova, Ángel Sandoval Álvarez, and Melike Lakadamyali. Correlative live-cell and superresolution microscopy reveals cargo transport dynamics at microtubule intersections. *Proceedings of the National Academy of Sciences*, 110(9):3375–3380, 2013. Same system, live cell widefield and STORM.
46. Viola Mönkemöller, Cristina Øie, Wolfgang Hübner, Thomas Huser, and Peter McCourt. Multimodal super-resolution optical microscopy visualizes the close connection between membrane and the cytoskeleton in liver sinusoidal endothelial cell fenestrations. *Scientific Reports*, 5(1):16279, 2015. Same system, 3D SIM and STORM.
47. Johnny Tam, Guillaume Alan Cordier, Štefan Bálint, Ángel Sandoval Álvarez, Joseph Steven Borbely, and Melike Lakadamyali. A microfluidic platform for correlative live-cell and super-resolution microscopy. *PloS one*, 9(12):e115512, 2014. Same system, Live cell imaging and STORM. Microfluidics to fix and stain, controller through software.
48. O Burri, T Laroche, R Guiet, and A Seitz. Correlative SIM-STORM Microscopy. *Methods in molecular biology (Clifton, N.J.)*, 1663:95–103, 2017. Same system.
49. David J. Crossman, Yufeng Hou, Isuru Jayasinghe, David Baddeley, and Christian Soeller. Combining confocal and single molecule localisation microscopy: A correlative approach to multi-scale tissue imaging. *Methods*, 88:98–108, 2015. Different system, manual alignment. Correlative microscopy though manual imaging and image registration to align. Used a "minimap" for orientation and image region selection.
50. Benjamin Flottmann, Manuel Gunkel, Tautvydas Lisauskas, Mike Heilemann, Vytaute Starkuviene, Jrgen Reymann, and Holger Erfle. Correlative light microscopy for high-content screening. *BioTechniques*, 55(5):243–252, 2013. Different systems.
51. Manuel Gunkel, Benjamin Flottmann, Mike Heilemann, Jürgen Reymann, and Holger Erfle. Integrated and correlative high-throughput and super-resolution microscopy. *Histochemistry and Cell Biology*, 141(6):597–603, 2014. Different systems, using SIFT to align reference images to the complete dataset and performing dSTORM.
52. Zheng Gong, Brandon K Chen, Jun Liu, Chao Zhou, Dave Anchel, Xiao Li, Ji Ge, David P Bazett-Jones, and Yu Sun. Fluorescence and SEM correlative microscopy for nanomanipulation of subcellular structures. *Light: Science & Applications*, 3(11):e224–e224, 11 2014. Different systems, using ASIFT to align the images.
53. Osamu Shimomura, Frank H. Johnson, and Yo Saiga. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, aequorea. *Journal of Cellular and Comparative Physiology*, 59(3):223–239, 1962.
54. O. SHIMOMURA. The discovery of aequorin and green fluorescent protein. *Journal of Microscopy*, 217(1):3–15, 2005.
55. Douglas C. Prasher, Virginia K. Eckenrode, William W. Ward, Frank G. Prendergast, and Milton J. Cormier. Primary structure of the aequorea victoria green-fluorescent protein. *Gene*, 111(2):229–233, 1992.
56. Martin Chalfie, Yuan Tu, Ghia Euskirchen, William W. Ward, and Douglas C. Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805, 1994.
57. Roger Heim, Andrew B. Cubitt, and Roger Y. Tsien. Improved green fluorescence. *Nature*, 373(6516):663–664, 1995.
58. Roger Y. Tsien. THE GREEN FLUORESCENT PROTEIN. *Biochemistry*, 67(1):509–544, 1998.
59. R Heim, D C Prasher, and R Y Tsien. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences*, 91(26):12501–12504, 1994.
60. Brendan P. Cormack, Raphael H. Valdivia, and Stanley Falkow. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, 173(1):33–38, 12 1996.
61. Fan Yang, Larry G. Moss, and George N. Phillips. The molecular structure of green fluorescent protein. *Nature Biotechnology*, 14(10):1246–1251, 10 1996.
62. David A. Zacharias, Jonathan D. Violin, Alexandra C. Newton, and Roger Y. Tsien. Partitioning of Lipid-Modified Monomeric GFPs into Membrane Microdomains of Live Cells. *Science*, 296(5569):913–916, 2002.
63. David von Stetten, Marjolaine Noirclerc-Savoye, Joachim Goedhart, Theodorus W. J. Gadella, and Antoine Royant. Structure of a fluorescent protein from Aequorea victoria bearing the obligate-monomer mutation A206K. *Acta Crystallographica Section F*, 68(8):878–882, 2012.
64. Jean-Denis Pédelacq, Stéphanie Cabantous, Timothy Tran, Thomas C Terwilliger, and Geoffrey S Waldo. Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology*, 24(1):79–88, 2006.
65. Mikhail V. Matz, Arkady F. Fradkov, Yulii A. Labas, Aleksandr P. Savitsky, Andrey G. Zaraisky, Mikhail L. Markelov, and Sergey A. Lukyanov. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature Biotechnology*, 17(10):969–973, 1999.
66. Geoffrey S. Baird, David A. Zacharias, and Roger Y. Tsien. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proceedings of the National Academy of Sciences*, 97(22):11984–11989, 2000.
67. Mark A. Wall, Michael Socolich, and Rama Ranganathan. The structural basis for red fluorescence in the tetrameric GFP homolog DsRed. *Nature Structural Biology*, 7(12):1133–1138, 2000.
68. Daniel Yarbrough, Rebekka M. Wachter, Karen Kallio, Mikhail V. Matz, and S. James Remington. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. *Proceedings of the National Academy of Sciences*, 98(2):462–467, 2001.
69. Alexey V Terskikh, Arkady F Fradkov, Andrey G Zaraisky, Andrey V Kajava, and Brigitte Angres. Analysis of DsRed Mutants. *Journal of Biological Chemistry*, 277(10):7633–7636, 3 2002.
70. B.J Bevis and B.S Glick. Erratum: Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed). *Nature Biotechnology*, 20(11):1159–1159, 2002.
71. Vladislav V. Verkhusha, Hideo Otsuna, Takeshi Awasaki, Hiroki Oda, Shoichiro Tsukita, and Kei Ito. An Enhanced Mutant of Red Fluorescent Protein DsRed for Double Labeling and Developmental Timer of Neural Fiber Bundle Formation*. *Journal of Biological Chemistry*, 276(32):29621–29624, 8 2001.
72. Roger Y. Tsien. Building and breeding molecules to spy on cells and tumors. *FEBS Letters*, 579(4):927–932, 2005.
73. Lei Wang, W. Coyt Jackson, Paul A. Steinbach, and Roger Y. Tsien. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proceedings of the National Academy of Sciences*, 101(48):16745–16749, 2004.
74. Nathan C Shaner, Robert E Campbell, Paul A Steinbach, Ben N G Giepmans, Amy E Palmer, and Roger Y Tsien. Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology*, 22(12):1567–1572, 2004.
75. Daphne S Bindels, Lindsay Haarbosch, Laura van Weeren, Marten Postma, Katrin E Wiese, Marieke Mastop, Sylvain Aumonier, Guillaume

Gotthard, Antoine Royant, Mark A Hink, and Theodorus W J Gadella. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nature Methods*, 14(1):53–56, 2017.
76. Stéphanie Cabantous, Thomas C Terwilliger, and Geoffrey S Waldo. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nature Biotechnology*, 23(1):102–107, 2005.
77. Stéphanie Cabantous, Hau B. Nguyen, Jean-Denis Pedelacq, Faten Koraïchi, Anu Chaudhary, Kumkum Ganguly, Meghan A. Lockard, Gilles Favre, Thomas C. Terwilliger, and Geoffrey S. Waldo. A New Protein-Protein Interaction Sensor Based on Tripartite Split-GFP Association. *Scientific Reports*, 3(1):2854, 10 2013.
