THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

COMPETITIVE-BINDING BASED OPTICAL DNA MAPPING

FROM BACTERIAL PLASMIDS TO THE HUMAN GENOME

VILHELM MÜLLER

Department of Biology and Biological Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2020

COMPETITIVE-BINDING BASED OPTICAL DNA MAPPING
FROM BACTERIAL PLASMIDS TO THE HUMAN GENOME
VILHELM MÜLLER
ISBN 978-91-7905-169-3

© VILHELM MÜLLER, 2020.

Doktorsavhandlingar vid Chalmers Tekniska Högskola
Ny serie nr 4636
ISSN 0346-718X
Denna doktorsavhandling är en bearbetning och utvidgning av tidigare publicerad licentiatavhandling (1)

Department of Biology and Biological Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover:
Illustration of DNA molecules labeled with YOYO-1 (blue) and netropsin (light gold) rendering a sequence-dependent emission intensity profile, before confinement inside nanofluidic channels.

Printed by Chalmers Reproservice
Gothenburg, Sweden 2020

COMPETITIVE-BINDING BASED OPTICAL DNA MAPPING

FROM BACTERIAL PLASMIDS TO THE HUMAN GENOME

VILHELM MÜLLER

Department of Biology and Biological Engineering
Chalmers University of Technology

Abstract

Significant advances within the field of DNA sequencing have allowed us to study DNA at a level of detail that was previously impossible. However, dynamic genomic regions with a high degree of structural variations, while being linked to disease in humans and increased resistance to antibiotics, are still challenging to characterize. Furthermore, DNA sequencing for bacterial diagnostics and detection of resistance genes is presently hampered by the excessive lead times associated with the overall complexity of the applied methods.

This Thesis describes the development of novel assays based on optical DNA mapping, which, although studying DNA at a lower resolution, is capable of rapid processing of significantly larger DNA fragments compared to sequencing. The fluorescent labeling in the assays presented here relies on competitive DNA binding between the emissive YOYO-1 and the sequence-specific, non-emissive, netropsin. The labeled DNA is then stretched in nanofluidic channels and imaged using fluorescence microscopy, enabling extraction of coarse-grained sequence information from ultralong DNA molecules at the single-molecule level.

The results demonstrate how competitive binding-based optical DNA mapping can be used to characterize and trace bacterial DNA, responsible for the spread of antibiotic resistance. The mapped bacterial DNA can also be used to identify bacterial species in complex mixtures and directly from clinical samples. Additionally, so-called long-range sequence information of the human genome can be obtained, with possible future applications including detection of disease-related structural variations and epigenetic profiling.

Keywords: optical DNA mapping, competitive binding, nanofluidics, DNA, fluorescence microscopy, single-molecule, genetics, antibiotic resistance, bacteria, plasmids.

List of Publications

This Thesis is based on the following papers, referred to in the text by Roman numerals:

I. Rapid tracing of resistance plasmids in a nosocomial outbreak using optical DNA mapping
Müller V, Karami N, Nyberg LK, Pichler C, Pedreschi PCT, Quaderi S, Fritzsche J, Ambjörnsson T, Åhrén C, and Westerlund F
ACS Infectious Disease, **2016**, 2, 322-328

II. Direct identification of antibiotic resistance genes on single plasmid molecules using CRISPR/Cas9 in combination with optical DNA mapping
Müller V, Rajer F, Frykholm K, Nyberg LK, Quaderi S, Fritzsche J, Kristiansson K, Ambjörnsson T, Sandegren L, and Westerlund F
Scientific Reports, **2016**, 6, 37938

III. Cultivation-free typing of bacteria using optical DNA mapping
Müller V*, Nyblom M*, Johnning A*, Wrande M, Dvirnas A, KK S, Giske CG, Ambjörnsson T, Sandegren L, Kristiansson E, and Westerlund F
Submitted Manuscript

IV. Enzyme-free optical DNA mapping of the human genome using competitive binding
Müller V, Dvirnas A, Andersson J, Singh V, KK S, Johansson P, Ebenstein Y, Ambjörnsson T, and Westerlund F
Nucleic Acid Research, **2019**, 47, e89

* *Authors contributed equally to the work*

Additional publications not included in this Thesis:

V. Enantioselective Cyclization of Photochromic Dithienylethenes Bound to DNA
Pace TCS, Müller V, Lincoln P, and Andreasson J
Angewandte Chemie International Edition, **2013**, 52, 4393-6

VI. Extension of nanoconfined DNA: Quantitative comparison between experiment and theory
Iarko V, Werner E, Nyberg L, Müller V, Fritzsche J, Ambjörnsson T, Beech J, Tegenfeldt J, Mehlig K, Westerlund F, and Mehlig B
Physical Review E, **2015**, 92, 062701

VII. Rapid identification of intact bacterial resistance plasmids via optical mapping of single DNA molecules
Nyberg LK, Quaderi S, Emilsson G, Karami N, Lagerstedt E, Müller V, Noble C, Hammarberg S, Nilsson AN, Sjöberg F, Fritzsche J, Kristiansson E, Sandegren L, Ambjörnsson T, and Westerlund F
Scientific Reports, **2016**, 6, 30410

VIII. Optical DNA Mapping in Nanofluidic Channels: Principles and Applications
Müller V and Westerlund F
Lab on a Chip, **2017**, 17, 579-590

IX. Noise Reduction in Single Time Frame Optical DNA Maps
Torche P, Müller V, Westerlund F, and Ambjörnsson T
PLoS ONE, **2017**, 12, 0179041

X. Facilitated sequence assembly using densely labeled optical DNA barcodes: A combinatorial auction approach
Dvirnas A, Pichler C, Quaderi S, Stewart C, Nyberg K, Müller V, Bikkarolla S, Kristiansson E, Sandegren L, Westerlund F, and Ambjörnsson T
PLoS ONE, **2018**, 13, 0193900

XI. Identification of pathogenic bacteria in complex samples using a smartphone based fluorescence microscope
Müller V, Sousa M, Koydemir H, Veli M, Tseng D, Cerqueira L, Ozcan A, Azevedo F, and Westerlund F
RSC Advances, **2018**, 8, 36493-36502

XII. The resistomes of six carbapenem-resistant pathogens – a critical genotype-phenotype analysis
Johnning A, Karami N, Hallbäck E, Müller V, Nyberg L, Pereira M, Stewart C, Ambjörnsson T, Westerlund F, Adlerberth I, and Kristiansson E
Microbial Genomics, **2018**, 4, 000233

XIII. Interspecies plasmid transfer appears rare in sequential infections with extended-spectrum Beta-lactamase (ESBL)-producing Enterobacteriaceae
Lindblom A, KK S, Müller V, Öz R, Sandström H, Åhren C, Westerlund F, and Nahid K
Diagnostic Microbiology and Infectious Disease, **2018**, 18, 30559-5

XIV. Optical DNA Mapping Combined with Cas9-Targeted Resistance Gene Identification for Rapid Tracking of Resistance Plasmids in a Neonatal Intensive Care Unit Outbreak
Bikkarolla SK, Nordberg V, Rajer F, Müller V, Kabir MH, KK S, Dvirnas A, Ambjörnsson T, Giske CG, Navér L, Sandegren L, and Westerlund F
mBio, **2019**, 10, e00347-19

Contribution Report

I. Planned and performed all the nanofluidic experiments, analyzed all the data, contributed to the conceptual design of the data analysis software, and wrote the paper together with FW.

II. Developed the assay together with FW, planned and performed the nanofluidic experiments, analyzed all the data, contributed to the conceptual design of the data analysis software, and wrote the paper together with FW.

III. Designed the study together with FW, planned and performed the experiments together with MN, analyzed the data, contributed to the conceptual design of the data analysis software, and wrote the paper together with MN and FW.

IV. Designed, planned, and performed the experiments, analyzed all the data, contributed to the conceptual design of the data analysis software, and wrote the paper together with FW.

V. Performed all the experiments, the preliminary data analysis, and contributed to writing the paper.

VI. Performed part of the nanofluidic experiments, the preliminary data analysis, and contributed to writing the paper.

VII. Performed part of the experiments, the data analysis and contributed to writing the paper and in the conceptual design of the data analysis software.

VIII. Performed the literature review and wrote the paper together with FW.

IX. Performed the experiments, contributed to the conceptual design of the data analysis software and contributed to writing the paper.

X. Performed part of the experiments, the preliminary data analysis, and contributed to the conceptual design of the data analysis software and in writing the paper.

XI. Contributed to the development of the smartphone-based microscope, designed, and performed experiments, analyzed all the data, and wrote the paper together with FW.

XII. Performed part of the nanofluidic experiments, analyzed the data and contributed to writing the paper.

XIII. Planned and supervised the experiments and wrote the paper together with AL, CÅ, NK, and FW.

XIV. Designed, planned, and performed part of the nanofluidic experiments and contributed to writing the paper.

Table of Contents

1 Thesis Introduction

"DNA neither cares nor knows. DNA just is. And we dance to its music."

– Richard Dawkins, 1994

Deoxyribonucleic acid (DNA), first isolated by Fredrich Miescher in 1869 (2), is often referred to as the *"blueprint of life."* Following the initial discovery, there were many significant scientific contributions before James Watson and Francis Crick in 1953 published the double-helix structure of the DNA molecule (3), based on the X-ray diffraction image taken by Rosalind Franklin and Raymond Goslings a year earlier. Methods have been developed to reveal the sequence of DNA, *i.e.,* the order in which the nucleobases comprising the DNA are put together, to achieve a greater understanding of how humans and other organisms function. The first breakthroughs of DNA sequencing, based on electrophoretic methods, came in 1977 (4, 5), which, after substantial optimization and refinement were used to sequence the first human genomes in 2001 (6, 7). During the next decade, second-generation DNA sequencing methods, based on massively parallel sequencing, radically increased the throughput and reduced the cost

(8, 9). Finally, so-called third-generation sequencing methods have recently been developed, allowing for real-time sequencing of single DNA molecules (10-12).

Although sequencing methods have been refined immensely since 1977, the process is still, in many cases, expensive, as well as time-consuming, and the final assembly, *i.e.*, arriving at the fully resolved sequence, is not always possible. To obtain the intact sequence of a genome, the genomic DNA first needs to be fragmented, each fragment sequenced, and the short sequences (the so-called reads) puzzled back together again. Even if third-generation sequencing methods have provided a drastic increase in read-lengths, the repetitive regions and structural variations commonly found in DNA hampers the assembly process of reads into a complete sequence (13). These structural variations, such as inversions, insertions, deletions, and translocations, are commonly linked to disease in humans (14), and to some extent, the spread of antibiotic resistance (15, 16). Additionally, the lead time and overall complexity limit the applicability of DNA sequencing for bacterial diagnostics and resistance gene detection.

This Thesis aims to describe the development of novel assays based on optical DNA mapping, where the DNA is studied at a lower resolution, but over significantly larger length scales, than with DNA sequencing. Pioneered by Schwartz *et al*. almost 30 years ago (17), optical DNA mapping has proven to be a versatile tool for extracting long-range sequence information with kilobase pair (kb) resolution (18, 19). For the assays described in this Thesis, fluorescently labeled DNA is stretched in nanofluidic channels and imaged using fluorescence microscopy. This approach allows coarse-grained long-range sequence information to be read at the single DNA molecule level, enabling the study of structural variations without averaging, as well as working with low amounts of sample. The developed methodology has many potentials applications, ranging from rapid identification of bacteria and tracing the spread of antibiotic resistance to retrieving new information about the human genome.

The three chapters that follow this general introduction provide a theoretical background to key topics and methodology related to the presented research (**Chapter 2-4**). The original work in this Thesis, which is summarized and discussed in **Chapter 5**, is based on the four appended papers, from here on referred to as **Paper I-IV**.

Paper I demonstrates how optical DNA mapping can be used to study intact bacterial plasmid DNA molecules extracted from clinical isolates originating from a nosocomial outbreak at a neonatal ward at Sahlgrenska University Hospital in Göteborg, Sweden.

The plasmids were traced between different bacterial strains, species, and in turn, between patients, by simultaneously extracting information about the DNA size and primary sequence. **Paper II** adds to **Paper I**, by introducing a CRISPR/Cas9-based approach to detect resistance genes directly from optical DNA maps of single plasmid DNA molecules. While **Paper I-II** focus on bacterial plasmids, **Paper III** utilizes the chromosomal DNA for bacterial identification. Extracting up to megabase pair (Mb) sized DNA molecules enables highly specific identification of bacteria at the species- or even sub-species level. Moreover, the low amount of DNA required for analysis allows for identification of bacteria in non-cultivated clinical urine samples.

Finally, in **Paper IV**, the approach used in **Paper III** to extract large DNA molecules from bacteria is instead applied to human DNA. This assay allows for the mapping of single DNA molecules to the human genome with high precision without the need for enzymatic reactions. The developed assay has the potential to aid in the sequence assembly of complex genomes, facilitate the study of structural variations, as well as to enable co-localization of sequence-based information with epigenetic marks or sites of DNA damage.

Finally, **Chapter 6** provides concluding remarks, as well as an indication about future directions of the research.

2 DNA

During the 20th century, the understanding of DNA and its role in biological systems started to grow. Avery and co-workers showed in 1944 that DNA, and not proteins as previously proposed, were responsible for inheritance (20). Still, the mechanism of genetic heredity remained a mystery until Watson and Crick published the structure of double-stranded DNA (dsDNA, Figure 1) in 1953 (3), followed by a suggestion for how genetic material could be replicated and transferred to the next generation (21). This chapter describes the DNA molecule in terms of its chemical structure and biological function. The chapter moreover briefly covers the physical properties of DNA both in solution and in confinement, as well as interactions between DNA and other molecules.

2.1 Chemical Structure

A DNA molecule consists of a series of nucleotides. A nucleotide is made up of three distinct units: a phosphate group, a deoxyribose sugar, and a nucleobase. In contrast to the phosphate and deoxyribose sugar, which remain the same, there are four different

nucleobases that make up the DNA structure: thymine (T), adenine (A), cytosine (C) and guanine (G), where T and C belong to the pyrimidine family, and A and G are purines. Thus, there are four possible nucleotide versions, each containing one of the four different nucleobases.

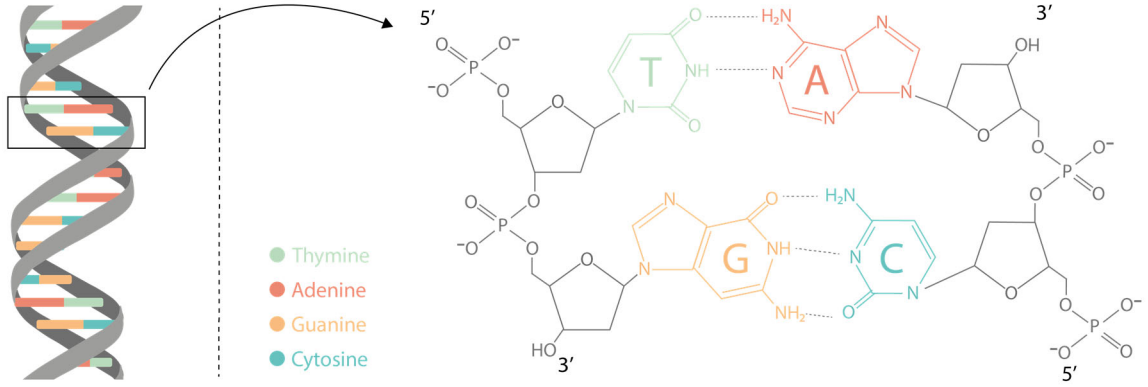

Figure 1: *The molecular structure of DNA. Phosphate and sugar groups (gray) constitute the DNA backbone. They make up the basic building blocks of DNA together with the four DNA bases thymine (green), adenine (red), guanine (orange), and cytosine (blue). Hydrogen bonds connect the two DNA strands (dashed lines), forming a double-stranded DNA helix.*

The nucleotides are connected by phosphodiester bonds between the phosphate and sugars, leading to the formation of single-stranded DNA (ssDNA). Each nucleotide is associated with its neighbor from the 5′ carbon on the sugar of the first nucleotide to the 3′ carbon on the subsequent nucleotide. Thus, ssDNA will have an inherent direction. The two DNA strands of double-stranded DNA are in turn connected by hydrogen bonds between the nucleobases to create the double-stranded helix, in which A base-pairs with T by forming two hydrogen bonds, and G with C, by forming three hydrogen bonds (Figure 1). Multiple natural forms of DNA are known. The B-form, which is by far the most abundant in nature, has an overall negative charge under physiological conditions and displays a right-handed helix structure.

2.2 Biological Function

Following the understanding of the DNA structure, the first version of the central dogma of molecular biology was proposed (Figure 2), explaining the relationship between DNA, ribonucleic acid (RNA), and proteins (22, 23). Proteins are long chains built up by twenty different amino acids linked *via* peptide bonds. RNA is transcribed from DNA, acting as a template for protein translation. Each triplet of DNA bases corresponds to a specific amino acid, and thus the DNA sequence sets the order of the amino acids in a protein.

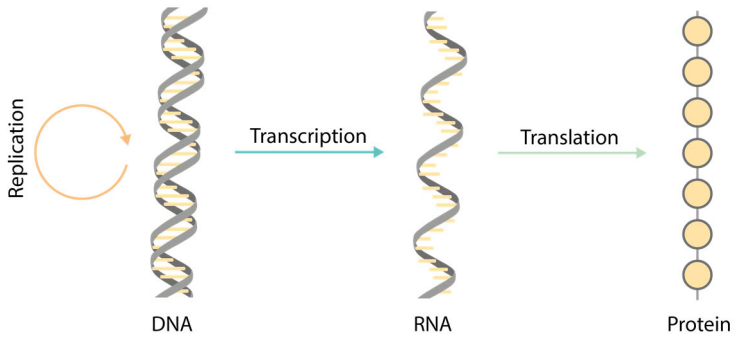

Figure 2: *The central dogma of molecular biology. The genetic information is stored in DNA molecules, which can be replicated. In addition to replication, the DNA can be transcribed into RNA, which in turn can act as a template in protein synthesis via the process of translation.*

Alterations in the DNA sequence will change the composition of amino acids in the resulting protein, which in turn can affect the function of the protein. There are many different types of alterations in DNA, which will be briefly discussed in the two following Sections (2.2.1-2.2.2). Changes in the DNA sequence can harm an organism and are often coupled to disease in humans (Section 2.2.2). However, changes can, in some cases, be beneficial, providing a selective advantage that drives evolution, such as in the case of bacteria and their development of resistance to antibiotics (Section 2.2.1).

2.2.1 Antibiotic Resistance

"A post-antibiotic era—in which common infections and minor injuries can kill—far from being an apocalyptic fantasy, is instead a very real possibility for the 21st century."

– Dr. Keiji Fukuda, World Health Organization, 2014

Antibiotics are antimicrobial drugs commonly used to treat and prevent bacterial infections. Starting with the discovery of penicillin by Fleming almost a century ago (24, 25), antibiotics have become a cornerstone of modern health care. However, severe overuse, both for humans and for livestock, in combination with no new major antibiotic discoveries the past decades, has led to a radical increase in the degree of resistant bacteria worldwide (26, 27). If the development continues, the prognosis is that one person will die every third second due to non-treatable bacterial infections by 2050, exceeding the death rates of cancer globally (27).

Bacteria

Bacteria are single-cell organisms that can cause life-threatening infections in humans. Bacteria belong to the family of prokaryotes, and thus do not harbor membrane-enclosed organelles. Only a few microns large, bacteria were one of the first life forms present on earth, and are found in a variety of different shapes and sizes (28). Bacteria reproduce by a process known as binary fission, where one mother cell divides to form two identical daughter cells.