78. Chiara Foglieni, Stéphanie Papin, Agnese Salvadè, Tariq Afroz, Sandra Pinton, Giona Pedrioli, Giorgio Ulrich, Magdalini Polymenidou, and Paolo Paganetti. Split GFP technologies to structurally characterize and quantify functional biomolecular interactions of FTD-related proteins. *Scientific Reports*, 7(1):14013, 2017.
79. Matthew G. Romei and Steven G. Boxer. Split Green Fluorescent Proteins: Scope, Limitations, and Outlook. *Annual review of biophysics*, 48(1):1–26, 5 2013.
80. Manuel D. Leonetti, Sayaka Sekine, Daichi Kamiyama, Jonathan S. Weissman, and Bo Huang. A scalable strategy for high-throughput GFP tagging of endogenous human proteins. *Proceedings of the National Academy of Sciences*, 113(25):E3501–E3508, 2016.
81. Daichi Kamiyama, Sayaka Sekine, Benjamin Barsi-Rhyne, Jeffrey Hu, Baohui Chen, Luke A. Gilbert, Hiroaki Ishikawa, Manuel D. Leonetti, Wallace F. Marshall, Jonathan S. Weissman, and Bo Huang. Versatile protein tagging in cells with split fluorescent protein. *Nature Communications*, 7(1):11046, 2016.
82. Tito Calì and Marisa Brini. Quantification of organelle contact sites by split-GFP-based contact site sensors (SPLICS) in living cells. *Nature Protocols*, 16(11):5287–5308, 11 2021.
83. Yuriko Kakimoto, Shinya Tashiro, Rieko Kojima, Yuki Morozumi, Toshiya Endo, and Yasushi Tamura. Visualizing multiple inter-organelle contact sites using the organelle-targeted split-GFP system. *Scientific Reports*, 8(1):6175, 4 2018.
84. Gloria I. Viboud and James B. Bliska. YERSINIA OUTER PROTEINS: Role in Modulation of Host Cell Signaling Responses and Pathogenesis. *Annual Review of Microbiology*, 59(1):69–89, 2005.
85. Julia Radics, Lisa Königsmaier, and Thomas C Marlovits. Structure of a pathogenic type 3 secretion system in action. *Nature Structural & Molecular Biology*, 21(1):82–87, 2014.
86. Antoine Loquet, Nikolaos G. Sgourakis, Rashmi Gupta, Karin Giller, Dietmar Riedel, Christian Goosmann, Christian Griesinger, Michael Kolbe, David Baker, Stefan Becker, and Adam Lange. Atomic model of the type III secretion system needle. *Nature*, 486(7402):276–279, 2012.
87. Jennifer E. Trosky, Amy D. B. Liverman, and Kim Orth. Yersinia outer proteins: Yops. *Cellular Microbiology*, 10(3):557–565, 2008.
88. Cedric N. Berger, Valerie F. Crepin, Kobi Baruch, Aurelie Mousnier, Ilan Rosenshine, and Gad Frankel. EspZ of Enteropathogenic and Enterohemorrhagic Escherichia coli Regulates Type III Secretion System Protein Translocation. *mBio*, 3(5):e00317–12, 2012.
89. Diana Munera, Valerie F Crepin, Olivier Marches, and Gad Frankel. N-Terminal Type III Secretion Signal of Enteropathogenic Escherichia coli Translocator Proteins. *Journal of Bacteriology*, 192(13):3534–3539, 2010.
90. Xavier Charpentier and Eric Oswald. Identification of the Secretion and Translocation Domain of the Enteropathogenic and Enterohemorrhagic Escherichia coli Effector Cif, Using TEM-1 β-Lactamase as a New Fluorescence-Based Reporter. *Journal of Bacteriology*, 186(16):5486–5495, 2004.
91. Jennifer C. Waters. Live-Cell Fluorescence Imaging. *Methods in Cell Biology*, 81:115–140, 2007.
92. Richard Cole. Live-cell imaging. *Cell Adhesion & Migration*, 8(5):452–459, 2014.
93. Michael A. Taylor, Jiri Janousek, Vincent Daria, Joachim Knittel, Boris Hage, Hans-A. Bachor, and Warwick P. Bowen. Biological measurement beyond the quantum limit. *Nature Photonics*, 7(3):229–233, 2013.
94. S. DEWITT and M.B. HALLETT. Optical complexities of living cytoplasm – implications for live cell imaging and photo-micromanipulation techniques. *Journal of Microscopy*, 241(3):221–224, 2011.
95. R.A. HOEBE, H.T.M. VAN DER VOORT, J. STAP, C.J.F. VAN NOORDEN, and E.M.M. MANDERS. Quantitative determination of the reduction of phototoxicity and photobleaching by controlled light exposure microscopy. *Journal of Microscopy*, 231(1):9–20, 2008.
96. Rajesh P. Rastogi, Richa, Ashok Kumar, Madhu B. Tyagi, and Rajeshwar P. Sinha. Molecular Mechanisms of Ultraviolet Radiation-Induced DNA Damage and Repair. *Journal of Nucleic Acids*, 2010:592980, 2010.
97. JOYE K. WILLCOX, SARAH L. ASH, and GEORGE L. CATIGNANI. Antioxidants and Prevention of Chronic Disease. *Critical Reviews in Food Science and Nutrition*, 44(4):275–295, 2004.
98. Pál Pacher, Joseph S. Beckman, and Lucas Liaudet. Nitric Oxide and Peroxynitrite in Health and Disease. *Physiological Reviews*, 87(1):315–424, 2007.
99. Marcelo Genestra. Oxyl radicals, redox-sensitive signalling cascades and antioxidants. *Cellular Signalling*, 19(9):1807–1819, 2007.
100. B Halliwell. Biochemistry of oxidative stress. *Biochemical Society Transactions*, 35(5):1147–1150, 2007.
101. I S Young and J V Woodside. Antioxidants in health and disease. *Journal of Clinical Pathology*, 54(3):176, 2001.
102. Wulf Dröge. Free Radicals in the Physiological Control of Cell Function. *Physiological Reviews*, 82(1):47–95, 2002.
103. Helmut Sies, Carsten Berndt, and Dean P. Jones. Oxidative Stress. *Annual Review of Biochemistry*, 86(1):1–34, 2016.
104. Felipe Mora-Bermúdez and Jan Ellenberg. Measuring structural dynamics of chromosomes in living cells by fluorescence microscopy. *Methods*, 41(2):158–167, 2007.
105. Elizabeth A. Specht, Esther Braselmann, and Amy E. Palmer. A Critical and Comparative Review of Fluorescent Tools for Live Cell Imaging. *Annual Review of Physiology*, 79(1):93–117, 2016.
106. Le Cong, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu, Xuebing Wu, Wenyan Jiang, Luciano A. Marraffini, and Feng Zhang. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, 339(6121):819–823, 2013.
107. JOHN M. MURRAY, PAUL L. APPLETON, JASON R. SWEDLOW, and JENNIFER C. WATERS. Evaluating performance in three-dimensional fluorescence microscopy. *Journal of Microscopy*, 228(3):390–405, 2007.
108. Ralph Gräf, Jens Rietdorf, and Timo Zimmermann. Live cell spinning disk microscopy. *Advances in Biochemical Engineering/Biotechnology*, 95:57–75, 2005.
109. E. WANG, C. M. BABBEY, and K. W. DUNN. Performance comparison between the high-speed Yokogawa spinning disc confocal system and single-point scanning confocal systems. *Journal of Microscopy*, 218(2):148–159, 2005.
110. IEEE Standard for Floating-Point Arithmetic, 2019.
111. Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98, 2017.

112. Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.
113. Derek Bradley and Gerhard Roth. Adaptive Thresholding using the Integral Image. *Journal of Graphics Tools*, 12(2):13–21, 2007.
114. J.-L. Starck and F. Murtagh. *Astronomical Image and Data Analysis*. Astronomy and Astrophysics Library. Springer Berlin Heidelberg, 2 edition, 8 2006.
115. Per-Erik Danielsson and Olle Seger. Machine Vision for Three-Dimensional Scenes. pages 347–379, 1990.
116. J Canny. A computational approach to edge detection. *IEEE transactions on pattern analysis and machine intelligence*, 8(6):679–98, 1986.
117. D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 207(1167):187–217, 1980.
118. R. Adams and L. Bischof. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6):641–647, 1994.
119. Andrew Mehnert and Paul Jackway. An improved seeded region growing algorithm. *Pattern Recognition Letters*, 18(10):1065–1071, 1997.
120. P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer Berlin Heidelberg, 2013.
121. S Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982. K-means.
122. Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer New York, NY, 1 edition, 8 2006.
123. Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.
124. Alden A. Dima, John T. Elliott, James J. Filliben, Michael Halter, Adele Peskin, Javier Bernal, Marcin Kociolek, Mary C. Brady, Hai C. Tang, and Anne L. Plant. Comparison of segmentation algorithms for fluorescence microscopy images of cells. *Cytometry Part A*, 79A(7):545–559, 2011.
125. Rintu Maria Thomas and Jisha John. A Review on Cell Detection and Segmentation in Microscopic Images. *2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT)*, pages 1–5, 2017.
126. Juan C. Caicedo, Jonathan Roth, Allen Goodman, Tim Becker, Kyle W. Karhohs, Matthieu Broisin, Csaba Molnar, Claire McQuin, Shantanu Singh, Fabian J. Theis, and Anne E. Carpenter. Evaluation of Deep Learning Strategies for Nucleus Segmentation in Fluorescence Images. *Cytometry Part A*, 95(9):952–965, 2019.
127. Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, 2012.
128. Michael D. Abramoff, Paulo J. Magalhaes, and Sunanda J. Ram. Image processing with imagej. *Biophotonics International*, 11(7):36–42, 2004.
129. Caroline A Schneider, Wayne S Rasband, and Kevin W Eliceiri. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7):671–675, 2012.
130. Anne E Carpenter, Thouis R Jones, Michael R Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David A Guertin, Joo Han Chang, Robert A Lindquist, Jason Moffat, Polina Golland, and David M Sabatini. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, 7(10):R100–R100, 2006.
131. Claire McQuin, Allen Goodman, Vasiliy Chernyshev, Lee Kamentsky, Beth A Cimini, Kyle W Karhohs, Minh Doan, Liya Ding, Susanne M Rafelski, Derek Thirstrup, Winfried Wiegraebe, Shantanu Singh, Tim Becker, Juan C Caicedo, and Anne E Carpenter. CellProfiler 3.0: Next-generation image processing for biology. *PLOS Biology*, 16(7):e2005970, 2018.
132. Nicholas Sofroniew, Talley Lambert, Abigail McGovern, et al. napari: a multi-dimensional image viewer for Python, November 2022.
133. Fabrice de Chaumont, Stéphane Dallongeville, Nicolas Chenouard, Nicolas Hervé, Sorin Pop, Thomas Provoost, Vannary Meas-Yedid, Praveen Pankajakshan, Timothée Lecomte, Yoann Le Montagner, Thibault Lagache, Alexandre Dufour, and Jean-Christophe Olivo-Marin. Icy: an open bioimage informatics platform for extended reproducible research. *Nature Methods*, 9(7):690–696, 2012.
134. Peter Bankhead, Maurice B. Loughrey, José A. Fernández, Yvonne Dombrowski, Darragh G. McArt, Philip D. Dunne, Stephen McQuaid, Ronan T. Gray, Liam J. Murray, Helen G. Coleman, Jacqueline A. James, Manuel Salto-Tellez, and Peter W. Hamilton. QuPath: Open source software for digital pathology image analysis. *Scientific Reports*, 7(1):16878, 2017.
135. Berk Ekmekci, Charles E. McAnany, and Cameron Mura. An Introduction to Programming for Bioscientists: A Python-Based Primer. *PLoS Computational Biology*, 12(6):e1004867, 2016.
136. Stefan Van der Walt, Johannes L Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. scikit-image: image processing in python. *PeerJ*, 2:e453, 2014.
137. Elisabeth Roesch, Joe G. Greener, Adam L. MacLean, Huda Nassar, Christopher Rackauckas, Timothy E. Holy, and Michael P. H. Stumpf. Julia for biologists. *Nature Methods*, pages 1–10, 2023.
138. Kota Miura and Nataša Sladoje, editors. *Bioimage Data Analysis Workflows*. Learning Materials in Biosciences. Springer Cham, 1 edition, 2019.
139. Jonathan M. Raser and Erin K. O'Shea. Noise in Gene Expression: Origins, Consequences, and Control. *Science*, 309(5743):2010–2013, 2005.
140. Simon V. Avery. Microbial cell individuality and the underlying sources of heterogeneity. *Nature Reviews Microbiology*, 4(8):577–587, 2006.
141. Arjun Raj and Alexander van Oudenaarden. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell*, 135(2):216–226, 2008.
142. Gábor Balázsi, Alexander van Oudenaarden, and James J. Collins. Cellular Decision Making and Biological Noise: From Microbes to Mammals. *Cell*, 144(6):910–925, 2011.
143. Alvaro Sanchez, Sandeep Choubey, and Jane Kondev. Regulation of Noise in Gene Expression. *Annual Review of Biophysics*, 42(1):469–491, 2013.
144. Mads Kærn, Timothy C. Elston, William J. Blake, and James J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*, 6(6):451–464, 2005.
145. William J. Blake, Gábor Balázsi, Michael A. Kohanski, Farren J. Isaacs, Kevin F. Murphy, Yina Kuang, Charles R. Cantor, David R. Walt, and James J. Collins. Phenotypic Consequences of Promoter-Mediated Transcriptional Noise. *Molecular Cell*, 24(6):853–865, 2006.
146. Gürol M. Süel, Rajan P. Kulkarni, Jonathan Dworkin, Jordi Garcia-Ojalvo, and Michael B. Elowitz. Tunability and Noise Dependence in Differentiation Dynamics. *Science*, 315(5819):1716–1719, 2007.
147. Dann Huh and Johan Paulsson. Non-genetic heterogeneity from stochastic partitioning at cell division. *Nature Genetics*, 43(2):95–100, 2011.
148. Purnananda Guptasarma. Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of Escherichia coli? *BioEssays*, 17(11):987–997, 1995.
149. Vahid Shahrezaei and Peter S Swain. The stochastic nature of biochemical networks. *Current Opinion in Biotechnology*, 19(4):369–374, 2008.
150. Berend Snijder, Raphael Sacher, Pauli Rämö, Eva-Maria Damm, Prisca Liberali, and Lucas Pelkmans. Population context determines cell-to-cell

variability in endocytosis and virus infection. *Nature*, 461(7263):520–523, 2009. Microenvironment and context features.
151. Peter Reuven and Avigdor Eldar. Macromotives and microbehaviors: the social dimension of bacterial phenotypic variability. *Current Opinion in Genetics & Development*, 21(6):759–767, 2011.
152. Alex Sigal, Ron Milo, Ariel Cohen, Naama Geva-Zatorsky, Yael Klein, Yuvalal Liron, Nitzan Rosenfeld, Tamar Danon, Natalie Perzov, and Uri Alon. Variability and memory of protein levels in human cells. *Nature*, 444(7119):643–646, 2006.
153. Naama Geva-Zatorsky, Nitzan Rosenfeld, Shalev Itzkovitz, Ron Milo, Alex Sigal, Erez Dekel, Talia Yarnitzky, Yuvalal Liron, Paz Polak, Galit Lahav, and Uri Alon. Oscillations and variability in the p53 system. *Molecular Systems Biology*, 2(1):2006.0033–2006.0033, 2006.
154. Ofer Feinerman, Joël Veiga, Jeffrey R. Dorfman, Ronald N. Germain, and Grégoire Altan-Bonnet. Variability and Robustness in T Cell Activation from Regulated Heterogeneity in Protein Levels. *Science*, 321(5892):1081–1084, 2008.