Figure 3A depicts the main components of a bacterial cell. The cell wall provides structural strength and functions as a barrier that separates the bacterium from its surroundings. On the cell wall surface of some types of bacteria there are small finger-like extensions, known as pili, aiding in the binding of the bacteria to surfaces or in the exchange of genetic material between bacterial cells. The main intracellular component, present in the cytoplasm of the bacteria, is the nucleoid, which besides small amounts of RNA and protein, mainly consists of chromosomal DNA. The cytoplasm also contains ribosomes for protein synthesis. Apart from the chromosomal DNA, responsible for maintaining essential cell functions, bacteria can also store genetic information in the form of extrachromosomal DNA molecules called plasmids.

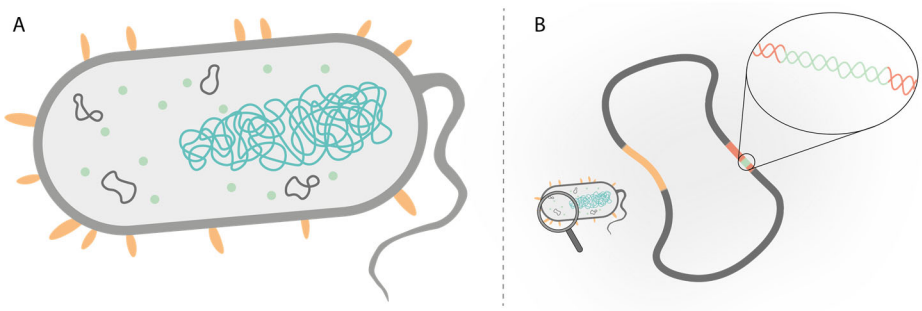

Figure 3: *Basic components and structure of a bacterium. (A) The Illustration shows the cell wall, and flagellum, in gray, the pili in orange, the ribosomes in light green, the nucleoid in green, and the circular plasmids in dark gray. (B) A circular plasmid DNA molecule marked with a region encoding antibiotic resistance in orange and the region initializing replication in red, including the origin of replication in light green.*

Plasmids

Termed by Lederberg in 1952 (29), plasmids are double-stranded, circular DNA molecules, separate from the chromosome, capable of replicating autonomously. Figure 3B shows a schematic representation of a plasmid, including the origin of replication (ORI). The ORI, as indicated by its name, is the DNA sequence at which the replication of the plasmid starts. The minimal portion of the plasmid required for replication, the replicon, includes, besides the ORI, genes encoding proteins that assist in the replication process, as well as in plasmid maintenance. Different plasmids that share one or more elements of the replication mechanism cannot co-exist in the same cell, a phenomenon referred to as plasmid incompatibility (30). The copy number, *i.e.*, the number of copies of a specific plasmid within the cell, can range between a single to over a hundred (31), and the size spans between 1 kb to over 1 Mb.

Even though plasmids are not required for a bacterium to regulate and maintain its essential functions, they play a significant role when a bacterium needs to adapt to a new environment, for instance, when exposed to antibiotics (32). Plasmids easily spread between different bacterial strains and species (33) and are very dynamic, capable of amplifying beneficial genes, as well as deleting those that no longer have any desired

function (15). Similarly, many plasmids can *via* transposons insert external DNA into its sequence, thus acquiring new genetic material (16).

Existing tools for characterizing plasmids allow measuring plasmid size, using gel electrophoresis (34), group plasmids based on specific genes, using polymerase chain reactions (PCR) (35, 36), and retrieve the entire sequence, using DNA sequencing (Section 2.2.2) (9, 37). However, at present, available methods for plasmid characterization that provide enough information for reliable diagnostics suffer from either being slow, complicated, or expensive, so far hampering their use in clinics (38-40).

Gene Transfer

There are two main modes of how bacteria can acquire resistance, vertically and horizontally (41). A vertical transfer is when a spontaneous mutation occurs, rendering a progeny slightly different than its ancestor. Far from all mutations will provide a selective advantage, but by chance a minor fraction of them may allow the bacteria to increase their fitness compared to their previous generation. In contrast, horizontal gene transfer is a process in which a bacterium acquires foreign genetic material. There are three main mechanisms of horizontal gene transfer: transformation, transduction, and conjugation (41, 42). Transformation is a process in which foreign material is taken up by a bacterium from the surrounding environment. In transduction, DNA transfers from one bacterium to another *via* bacteriophages, which are bacterial viruses. Lastly, entire plasmids can be transferred from one bacterium to another in a process known as conjugation (Figure 4A) (32, 41-43).

Conjugation is a contact-dependent process, aided by a special pilus on the donor bacterium, which attaches to the cell wall of the recipient. DNA polymerase replicates the plasmid, and one of the plasmids is transferred to the recipient cell, ending up with one copy in both the donor and recipient. Additionally, two protein complexes, transferasomes and relaxomes, assist in the simultaneous transfer mechanism (43). The machinery for conjugation is encoded on the plasmid, and due to the number of genes required to transfer a plasmid, conjugative plasmids tend to be larger than 30 kb (33). Since the entire plasmid transfers, a plasmid harboring resistance genes against a variety of antibiotics can make a fully susceptible bacterium multiresistant in a single conjugation event.

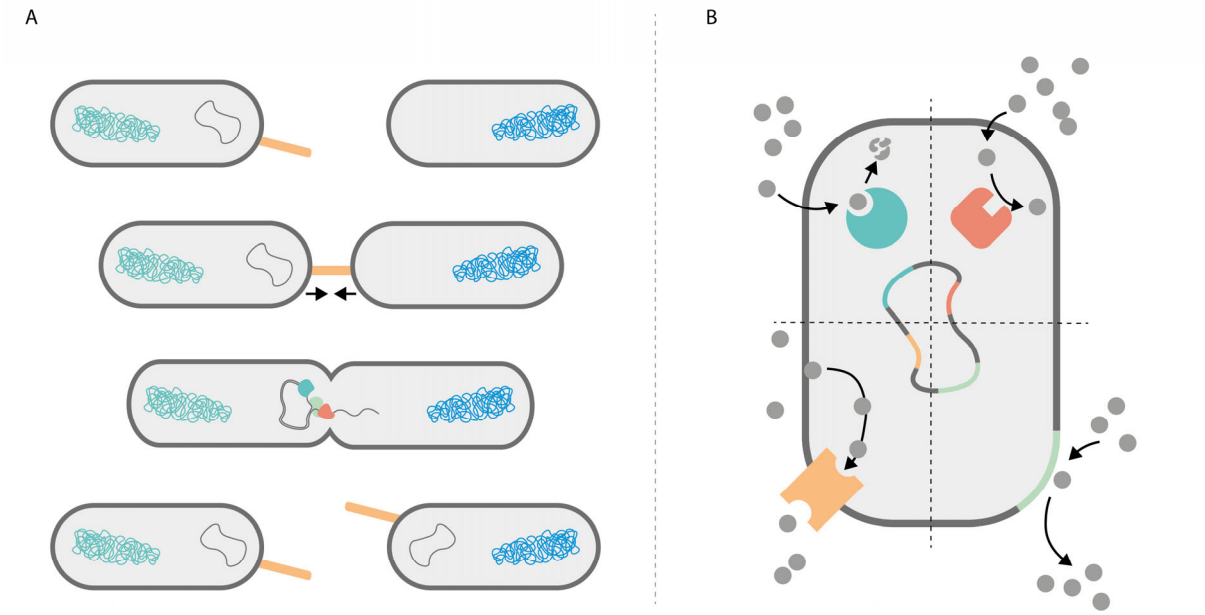

Figure 4: *(A) Plasmid transfer via conjugation. Transfer of a plasmid (dark gray) from a donor cell (green nucleoid) to a recipient cell (blue nucleoid) using pili (orange) for the cell to cell contact and key enzymes DNA-polymerase (green), transferase (light green), and relaxase (red) for plasmid replication and transfer. (B) Bacterial defense mechanisms. A plasmid (dark gray, center) encoding antibiotic resistance via enzymatic degradation (green), modification of drug target (red), altered cell wall structure (light green), and utilization of protein pumps positioned in the cell wall (orange).*

Bacterial Defense Mechanisms

Once a bacterium has acquired a resistance gene against a specific antibiotic, the gene will remove the effect of the antibiotic. However, the mechanism which eliminates the effect of the antibiotic can differ. Figure 4B shows four main defense mechanisms of bacteria to antibiotics (31, 41, 44). The antibiotic usually targets a structural component within the bacterial cell, such as a protein, altering its function, and hence altering the process in which the protein is involved. Targeted cell processes include cell wall synthesis, protein synthesis, RNA/DNA synthesis, or vital metabolic pathways. By changing the antibiotic binding motif of the target without changing the function of the bacterial component itself, the effect of the antibiotic can be neutralized.

Other approaches the bacteria can use as defense include alteration of the cell envelope, hindering the antibiotic from entering, as well as utilizing pumps capable of removing the antibiotic once it has entered the cell. Finally, the bacteria can also produce enzymes capable of chemically altering, or degrading, the antibiotic. One example of such an enzyme is β-lactamase, with the ability to degrade antibiotics belonging to the group of β-lactams, including penicillin. β-lactam antibiotics interfere with the cell wall synthesis of the bacteria by inhibiting the synthesis process (45). However, if the β-lactamase enzyme cleaves the β-lactam ring of the antibiotic, the activity is lost, and the bacteria is left unharmed.

Tackling Antibiotic Resistance

To reverse the trend of increasing resistance to antibiotics worldwide, the problem needs to be tackled on multiple fronts simultaneously (27). New antimicrobial drugs need to be developed, either *via* compounds similar to the traditional antibiotics used today or by novel strategies, including phage therapy, antibodies, and antimicrobial peptides (46). Decreasing the use of existing antibiotics *via* actions such as increased public awareness and new policies for the use of antibiotics in agriculture, together with reducing the spread of resistant bacteria by improved hygiene and sanitary routines, is drastically needed. Finally, there is an urgent need for tools capable of rapid diagnostics of resistant bacteria.

At present, the basis of bacterial diagnostics is mainly empirical observations and culture-based approaches (Figure 5) (27, 40). Diagnostics *via* bacterial cultivation in the presence of different types of antibiotics is an inherently time-consuming process, spanning over more than 36 hours, and is further hampered by pathogenic bacteria frequently failing to grow in samples acquired from patients with clinical signs of infection (40). Rapid diagnostics would be beneficial for patients with severe bacterial infections, allowing the medical doctor to prescribe the most efficient antibiotic immediately. Moreover, rapid diagnostics aids the medical doctor in selecting the narrowest spectrum, while still effective, antibiotic, sparing the use of last-resort antibiotics to when they are genuinely needed.

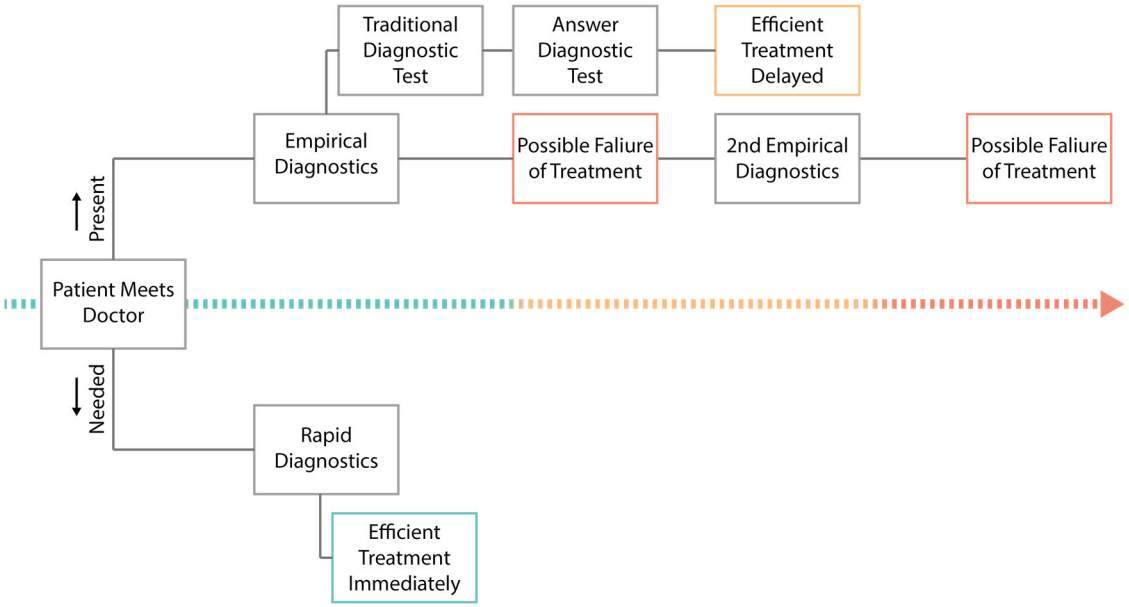

Figure 5: *Present and desired workflow for diagnosis of bacterial infections in the clinical setting.*

One of the main reasons behind the rapid spread of antibiotic resistance worldwide is bacterial plasmids (16, 38, 47). The ability of a plasmid to transfer between different strains and species *via* horizontal gene transfer, as well as their very dynamic nature, have made plasmids ideal for acquiring antibiotic resistance genes. Therefore, not surprisingly, plasmid-mediated resistance has recently been reported for last-resort antibiotics, such as polymyxins (48) and carbapenems (47), emphasizing the need for efficient methods to monitor and study plasmids.

2.2.2 The Human Genome

The human genome consists of 3.23 gigabase pairs (Gb) of DNA, divided into 23 pairs of chromosomes. The DNA sequence consists of both coding and non-coding regions, where coding regions, so-called genes, can be defined as DNA sequences that are transcribed and translated into proteins. Additionally, alterations made to the DNA which do not affect the DNA sequence, so-called epigenetic marks, can have a significant impact on which parts of the coding sequences that in the end are transcribed (49).

Each time a cell divides, the DNA replicates in order for both the mother and daughter cell to have one copy each of the genome, a procedure which is the basis of biological inheritance. DNA replication is a complex process involving a variety of different proteins, such as DNA topoisomerase for unwinding of the dsDNA, DNA helicase for breaking the hydrogen bonds between the base pairs, and DNA polymerase for synthesizing the new DNA strand. Even if the replication of DNA is a highly precise process, errors will occur over time, which, in the worst case, can cause genetically linked disorders (50). Moreover, the DNA in our cells is damaged frequently by metabolic compounds and byproducts, such as reactive oxygen species, or *via* external factors, such as UV light and ionizing radiation (51). Damages to the DNA include altered DNA bases, as well as single- and double-stranded breaks. In many cases, the DNA is repaired by proteins in the cell, but the repair process is far from trivial and can result in permanent alterations to the DNA sequence (51).

While sequence variation many times occurs at the single nucleotide level, referred to as single nucleotide polymorphism (SNPs), large scale structural variations in the human genome are also frequently discovered and are many times linked to genetic disease (14). Structural variations include insertions, deletions, and inversions of sequences, as well as translocations, in which a part of the DNA sequence shifts place within the genome (Figure 6). Moreover, alterations in the number of copies of a gene, known as copy number variations (CNVs), can be observed when comparing human genomes (14).

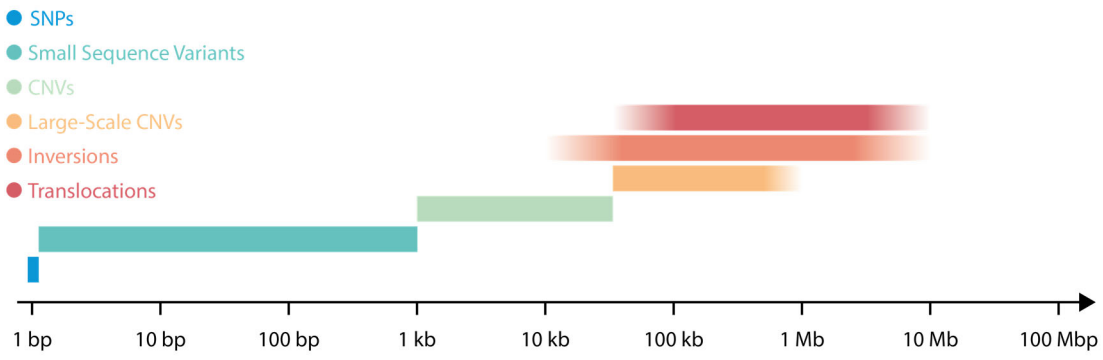

Figure 6: *The size of different structural variations in the human genome (14). Including single-nucleotide polymorphism (SNPs), small sequence variants, copy number variations (CNVs), inversions, and translocations.*

Methods to reveal information about the DNA sequence are needed to gain knowledge about the human genome, as well as diagnose diseases based on genetic alterations and rearrangements. The sequencing of DNA has developed immensely during the past decades, reaching a significant milestone back in 2001 when the so-called *Human Genome Project* resulted in the first human genome sequences (6, 7).

Reading the DNA Sequence

The first significant breakthrough of DNA Sequencing came in 1977, where two methods, both based on electrophoretic size determination, were introduced. Sanger and Coulson developed the chain terminator procedure (4), while Maxam and Gilbert used a chemical cleavage approach (5). These basic principles were after substantial optimization and refinement during the next two decades used in the *Human Genome Project* (6, 7).

At present, next-generation sequencing (NGS) approaches, many times referred to as second-generation DNA sequencing, are most extensively used for sequencing DNA (9, 52, 53), and even if multiple strategies exist, most of them rely on the same basic principle. First, the DNA is fragmented into smaller pieces, generally from ten up to a couple of hundreds of bp in size (13). The fragments are then amplified using PCR and later sequenced individually. In order to obtain the full DNA sequence, these individual fragments, termed reads, need to be assembled. By generating many copies of the DNA sequence of interest, and relying on the fragmentation process to be random, the idea is that there will be an overlap between different reads, making them possible to puzzle back together. New so-called third-generation DNA sequencing technologies (11, 12), relying on high-quality samples, have made it possible to generate longer read lengths, aiding in the assembly process of complex sequences. However, the process is, unfortunately, in many cases, still far from trivial due to the high variability of the DNA sequence.

For reliable sequencing results, a general guideline is that each nucleotide should be sequenced several times, referred to as coverage. This requires multiple copies of each genome in order to assemble the sequence with high accuracy. The need for averaging to remove potential errors in the obtained nucleotide sequence reduces the possibility to study heterogeneity between different cells.

An alternative method for studying large rearrangements in a genome is fluorescence *in-situ* hybridization (FISH) (54). Using fluorescently labeled probes designed to hybridize to specific regions of a chromosome, targeted rearrangements can be studied with a conventional fluorescence microscope (Section 3.3). Even if FISH provides advantages, such as whole-genome coverage and high sensitivity and specificity, the resolution (~ 100 kb) is often too low, and the analysis usually time-consuming.

2.3 Physical Properties

When studying the behavior of DNA molecules at larger length scales, it is not enough to consider only the basic chemistry of the nucleotides (Section 2.1). Instead, a model of the DNA is used where the large polymer is built up by many monomeric subunits and is characterized by its size, shape, and molar mass. The size and shape of the DNA polymer depends on the surrounding environment, while the number of monomers determines the molar mass. Under physiological conditions, the net charge of DNA is negative as a consequence of the phosphate groups that make up the DNA backbone. B-form DNA is a right-handed helical structure with a helix width of 2 nm, and a twist angle of 36 degrees and a rise of 0.34 nm per base pair (55).