155. Jianjiang Hu, Xavier Serra-Picamal, Gert-Jan Bakker, Marleen Van Troys, Sabina Winograd-Katz, Nil Ege, Xiaowei Gong, Yuliia Didan, Inna Grosheva, Omer Polansky, Karima Bakkali, Evelien Van Hamme, Merijn Erp, Manon Vullings, Felix Weiss, Jarama Clucas, Anna M Dowbaj, Erik Sahai, Christophe Ampe, Benjamin Geiger, Peter Friedl, Matteo Bottai, and Staffan Strömblad. Multisite assessment of reproducibility in high-content cell migration imaging data. *Molecular Systems Biology*, page e11490, 2023.
156. Harley H. McAdams and Adam Arkin. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, 94(3):814–819, 1997.
157. William J. Blake, Mads KÆrn, Charles R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422(6932):633–637, 2003.
158. Gürol M. Süel, Rajan P. Kulkarni, Jonathan Dworkin, Jordi Garcia-Ojalvo, and Michael B. Elowitz. Tunability and Noise Dependence in Differentiation Dynamics. *Science*, 315(5819):1716–1719, 2007.
159. Peter Lenz and Lotte Søgaard-Andersen. Temporal and spatial oscillations in bacteria. *Nature Reviews Microbiology*, 9(8):565–577, 2011.
160. Yuichi Wakamoto, Neeraj Dhar, Remy Chait, Katrin Schneider, François Signorino-Gelo, Stanislas Leibler, and John D. McKinney. Dynamic Persistence of Antibiotic-Stressed Mycobacteria. *Science*, 339(6115):91–95, 2013.
161. Joe H. Levine, Michelle E. Fontes, Jonathan Dworkin, and Michael B. Elowitz. Pulsed Feedback Defers Cellular Differentiation. *PLoS Biology*, 10(1):e1001252, 2012.
162. James C. W. Locke, Jonathan W. Young, Michelle Fontes, María Jesús Hernández Jiménez, and Michael B. Elowitz. Stochastic Pulse Regulation in Bacterial Stress Response. *Science*, 334(6054):366–369, 2011.
163. Albert Goldbeter and M. J. Berridge. *Biochemical Oscillations and Cellular Rhythms: The Molecular Bases of Periodic and Chaotic Behaviour*. Cambridge University Press, 1996.
164. Jay C Dunlap and Jennifer J Loros. Making Time: Conservation of Biological Clocks from Fungi to Animals. *Microbiology Spectrum*, 5(3), 2017.
165. Joe H. Levine, Yihan Lin, and Michael B. Elowitz. Functional Roles of Pulsing in Genetic Circuits. *Science*, 342(6163):1193–1200, 2013.
166. Jeremy E. Purvis, Kyle W. Karhohs, Caroline Mock, Eric Batchelor, Alexander Loewer, and Galit Lahav. p53 Dynamics Control Cell Fate. *Science*, 336(6087):1440–1444, 2012.
167. Eric Batchelor, Caroline S. Mock, Irun Bhan, Alexander Loewer, and Galit Lahav. Recurrent Initiation: A Mechanism for Triggering p53 Pulses in Response to DNA Damage. *Molecular Cell*, 30(3):277–289, 2008.
168. Eric Batchelor, Alexander Loewer, Caroline Mock, and Galit Lahav. Stimulus-dependent dynamics of p53 in single cells. *Molecular Systems Biology*, 7(1):488–488, 2011.
169. Louise Ashall, Caroline A. Horton, David E. Nelson, Pawel Paszek, Claire V. Harper, Kate Sillitoe, Sheila Ryan, David G. Spiller, John F. Unitt, David S. Broomhead, Douglas B. Kell, David A. Rand, Violaine Sée, and Michael R. H. White. Pulsatile Stimulation Determines Timing and Specificity of NF-κB-Dependent Transcription. *Science*, 324(5924):242–246, 2009.
170. D. E. Nelson, A. E. C. Ihekwaba, M. Elliott, J. R. Johnson, C. A. Gibney, B. E. Foreman, G. Nelson, V. See, C. A. Horton, D. G. Spiller, S. W. Edwards, H. P. McDowell, J. F. Unitt, E. Sullivan, R. Grimley, N. Benson, D. Broomhead, D. B. Kell, and M. R. H. White. Oscillations in NF-κB Signaling Control the Dynamics of Gene Expression. *Science*, 306(5696):704–708, 2004.
171. Savaş Tay, Jacob J. Hughey, Timothy K. Lee, Tomasz Lipniacki, Stephen R. Quake, and Markus W. Covert. Single-cell NF-κB dynamics reveal digital activation and analogue information processing. *Nature*, 466(7303):267–271, 2010.
172. Carlos López-Otín, Maria A. Blasco, Linda Partridge, Manuel Serrano, and Guido Kroemer. The Hallmarks of Aging. *Cell*, 153(6):1194–1217, 2013.
173. Fabrice Caudron and Yves Barral. A Super-Assembly of Whi3 Encodes Memory of Deceptive Encounters by Single Cells during Yeast Courtship. *Cell*, 155(6):1244–1257, 2013.
174. Sasha F. Levy, Naomi Ziv, and Mark L. Siegal. Bet Hedging in Yeast by Heterogeneous, Age-Correlated Expression of a Stress Protectant. *PLoS Biology*, 10(5):e1001325, 2012.
175. Tobias Bergmiller and Martin Ackermann. Pole Age Affects Cell Size and the Timing of Cell Division in Methylobacterium extorquens AM1. *Journal of Bacteriology*, 193(19):5216–5221, 2011.
176. Rolf Lood, Kristofer Wollein Waldetoft, and Pontus Nordenfelt. Localization-triggered bacterial pathogenesis. *Future Microbiology*, 10(10):1659–1668, 10 2015.
177. Nathalie Q. Balaban, Jack Merrin, Remy Chait, Lukasz Kowalik, and Stanislas Leibler. Bacterial Persistence as a Phenotypic Switch. *Science*, 305(5690):1622–1625, 2004.
178. Kim Lewis. Persister Cells. *Microbiology*, 64(1):357–372, 2010.
179. Orit Gefen and Nathalie Q. Balaban. The importance of being persistent: heterogeneity of bacterial populations under antibiotic stress. *FEMS Microbiology Reviews*, 33(4):704–717, 2009.
180. JosephW. Bigger. TREATMENT OF STAPHYLOCOCCAL INFECTIONS WITH PENICILLIN BY INTERMITTENT STERILISATION. *The Lancet*, 244(6320):497–500, 1944.
181. Oliver Kotte, Benjamin Volkmer, Jakub L Radzikowski, and Matthias Heinemann. Phenotypic bistability in Escherichia coli's central carbon metabolism. *Molecular Systems Biology*, 10(7):736–736, 2014.
182. Markus Arnoldini, Ima Avalos Vizcarra, Rafael Peña-Miller, Nicolas Stocker, Médéric Diard, Viola Vogel, Robert E. Beardmore, Wolf-Dietrich Hardt, and Martin Ackermann. Bistable Expression of Virulence Genes in Salmonella Leads to the Formation of an Antibiotic-Tolerant Subpopulation. *PLoS Biology*, 12(8):e1001928, 2014.
183. Eitan Rotem, Adiel Loinger, Irine Ronin, Irit Levin-Reisman, Chana Gabay, Noam Shoresh, Ofer Biham, and Nathalie Q. Balaban. Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proceedings of the National Academy of Sciences*, 107(28):12541–12546, 2010.
184. Mary K. Stewart, Lisa A. Cummings, Matthew L. Johnson, Alex B. Berezow, and Brad T. Cookson. Regulation of phenotypic heterogeneity

permits Salmonella evasion of the host caspase-1 inflammatory response. *Proceedings of the National Academy of Sciences*, 108(51):20742–20747, 2011.
185. T. C. Barnett, A. C. Bowen, and J. R. Carapetis. The fall and rise of Group A Streptococcus diseases. *Epidemiology and Infection*, 147:e4, 2019.
186. Mark J Walker, Timothy C Barnett, Jason D McArthur, Jason N Cole, Christine M Gillen, Anna Henningham, K S Sriprakash, Martina L Sanderson-Smith, and Victor Nizet. Disease Manifestations and Pathogenic Mechanisms of Group A Streptococcus. *Clinical Microbiology Reviews*, 27(2):264–301, 2014.
187. Debra E. Bessen. Population biology of the human restricted pathogen, Streptococcus pyogenes. *Infection, Genetics and Evolution*, 9(4):581–593, 2009.
188. Sam P. Brown, Daniel M. Cornforth, and Nicole Mideo. Evolution of virulence in opportunistic pathogens: generalism, plasticity, and control. *Trends in Microbiology*, 20(7):336–342, 2012.
189. Carol Berkower, Miriam Ravins, Allon E. Moses, and Emanuel Hanski. Expression of different group A streptococcal M proteins in an isogenic background demonstrates diversity in adherence to and invasion of eukaryotic cells. *Molecular Microbiology*, 31(5):1463–1475, 1999.
190. Simon Döhrmann, Christopher N. LaRock, Ericka L. Anderson, Jason N. Cole, Brinda Ryali, Chelsea Stewart, Poochit Nonejuie, Joe Pogliano, Ross Corriden, Partho Ghosh, and Victor Nizet. Group A Streptococcal M1 Protein Provides Resistance against the Antimicrobial Activity of Histones. *Scientific Reports*, 7(1):43039, 2017.
191. Beinan Wang and P Patrick Cleary. Intracellular Invasion by Streptococcus pyogenes: Invasins, Host Receptors, and Relevance to Human Disease. *Microbiology spectrum*, 7(4), 2019.
192. R. Facklam, B. Beall, A. Efstratiou, V. Fischetti, D. Johnson, E. Kaplan, P. Kriz, M. Lovgren, D. Martin, B. Schwartz, A. Totolian, D. Bessen, S. Hollingshead, F. Rubin, J. Scott, and G. Tyrrell. emm Typing and Validation of Provisional M Types for Group A Streptococci. *Emerging Infectious Diseases*, 5(2):247–253, 1999.
193. Yuan Li, Joy Rivers, Saundra Mathis, Zhongya Li, Srinivasan Velusamy, Srinivas A. Nanduri, Chris A. Van Beneden, Paula Snippes-Vagnone, Ruth Lynfield, Lesley McGee, Sopio Chochua, Benjamin J. Metcalf, and Bernard Beall. Genomic Surveillance of Streptococcus pyogenes Strains Causing Invasive Disease, United States, 2016–2017. *Frontiers in Microbiology*, 11:1547, 2020.
194. H R Frost, M R Davies, S Velusamy, V Delforge, A Erhart, S Darboe, A Steer, M J Walker, B Beall, A Botteaux, and P R Smeesters. Updated emm-typing protocol for Streptococcus pyogenes. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 26(7):946.e5–946.e8, 2020.
195. D G Heath and P P Cleary. Fc-receptor and M-protein genes of group A streptococci are products of gene duplication. *Proceedings of the National Academy of Sciences*, 86(12):4741–4745, 1989.
196. Pontus Nordenfelt, Sofia Waldemarson, Adam Linder, Matthias Mörgelin, Christofer Karlsson, Johan Malmström, and Lars Björck. Antibody orientation at bacterial surfaces is related to invasive infection. *Journal of Experimental Medicine*, 209(13):2367–2381, 2012.
197. M W Cunningham, J M McCormack, P G Fenderson, M K Ho, E H Beachey, and J B Dale. Human and murine antibodies cross-reactive with streptococcal M protein and myosin recognize the sequence GLN-LYS-SER-LYS-GLN in M protein. *The Journal of Immunology*, 143(8):2677–2683, 1989.
198. Donald N Forthal. Functions of Antibodies. *Microbiology Spectrum*, 2(4):AID–0019–2014, 2014.
199. Lenette L. Lu, Todd J. Suscovich, Sarah M. Fortune, and Galit Alter. Beyond binding: antibody effector functions in infectious diseases. *Nature Reviews Immunology*, 18(1):46–61, 2018.
200. Wael Bahnan, Lotta Happonen, Hamed Khakzad, Vibha Kumra Ahnlide, Therese Neergaard, Sebastian Wrighton, Oscar André, Eleni Bratanis, Di Tang, Thomas Hellmark, Lars Björck, Oonagh Shannon, Lars Malmström, Johan Malmström, and Pontus Nordenfelt. A human monoclonal antibody bivalently binding two different epitopes in streptococcal M protein mediates immune function. *EMBO Molecular Medicine*, 15(2):e16208, 2023.
201. P Åkesson, K H Schmidt, J Cooney, and L Björck. M1 protein and protein H: IgGFc- and albumin-binding streptococcal surface proteins encoded by adjacent genes. *Biochemical Journal*, 300(3):877–886, 1994.
202. The IMpact-RSV Study Group. Palivizumab, a Humanized Respiratory Syncytial Virus Monoclonal Antibody, Reduces Hospitalization From Respiratory Syncytial Virus Infection in High-risk Infants. *Pediatrics*, 102(3):531–537, 09 1998.
203. Sabue Mulangu, Lori E Dodd, Richard T Davey, Olivier Tshiani Mbaya, Michael Proschan, Daniel Mukadi, Mariano Lusakibanza Manzo, Didier Nzolo, Antoine Tshomba Oloma, Augustin Ibanda, Rosine Ali, Sinaré Coulibaly, Adam C Levine, Rebecca Grais, Janet Diaz, H Clifford Lane, Jean-Jacques Muyembe-Tamfum, PALM Writing Group, Billy Sivahera, Modet Camara, Richard Kojan, Robert Walker, Bonnie Dighero-Kemp, Huyen Cao, Philippe Mukumbayi, Placide Mbala-Kingebeni, Steve Ahuka, Sarah Albert, Tyler Bonnett, Ian Crozier, Michael Duvenhage, Calvin Proffitt, Marc Teitelbaum, Thomas Moench, Jamila Aboulhab, Kevin Barrett, Kelly Cahill, Katherine Cone, Risa Eckes, Lisa Hensley, Betsey Herpin, Elizabeth Higgs, Julie Ledgerwood, Jerome Pierson, Mary Smolskis, Ydrissa Sow, John Tierney, Sumathi Sivapalasingam, Wendy Holman, Nikki Gettinger, David Vallée, Jacqueline Nordwall, and PALM Consortium Study Team. A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. *New England Journal of Medicine*, 381(24):2293–2303, 2019.
204. Vivek Shinde, Sutika Bhikha, Zaheer Hoosain, Moherndran Archary, Qasim Bhorat, Lee Fairlie, Umesh Lalloo, Mduduzi S L Masilela, Dhayendre Moodley, Sherika Hanley, Leon Fouche, Cheryl Louw, Michele Tameris, Nishanta Singh, Ameena Goga, Keertan Dheda, Coert Grobbelaar, Gertruida Kruger, Nazira Carrim-Ganey, Vicky Baillie, Tulio de Oliveira, Anthonet Lombard Koen, Johan J Lombaard, Rosie Mngqibisa, As'ad E Bhorat, Gabriella Benadé, Natasha Lalloo, Annah Pitsi, Pieter-Louis Vollgraaff, Angelique Luabeya, Aliasgar Esmail, Friedrich G Petrick, Aylin Oommen-Jose, Sharne Foulkes, Khatija Ahmed, Asha Thombrayil, Lou Fries, Shane Cloney-Clark, Mingzhu Zhu, Chijioke Bennett, Gary Albert, Emmanuel Faust, Joyce S Plested, Andreana Robertson, Susan Neal, Iksung Cho, Greg M Glenn, Filip Dubovsky, Shabir A Madhi, and 2019nCoV-501 Study Group. Efficacy of NVX-CoV2373 Covid-19 Vaccine against the B.1.351 Variant. *New England Journal of Medicine*, 384(20):1899–1909, 2021.
205. Nick Andrews, Julia Stowe, Freja Kirsebom, Samuel Toffa, Tim Rickeard, Eileen Gallagher, Charlotte Gower, Meaghan Kall, Natalie Groves, Anne-Marie O'Connell, David Simons, Paula B. Blomquist, Asad Zaidi, Sophie Nash, Nurin Iwani Binti Abdul Aziz, Simon Thelwall, Gavin Dabrera, Richard Myers, Gayatri Amirthalingam, Saheer Gharbia, Jeffrey C. Barrett, Richard Elson, Shamez N. Ladhani, Neil Ferguson, Maria Zambon, Colin N.J. Campbell, Kevin Brown, Susan Hopkins, Meera Chand, Mary Ramsay, and Jamie Lopez Bernal. Covid-19 Vaccine Effectiveness against the Omicron (B.1.1.529) Variant. *New England Journal of Medicine*, 386(16):1532–1546, 2022.