In order to describe the behavior of DNA, both free in solution and under confinement, some basic parameters need to be introduced. The contour length (L) of DNA is defined as the end to end distance of the DNA when fully stretched and is calculated by multiplying the number base pairs with the rise per base pair. The persistence length (P) of B-DNA, defined as the length of the unit over which the DNA cannot be bent, is approximately 50 nm under physiological conditions (56, 57) and varies with the ionic strength. A DNA molecule free in solution will coil to minimize its free energy. The size of the coiled DNA is described by the radius of gyration (R_G), which is a measure of the root mean square distance of the segments to the center of the coil.

2.3.1 Confined DNA

The work in this Thesis is based on confining DNA inside nanochannels (Section 4.1 and 5.1.1), in which the length of the channels is assumed to be infinitely long, while the width and height of the channels are restricted. The dimensions are generally expressed as the average cross-section area (D) of the channel width (D_w) and height (D_h), and calculated according to Equation (1).

$$D = \sqrt{D_w \cdot D_h} \quad (1)$$

The degree of extension in the channels will depend on the persistence length (P), channel dimensions (D), and the radius of gyration (R_G) of the DNA molecule (Figure 7). If the dimensions of the channels are larger than the radius of gyration ($R_G < D$), the DNA will act as when free in solution. If the channel dimensions are smaller than the radius of gyration, but still much larger than the persistence length ($P \ll D < R_G$), the DNA will be weakly confined and can be seen as a series of non-interacting blobs with diameters similar to the cross-section area, as explained by the theory of de Gennes (58). If the channels instead have dimensions smaller than the persistence length ($D < P$), the DNA molecule will behave according to the theory described by Odijk (59, 60). Here the DNA cannot coil up anymore, and its movement is restricted to only small undulations (Figure 7).

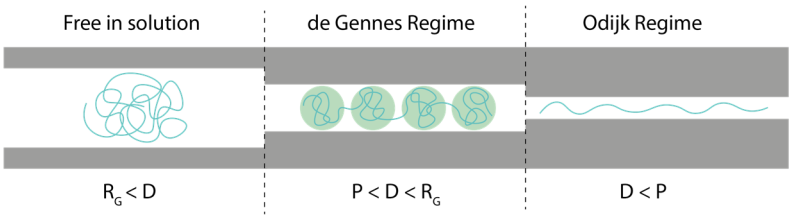

Figure 7: *Illustration of the behavior of DNA under different degrees of confinement. The predictions are based on the radius of gyration (R_G), the average cross-section of the confining environment (D), and the persistence length (P).*

For nanofluidic experiments involving DNA molecules, as for the work in this Thesis, the channel dimensions are many times similar to the persistence length ($P \leq D$), resulting in experiments conducted in a regime between the ones described by de Gennes and Odijk, commonly referred to as the extended de Gennes regime (61-65). Notably, the extension in the regimes described by both de Gennes and Odijk is directly proportional to the contour length of the DNA. Hence, sequence-specific information positioned X% into the sequence can, on average, be observed X% into the nanoconfined DNA molecule.

Since the physical properties of DNA are sensitive to the concentration of ions, the degree of the extension will depend on the ionic strength of the solution (66). For example, a lower concentration of ions in the solution will provide less shielding for the DNA from its own negatively charged backbone, increasing the stiffness, and thus increase the degree of extension.

2.4 DNA – Molecule Interactions

Small molecules and proteins, such as enzymes, are essential for the biological function of DNA. Thus, the interaction between DNA and other molecules is crucial for all organisms. The versatile chemical and structural properties of DNA provide a variety of possibilities for different molecules to interact.

2.4.1 Binding Modes

There are four main ways a molecule, here referred to as a ligand, can bind to DNA: external biding, binding in the minor or major groove of the DNA helix, and intercalation between the DNA base pairs (Figure 8A). Typically, electrostatic interactions between a positively charged ligand and the negatively charged DNA backbone facilitate the external binding. Groove binders, on the other hand, rely on forming hydrogen bonds with the exposed bases in either the minor or major groove. The minor groove allows for smaller, often curved, molecules, to bind along the twisted DNA helix. In contrast, larger molecules, such as proteins, are instead more frequently found to bind in the major groove, which can harbor their larger structural motives. Lastly, a ligand can insert one or several parts of its structure in the hydrophobic space created in-between the stacked DNA base pairs, a process referred to as intercalation.

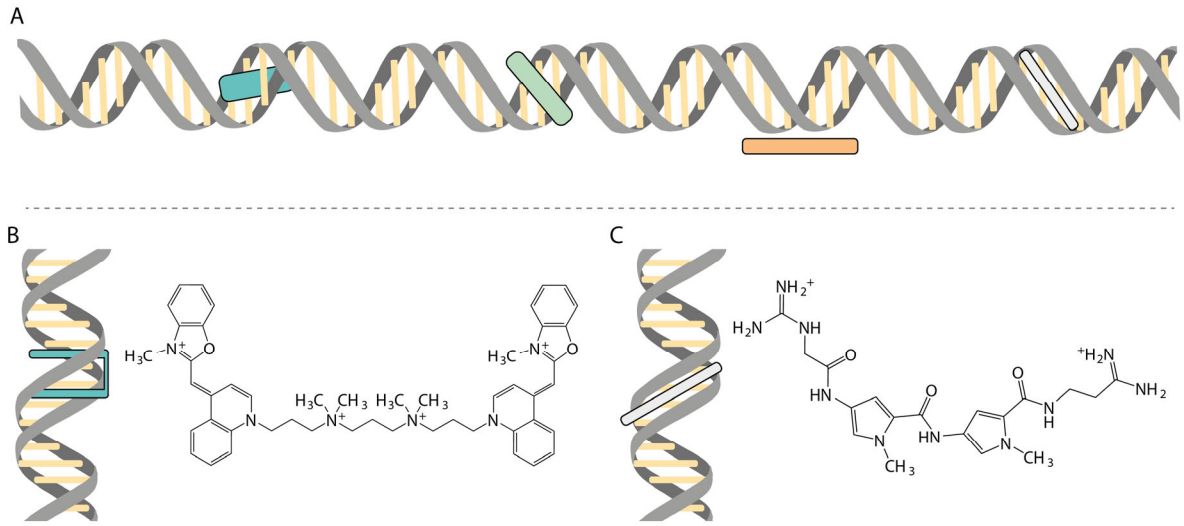

Figure 8: *(A) Modes of binding to DNA. Intercalation (green), major (light green) and minor (light gray) groove binding, and external binding (orange). Binding mode and molecular structure of YOYO-1 (B) and netropsin (C).*

For the work presented in this thesis, two DNA binding molecules have been used, the bis-intercalating cyanine dye YOYO-1 (YOYO), and the AT-selective groove binder netropsin (Figure 8B-C). When studying DNA with fluorescence microscopy (Section 3.3), the inherently non-fluorescent DNA needs to be stained with a dye to enable visualization. For this purpose, cyanine dyes, such as YOYO, provide high binding constants and display favorable photophysical properties, including high emission quantum yields, and low background fluorescence. Details on the emission of light, including fluorescence, can be found in Chapter 3.

YOYO binds to DNA by inserting its two aromatic ring structures in-between the DNA base pairs, a process referred to as bis-intercalation (67, 68). When free in solution, YOYO is virtually non-fluorescent, but once bound to DNA, the rotation around the methine bridge is sterically hindered, increasing the emission more than a thousand-fold (67), making YOYO ideal for fluorescence microscopy experiments. The maximum amount of YOYO bound to DNA is approximately one YOYO molecule per four base pairs of DNA. External (electrostatic) binding to the negatively charged DNA backbone is still possible; however, at much lower affinities (68, 69). The binding of YOYO will also affect the structural properties of DNA. When inserting the two aromatic moieties between the

DNA base pairs, YOYO will increase the contour length by 0.51 nm, as well as unwind the DNA helix by approximately 24 degrees (70, 71). Also, the positive charge (+4) of YOYO will reduce the overall negative charge of DNA.

Netropsin is a non-fluorescent, natural antibiotic, which binds selectively to AT base pairs in the minor groove of the DNA helix. While the molecular structure of netropsin (Figure 8C) allows it to form hydrogen bonds with adenine and thymine, giving rise to a site-specific binding, the amino group present in guanine will sterically hinder netropsin from forming hydrogen bonds to GC base pairs (72, 73). The strength of binding between netropsin and DNA will depend on a 5bp site to which netropsin binds. The sequence motif 5′-AAAAA-3′ provides the strongest interaction, while the presence of a guanine or cytosine base in a central position will reduce the binding affinity more than by a factor 100 (74-76).

2.4.2 Enzymes

As previously mentioned, proteins are long chains of amino acids linked *via* peptide bonds. Proteins function as the building blocks of the cell, capable of performing a variety of functions. Different groups of proteins include transporter proteins, receptor proteins, structural proteins, and enzymes. The latter function by accelerating chemical reactions, regulating a myriad of processes in a highly selective manner within a living cell. For instance, enzymes are essential in the process of replication, repair, and modification of DNA.

One example of an enzyme involved in DNA modification is the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) - associated protein-9 nuclease (Cas9) (77, 78). Cas9 has recently been developed into a gene-editing tool, capable of catalyzing highly sequence-specific double-stranded breaks, *via* the cleavage of the phosphodiester bonds in the DNA backbone (79, 80). The enzyme is involved in the adaptive immune system of bacteria and archaea (77, 81) and has found numerous scientific applications, including gene editing, targeting, and regulation (82).

In order for the Cas9 enzyme to find the specific DNA sequence at which it will produce a double-stranded break, it utilizes an RNA molecule with a 20 bp motif complementary to the site of interest (Figure 9). This RNA molecule is called guide RNA (gRNA), and can be further dived into two parts, termed tracrRNA and crRNA. The tracrRNA is responsible for the association to the Cas9 enzyme, while the crRNA provides the

complementary 20 bp sequence to the targeted site on the DNA molecule. Thus, the sequence of the tracrRNA will stay the same, while the crRNA sequence can be selected depending on the site of interest on the DNA molecule. The only constraint when selecting the crRNA sequence is that it has to be followed by a three-base sequence (NGG), termed protospacer adjacent motif (PAM). The flexibility in the design of the crRNA allows targeting of nearly any gene with high precision.

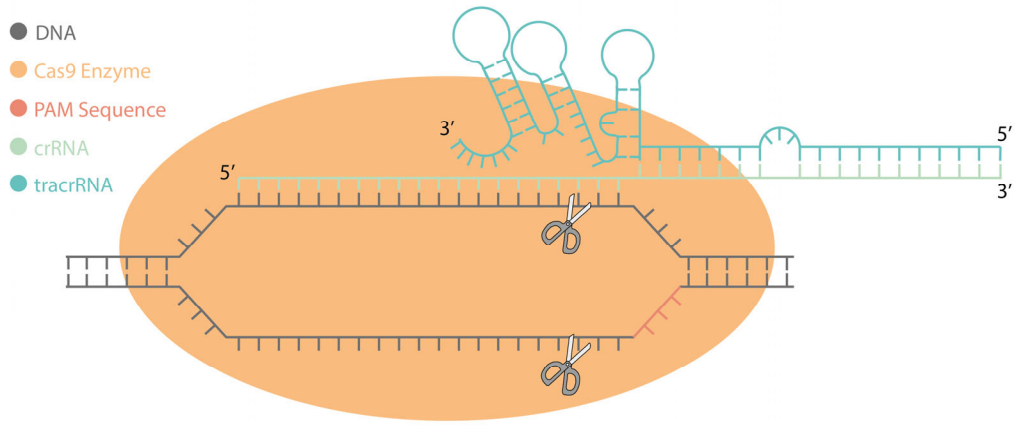

Figure 9: *Sequence-specific double-stranded DNA break mediated by Cas9. The Cas9 protein, together with a guide RNA (crRNA + tracrRNA), scans the DNA for a sequence complementary to the crRNA and cuts the DNA on both strands approximately three nucleotides upstream of the protospacer adjacent motif (PAM) sequence.*

Forms of Cas9 have also been engineered that selectively bind to DNA without producing a double-strand break (83) or those only capable of producing a single-strand break. The first has been used in order to visualize sequence-specific sites (84, 85), and the latter to induce double-stranded breaks with higher precision (79, 86), or to mark the site of the single-strand break by incorporating a fluorescent dye (87).

Other enzymes capable of altering DNA, but with shorter recognition sites, such as nicking-enzymes, are also readily available and have found several applications, for instance, within the field of optical DNA mapping, further discussed in Chapter 4.

3 Fluorescence

The phenomenon of fluorescence was described already in 1845 by Sir John Frederick William Herschel (88) when observing a solution containing quinine irradiated by sunlight. However, scientific knowledge at the time was not enough to explain the phenomenon observed. Now we know that fluorescence is the emission of light from a molecule when returning from a higher energy state to its natural ground state. In this chapter, the principles of light-matter interactions are briefly described (Section 3.1) alongside the properties of a useful fluorescent molecule, a fluorophore. (Section 3.2). Furthermore, an overview of the basic principles behind, and components of, a fluorescence microscope is provided (Section 3.3).

3.1 Emission of Light

The interaction between light (photons) and matter will depend on the electronic configuration of the irradiated molecules in the sample. The energy of the photons, described by their wavelength, can be absorbed by a molecule *via* the excitation into a

higher molecular state. For excitation to occur, the energy of the photon needs to match the energy gap between an excited state of the molecule (S_n) and the molecule's ground state (S_0). A Jablonski diagram, shown in Figure 10, illustrates the different excitation and de-activation processes occurring when a molecule interacts with light.

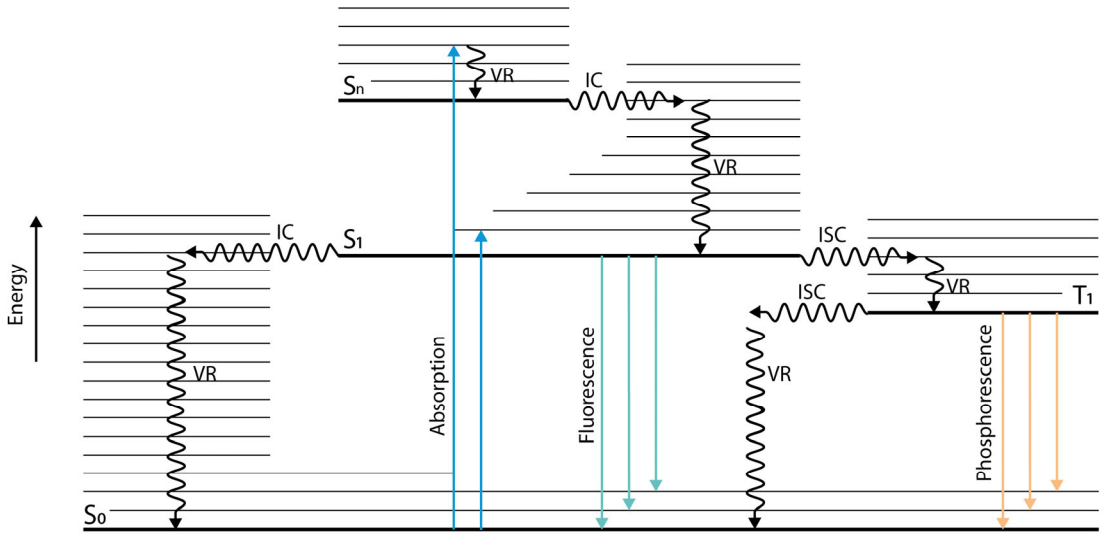

Figure 10: *A Jablonski diagram depicting the excited-state deactivation pathways of a molecule. Singlet (S_0, S_1, S_n) and triplet (T_1) states, together with corresponding vibronic states, are shown as straight horizontal lines. Undulating arrows show the non-radiative pathways vibrational relaxation (VR), internal conversion (IC), and intersystem crossing (ISC). The straight vertical arrows show the radiative processes absorption (blue), fluorescence (green), and phosphorescence (orange).*

When a molecule absorbs a photon, the absorbed energy will lead to a reconfiguration of the electrons, resulting in an excited state of the molecule. The excited states, both singlet (S_n), with paired electrons, and triplet (T_n), with unpaired electrons, are numbered in a consecutive order based on their energy gap relative to the ground state. Each electronic state is also associated with several vibrational states, which are typically closer in energy.

Following excitation to S_n, the molecule will undergo rapid relaxation to the lowest vibrational energy of the first excited state (S_1), *via* a combination of internal conversion and vibrational relaxation. Hence, excited-state processes can be assumed to only occur from the lowest vibrational state of S_1, a phenomenon referred to as Kasha's rule (89).

Once in S_1, the molecule can return to the ground state *via* three main routes: non-radiative relaxation, emission from the first excited singlet state (S_1) and decay through the first excited triplet state (T_1). For non-radiative relaxation, the molecule undergoes internal conversion from S_1 to S_0, followed by vibrational relaxation (Figure 10), dissipating the excess energy to the surrounding environment. Deactivation *via* radiative relaxation from in S_1 to S_0 is termed fluorescence, and results in an emitted photon. Deactivation *via* T_1 involves intersystem crossing, where the molecule undergoes a spin-forbidden transition, rendering a state with unpaired electrons. From T_1 the molecule can either return to its ground state by emitting a photon, known as phosphorescence, or by relaxing non-radiatively (Figure 10). Which deactivation pathway that dominates will depend on the structure of the molecule, temperature as well as the surrounding environment.

The wavelength of an emitted photon is longer than the wavelength used for excitation. For fluorescence, this difference in wavelength is referred to as the Stokes shift (90). The Stokes shift can be explained by the rapid relaxation from S_n to S_1, the tendency of a molecule to return to an excited vibrational state of S_0 upon emission, as well as excited-state reactions and solvent relaxation effects. The number of emitted photons will depend on the ability of a molecule to absorb light, which is governed by the molar absorptivity (ε) at each wavelength. The number of emitted photons will, moreover, depend on the fraction of absorbed photons that result in an emission event, measured as the fluorescence quantum yield (ϕ_f).

3.2 Fluorophores

The preferred properties of a fluorescent molecule, a so-called fluorophore, vary depending on the application. In general, a high quantum yield, in combination with a high molar absorptivity, and a high photostability, is desirable, providing a bright and stable fluorophore. Besides brightness, a significant difference in the emission of a fluorophore when bound to a target molecule compared to when free in solution is advantageous in many applications. Furthermore, a distinct Stokes shift allows the separation of excitation and emission more efficiently. For instance, the fluorophore YOYO, used in the work presented in this Thesis, has a quantum yield of 0.52, a molar absorptivity of 10^5 $M^{-1}cm^{-1}$ at 491 nm, and an emission maximum at 509 nm when bound to dsDNA (67). When selecting excitation wavelength and optical filters for fluorescence

microscopy (Section 3.3), it is essential to know the excitation and emission spectra of a fluorophore. Figure 11 shows the excitation and emission spectra of YOYO.

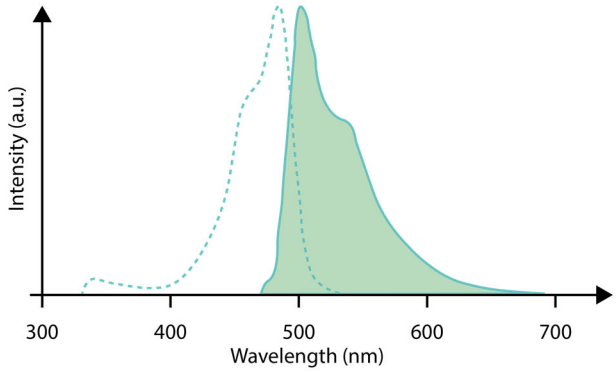

Figure 11: *Absorption and emission spectra. The absorption (dashed line) and emission (solid line, filled) spectra of the fluorescent molecule YOYO-1 when bound to DNA.*

3.3 Microscopy

Fluorescence microscopy has been extensively used throughout the work presented in this Thesis. In contrast to other optical microscopy techniques, relying on reflection and absorption, fluorescence microscopy utilizes emission to visualize a substrate. For inherently non-fluorescent molecules, such as DNA, a fluorophore can be used to enable visualization. Figure 12 shows the principle of an inverted fluorescence microscope, such as the one used for the work presented in this Thesis.