206. Pragya D Yadav, Gajanan N Sapkal, Priya Abraham, Raches Ella, Gururaj Deshpande, Deepak Y Patil, Dimpal A Nyayanit, Nivedita Gupta, Rima R Sahay, Anita M Shete, Samiran Panda, Balram Bhargava, and V Krishna Mohan. Neutralization of Variant Under Investigation B.1.617.1 With Sera of BBV152 Vaccinees. *Clinical Infectious Diseases*, 74(2):366–368, 2021.
207. Nicholas G. Davies, Sam Abbott, Rosanna C. Barnard, Christopher I. Jarvis, Adam J. Kucharski, James D. Munday, Carl A. B. Pearson, Timothy W.

Russell, Damien C. Tully, Alex D. Washburne, Tom Wenseleers, Amy Gimma, William Waites, Kerry L. M. Wong, Kevin van Zandvoort, Justin D. Silverman, CMMID COVID-19 Working Group1‡, COVID-19 Genomics UK (COG-UK) Consortium‡, Karla Diaz-Ordaz, Ruth Keogh, Rosalind M. Eggo, Sebastian Funk, Mark Jit, Katherine E. Atkins, and W. John Edmunds. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science*, 372(6538):eabg3055, 2021.

208. Delphine Planas, David Veyer, Artem Baidaliuk, Isabelle Staropoli, Florence Guivel-Benhassine, Maaran Michael Rajah, Cyril Planchais, Françoise Porrot, Nicolas Robillard, Julien Puech, Matthieu Prot, Floriane Gallais, Pierre Gantner, Aurélie Velay, Julien Le Guen, Najiby Kassis-Chikhani, Dhiaeddine Edriss, Laurent Belec, Aymeric Seve, Laura Courtellemont, Hélène Péré, Laurent Hocqueloux, Samira Fafi-Kremer, Thierry Prazuck, Hugo Mouquet, Timothée Bruel, Etienne Simon-Lorière, Felix A. Rey, and Olivier Schwartz. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature*, 596(7871):276–280, 2021.
209. Pengfei Wang, Manoj S. Nair, Lihong Liu, Sho Iketani, Yang Luo, Yicheng Guo, Maple Wang, Jian Yu, Baoshan Zhang, Peter D. Kwong, Barney S. Graham, John R. Mascola, Jennifer Y. Chang, Michael T. Yin, Magdalena Sobieszczyk, Christos A. Kyratsous, Lawrence Shapiro, Zizhang Sheng, Yaoxing Huang, and David D. Ho. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature*, 593(7857):130–135, 2021.
210. Wilfredo F. Garcia-Beltran, Evan C. Lam, Kerri St. Denis, Adam D. Nitido, Zeidy H. Garcia, Blake M. Hauser, Jared Feldman, Maia N. Pavlovic, David J. Gregory, Mark C. Poznansky, Alex Sigal, Aaron G. Schmidt, A. John Iafrate, Vivek Naranbhai, and Alejandro B. Balazs. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. *Cell*, 184(9):2372–2383.e9, 2021.
211. David Ellinghaus, Tom H Degenhardt, Karlsen, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *New England Journal of Medicine*, 383(16):1522–1534, 2020.
212. Erola Pairo-Castineira, J. Kenneth Clohisey, Baillie, et al. Genetic mechanisms of critical illness in COVID-19. *Nature*, 591(7848):92–98, 2021.
213. Xiaonan Zhang, Yun Tan, Yun Ling, Gang Lu, Feng Liu, Zhigang Yi, Xiaofang Jia, Min Wu, Bisheng Shi, Shuibao Xu, Jun Chen, Wei Wang, Bing Chen, Lu Jiang, Shuting Yu, Jing Lu, Jinzeng Wang, Mingzhu Xu, Zhenghong Yuan, Qin Zhang, Xinxin Zhang, Guoping Zhao, Shengyue Wang, Saijuan Chen, and Hongzhou Lu. Viral and host factors related to the clinical outcome of COVID-19. *Nature*, 583(7816):437–440, 2020.
214. Jérôme Hadjadj, Nader Yatim, Laura Barnabei, Aurélien Corneau, Jeremy Boussier, Nikaïa Smith, Hélène Péré, Bruno Charbit, Vincent Bondet, Camille Chenevier-Gobeaux, Paul Breillat, Nicolas Carlier, Rémy Gauzit, Caroline Morbieu, Frédéric Pène, Nathalie Marin, Nicolas Roche, Tali-Anne Szwebel, Sarah H. Merkling, Jean-Marc Treluyer, David Veyer, Luc Mouthon, Catherine Blanc, Pierre-Louis Tharaux, Flore Rozenberg, Alain Fischer, Darragh Duffy, Frédéric Rieux-Laucat, Solen Kernéis, and Benjamin Terrier. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science*, 369(6504):718–724, 2020.
215. Daniel Blanco-Melo, Benjamin E. Nilsson-Payant, Wen-Chun Liu, Skyler Uhl, Daisy Hoagland, Rasmus Møller, Tristan X. Jordan, Kohei Oishi, Maryline Panis, David Sachs, Taia T. Wang, Robert E. Schwartz, Jean K. Lim, Randy A. Albrecht, and Benjamin R. tenOever. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*, 181(5):1036–1045.e9, 2020.
216. Mingfeng Liao, Yang Liu, Jing Yuan, Yanling Wen, Gang Xu, Juanjuan Zhao, Lin Cheng, Jinxiu Li, Xin Wang, Fuxiang Wang, Lei Liu, Ido Amit, Shuye Zhang, and Zheng Zhang. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nature Medicine*, 26(6):842–844, 2020.
217. Nico Battich, Thomas Stoeger, and Lucas Pelkmans. Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nature methods*, 10(11):1127–33, 2013.
218. Beate Neumann, Michael Held, Urban Liebel, Holger Erfle, Phill Rogers, Rainer Pepperkok, and Jan Ellenberg. High-throughput RNAi screening by time-lapse imaging of live human cells. *Nature Methods*, 3(5):385–390, 2006.
219. Beate Neumann, Thomas Walter, Jean-Karim Hériché, Jutta Bulkescher, Holger Erfle, Christian Conrad, Phill Rogers, Ina Poser, Michael Held, Urban Liebel, Cihan Cetin, Frank Sieckmann, Gregoire Pau, Rolf Kabbe, Annelie Wünsche, Venkata Satagopam, Michael H. A. Schmitz, Catherine Chapuis, Daniel W. Gerlich, Reinhard Schneider, Roland Eils, Wolfgang Huber, Jan-Michael Peters, Anthony A. Hyman, Richard Durbin, Rainer Pepperkok, and Jan Ellenberg. Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature*, 464(7289):721–727, 2010.
220. Georgiana Crainiciuc, Miguel Palomino-Segura, Miguel Molina-Moreno, Jon Sicilia, David G. Aragones, Jackson Liang Yao Li, Rodrigo Madurga, José M. Adrover, Alejandra Aroca-Crevillén, Sandra Martin-Salamanca, Alfonso Serrano del Valle, Sandra D. Castillo, Heidi C. E. Welch, Oliver Soehnlein, Mariona Graupera, Fátima Sánchez-Cabo, Alexander Zarbock, Thomas E. Smithgall, Mauro Di Pilato, Thorsten R. Mempel, Pierre-Louis Tharaux, Santiago F. González, Angel Ayuso-Sacido, Lai Guan Ng, Gabriel F. Calvo, Iván González-Díaz, Fernando Díaz-de María, and Andrés Hidalgo. Behavioural immune landscapes of inflammation. *Nature*, 601(7893):415–421, 2022.
221. Mark-Anthony Bray, Shantanu Singh, Han Han, Chadwick T Davis, Blake Borgeson, Cathy Hartland, Maria Kost-Alimova, Sigrun M Gustafsdottir, Christopher C Gibson, and Anne E Carpenter. Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nature Protocols*, 11(9):1757–1774, 2016.
222. Carlos Aguilar-Avelar, Brenda Soto-García, Diana Aráiz-Hernández, Juan F. Yee-de León, Miguel Esparza, Franco Chacón, Jesús Rolando Delgado-Balderas, Mario M. Alvarez, Grissel Trujillo-de Santiago, Lauro S. Gómez-Guerra, Liza P. Velarde-Calvillo, Alejandro Abarca-Blanco, and J. D. Wong-Campos. High-Throughput Automated Microscopy of Circulating Tumor Cells. *Scientific Reports*, 9(1):13766, 2019.
223. Dora Mahecic, Willi L. Stepp, Chen Zhang, Juliette Griffié, Martin Weigert, and Suliana Manley. Event-driven acquisition for content-enriched microscopy. *Nature Methods*, 19(10):1262–1267, 2022.
224. Jonatan Alvelid, Martina Damenti, Chiara Sgattoni, and Ilaria Testa. Event-triggered STED imaging. *Nature Methods*, 19(10):1268–1275, 2022.
225. Oscar André, Johannes Kumra Ahnlide, Nils Norlin, Vinay Swaminathan, and Pontus Nordenfelt. Data-driven microscopy allows for automated context-specific acquisition of high-fidelity image data. *Cell Reports Methods*, 3(3):100419, 2023.
226. Colleen M. Garvey, Erin Spiller, Danika Lindsay, Chun-Te Chiang, Nathan C. Choi, David B. Agus, Parag Mallick, Jasmine Foo, and Shannon M. Mumenthaler. A high-content image-based method for quantitatively studying context-dependent cell population dynamics. *Scientific Reports*, 6(1):29752, 2016.
227. Christian Conrad, Annelie Wünsche, Tze Heng Tan, Jutta Bulkescher, Frank Sieckmann, Fatima Verissimo, Arthur Edelstein, Thomas Walter, Urban Liebel, Rainer Pepperkok, and Jan Ellenberg. Micropilot: automation of fluorescence microscopy–based imaging for systems biology. *Nature Methods*, 8(3):246–249, 2011.
228. Fabian Zanella, James B. Lorens, and Wolfgang Link. High content screening: seeing is believing. *Trends in Biotechnology*, 28(5):237–245, 2010.
229. Mojca Mattiazzi Usaj, Erin B. Styles, Adrian J. Verster, Helena Friesen, Charles Boone, and Brenda J. Andrews. High-Content Screening for Quantitative Cell Biology. *Trends in Cell Biology*, 26(8):598–611, 2016.
230. Michael Boutros, Florian Heigwer, and Christina Laufer. Microscopy-Based High-Content Screening. *Cell*, 163(6):1314–1325, 2015.
231. Gaudenz Danuser. Computer Vision in Cell Biology. *Cell*, 147(5):973–978, 2011.

232. Thiago FA França and José Maria Monserrat. Reproducibility crisis in science or unrealistic expectations? *EMBO reports*, 19(6), 2018.
233. Indhupriya Subramanian, Srikant Verma, Shiva Kumar, Abhay Jere, and Krishanpal Anamika. Multi-omics Data Integration, Interpretation, and Its Application. *Bioinformatics and Biology Insights*, 14:1177932219899051, 2020.
234. Yehudit Hasin, Marcus Seldin, and Aldons Lusis. Multi-omics approaches to disease. *Genome Biology*, 18(1):83, 2017.
235. Konrad J. Karczewski and Michael P. Snyder. Integrative omics for health and disease. *Nature Reviews Genetics*, 19(5):299–310, 2018.
236. Ole Seehausen, Roger K. Butlin, Irene Keller, Catherine E. Wagner, Janette W. Boughman, Paul A. Hohenlohe, Catherine L. Peichel, Glenn-Peter Saetre, Claudia Bank, Åke Brännström, Alan Brelsford, Chris S. Clarkson, Fabrice Eroukhmanoff, Jeffrey L. Feder, Martin C. Fischer, Andrew D. Foote, Paolo Franchini, Chris D. Jiggins, Felicity C. Jones, Anna K. Lindholm, Kay Lucek, Martine E. Maan, David A. Marques, Simon H. Martin, Blake Matthews, Joana I. Meier, Markus Möst, Michael W. Nachman, Etsuko Nonaka, Diana J. Rennison, Julia Schwarzer, Eric T. Watson, Anja M. Westram, and Alex Widmer. Genomics and the origin of species. *Nature Reviews Genetics*, 15(3):176–192, 2014.
237. R. David Hawkins, Gary C. Hon, and Bing Ren. Next-generation genomics: an integrative approach. *Nature Reviews Genetics*, 11(7):476–486, 2010.
238. J. Craig Venter, Mark D. Adams, Xiaohong Myers, Zhu, et al. The Sequence of the Human Genome. *Science*, 291(5507):1304–1351, 2001.
239. Lambda Moses and Lior Pachter. Museum of spatial transcriptomics. *Nature Methods*, 19(5):534–546, 2022.
240. ZhiCheng Dong and Yan Chen. Transcriptomics: Advances and approaches. *Science China Life Sciences*, 56(10):960–967, 2013.
241. Zhong Wang, Mark Gerstein, and Michael Snyder. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1):57–63, 2009.
242. Nishant Pappireddi, Lance Martin, and Martin Wühr. A Review on Quantitative Multiplexed Proteomics. *ChemBioChem*, 20(10):1210–1224, 2019.
243. Winston Timp and Gregory Timp. Beyond mass spectrometry, the next step in proteomics. *Science Advances*, 6(2):eaax8978, 2020.
244. John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021.
245. A. F. Maarten Altelaar, Javier Munoz, and Albert J. R. Heck. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nature Reviews Genetics*, 14(1):35–48, 2013.
246. Mihaly Varadi, Stephen Anyango, Mandar Deshpande, Sreenath Nair, Cindy Natassia, Galabina Yordanova, David Yuan, Oana Stroe, Gemma Wood, Agata Laydon, Augustin Žídek, Tim Green, Kathryn Tunyasuvunakool, Stig Petersen, John Jumper, Ellen Clancy, Richard Green, Ankur Vora, Mira Lutfi, Michael Figurnov, Andrew Cowie, Nicole Hobbs, Pushmeet Kohli, Gerard Kleywegt, Ewan Birney, Demis Hassabis, and Sameer Velankar. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, 50(D1):D439–D444, 2021.
247. David S. Wishart. Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery*, 15(7):473–484, 2016.
248. J. K. Nicholson, J. C. Lindon, and E. Holmes. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, 29(11):1181–1189, 1999.
249. Caroline H. Johnson, Julijana Ivanisevic, and Gary Siuzdak. Metabolomics: beyond biomarkers and towards mechanisms. *Nature Reviews Molecular Cell Biology*, 17(7):451–459, 2016.
250. Andrea D. Weston and Leroy Hood. Systems Biology, Proteomics, and the Future of Health Care: Toward Predictive, Preventative, and Personalized Medicine. *Journal of Proteome Research*, 3(2):179–196, 2004.
251. Hiroaki Kitano. Systems Biology: A Brief Overview. *Science*, 295(5560):1662–1664, 2002.
252. Alan Aderem. Systems Biology: Its Practice and Challenges. *Cell*, 121(4):511–513, 2005.
253. Edison T. Liu. Systems Biology, Integrative Biology, Predictive Biology. *Cell*, 121(4):505–506, 2005.
254. Farhana R. Pinu, David J. Beale, Amy M. Paten, Konstantinos Kouremenos, Sanjay Swarup, Horst J. Schirra, and David Wishart. Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. *Metabolites*, 9(4):76, 2019.
255. Rick Horwitz. Integrated, multi-scale, spatial–temporal cell biology – A next step in the post genomic era. *Methods*, 96:3–5, 2016.