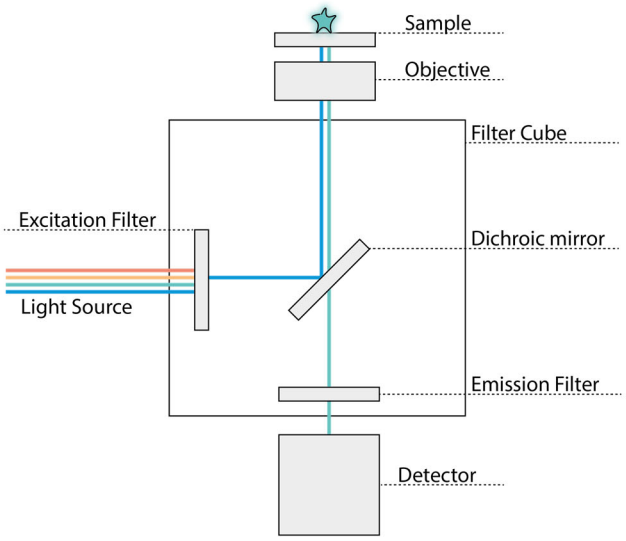

Figure 12: *The main components of an inverted fluorescence microscope.*

The main components of a fluorescence microscope include an excitation light source, two filters, one for excitation and one for emission, an objective to focus the light onto the sample, as well as a detector to record the emitted photons. The light source can either be monochromatic or contain a broad spectrum of wavelengths. Excitation and emission filters are usually combined with a dichroic mirror, possessing the ability to reflect excitation light and transmit emitted light. These three components make up a so-called filter cube, which can be changed to match the maximum excitation and emission wavelengths of the selected fluorophore.

In short, light is sent through the excitation filter to match the excitation wavelength of the fluorophore of interest. The transmitted light is reflected on a dichroic mirror and focused through an objective before reaching the sample. The back-scattered fluorescence emission is then collected through the same objective and transmitted through the dichroic mirror and emission filter before reaching the detector. Detection is typically enabled *via* a sensitive camera, such as a complementary metal-oxide-semiconductor (CMOS) camera, or an electron-multiplying charged coupled device (EMCCD) camera.

4 Optical DNA Mapping

Optical DNA mapping is based on visualizing long-range sequence information on individual DNA molecules with a resolution on the kilo base-pair level. The DNA is first labeled in a sequence-specific manner, stretched, and then imaged, generally using a fluorescence microscope. Since the pioneering work by Schwartz *et al*. (17), who utilized restriction enzymes to cut the DNA molecules at specific sites and visualized the length of the resulting fragments in an agarose gel, the field of optical DNA mapping has advanced significantly. In 2013, *Bionano Genomics*® was the first company to commercialize optical DNA mapping, using labels produced by nicking enzymes and DNA stretched and visualized in nanofluidic channels.

The advantages of using optical mapping to study DNA include remarkable read lengths of up to several Mbp, as well as extracting information on single-molecule variations by avoiding averaging. The length scales covered by optical DNA mapping bridge the gap between sequencing-based techniques and FISH (Section 2.2.2). Below follows a brief overview of the progress which has been made within the field, including strategies for

stretching (Section 4.1) and labeling (Section 4.2) of DNA. For a more detailed description, the reader is referred to the following reviews (18, 19, 91-95).

4.1 Stretching DNA

There are different approaches for stretching individual DNA molecules, including the use of magnetic or optical tweezers, flow, molecular combing, and nanofluidic channels (96). For optical DNA mapping, stretching DNA on modified glass surfaces, known as molecular combing, or extending DNA in nanofluidic channels, are the most common approaches. Molecular combing (97, 98) utilizes functionalized glass surfaces to which the ends of DNA molecules can attach. In the presence of flow, the forces at the air-water interface will stretch the DNA along the flow direction. A variety of different approaches have been developed to achieve a uniform flow, including, retracting a glass slide from a solution (99, 100), dragging a droplet over a glass surface (101), utilizing the gravitational flow of a droplet on a tilted glass slide (102), and droplet evaporation directly on a functionalized glass surface (103).

While molecular combing offers a rapid overview of many molecules, in combination with being readily accessible, uniform stretching is not guaranteed. Instead, DNA can be confined in nanofluidic channels (104), where the ends of the DNA are free, and the DNA uniformly stretched. Nanofluidic devices also offer the possibility of automation, allowing for high throughput applications. Techniques for fabrication of nanofluidics chips have been established utilizing the advances made in the field of microelectronics (91). Generally, the nanofluidic chips are fabricated from oxidized silicon, also known as silica, due to its great availability and favorable properties. The low surface roughness and negative charge of the silica prevent DNA molecules from adhering to the channel walls. Also, the hydrophilic nature of the material promotes wetting once buffer is added to the chip.

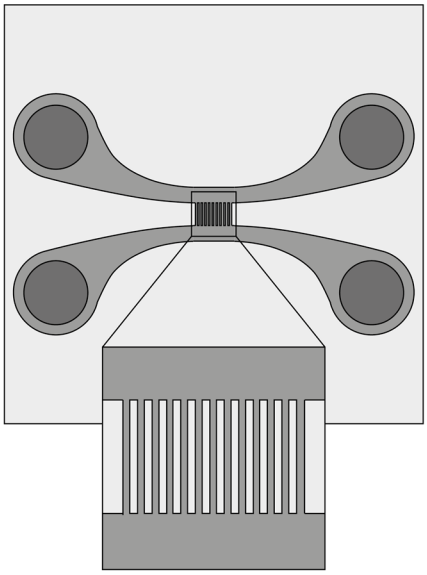

Figure 13: *Sketch of a typical nanofluidic chip. The four circular loading wells are connected two and two by microchannels, which in turn are spanned by a set of nanochannels.*

The nanofluidic chip used for the work presented in this Thesis (shown schematically in Figure 13) has four wells for sample loading, connected pairwise *via* microchannels. The microchannels are, in turn, spanned by a set of typically 100 nanochannels with dimensions of 100 x 150 nm (width x height) and a length of 500 μm. The nanofluidic chips are made from a 500 μm thick silicon wafer using standard cleanroom procedures including thermal oxidation, photolithography, E-beam lithography, reactive ion etching (RIE), and fusion bonding. In order to allow visualization of the DNA inside the nanofluidic channels, the chip is sealed with a transparent Pyrex glass lid.

Over the years, various designs of nanofluidic chips have been demonstrated, tailor-made for their specific applications. Approaches include, tapered nanochannels which allow to measure the extension of a DNA molecule at different degrees of confinement (105), entropic gradients lowering the entropy barrier by pre-stretching the DNA prior to entering the nanochannel (106), as well as meandering channels in order to visualize Mb sized DNA molecules in one single field of view (107). Chips with pressure-based tunable confinement, allowing combined single-cell manipulation and DNA molecule analysis, have also recently been reported (108). Moreover, post-processing procedures, such as lipid passivation (109), are available and can prevent proteins and other biological

molecules from adhering to the channel walls. The DNA can be controlled inside the nanofluidic chip *via* an electric field or by pressure-driven flow. For more details on the fabrication process of nanofluidic chips and fluid control, the reader is referred to the following reviews (56, 91, 110).

An experimental set-up based on fluorescence microscopy was used for the work presented in this Thesis to visualize and image the DNA when confined in the nanofluidic channels. The chip is mounted onto a holder, which, in turn, is placed on the microscope (Section 3.3). Besides functioning as a support for placing the chip in contact with the objective of the microscope, the holder provides access to the loading wells of the chip (Figure 13), which can be filled with a sample and pressurized with N_2 to flow DNA inside the chip.

4.2 Labeling Strategies

Different approaches for labeling DNA for optical mapping have been developed. The labeling strategies can be divided into two main categories, enzyme-based labeling, and affinity-based labeling, illustrated in Figure 14. Introduced by Jo *et al*. (111), nick labeling is based on incorporating fluorescently labeled nucleotides at sequence-specific sites (Figure 14). First, a nicking enzyme generates a single-stranded break, referred to as a nick, at its recognition site, typically 4-7 bp long. Next, a DNA polymerase is used to remove the original nucleotides while subsequently incorporating new fluorescently labeled nucleotides along the strand. To determine the start and end of each molecule in the fluorescence images, YOYO is used to stain the full contour of the DNA molecule. Prior to labeling, existing nicks can be repaired, avoiding false positives. Moreover, the excess nucleotides can be removed after the labeling reaction to decrease the background fluorescence. If two nicking sites at opposite strands are nearby, a dsDNA break will occur. Thus, dense labeling will come at the cost of mapping shorter DNA fragments.

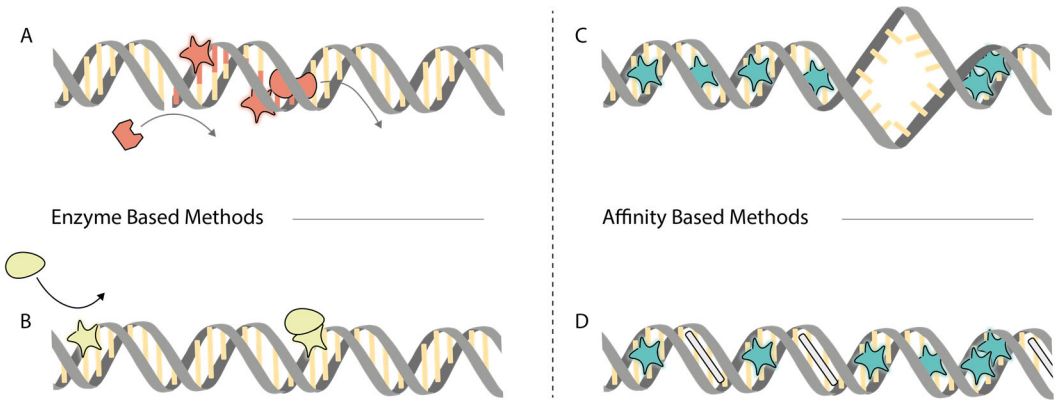

Figure 14: *Overview of the main labeling strategies used in optical DNA mapping. (A) Nick labeling, where a nicking enzyme creates a single-strand break at its recognition site, ~ 4-7 bp in size. A DNA polymerase then removes the original nucleotides and incorporates new fluorescently labeled nucleotides to visualize the site. (B) Methyltransferase based labeling, where an enzyme transfers a co-factor to a methylation site. A fluorescent label is either pre-attached to the co-factor or attached after the co-factor transfer. (C) Denaturation mapping, where the lower melting temperature of AT base pairs compared to GC base pairs, is utilized. DNA is stained with the fluorescent dye YOYO-1, which prefers dsDNA over ssDNA. A careful increase in temperature results in the dissociation of YOYO-1 from the locally melted AT base pairs. (D) Competitive binding based labeling, where YOYO-1 and the AT-specific, non-fluorescent, molecule netropsin compete in binding to DNA. Netropsin will bind to AT sites, blocking these from YOYO-1, creating an emission intensity profile along the DNA where AT-rich regions are dark, and GC-rich regions bright.*

Das *et al*. later extended the nick labeling assay, introducing a nick-flap labeling scheme (112). By using a DNA polymerase lacking exonuclease activity, the nucleotides following the nicking site are displaced rather than removed, leaving a single-stranded DNA flap. The displacement allows for dual-labeling, including the nicking site, as well as the flap, using a complementary, labeled oligonucleotide. Hastie *et al*. used a different dual-labeling approach, using two nicking enzymes with different recognition sites (113). Performing the reactions consecutively, with an intermediate purification step, allowed the use of differently labeled nucleotides and, thus, a dual-color readout. To address the limited specificity (4-7 bp) of the available nicking enzymes, McCaffery *et al*. further extended the nick labeling assay using a CRISPR/Cas9 based approach, enabling the use of a highly adaptable 20 bp long recognition site (Section 2.4.2) (87, 114). The same specificity is also possible to obtain by instead utilizing a dead, fluorescently labeled, Cas9

protein which only binds to the sequence of interest (115). Nick labeling, in combination with nanochannel based DNA extension, has been commercialized *via Bionano Genomics*®, allowing for a range of applications, including the study of genetically linked disease and de novo assembly of complex genomes, where the optical maps are used as templates (106, 116-135).

Similar to nick labeling, methyltransferases can be used to label DNA in a sequence-specific manner (136, 137). The enzyme recognizes a methylation site, typically four bp long, and transfers a co-factor to the DNA. A fluorescent label is either pre-attached to the co-factor or attached after the co-factor transfer (Figure 14). Vranken *et al*. used this approach to achieve optical maps with nanometer resolution (138). Methyltransferase labeling has also been used to generate densely labeled DNA profiles, so-called amplitude modulation profiles, taking advantage of the four bp recognition site of the enzyme used, which facilitates the dense labeling (139, 140). In addition to nickase and methyltransferase based labeling, there is also the possibility to extract information beyond the DNA sequence. Applications include epigenetic marks in the form of methylations (141), and hydroxymethylations (142-145), as well as histone modifications (146). As an example, Sharim *et al*. recently presented a reduced representation optical methylation mapping assay to report on the methylation status of a genome (147). Furthermore, DNA damage sites have been successfully fluorescently labeled and visualized after stretching the DNA molecules on a glass surface (148).

Optical mapping can also be performed without the use of any enzymes. Denaturation mapping, developed by Reisner *et al*. (149), utilizes the lower melting temperature of AT base pairs, which are connected by two hydrogen bonds, compared to GC base pairs, which have three hydrogen bonds. By staining the entire DNA contour with the fluorescent dye YOYO and carefully increasing the temperature, AT base pairs will melt, causing YOYO, which prefers dsDNA over ssDNA, to dissociate. The dissociation will create an emission intensity profile along the contour of the DNA, where AT-rich regions will be dark, and GC-rich regions bright. The DNA can be heated either on-chip (149, 150) or before sample loading (151, 152). Denaturation mapping was recently combined with on-chip single-cell lysis and DNA extraction, revealing structural variations at the single-cell, single-molecule level (153).

Finally, a complementary approach for labeling DNA, based on competitive binding, has been established (154). When simultaneously adding the AT specific, non-fluorescent, molecule netropsin, and the fluorescent dye YOYO, the molecules will compete for the

binding sites on the DNA. If added in a correct ratio, netropsin will block many of the AT-sites, preventing YOYO from binding, rendering an emission intensity profile along the DNA, where AT-rich regions will appear dark, and GC-rich regions bright. With the competitive binding assay, which has been used throughout the work presented in this Thesis, there is no need for repairing the DNA before labeling, performing additional purification steps, or precise temperature control.

5 Results and Discussion

In this chapter, the work on which this Thesis is based will be discussed and summarized. An overview of the results presented in the appended papers is provided, starting with a brief introduction to the developed assays, followed by some general information on the experimental procedure and subsequent data analysis. For further details on the presented research, the reader is referred to the appended papers.

5.1 Method Development

The work presented in this Thesis has resulted in three different assays focusing on characterizing plasmids (**Paper I-II**), typing bacteria (**Paper III**), and studying genetic variations in the human genome (**Paper IV**) (Figure 15). The assays are based on the previously established competitive binding scheme using YOYO and netropsin (154), and all rely on mapping single-DNA molecules in nanofluidic channels.

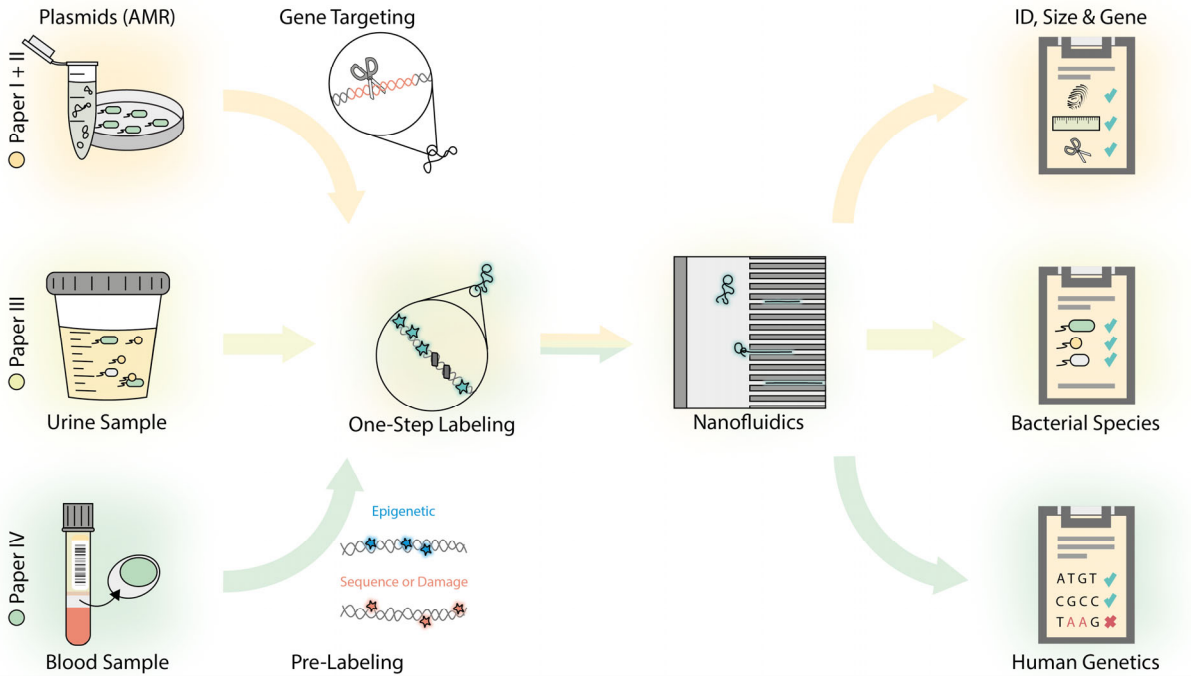

Figure 15: *Illustration of key concepts of the assays used in Paper I-IV. Starting from plasmid DNA extracted from pre-cultivated bacteria (Paper I-II, orange), bacterial DNA from uncultivated urine samples (Paper III, light green) or human DNA from mononuclear cells in a blood sample (Paper IV, green), the DNA is labeled with YOYO-1 (green star) and netropsin (light gray rectangle) rendering an emission intensity profile along the DNA reflecting the underlying sequence. Nanofluidic channels are used to stretch the DNA for subsequent visualization of the sequence-specific emission intensity profile along the DNA contour using fluorescence microscopy. For the purified plasmid samples, resistance genes can be targeted using Cas9 (gray scissor, Paper II). The human DNA can be pre-labeled to target sequence-specific motifs (red star), epigenetic marks (blue star), or DNA-damage sites (red star, Paper IV).*

In **Paper I** (details in Section 5.2.1), we presented an assay that allowed us to identify and trace plasmids in clinical samples from a nosocomial outbreak at a neonatal ward at Sahlgrenska University Hospital. Using an approach based on CRISPR/Cas9, we demonstrated in **Paper II** (details in Section 5.2.2) how the assay could be expanded to simultaneously also detect resistance genes on individual plasmids. Utilizing and optimizing methods for extracting DNA molecules up to megabase pairs in size, allowed us to type bacteria both in complex mixtures and from uncultivated clinical urine samples in **Paper III** (details in Section 5.2.3). In **Paper IV** (details in Section 5.3), we adapted the assay for enzyme-free mapping of single DNA molecules to the human genome.