256. John G. Lock and Staffan Strömblad. Systems microscopy: An emerging strategy for the life sciences. *Experimental Cell Research*, 316(8):1438–1444, 2010.
257. Peter J. Verveer and Philippe I. H. Bastiaens. Quantitative microscopy and systems biology: seeing the whole picture. *Histochemistry and Cell Biology*, 130(5):833, 2008.
258. Laura Antonelli, Mario Rosario Guarracino, Lucia Maddalena, and Mara Sangiovanni. Integrating imaging and omics data: A review. *Biomedical Signal Processing and Control*, 52:264–280, 2019.
259. Jason R. Swedlow, Pasi Kankaanpää, Ugis Sarkans, Wojtek Goscinski, Graham Galloway, Leonel Malacrida, Ryan P. Sullivan, Steffen Härtel, Claire M. Brown, Christopher Wood, Antje Keppler, Federica Paina, Ben Loos, Sara Zullino, Dario Livio Longo, Silvio Aime, and Shuichi Onami. A global view of standards for open image data formats and repositories. *Nature Methods*, 18(12):1440–1446, 2021.
260. Jean-Karim Hériché, Stephanie Alexander, and Jan Ellenberg. Integrating Imaging and Omics: Computational Methods and Challenges. *Annual Review of Biomedical Data Science*, 2(1):175–197, 2019.
261. Ilya G Goldberg, Chris Allan, Jean-Marie Burel, Doug Creager, Andrea Falconi, Harry Hochheiser, Josiah Johnston, Jeff Mellen, Peter K Sorger, and Jason R Swedlow. The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging. *Genome Biology*, 6(5):R47–R47, 2005.
262. Chris Allan, Jean-Marie Burel, Josh Moore, Colin Blackburn, Melissa Linkert, Scott Loynton, Donald MacDonald, William J Moore, Carlos Neves, Andrew Patterson, Michael Porter, Aleksandra Tarkowska, Brian Loranger, Jerome Avondo, Ingvar Lagerstedt, Luca Lianas, Simone Leo, Katherine Hands, Ron T Hay, Ardan Patwardhan, Christoph Best, Gerard J Kleywegt, Gianluigi Zanetti, and Jason R Swedlow. OMERO: flexible, model-driven data management for experimental biology. *Nature Methods*, 9(3):245–253, 2012.
263. Simon Li, Sébastien Besson, Colin Blackburn, Mark Carroll, Richard K. Ferguson, Helen Flynn, Kenneth Gillen, Roger Leigh, Dominik Lindner, Melissa Linkert, William J. Moore, Balaji Ramalingam, Emil Rozbicki, Gabriella Rustici, Aleksandra Tarkowska, Petr Walczysko, Eleanor Williams, Chris Allan, Jean-Marie Burel, Josh Moore, and Jason R. Swedlow. Metadata management for high content screening in OMERO. *Methods*, 96:27–32, 2016.
264. Jean-Marie Burel, Sébastien Besson, Colin Blackburn, Mark Carroll, Richard K. Ferguson, Helen Flynn, Kenneth Gillen, Roger Leigh, Simon Li,

Dominik Lindner, Melissa Linkert, William J. Moore, Balaji Ramalingam, Emil Rozbicki, Aleksandra Tarkowska, Petr Walczysko, Chris Allan, Josh Moore, and Jason R. Swedlow. Publishing and sharing multi-dimensional image data with OMERO. *Mammalian Genome*, 26(9-10):441–447, 2015.

265. Mojca Mattiazzi Usaj, Nil Sahin, Helena Friesen, Carles Pons, Matej Usaj, Myra Paz D Masinas, Ermira Shuteriqi, Aleksei Shkurin, Patrick Aloy, Quaid Morris, Charles Boone, and Brenda J Andrews. Systematic genetics and single-cell imaging reveal widespread morphological pleiotropy and cell-to-cell variability. *Molecular Systems Biology*, 16(2):e9243, 2020.
266. Staffan Strömblad and John G Lock. Using Systems Microscopy to Understand the Emergence of Cell Migration from Cell Organization. *Methods in molecular biology (Clifton, N.J.)*, 1749:119–134, 2018.
267. Christian Tischer, Volker Hilsenstein, Kirsten Hanson, and Rainer Pepperkok. Adaptive fluorescence microscopy by online feedback image analysis. *Methods in cell biology*, 123:489–503, 2014.
268. Pedro Almada, Pedro M Pereira, Siân Culley, Ghislaine Caillol, Fanny Boroni-Rueda, Christina L Dix, Guillaume Charras, Buzz Baum, Romain F Laine, Christophe Leterrier, and Ricardo Henriques. Automating multimodal microscopy with NanoJ-Fluidics. *Nature communications*, 10(1):1223, 2018.
269. Sébastien Tosi, Anna Lladó, Lídia Bardia, Elena Rebollo, Anna Godo, Petra Stockinger, and Julien Colombelli. AutoScanJ: A Suite of ImageJ Scripts for Intelligent Microscopy. *Frontiers in Bioinformatics*, 1:627626, 2021.
270. Zachary R. Fox, Steven Fletcher, Achille Fraisse, Chetan Aditya, Sebastián Sosa-Carrillo, Julienne Petit, Sébastien Gilles, François Bertaux, Jakob Ruess, and Gregory Batt. Enabling reactive microscopy with MicroMator. *Nature Communications*, 13(1):2199, 2022.
271. Stefan Leutenegger, Margarita Chli, and Roland Y. Siegwart. BRISK: Binary Robust Invariant Scalable Keypoints. *2011 International Conference on Computer Vision*, pages 2548–2555, 2011.
272. David G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
273. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. SURF: Speeded Up Robust Features. *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 404–417, 2008.
274. Jean-Michel Morel and Guoshen Yu. ASIFT: A New Framework for Fully Affine Invariant Image Comparison. *SIAM Journal on Imaging Sciences*, 2(2):438–469, 2009.
275. Shamus P. Keeler and Julie M. Fox. Requirement of Fc-Fc Gamma Receptor Interaction for Antibody-Based Protection against Emerging Virus Infections. *Viruses*, 13(6):1037, 2021.
276. Victor C. Huber, Joyce M. Lynch, Doris J. Bucher, Jianhua Le, and Dennis W. Metzger. Fc Receptor-Mediated Phagocytosis Makes a Significant Contribution to Clearance of Influenza Virus Infections. *The Journal of Immunology*, 166(12):7381–7388, 2001.
277. Haruo Fujisawa. Neutrophils Play an Essential Role in Cooperation with Antibody in both Protection against and Recovery from Pulmonary Infection with Influenza Virus in Mice. *Journal of Virology*, 82(6):2772–2783, 2008.
278. Masaya Yamaguchi, Yutaka Terao, and Shigetada Kawabata. Multiple role of S. pyogenes Fn-binding proteins. *Cellular Microbiology*, 15(4):503–511, 2013.
279. Anja Ochel, Manfred Rohde, Gursharan S. Chhatwal, and Susanne R. Talay. The M1 Protein of Streptococcus pyogenes Triggers an Innate Uptake Mechanism into Polarized Human Endothelial Cells. *Journal of Innate Immunity*, 6(5):585–596, 2014.
280. Pietro Speziale, Carla Renata Arciola, and Giampiero Pietrocola. Fibronectin and Its Role in Human Infective Diseases. *Cells*, 8(12):1516, 2019.
281. Stefan Steinerberger. On the number of positions in chess without promotion. *International Journal of Game Theory*, 44(3):761–767, 2015.
282. John Tromp and Gunnar Farnebäck. Combinatorics of go, 2006. Accessed: 2023-04-24.
283. David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.
284. Fabio Urbina, Filippa Lentzos, Cédric Invernizzi, and Sean Ekins. Dual use of artificial-intelligence-powered drug discovery. *Nature Machine Intelligence*, 4(3):189–191, 2022.

Scientific publications

Printed by Media-Tryck, Lund 2023 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

FACULTY OF MEDICINE

Division of Infection Medicine
Department of Clinical Sciences, Lund

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2023:78
ISBN 978-91-8021-418-6
ISSN 1652-8220