5.1.1 Experimental Work

In this section, some general considerations of the experimental procedures in the four appended papers are provided. In **Paper I-IV**, different sources of DNA were used; plasmid and chromosomal DNA extracted from bacteria, as wells as human DNA extracted from white blood cells. Plasmid DNA was isolated after cultivation of clinically relevant bacteria using standard commercial kits. The protocol allows separating plasmid DNA not only from other cellular components, but also from the chromosomal DNA of the bacteria. To obtain large chromosomal DNA molecules the DNA needs to be carefully extracted from the bacterial or human cell. The approach used in **Paper III-IV** is based on trapping cells in agarose gel plugs. The agarose matrix allows efficient exchange of the surrounding buffer while hindering the large DNA molecules from diffusing out of the gel plug. The cells can be lysed, residues removed, and the remaining DNA washed without any direct interaction with the DNA, minimizing DNA fragmentation.

The DNA was stained using an approach based on competitive binding between YOYO and netropsin, as described in Section 4.2 (154). In short, the non-fluorescent, AT-specific, molecule netropsin competes with the fluorescent dye YOYO for AT-rich regions, resulting in an overall lower emission from AT-rich regions than from GC-rich regions along the extension of the DNA molecule (Figure 16). Netropsin was generally added at a ratio of 100:1 to YOYO to compensate for the lower overall binding affinity to DNA. The obtained intensity profile serves as a "fingerprint" of the DNA molecule, reflecting the underlying sequence (Figure 16).

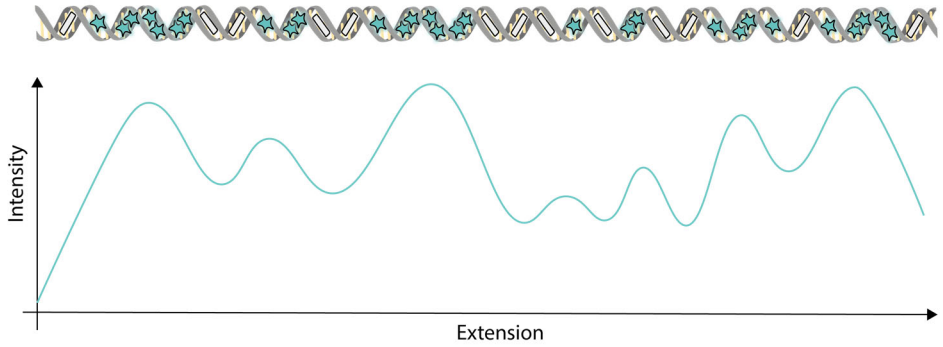

Figure 16: *Illustration of the emission intensity profile created by competitive binding of YOYO-1 (green star) and netropsin (light gray rectangle) to DNA. The extended DNA molecules will have a larger fraction of the fluorescent molecule YOYO-1 bound in GC-rich regions compared to AT-rich regions, rendering an emission intensity profile based on the underlying DNA sequence.*

In addition to the DNA of interest, lambda DNA was added to all samples as an internal size reference. The size and intensity profile of the linear lambda DNA molecule (48 502 bp) is known, allowing a conversion factor to be calculated, relating the extension in pixels to the number of bp of any observed DNA molecule. Since lambda DNA was added to each sample, it was possible to compensate for minor deviations in the experimental conditions, including factors such as ionic strength and nanochannel dimensions.

The DNA was stained with YOYO and netropsin at high ionic strength at 50 °C to facilitate rapid equilibration (155). The high ionic strength shields the negatively charged DNA backbone, increasing the rate of association and dissociation of the positively charged YOYO molecule, facilitating rapid rearrangement along the DNA. However, as discussed in Section 2.3.1, a high ionic strength will decrease the degree of extension of DNA when confined in nanofluidic channels (91). Therefore, after the labeling, the sample was diluted approximately ten times with ultrapure water, providing a solution of much lower ionic strength.

When exposing YOYO to light in the presence of oxygen, oxygen radicals will be formed, which are capable of producing single-strand breaks in the DNA (156), a process referred to as photo-nicking. If two single-strand breaks occur close enough on opposite DNA strands, the dsDNA will break, rendering shorter DNA molecules that will contain less

information. The oxygen radical scavenger β-mercaptoethanol (BME) was used to reduce the amount of photo-nicking.

As briefly mentioned in Section 3.3, an epifluorescence microscope, with the objective facing the glass lid of the nanofluidic chip, was used for the work presented in this Thesis. The use of a 100x oil immersion objective with a numerical aperture of 1.46, in combination with an EMCCD camera, sets a resolution limit on the optical DNA maps of around 1 kb. The resolution is limited by the dimensions of the nanochannels, governing the degree of the extension, as well as the fundamental resolution limit of optical microscopy. An image of plasmid DNA molecules (50-95 kb in size) inside the micro and nanochannels of the chip, visualized with fluorescence microscopy, is shown in Figure 17.

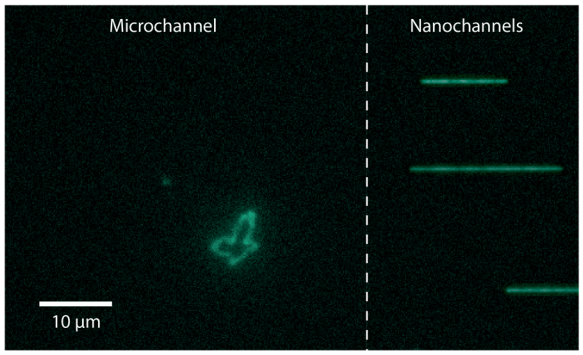

Figure 17: *Fluorescence microscopy image of DNA in a nanofluidic chip. The image shows one DNA molecule (circular plasmid) in the microchannel (50 µm x 850 nm) and three DNA molecules (linearized plasmids) of different size (50-95 kb) confined in the nanochannels (100 nm x 150 nm).*

In order to control the movement of the DNA inside the nanofluidic chip, an N_2 flow system was used. The rationale for using N_2 in the flow system, rather than air, was to reduce the overall concentration of O_2 in the sample, preventing extensive photo-nicking of the DNA. By connecting one switch to each loading well, DNA can be flushed inside the microchannel by applying pressure to one of the wells, and subsequently forced into the nanochannels, by applying pressure to both sides of the microchannel (Figure 13 and 17). Once confined, a series of typically 100 images, with an exposure time of 100 ms, was recorded of the extended DNA molecules.

5.1.2 Data Processing

The data analysis tools used for the work presented in this Thesis have been developed in collaboration with Tobias Ambjörnsson's research group at Lund University. The Matlab-based data and image analysis software, written in Lund, has been continuously improved, adding new features and novelties for each appended paper. An overview of the data analysis workflow used in **Paper I-IV** can be seen in Figure 18.

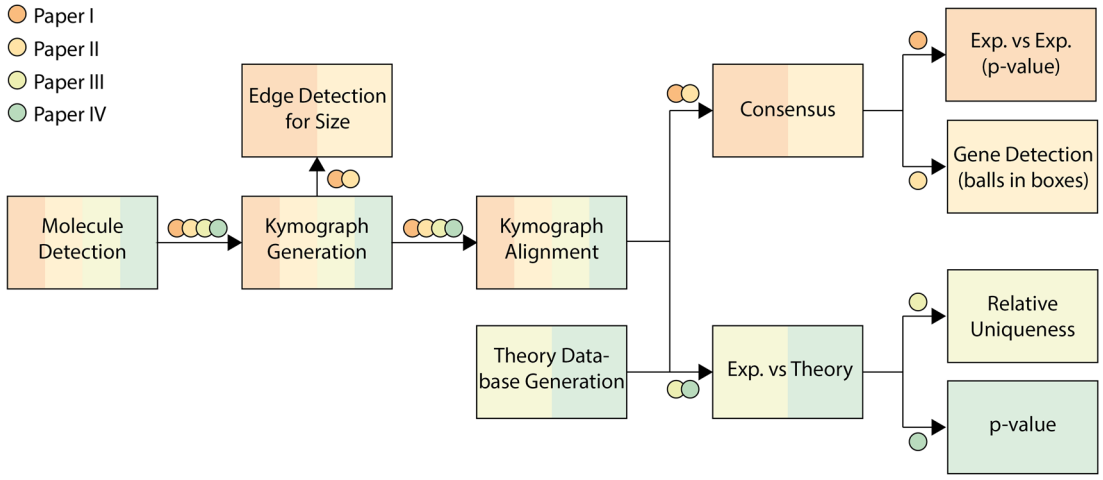

Figure 18: *Overview of the data analysis workflow used in Paper I-IV.*

General Data Processing – From Movies to Intensity Profiles

Following data acquisition, the DNA molecule is detected in each time-frame and the individual time-frames collapsed into a one-dimensional (1D) image, containing the mean intensity value for each pixel along the extension of the detected molecule. Once the DNA molecule is detected for each time frame, the obtained 1D images are stacked underneath each other, creating a kymograph (Figure 19). The extension of the DNA molecule can be calculated by detecting the edges of the DNA molecule for each time frame in the kymograph and calculating the average.

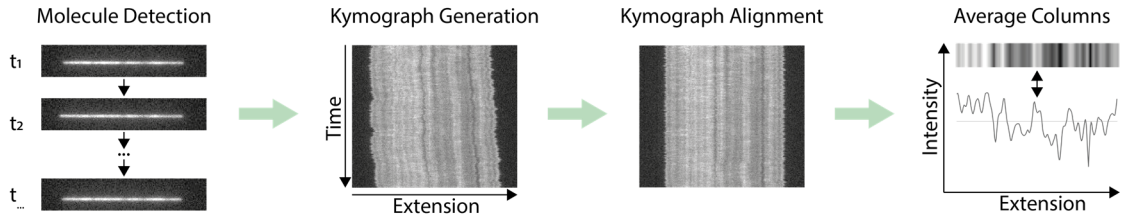

Figure 19: *Processing of raw data. First, the DNA molecule is detected in each time frame, and a one-dimensional emission intensity profile, based on the mean intensity along the extended molecule, is calculated and saved. Second, the one-dimensional intensity profiles are stacked underneath each other to create a kymograph. Third, the kymograph is aligned to account for local and global molecule fluctuations and drift inside the nanofluidic channels. Finally, the mean value of each column is computed, rendering an average emission intensity profile along the extended molecule that is ready for subsequent applications.*

Molecule fluctuations and drift in the nanofluidic channels need to be considered before extracting the sequence-specific information from the kymographs. Therefore, an algorithm is applied, allowing the features of the kymograph to be aligned (157). In short, the most pronounced bright, or dark, feature, is detected and aligned by stretching each row in the kymograph separately. The newly aligned feature is fixed, and the process is repeated until the regions separating the fixed pixels are considered small enough. In order to obtain the final average intensity profile, the mean value of each column in the aligned kymograph is calculated and plotted versus the extension (Figure 19). The intensity traces are normalized, with the average value set to 0, and the standard deviation to 1, to account for experimental differences in overall kymograph intensities.

Pearson Correlation Coefficient – Comparing Two Intensity Profiles

A Pearson Correlation Coefficient (PCC) can be calculated to compare the similarity between two intensity profiles (Figure 20). The PCC score between two intensity profiles, x, and y, with n pixels, is calculated according to Equation (2).

$$PCC = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (2)$$

The PCC value ranges between +1 and -1, where a value of +1 implies that the intensity profiles are identical, and -1 implies a total negative linear correlation (mirror images). When comparing two intensity profiles, one will be fixed, and the other one shifted one pixel at the time, allowing the PCC score for each possible starting position to be calculated. The procedure is repeated for the flipped version of one of the intensity profiles before the overall maximum PCC score for a pair of intensity profiles can be found.

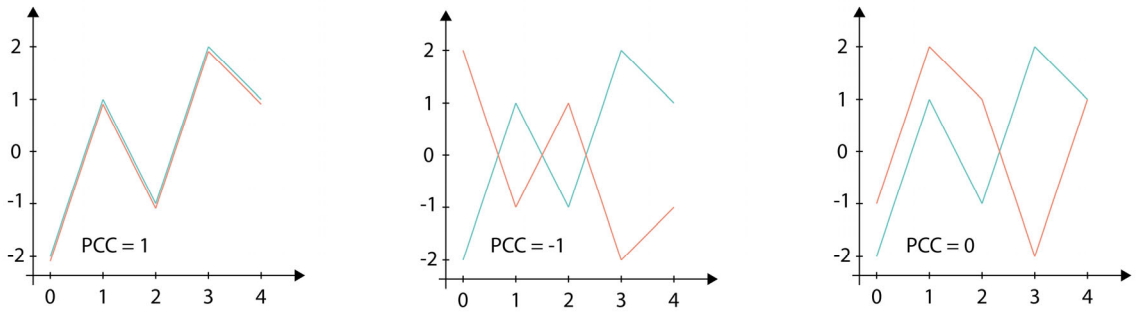

Figure 20: *Examples of three pairs of intensity profiles and their corresponding Pearson Correlation Coefficient scores.*

Consensus Generations – Creating Averages of Intensity Profiles

When multiple copies of a DNA molecule with identical sequences are mapped, an average intensity profile, *i.e.*, a consensus intensity profile, can be created to minimize the experimental noise. The process of generating a consensus intensity profile can be seen in Figure 21. First, all intensity profiles are stretched to the same length, followed by calculating the maximum PCC value for all possible pairs of intensity profiles. The two intensity profiles that display the highest PCC value will then be merged at the position of the optimal shift and possible flip, representing an average of the two individual intensity profiles. Next, the process is repeated, with the new average intensity profile replacing the two individual ones. Once again, the two intensity profiles showing the highest degree of similarity will be merged, and the process is repeated until only one consensus intensity profile remains.

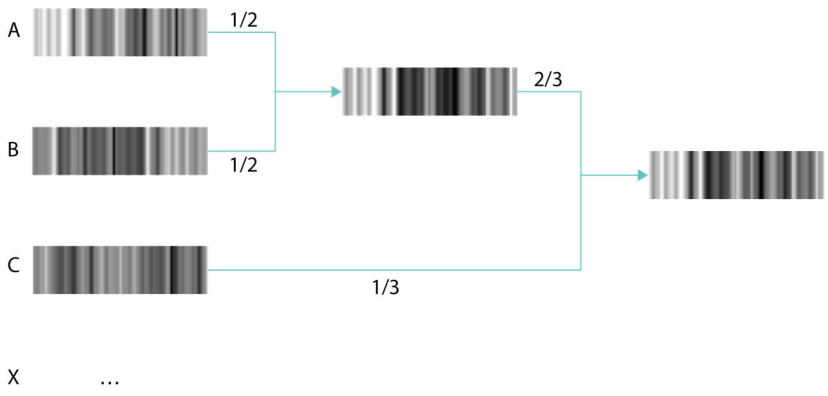

Figure 21: *Process of creating a consensus intensity profile from individual profiles. Individual intensity profiles are successively merged based on their degree of similarity, weighted as shown at the corresponding path.*

The obtained consensus intensity profiles were used to trace the spread of plasmids in an outbreak (**Paper I**) and to detect the presence of a resistance gene (**Paper II**). Moreover, consensus intensity profiles of bacterial artificial chromosomes (BACs) were used in **Paper IV**.

Theoretical Intensity Profiles – From Base Pair Sequence to Optical DNA Map

Theoretical intensity profiles can be generated to compare an experimental intensity profile to a known DNA sequence (158, 159). Theoretical intensity profiles are created by considering the likelihood of YOYO and netropsin to bind a four bp site along the DNA sequence and compensating for the resolution limit set by the experimental setup. Comparisons between experimental and theoretical intensity profiles were done in **Paper II-IV**.

p-Values – Experimental vs. Experimental and Theoretical vs. Experimental

Methods for computing p-values were developed to obtain a statistical measure for the similarity between two intensity profiles. The characteristic features from all sequenced plasmids in the NCBI RefSeq database (**Paper I**) or the human genome (**Paper IV**) were

extracted. Based on the extracted information, randomized intensity profiles were generated, acting as a reference in order to convert the obtained PCC value, which is size-dependent, to a p-value. P-values were used in **Paper I,** when comparing pairs of experimental intensity profiles, and, in **Paper IV,** when comparing experimental intensity profiles to theoretical intensity profiles.

Balls in Boxes Statistics – Determining the Presence of a (Resistance) Gene

To determine if a (resistance) gene, targeted by Cas9, is present on a plasmid (**Paper II**), a "balls in boxes" approach was used. From the consensus intensity profile, the position of the double-stranded break on each of the plasmid molecules can be obtained (details in Section 5.2.2). The number of double-strand breaks that by chance would be expected to occur at the same position along the intensity profile is simulated, and a detection limit of + 3 STD of the obtained mean value is applied as a threshold for the detection of a gene. If the observed number of double-stranded breaks at a specific position along the DNA contour is above the set threshold, the targeted gene is deemed present.

Discriminative Intensity Profiles – Determining the Relative Uniqueness of a Match

In **Paper III** (details in Section 5.2.3), individual intensity profiles from mapped DNA molecules are evaluated based on if they are discriminative to a specific species when comparing against the NCBI database with over 10 000 sequenced bacterial genomes. An intensity profile is deemed discriminative if only matches to a specific species remain after applying a set of thresholds. All thresholds are based on the PCC scores obtained when comparing an experimental intensity profile to the theoretical intensity profiles of all bacterial genomes found in the NCBI database.

5.2 Mapping of Bacterial DNA – Fighting Antibiotic Resistance

As previously discussed, one of the main reasons for the rapid spread of antibiotic resistance is horizontal gene transfer of mobile genetic elements, such as plasmids. Therefore, tools allowing studies of epidemiology and evolution of bacteria are of great interest. Moreover, there is an urgent need for rapid diagnostic tools to enable more efficient treatment of acute bacterial infections, as well as to spare last-resort antibiotics for when needed.

Two assays were developed to meet this need, one for rapid characterization of intact plasmid molecules, and one for bacterial typing, both based on the previously established optical DNA mapping method using the competitive binding scheme of YOYO and netropsin (154). The advantage of optical DNA mapping compared to existing strategies for bacterial DNA analysis include the low amount of sample required, as well as obtaining sequence information from chromosomal DNA or intact plasmids on length scales that are not easily accessible with other techniques. An overview of the current version of the two assays, developed in **Paper I-III**, can be seen in Figure 15.

5.2.1 Tracing Plasmids During a Resistance Outbreak

In **Paper I**, we investigate how the size of a plasmid, in combination with its intensity profile, could be used to trace plasmids originating from a nosocomial outbreak at Sahlgrenska University Hospital (160). The previously established competitive binding protocol (154) was adapted to plasmids and evaluated on clinical samples. The circular nature of the plasmids provided both advantages and challenges when designing the assay. The ability to discriminate a circular plasmid from linear pieces of DNA (161, 162) ensured that only intact plasmids were studied. However, in order to reveal the sequence-specific pattern, the circular plasmids needed to be linearized. Reactive oxygen species were created (Section 5.1.1), capable of causing single-strand breaks on the confined circular DNA, using intense light exposure. Once two single-stranded breaks occurred close enough to each other, the plasmid spontaneously unfolded to its linear configuration.

During a four-month-long outbreak in 2008 at Sahlgrenska University Hospital, over 20 neonates were infected or colonized with multi-resistant bacteria, harboring extended-spectrum beta-lactamase (ESBL) genes. The principal strain of the outbreak was found to be an *E. coli* of sequence type (ST) 131 (160). However, two additional *E. coli* strains, ST73, and ST1444 were identified, as well as a *K. Pneumoniae* strain. Both *E. coli* and *K. Pneumoniae* belong to the family of *Enterobacteriaceae*, which are well known to express ESBLs (47, 163, 164). Moreover, ESBL-mediated resistance is frequently located on plasmids (165-167). Thus, the resistance gene was suspected to be plasmid-born. However, it was not known if there, in fact, were multiple outbreaks co-occurring, indicated by the presence of different types of bacterial strains and species, or if the outbreak was mediated by the transfer of a single plasmid between the different types of bacteria.

The aim of the study in **Paper I** was to trace plasmids from the outbreak by both monitoring plasmid content over time, and detecting the spread of plasmids between bacterial strains and species, as well as between patients. A method for comparing two experimental intensity profiles with each other was developed, which allowed converting the obtained PCC value, which is size-dependent, to a p-value.

In total, nine isolates, originating from four different neonates, were selected for the study to include isolates from all different sequence types and both species of the outbreak, as well as isolates collected at different time points (details in **Paper I**). The detected plasmids were characterized by size and intensity profiles. The size of all the plasmids found in the nine different isolates was obtained by measuring the extension of the plasmids in their circular form (Figure 22A) (161, 162).

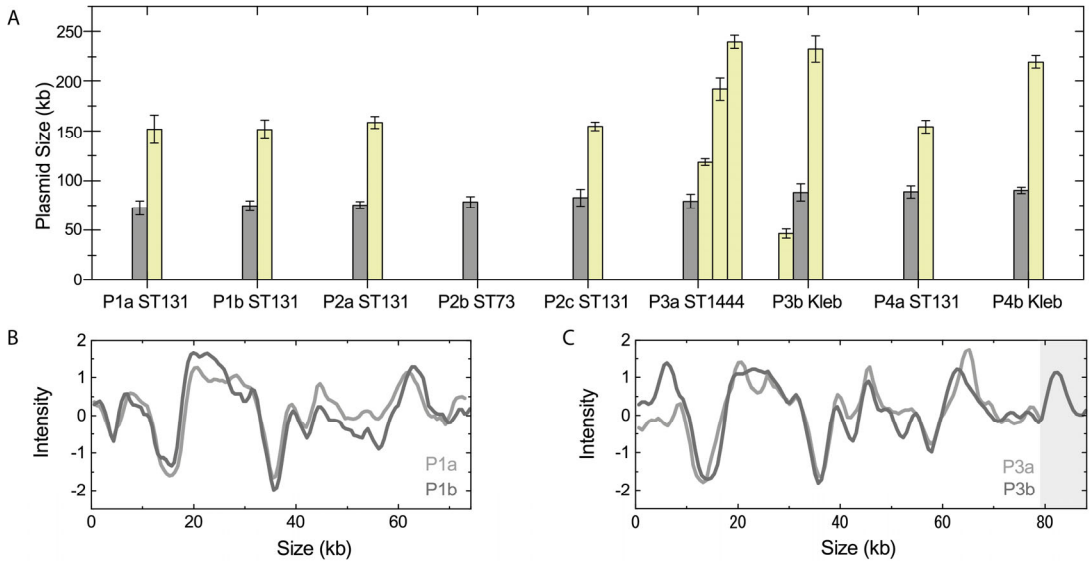

Figure 22: *Key results presented in Paper I. (A) Histogram showing the number of different plasmids detected in each isolate and their corresponding size. The isolates are indexed by the patient (P) number, and the isolates within a patient are separated by the letters a-c. The plasmids originated from E. coli (sequence types 73, 131, and 1444) or K. pneumoniae (sequence type not specified). (B-C) Comparison of experimental consensus emission intensity profiles from the ~80 kb plasmid in isolate P1a and P1b (isolates acquired at different time points, B), and P3a and P3b (isolates from different bacterial species, C). In (C), the ~80 kb plasmid from the K. pneumoniae isolate P3b has an insert of an extra piece of DNA into its sequence (light gray box), explaining the slightly larger measured size in A.*

The size measurements revealed a plasmid of similar size shared by all isolates, ranging from 73 to 90 kb in size, indicating that this plasmid was responsible for the spread of resistance. Moreover, by computing the p-value for all different pairs of plasmids, the results showed that the plasmids indeed displayed a statistically significant degree of similarity (Figure 22B-C). Using the p-value tool when comparing two experimental consensus intensity profiles, furthermore allowed the shared plasmid to be divided into two different subgroups, one displaying an extra insert at a specific position along the sequence (light gray box, Figure 22C), and one group without this insert. The extra piece in the DNA sequence in some of the plasmids also explains the differences in the measured size. An overview of the findings when studying the plasmid content in all selected isolates can be seen in Figure 23.

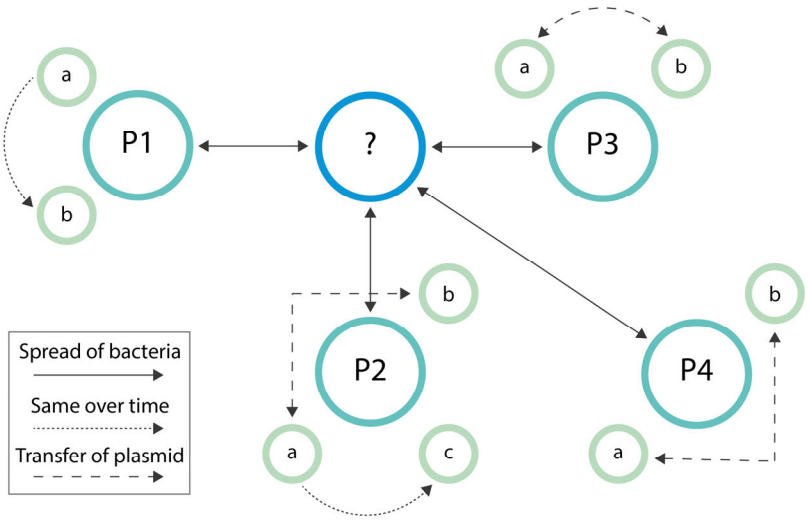

Figure 23: *Overview of the results in Paper I. The different patients are represented by green circles, individual isolates by light green circles, and the unknown rout of transmission by the blue circle.*

In summary, a shared plasmid of about 80 kb in size with a highly similar intensity profile ($p < 0.01$) was found in each of the studied isolates, indicating that this plasmid was responsible for the spread of resistance. The results demonstrate that the developed assay

can monitor plasmid content over time, detect the spread of plasmids between different bacterial strains and species, as well as indicate the transfer of bacteria between patients. The study shows how conclusions regarding the spread of resistance can be made without knowing the sequence of the plasmid. Hence, the method shows great potential to guide at an early stage during an outbreak situation, pinpointing which isolates that subsequently require a more detailed characterization. As a final note, the resolution of the optical setup sets a limit on how small plasmids that can be accurately characterized using the assay presented in **Paper I**. The size of a plasmid can be accurately measured down to at least 8 kb (161), while, in general, a size of above 30 kb is desirable in order to obtain a unique intensity profile.

In parallel with the work presented in **Paper I**, the assay was also adapted for the identification of plasmids by comparing the obtained intensity profile to a database of over 3000 sequenced plasmids (**Paper VII**) (168). Following the work in **Paper I** and **VII**, the developed assay was used to aid and confirm the assembly of plasmid sequencing data (**Paper X and XII**) (159, 169).

5.2.2 Detection of Resistance Genes on Plasmids

One of the main limitations of the assay presented in **Paper I** was the inability to detect the presence of resistance genes. For the outbreak studied in **Paper I**, a modified assay that allows pinpointing the resistance gene directly to the shared plasmid, in **Paper I** confirmed *via* subsequent PCR, would be highly desirable. Moreover, using the intensity profile to detect the presence of a resistance gene would allow the assignment of the gene to a specific plasmid, and also to a precise location along its sequence.

With that goal in mind, **Paper II** demonstrates the development and evaluation of a novel assay for detecting resistance genes on single plasmids, utilizing the high sequence specificity of the Cas9 enzyme. By designing a crRNA sequence complementary to the gene of interest (Section 2.4.2), a plasmid harboring the gene will be "cut" at a specific location. Studying the positions of the double-strand breaks along the intensity profiles of individual plasmids enables determination of if a vast majority of the cuts have occurred at the same position along the sequence (Figure 24). By introducing a "balls in boxes" approach (Section 5.1.2), the number of double-strand breaks that by random chance would be expected to occur at the same position along the intensity profile can be simulated, and from this a detection limit set.

In contrast to **Paper I**, circular plasmids were in **Paper II** not broken with light in order to reveal the intensity profile. Instead, intact circular plasmids were used solely for size determination and pre-linearized plasmids for obtaining the intensity profiles. For analysis, only linear fragments +/- 20% in size of the detected circular plasmids were analyzed. The positions of the double-strand breaks of the corresponding consensus intensity profiles were used to detect the presence, or absence, of the gene of interest. In short, for a plasmid that harbors the gene targeted by Cas9, the overall number of linearized plasmids will be high, mainly consisting of Cas9 cut plasmids, but also some plasmids which have been randomly linearized due to mechanical stress or light exposure. In contrast, the overall number of linearized plasmids for a plasmid not harboring the targeted gene will be low since it will only contain the randomly linearized plasmids (Figure 24).

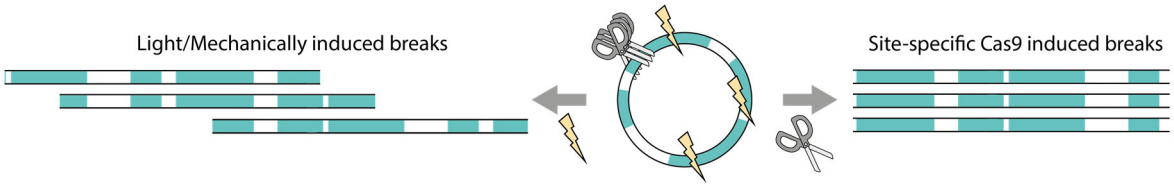

Figure 24: *The principle behind Cas9-assisted detection of resistance genes on plasmids. A double-stranded break induced by Cas9 (scissors) will be site-specific. In contrast, a double-stranded break caused by light or mechanical stress (lightning bolt) will appear at random sites along the DNA contour. Thus, in the presence of a resistance gene (right), a majority of the emission intensity profiles will be identical. If there are no target genes present (left), the emission intensity profiles (green and white) will instead be circularly permutated.*

The principle of the assay was demonstrated on the previously sequenced plasmid pUUH239.2 (pUUH) (220.8 kb), harboring the common ESBL gene *bla*$_{\text{CTX-M-15}}$. When adding crRNA with a sequence complementary to the *bla*$_{\text{CTX-M-15}}$ gene, a vast majority of the double-strand breaks occurred at the predicted position along the sequence, thus confirming the presence of the gene. Similar results were obtained for other plasmid isolates, containing either one or multiple plasmids, demonstrating the assay's ability to not only detect the presence of resistance, but also to pinpoint it to a specific location on a single plasmid.

As a final test, two previously uncharacterized bacterial isolates, DA28170 and D49173, were studied. Adding a cocktail consisting of two specially designed crRNA molecules to the samples, it was possible to simultaneously target two of the primary gene groups, *bla*KPC and *bla*NDM, encoding resistance against the last-line group of antibiotics, carbapenems. For isolate DA49173, the results revealed the presence of two plasmids in the sample, where the smaller encoded the resistance (Figure 25). By subsequently testing each of the crRNAs separately, it could be concluded that the gene mediating the resistance belonged to the *bla*KPC group. Similarly, a *bla*NDM gene was detected on one of the plasmids in isolate DA28170 (see **Paper II**).

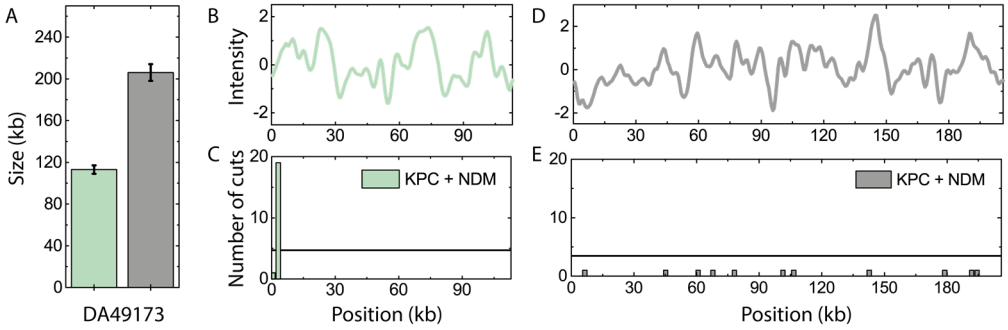

Figure 25: *Key results in Paper II. (A) The size of the two plasmids detected in isolate DA49173. The color of the intensity profile indicates the presence (light green) or absence (gray) of the resistance gene. (B and D) Experimental consensus emission intensity profiles of the smaller (light green) and larger (gray) plasmid in isolate DA49173. (C and E) Histograms of the locations of double-strand breaks along the contour of the small (light green) and large (gray) plasmid when simultaneously targeting the bla*KPC *and bla*NDM *gene families. The horizontal lines in (C) and (E) correspond to three standard deviations above the mean calculated value using the "balls in boxes" statistics. The results show that a vast majority of the double-strand breaks occurred at the same position along the contour for the smaller plasmid, indicating the presence of a resistance gene. For the larger plasmid, the double-stranded breaks occurred at random positions along the contour, and thus no resistance gene was detected.*

To summarize, with the progress presented in **Paper II,** the assay is capable of (i) detecting the size of the plasmids in a sample (ii) providing an intensity profile which can be used for plasmid tracing or direct identification *via* a plasmid sequence database, and (iii) detecting the presence, or absence, of a (resistance) gene, assigning it to a precise position

along a specific plasmid. Also, circumventing the need to manually break plasmids with light increases the speed of the method, as well as opens up for future automation.

The assay in its current format (**Paper II)** has been successfully applied to study the role of plasmid transfer in recurring infections (**Paper XIII**) (170), and to study an outbreak of resistant bacteria at the neonatal intensive care unit at Karolinska University Hospital in Stockholm (**Paper XIV**) (171).

5.2.3 Cultivation-Free Bacterial Identification

Whereas **Paper I** and **Paper II** focused on bacterial plasmids, the study in **Paper III** uses the chromosomal DNA of bacteria to identify the bacterial species from which the DNA originates. Many methods for typing bacteria exist, such as culture-based phenotypic methods, MALDI-TOF, and 16S rRNA gene sequencing (172, 173). However, the identification process is many times either expensive or time-consuming. While continuous improvements have allowed for the introduction of whole-genome sequencing in healthcare (174), with promising strategies on their way (175), the high cost in combination with extensive preparation protocols and complex analysis has so far hampered the use in clinics (176).

Correct and rapid identification is critical for patient outcomes, to minimize the spread of disease, and to control outbreaks. With this in mind, we developed an assay (Figure 15), based on extracting and mapping large DNA molecules, up to megabase pairs in size, to identify bacteria with high sensitivity and selectivity. The intensity profiles from the mapped chromosomal DNA molecules were individually compared to a reference database consisting of over 10 000 sequenced bacterial genomes and evaluated based on their relative uniqueness to a specific species. An intensity profile was deemed discriminative, *i.e.*, providing useful information, if a high match score was obtained only for genomes of the same bacterial species. At least three intensity profiles were required to be discriminative to the same bacterial species to state the presence of a bacterium, increasing the specificity of the assay.

First, the data analysis parameters were optimized in order to maximize the performance of the assay. The parameter which influenced the results the most was C_{diff}, which affects how many genomes in the database that obtain a high enough PCC score to be considered when determining if a fragment is discriminative. For an intensity profile to be deemed discriminative, all the genomes which remain after applying the C_{diff} threshold are

required to belong to the same bacterial species. A smaller C_{diff} will lower the requirements to obtain a discriminative intensity profile. Thus, a small C_{diff} will increase the number of intensity profiles that will be deemed discriminative. However, the reduced stringency of a small C_{diff} will also lead to a decrease in the true positive rate (TPR). In order to obtain a high TPR without reducing the throughput of the assay to a large extent, C_{diff} was set to 0.05 (Figure 26A).

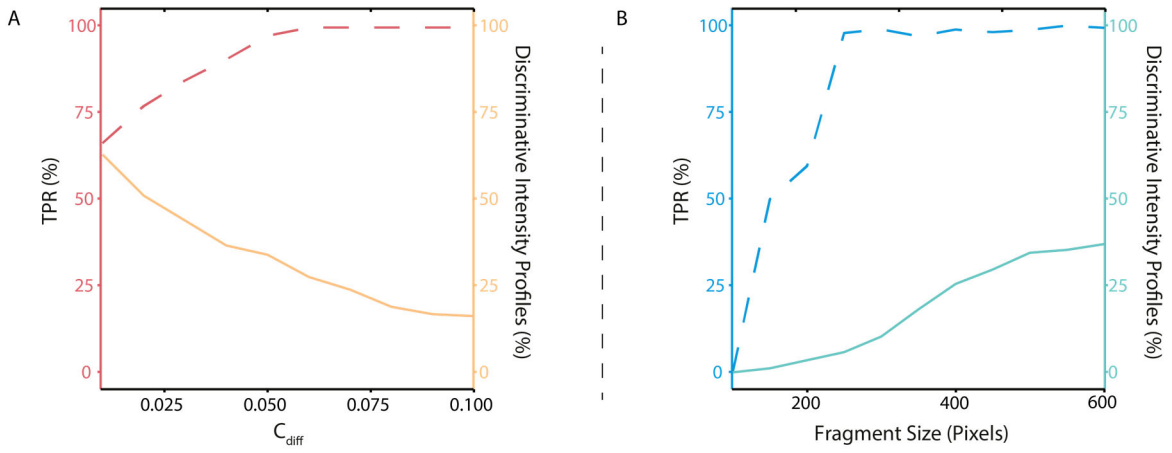

Figure 26: *Optimization of thresholds for bacterial identification. Effect of (A) C_{diff} and (B) fragment size on the true positive rate (TPR) (dashed lines), and fraction (%) of discriminative profiles out of the total number of experimental profiles (solid lines). One pixel corresponds to ~ 500 bp.*

The size of the mapped DNA molecules also has a significant effect on the performance of the assay (Figure 26B). The larger the molecule, the more information the resulting intensity profile will contain. Thus, the specificity of the assay in general increases with an increased size of the DNA molecules. Therefore, larger molecules will, in general, have a higher probability of being discriminative. Moreover, using molecules smaller than ~ 125 kb will decrease the specificity of the assay, decreasing the TPR.

The assay was validated using a wide range of samples, starting with samples containing a single bacterial species, followed by mixtures of up to four species, as well as non-cultivated urine samples, positive for urinary tract infection (Figure 27). For all of the

different samples, the assay performed remarkably well, with an overall TPR of 99%, and identifying only the correct species in all of the evaluated samples.

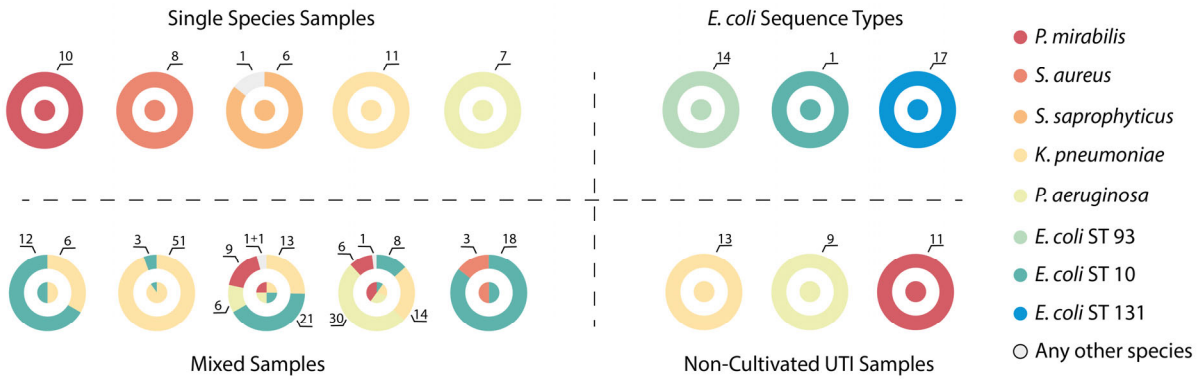

Figure 27: *Key results presented in Paper III. Outcomes of bacterial identification from samples including a single species, a mix of different species, E. coli samples of different sequence types, and non-cultivated urine samples. Each chart represents a separate sample. The inner circle shows the expected distribution of species in the sample, while the outer circle shows the obtained distribution with the number of discriminative profiles specified.*

We also investigated the potential of using the intensity profiles of individual DNA molecules to discriminate bacteria at the sub-species level. An initial evaluation of three *E. coli* isolates belonging to different sequence types showed that the assay, in fact, was able to assign the bacteria to the correct sequence type (Figure 27). The potential of identifying bacteria at a sub-species level will be evaluated further in future studies.

To summarize, we developed an assay capable of identifying bacteria based on sequence-specific labeling of YOYO and netropsin. The DNA extraction protocol is general and fast and allows the extraction of megabase pair large DNA molecules. The high specificity allowed us to identify bacteria in complex mixtures. The low amount of DNA needed for analysis allowed us to identify bacteria in non-cultivated samples from patients with urinary tract infections. Finally, the assay has the potential to type bacteria at the subspecies level. Compared to previous approaches for identifying bacteria using different optical DNA mapping strategies (158, 177-183), the current assay is general for all bacteria, highly specific, and has a streamlined and straightforward workflow.

5.2.4 Designing a Complete Assay for Bacterial Diagnostics

The assay described in **Paper I-II** allows us to characterize plasmids based on their size, their underlying DNA sequence and enables pinpointing of resistance genes to a specific plasmid. In turn, the assay presented in **Paper III** facilitates the identification of bacteria not only in monocultures but also in complex mixtures and clinical urine samples. Importantly, the extract of large chromosomal DNA molecules used for bacterial identification in **Paper III** will also include bacterial plasmids. Thus, there is a possibility to combine the assay in **Paper I-II** with the assay in **Paper III** without compromising the simplicity of the workflow.

Such an assay can be used both for tracing plasmids and for identifying bacteria in the same sample. The assay could either function as a complement to sequencing-based strategies, or on its own for applications were base-pair resolution is not required. Compared to electrophoretic methods, such as PFGE combined with restriction enzymes, which provide information about plasmid size and bacterial strain, the assay, following DNA extraction, can be completed within hours instead of days. Moreover, for plasmids, the assay can separate plasmid molecules of the same size based on their intensity profiles. Similar to PCR-based techniques for plasmid characterization, the progress in **Paper II** allows the assay presented here to target specific genes. However, in contrast to PCR-based techniques, the developed assay can pinpoint which particular plasmid in the sample that harbors the gene of interest. Also, if the targeted gene is not present, PCR-based methods would not reveal any information about the sample, while the assay presented in this Thesis would still provide details of plasmid sizes, corresponding intensity profiles, as well as allow for bacterial identification.

The assay, in its current form, provides valuable insights and shows great potential to be used as a research tool when studying plasmids and identifying bacteria. However, several key points still need to be addressed in order for the assay to function as a reliable diagnostic tool in clinics. First, avoiding the cultivation of bacteria is critical in order to reduce the lead time from sample collection to diagnosis. As shown in **Paper III**, urine samples are already possible to analyze. The end goal, however, is to analyze even more challenging clinical samples, such as blood samples, which will increase the demands on sample prep and data analysis even further. Second, detecting resistance genes not only on plasmids but also on chromosomes is central if the assay is going to be reliable and useful in a clinical setting. Third, since the resolution of the assay is limited to the kilo base-pair level, additional sequence-specific marks would allow detecting resistance

genes with high precision also on smaller plasmids (> 30 kb). Finally, further simplifying the workflow, reducing the cost, and decreasing the lead time of the assay is important to maximize its future impact.

Even if successful in addressing all the remaining challenges raised above, the competition is intense, and the chances of the assay reaching the market rather slim. So far, many impressive strategies for diagnostics of resistant bacteria have been presented, including methods based on digital real-time loop-mediated isothermal amplification (dLAMP) (184), optical detection of bacterial growth rates (185), and sequencing-based methods (175). For the latter, it should be noted, however, that even if whole-genome DNA sequencing has a vast potential, the use of sequencing-based assays in clinical practice is, so far, still hampered by the significant cost, a long turnover time, lack of standardized databases, and the challenges of predicting the phenotypic outcome from genotypic data (176, 186).

5.3 Mapping the Human Genome Without Enzymes

While **Paper I-III** focuses on studying bacteria, **Paper IV** targets the challenges of studying human DNA. In recent years, optical DNA mapping has been extensively used to study structural variations within the human genome, as well as to aid in the assembly of human genomes from DNA sequencing data (Section 4.2) (106, 116-135). A vast majority of optical DNA mapping approaches have so far been based on enzymatic labeling, which has also been commercialized by *Bionano Genomics*®. For complex genomic regions, dual enzymatic labels have been used, and enzymatic labels have also been combined with epigenetic marks (113, 147). However, enzymatic labeling protocols many times require extensive labeling schemes, including washing steps to remove any unbound fluorophores. In addition to the imaging channel(s) used to observe the enzymatic labels, the DNA contour needs to be detected, which is traditionally done in a separate imaging channel, after staining the DNA with YOYO.

In **Paper IV**, we aimed to develop an assay in which individual DNA molecules, stained with YOYO and netropsin, could be mapped to the human genome with high precision without the use of any enzymatic reactions. The assay would have the potential to be used for stand-alone applications or to be combined with already commercialized systems, utilizing the YOYO channel not only to detect the DNA contour but to obtain sequence information simultaneously.

Five different Bacterial Artificial Chromosomes (BACs) with known sequences (details in **Paper IV**) were used to optimize the assay in terms of the number of frames used during image acquisition (Figure 28A-B), corrections for varying DNA extension (Figure 28C), as well as by selecting a p-value threshold (Figure 28D). As expected, the maximum PCC score (C_{max}) when comparing an experimental intensity profile to a theoretical intensity profile increased with averaging over an increased number of frames. However, increasing the number of frames will also decrease the overall throughput of the assay.

While the conversion factor from pixels to base pairs, obtained from the reference molecule lambda DNA, is precise, small variations in the experimental conditions, as well as nanochannel dimensions, will still affect the determined size. Therefore, allowing the measured size of a DNA molecule to vary when compared to the theoretical profile will increase the likelihood of obtaining a high C_{max} score to the correct position within the human genome. However, if a too large variation is allowed, the experimental intensity profile will also match to other, incorrect, positions with a high C_{max} score. Similarly, a strict p-value threshold will remove a vast majority of false positives, *i.e.* experimental intensity profiles, which are mapped to the incorrect position within the human genome. However, a stringent p-value threshold also reduces the amount of data that is used, decreasing the throughput of the assay. Based on the results in Figure 28, ten time-frames, 5% size variation, and a p-value threshold of 10^{-6} was selected. Figure 28E shows a visual comparison of the fit of an experimental intensity profile to its corresponding position on the theoretical version of the human genome for each of the five different BACs.

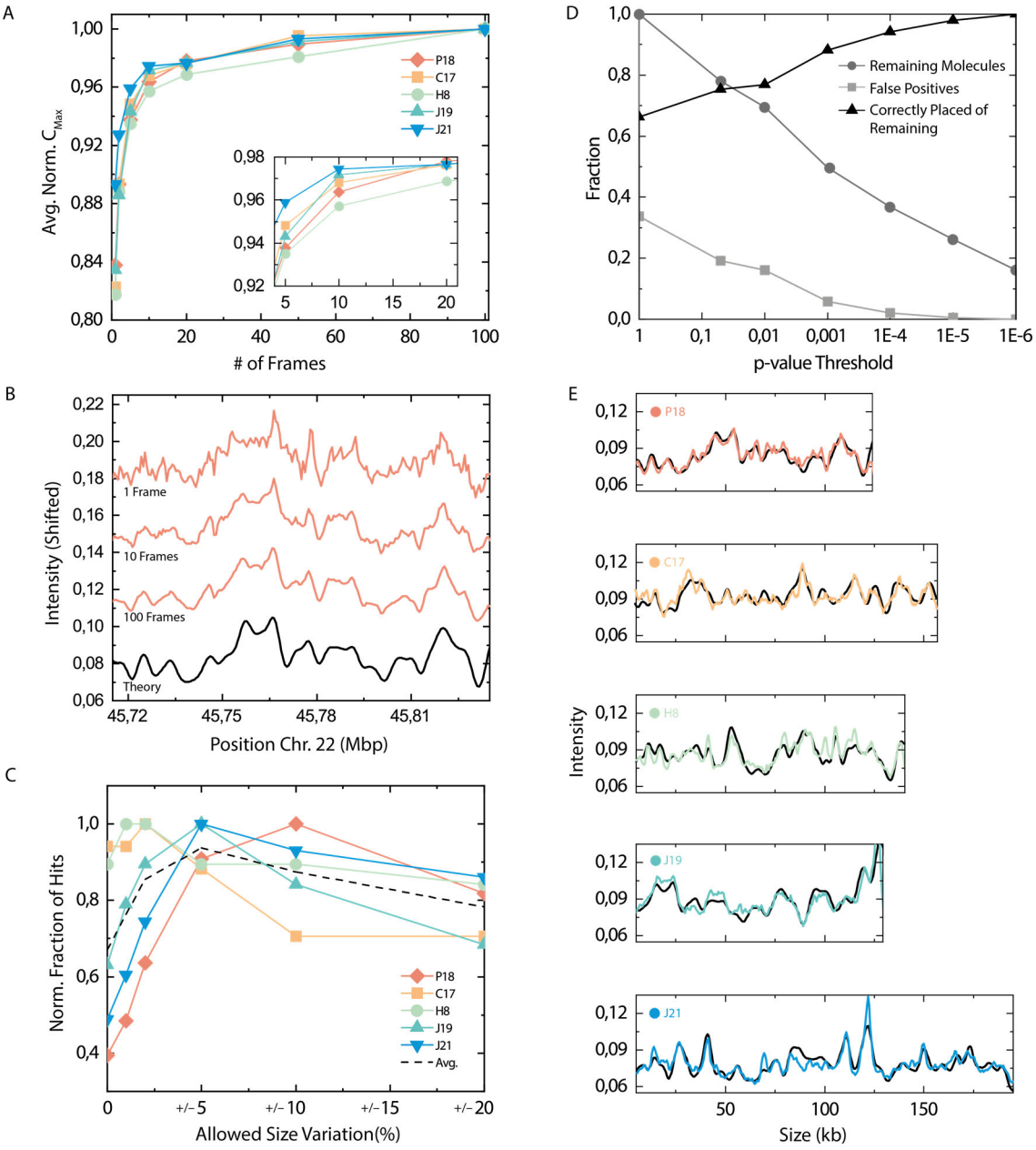
A
Avg. Norm. C_{Max}
# of Frames
P18
C17
H8
J19
J21
B
Intensity (Shifted)
1 Frame
10 Frames
100 Frames
Theory
Position Chr. 22 (Mbp)
C
Norm. Fraction of Hits
Allowed Size Variation(%)
Avg.
D
Fraction
p-value Threshold
Remaining Molecules
False Positives
Correctly Placed of Remaining
E
Intensity
Size (kb)

Figure 28: *Optimization of the competitive binding based assay with Bacterial Artificial Chromosomes (BACs). (A) The average normalized C_{max} values for a varying number of time frames for the five BACs (P18, C17, H8, J19, and J21) when compared to their corresponding theoretical intensity profile. (B) Visual comparison of an experimental intensity profile from BAC P18 to its corresponding theoretical intensity profile (black) using a different number of frames (the intensity profiles have been horizontally shifted for clarity). (C) The normalized fraction of correct hits when experimental intensity profiles (10 frames) from each BAC were mapped to the human genome with different degrees of size variations allowed (1% steps). (D) The fraction of remaining experimental intensity profiles (gray circles) and the number of false positives (light gray rectangles), after applying different p-value thresholds (10 frames, 5% stretch, 1% steps). The fraction of correctly placed intensity profiles on the human genome out of the profiles that pass a specified p-value threshold is also shown (dark gray triangles). (E) Visual comparison of an experimental intensity profile (10 time frames) for each of the five BACs to their corresponding theoretical intensity profile (black).*

The information contained in the experimental intensity profiles is strongly correlated to its size. As an example, for the smallest BAC, P18 (123 kb), only 11% of the mapped molecules passed the p-value threshold of 10^{-6}, compared to 53% for the largest BAC, J21 (203 kb). Thus, it is clear that extracting large DNA molecules is of great importance in order to take full advantage of the assay. As in **Paper III**, large DNA molecules were in the next step extracted from cells trapped in agarose plugs.

DNA was extracted from white blood cells isolated from human blood samples and mapped to the human genome. The results showed that 95 % of the mapped DNA molecules (average size of 333 kb) obtained a p-value below the threshold of 10^{-6}. Two examples of fits between an experimental intensity profile and the human genome are shown in Figure 29A-B, where the DNA molecule in Figure 29B displayed a highly repetitive region, shown in detail in Figure 29C. The repetitive region in Figure 29B highlights the advantage of optical DNA mapping compared to traditional sequencing techniques, which is to acquire long-range sequence information, where the number of repeats can be observed directly from the intensity profile.

An *in-silico* simulation based on the experimental results demonstrated that 97% of the human genome should be possible to map using our assay. The simulations were based on an average experimental DNA molecule size of 333 kb. With an increased average size of mapped molecules, currently limited by the chip design, the mappable parts of the

human genome would approach 100%, and the potential of the assay to detect structural variations increase.

One advantage of using the YOYO imaging channel to gain sequence-specific information, and not only to detect the DNA contour, is that an additional imaging channel is made available. The second (or third) imaging channel can then be used for applications such as detecting additional sequence-specific marks, reveling epigenetic information, or detecting DNA damage sites. The latter was used as a proof of principle experiment in our study.

Human white blood cells were treated with the cancer drug etoposide and the single-stranded breaks repaired by adding fluorescently labeled nucleotides together with DNA polymerase. The protocol to label sites of DNA damage was adapted from previous work by Zirkin *et al*. (148). The intensity profile created by YOYO and netropsin was used to map the DNA molecule to a specific position within the human genome, and the damage sites at this position were simultaneously visualized (Figure 29D).

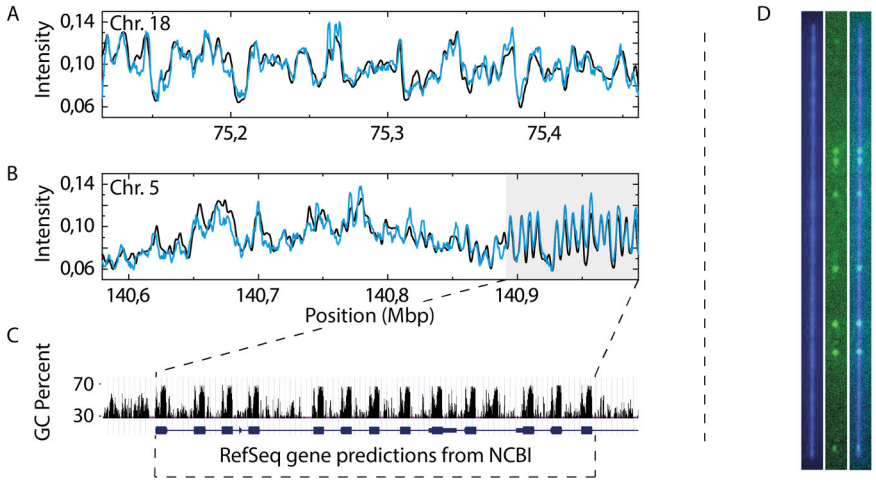

Figure 29: *Key results presented in Paper IV. (A-B) Visual comparison of experimental intensity profiles (blue) when mapped to a specified position ($p > 10^{-6}$) on the human genome (black). In (B), a highly repetitive region (gray box) was found, which is shown in greater detail in (C) with RefSeq genes marked in blue and the GC content (5 bp window) shown in black. (D) An example of a DNA molecule (450 kb) with labeled DNA damage sites (green channel) and a sequence-specific intensity profile (blue channel).*

To summarize, the developed assay has the potential to be used either as a standalone application, or combined with commercialized setups based on enzymatic labeling, to study structural variations in complex genomic regions, or to aid genome assembly from sequencing data. The assay is enzyme-free and based only on two commercially available molecules, YOYO and netropsin, added in a single step. The assay does not require any pre-repair of single-stranded breaks before labeling, removal of excess fluorophores, or strict temperature control. The functionality and throughput of commercial enzyme-based assays can be increased, and dual epigenetic marks studied simultaneously, using the YOYO channel to position the DNA molecules along the human genome.

6 Concluding Remarks and Outlook

This Thesis aimed to describe the development of novel assays based on optical DNA mapping, where the DNA is studied at a lower resolution, but over significantly larger length scales, than with DNA sequencing. Even if third-generation sequencing methods have provided a drastic increase in read-lengths, the assembly process of reads into a complete sequence is still hampered by repetitive regions and structural variations commonly found in DNA. These structural variations, such as inversions, insertions, deletions, and translocations, are commonly linked to disease in humans, and, to some extent, the spread of antibiotic resistance. Additionally, the excess lead time and overall complexity limit the applicability of DNA sequencing for bacterial diagnostics and resistance gene detection.

With the rates of bacteria acquiring resistance to antibiotics increasing drastically throughout the world, new methods allowing rapid detection of resistant bacteria will be of paramount importance. The two assays developed in **Paper I-III** combined provide detailed information about plasmids and bacteria and overall show great potential as the first step of bacterial characterization. In order for the assays to be a reliable alternative as a clinical diagnostics tool, a few key issues remain to be solved. Based on the presented

work on bacteria, current and future efforts include (i) automating the nanofluidic experiments, increasing the throughput and the simplicity of the workflow (ii) detecting resistance genes on chromosomes and smaller plasmids (> 30 kb), and (iii) further improving the DNA extraction protocol to enable diagnostics directly from additional clinical samples, such as fecal samples and blood.

Also, the assay developed in **Paper IV** has the potential to function as a stand-alone tool to study structural variations and to map epigenetic information across the human genome. The assay could moreover improve commercially available enzyme-based assays and aid in the assembly of complex genomic regions. To conclude, the work presented in this Thesis demonstrates that competitive binding based optical DNA mapping can function as a complement to DNA sequencing, when a base-pair resolution is desired, as well as be used as a versatile, stand-alone, method, to characterize DNA over large length scales

7 Acknowledgments

I would like to sincerely thank the following people for their direct or indirect support.

My supervisor, **Fredrik,** for all your support, encouragement, and commitment. I appreciate your trust in me as a student, and that you have allowed me to make many decisions on my own, even if you have not always agreed with them. It has been a great joy working together with you, and I hope we will continue as friends!

My examiner **Pernilla** for your positive energy, passion for science, and strong leadership.

Lena for teaching me all the secrets in the lab.

Previous and current members of the Westerlund group; **Lena**, **Karolin**, **Reza**, **Robin**, **Santosh**, **Sriram**, **Sune**, **Kai**, **Vandana**, **Rick**, **Rajhans**, **John**, **My**, **Gaurav**, **Vinoth**, **Elina**, and **Miriam**.

Tobias and his research group at Lund University, especially **Albertas**, **Saair**, **Jens**, **Christoffer**, and **Paola**. Thanks for all the fantastic work on the data analysis software and for responding to all my endless lists *via* email.

The experts on bacteria and antibiotic resistance, especially **Nahid, Anna,** and **Tinna** from Sahlgrenska Academy**, Linus, Marie**, and **Fredrika** from Uppsala University, **Christian** from Karolinska Institutet, as well as **Anna** and **Erik** from Chalmers University of Technology. Thanks for all the exciting discussions and for sharing your knowledge.

All additional **co-authors** for exciting collaborations.

My roommates **Hoda**, **Jesper**, **Sangamesh, Pegah**, **Gerard**, and **Rasmus**. Thanks for all the fun moments.

My mentors **David**, **Jesper,** and **Lena**.

My work BFF **Robin**, for all the arguments, disagreements, and occasional moments of harmony, it would not have been as fun without you. A special thanks also to **My**, **Sune**, **Vandana,** and **Emelie**.

All thoughtful **colleagues**, you are many**,** at Chemical Biology and Physical Chemistry, for the friendly atmosphere and the many innebandy Tuesdays.

Finally, to the most important people in my life, **Sanna** and **Stella**, for all your support and for being so lovable ♥.

8 References

1. Müller V. Optical Mapping of Bacterial Plasmids. Göteborg: Chalmers University of Technology; 2017.

2. Dahm R. Friedrich Miescher and the discovery of DNA. Developmental biology. 2005, 278.

3. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953, 171.

4. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America. 1977, 74.

5. Maxam AM, Gilbert W. A new method for sequencing DNA. Proceedings of the National Academy of Sciences of the United States of America. 1977, 74.

6. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001, 409.

7. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science (New York, NY). 2001, 291.

8. Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, et al. DNA sequencing at 40: past, present and future. Nature. 2017, 550.

9. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. Genomics. 2016, 107.

10. Bayley H. Nanopore sequencing: from imagination to reality. Clinical chemistry. 2015, 61.

11. Laszlo AH, Derrington IM, Ross BC, Brinkerhoff H, Adey A, Nova IC, et al. Decoding long nanopore sequencing reads of natural DNA. Nature biotechnology. 2014, 32.

12. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, et al. Real-time DNA sequencing from single polymerase molecules. Science (New York, NY). 2009, 323.

13. Shendure J, Ji H. Next-generation DNA sequencing. Nature biotechnology. 2008, 26.

14. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. Nature reviews Genetics. 2006, 7.

15. Sandegren L, Andersson DI. Bacterial gene amplification: implications for the evolution of antibiotic resistance. Nature reviews Microbiology. 2009, 7.

16. Carattoli A. Plasmids and the spread of resistance. International journal of medical microbiology. 2013, 303.

17. Schwartz DC, Li X, Hernandez LI, Ramnarain SP, Huff EJ, Wang YK. Ordered restriction maps of Saccharomyces cerevisiae chromosomes constructed by optical mapping. Science (New York, NY). 1993, 262.

18. Neely RK, Deen J, Hofkens J. Optical mapping of DNA: single-molecule-based methods for mapping genomes. Biopolymers. 2011, 95.

19. Levy-Sakin M, Ebenstein Y. Beyond sequencing: optical mapping of DNA in the age of nanotechnology and nanoscopy. Current opinion in biotechnology. 2013, 24.

20. Avery OT, MacLeod CM, McCarty M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus typE III. The Journal of Experimental Medicine. 1944, 79.

21. Watson JD, Crick FH. Genetical implications of the structure of deoxyribonucleic acid. Nature. 1953, 171.

22. Crick FH. On protein synthesis. Symposia of the Society for Experimental Biology. 1958, 12.

23. Crick FH. Central dogma of molecular biology. Nature. 1970, 227.

24. Fleming A. On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of B. influenzæ. British journal of experimental pathology. 1929, 10.

25. Chain E, Florey HW, Adelaide MB, Gardner AD, Heatley NG, Jennings MA, et al. Penicillin as a chemoterapeutic agent. The Lancet Infectious Diseases. 2014, 236.

26. World Health Organization. Antimicrobial resistance: global report on surveillance. 2014.

27. The Review on Antimicrobial Resistance. Tackling drug-resistant infections globally: final report and recommendations. 2016.

28. Wassenaar TM. Bacteria: The Benign, the Bad, and the Beautiful: Wiley-Blackwell; 2011.

29. Lederberg J. Cell genetics and hereditary symbiosis. Physiological reviews. 1952, 32.

30. Novick RP. Plasmid incompatibility. Microbiological reviews. 1987, 51.

31. Madigan MTM, John M, Brock TD. Brock - Biology of Microorganisms: Pearson Prentice Hall; 2006.

32. Srivastava S. Genetics of Bacteria: Springer; 2013.

33. Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EPC, de la Cruz F. Mobility of Plasmids. Microbiology and Molecular Biology Reviews. 2010, 74.

34. Barton BM, Harding GP, Zuccarelli AJ. A general method for detecting and sizing large plasmids. Analytical biochemistry. 1995, 226.

35. Carattoli A. Plasmids in Gram negatives: molecular typing of resistance plasmids. International journal of medical microbiology. 2011, 301.

36. Carloni E, Andreoni F, Omiccioli E, Villa L, Magnani M, Carattoli A. Comparative analysis of the standard PCR-Based Replicon Typing (PBRT) with the commercial PBRT-KIT. Plasmid. 2017, 90.

37. Conlan S, Thomas PJ, Deming C, Park M, Lau AF, Dekker JP, et al. Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing Enterobacteriaceae. Science translational medicine. 2014, 6.

38. Brolund A, Sandegren L. Characterization of ESBL disseminating plasmids. Infectious diseases. 2016, 48.

39. Köser CU, Ellington MJ, Peacock SJ. Whole-genome sequencing to control antimicrobial resistance. Trends in Genetics. 2014, 30.

40. Livermore DM, Wain J. Revolutionising bacteriology to improve treatment outcomes and antibiotic stewardship. Infection & chemotherapy. 2013, 45.

41. Dantas G, Sommer MAO. How to Fight Back Against Antibiotic Resistance. American Scientist. 2014, 102.

42. Sorensen SJ, Bailey M, Hansen LH, Kroer N, Wuertz S. Studying plasmid horizontal transfer in situ: a critical review. Nature reviews Microbiology. 2005, 3.

43. Funnell BE, Phillips GJ. Plasmid Biology: ASM Press; 2004.

44. Blair JMA, Webber MA, Baylay AJ, Ogbolu DO, Piddock LJV. Molecular mechanisms of antibiotic resistance. Nature reviews Microbiology. 2015, 13.

45. Cho H, Uehara T, Bernhardt TG. Beta-lactam antibiotics induce a lethal malfunctioning of the bacterial cell wall synthesis machinery. Cell. 2014, 159.

46. Czaplewski L, Bax R, Clokie M, Dawson M, Fairhead H, Fischetti VA, et al. Alternatives to antibiotics - a pipeline portfolio review. The Lancet Infectious Diseases. 2016, 16.

47. Nordmann P, Dortet L, Poirel L. Carbapenem resistance in Enterobacteriaceae: here is the storm! Trends in molecular medicine. 2012, 18.

48. Liu Y-Y, Wang Y, Walsh TR, Yi L-X, Zhang R, Spencer J, et al. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. The Lancet Infectious Diseases. 2016, 16.

49. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nature reviews Genetics. 2016, 17.

50. Ganai Rais A, Johansson E. DNA Replication—A Matter of Fidelity. Molecular Cell. 2016, 62.

51. Jackson SP, Bartek J. The DNA-damage response in human biology and disease. Nature. 2009, 461.

52. Metzker ML. Sequencing technologies - the next generation. Nature reviews Genetics. 2010, 11.

53. Bahassi el M, Stambrook PJ. Next-generation sequencing technologies: breaking the sound barrier of human genetics. Mutagenesis. 2014, 29.

54. Cui C, Shu W, Li P. Fluorescence In situ Hybridization: Cell-Based Genetic Diagnostic and Research Applications. Front Cell Dev Biol. 2016, 4.

55. Bloomfield VAC, Donald M., Tinoco I. Nucleic Acids: Structure, Properties, and Functions: University Science Books; 2000.

56. Reisner W, Pedersen JN, Austin RH. DNA confinement in nanochannels: physics and biological applications. Reports on progress in physics. 2012, 75.

57. Baumann CG, Smith SB, Bloomfield VA, Bustamante C. Ionic effects on the elasticity of single DNA molecules. Proceedings of the National Academy of Sciences of the United States of America. 1997, 94.

58. Daoud M, De Gennes PG. Statistics of macromolecular solutions trapped in small pores. J Phys France. 1977, 38.

59. Odijk T. The statistics and dynamics of confined or entangled stiff polymers. Macromolecules. 1983, 16.

60. Odijk T. Similarity applied to the statistics of confined stiff polymers. Macromolecules. 1984, 17.

61. Tree DR, Wang Y, Dorfman KD. Mobility of a semiflexible chain confined in a nanochannel. Physical review letters. 2012, 108.

62. Wang Y, Tree DR, Dorfman KD. Simulation of DNA Extension in Nanochannels. Macromolecules. 2011, 44.

63. Werner E, Mehlig B. Confined polymers in the extended de Gennes regime. Physical review E. 2014, 90.

64. Iarko V, Werner E, Nyberg LK, Muller V, Fritzsche J, Ambjornsson T, et al. Extension of nanoconfined DNA: Quantitative comparison between experiment and theory. Physical review E. 2015, 92.

65. Gupta D, Miller JJ, Muralidhar A, Mahshid S, Reisner W, Dorfman KD. Experimental Evidence of Weak Excluded Volume Effects for Nanochannel Confined DNA. ACS Macro Letters. 2015, 4.

66. Reisner W, Beech JP, Larsen NB, Flyvbjerg H, Kristensen A, Tegenfeldt JO. Nanoconfinement-Enhanced Conformational Response of Single DNA Molecules to Changes in Ionic Environment. Physical Review Letters. 2007, 99.

67. Rye HS, Yue S, Wemmer DE, Quesada MA, Haugland RP, Mathies RA, et al. Stable fluorescent complexes of double-stranded DNA with bis-intercalating asymmetric cyanine dyes: properties and applications. Nucleic Acids Research. 1992, 20.

68. Larsson A, Carlsson C, Jonsson M, Albinsson B. Characterization of the Binding of the Fluorescent Dyes YO and YOYO to DNA by Polarized Light Spectroscopy. Journal of the American Chemical Society. 1994, 116.

69. Paik DH, Perkins TT. Dynamics and Multiple Stable Binding Modes of DNA Intercalators Revealed by Single-Molecule Force Spectroscopy. Angewandte Chemie International Edition. 2012, 51.

70. Gunther K, Mertig M, Seidel R. Mechanical and structural properties of YOYO-1 complexed DNA. Nucleic Acids Research. 2010, 38.

71. Kundukad B, Yan J, Doyle PS. Effect of YOYO-1 on the mechanical properties of DNA. Soft matter. 2014, 10.

72. Berman HM, Neidle S, Zimmer C, Thrum H. Netropsin, a DNA-binding oligopeptide structural and binding studies. Biochim Biophys Acta. 1979, 561.

73. Bailly C, Chaires JB. Sequence-specific DNA minor groove binders. Design and synthesis of netropsin and distamycin analogues. Bioconjugate chemistry. 1998, 9.

74. Marky LA, Breslauer KJ. Origins of netropsin binding affinity and specificity: correlations of thermodynamic and structural data. Proceedings of the National Academy of Sciences of the United States of America. 1987, 84.

75. Boger DL, Fink BE, Brunette SR, Tse WC, Hedrick MP. A Simple, High-Resolution Method for Establishing DNA Binding Affinity and Sequence Selectivity. Journal of the American Chemical Society. 2001, 123.

76. Zimmer C, Marck C, Schneider C, Guschlbauer W. Influence of nucleotide sequence on dA.dT-specific binding of Netropsin to double stranded DNA. Nucleic Acids Research. 1979, 6.

77. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. Science (New York, NY). 2010, 327.

78. Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nature reviews Genetics. 2010, 11.

79. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science (New York, NY). 2012, 337.

80. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science (New York, NY). 2013, 339.

81. Rath D, Amlinger L, Rath A, Lundgren M. The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie. 2015, 117.

82. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nature biotechnology. 2014, 32.

83. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation of endogenous human genes. Nature methods. 2013, 10.

84. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, et al. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. Cell. 2013, 155.

85. Anton T, Bultmann S, Leonhardt H, Markaki Y. Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. Nucleus (Austin, Tex). 2014, 5.

86. Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C, Yusa K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nature biotechnology. 2014, 32.

87. McCaffrey J, Sibert J, Zhang B, Zhang Y, Hu W, Riethman H, et al. CRISPR-CAS9 D10A nickase target-specific fluorescent labeling of double strand DNA for whole genome mapping and structural variation analysis. Nucleic Acids Research. 2016, 44.

88. Herschel SJFW. On a case of superficial colour presented by a homogeneous liquid internally colourless. Phil Trans Roy Soc. 1845, 135.

89. Kasha M. Characterization of electronic transitions in complex molecules. Discussions of the Faraday Society. 1950, 9.

90. Lakowicz JR. Principles of Fluorescence Spectroscopy: Springer; 2006.

91. Persson F, Tegenfeldt JO. DNA in nanochannels-directly visualizing genomic information. Chemical Society Reviews. 2010, 39.

92. Levy SL, Craighead HG. DNA manipulation, sorting, and mapping in nanofluidic systems. Chemical Society Reviews. 2010, 39.

93. Friedrich SM, Zec HC, Wang T-H. Analysis of single nucleic acid molecules in micro- and nano-fluidics. Lab on a Chip. 2016, 16.

94. Marie R, Kristensen A. Nanofluidic devices towards single DNA molecule sequence mapping. Journal of biophotonics. 2012, 5.

95. Muller V, Westerlund F. Optical DNA mapping in nanofluidic devices: principles and applications. Lab on a Chip. 2017, 17.

96. Dorfman KD, King SB, Olson DW, Thomas JDP, Tree DR. Beyond Gel Electrophoresis: Microfluidic Separations, Fluorescence Burst Analysis, and DNA Stretching. Chemical Reviews. 2013, 113.

97. Bensimon A, Simon A, Chiffaudel A, Croquette V, Heslot F, Bensimon D. Alignment and sensitive detection of DNA by a moving interface. Science (New York, NY). 1994, 265.

98. Bensimon D, Simon AJ, Croquette VV, Bensimon A. Stretching DNA with a receding meniscus: Experiments and models. Phys Rev Lett. 1995, 74.

99. Michalet X, Ekong R, Fougerousse F, Rousseaux S, Schurra C, Hornigold N, et al. Dynamic molecular combing: stretching the whole human genome for high-resolution studies. Science (New York, NY). 1997, 277.

100. Kaykov A, Taillefumier T, Bensimon A, Nurse P. Molecular Combing of Single DNA Molecules on the 10 Megabase Scale. Scientific reports. 2016, 6.

101. Otobe K, Ohtani T. Behavior of DNA fibers stretched by precise meniscus motion control. Nucleic Acids Research. 2001, 29.

102. Oshige M, Yamaguchi K, Matsuura S, Kurita H, Mizuno A, Katsura S. A new DNA combing method for biochemical analysis. Analytical biochemistry. 2010, 400.

103. Jing J, Reed J, Huang J, Hu X, Clarke V, Edington J, et al. Automated high resolution optical mapping using arrayed, fluid-fixed DNA molecules. Proceedings of the National Academy of Sciences of the United States of America. 1998, 95.

104. Tegenfeldt JO, Prinz C, Cao H, Chou S, Reisner WW, Riehn R, et al. The dynamics of genomic-length DNA molecules in 100-nm channels. Proceedings of the National Academy of Sciences of the United States of America. 2004, 101.

105. Persson F, Utko P, Reisner W, Larsen NB, Kristensen A. Confinement spectroscopy: probing single DNA molecules with tapered nanochannels. Nano letters. 2009, 9.

106. Lam ET, Hastie A, Lin C, Ehrlich D, Das SK, Austin MD, et al. Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. Nature biotechnology. 2012, 30.

107. Freitag C, Noble C, Fritzsche J, Persson F, Reiter-Schad M, Nilsson AN, et al. Visualizing the entire DNA from a chromosome in a single frame. Biomicrofluidics. 2015, 9.

108. Yu M, Hou Y, Song R, Xu X, Yao S. Tunable Confinement for Bridging Single-Cell Manipulation and Single-Molecule DNA Linearization. Small (Weinheim an der Bergstrasse, Germany). 2018, 14.

109. Persson F, Fritzsche J, Mir KU, Modesti M, Westerlund F, Tegenfeldt JO. Lipid-Based Passivation in Nanofluidics. Nano letters. 2012, 12.

110. Duan C, Wang W, Xie Q. Review article: Fabrication of nanofluidic devices. Biomicrofluidics. 2013, 7.

111. Jo K, Dhingra DM, Odijk T, de Pablo JJ, Graham MD, Runnheim R, et al. A single-molecule barcoding system using nanoslits for DNA analysis. Proceedings of the National Academy of Sciences of the United States of America. 2007, 104.

112. Das SK, Austin MD, Akana MC, Deshpande P, Cao H, Xiao M. Single molecule linear analysis of DNA in nano-channel labeled with sequence specific fluorescent probes. Nucleic Acids Research. 2010, 38.

113. Hastie AR, Dong L, Smith A, Finklestein J, Lam ET, Huo N, et al. Rapid Genome Mapping in Nanochannel Arrays for Highly Complete and Accurate De Novo Sequence Assembly of the Complex Aegilops tauschii Genome. PLoS ONE. 2013, 8.

114. McCaffrey J, Young E, Lassahn K, Sibert J, Pastor S, Riethman H, et al. High-throughput single-molecule telomere characterization. Genome research. 2017, 27.

115. Zhang D, Chan S, Sugerman K, Lee J, Lam ET, Bocklandt S, et al. CRISPR-bind: a simple, custom CRISPR/dCas9-mediated labeling of genomic DNA for mapping in nanochannel arrays. bioRxiv. 2018.

116. Pendleton M, Sebra R, Pang AW, Ummat A, Franzen O, Rausch T, et al. Assembly and diploid architecture of an individual human genome via single-molecule technologies. Nature methods. 2015, 12.

117. Mostovoy Y, Levy-Sakin M, Lam J, Lam ET, Hastie AR, Marks P, et al. A hybrid approach for de novo human genome sequence assembly and phasing. Nature methods. 2016, 13.

118. Xiao S, Li J, Ma F, Fang L, Xu S, Chen W, et al. Rapid construction of genome map for large yellow croaker (Larimichthys crocea) by the whole-genome mapping in BioNano Genomics Irys system. BMC genomics. 2015, 16.

119. Usher CL, Handsaker RE, Esko T, Tuke MA, Weedon MN, Hastie AR, et al. Structural forms of the human amylase locus and their relationships to SNPs, haplotypes and obesity. Nature genetics. 2015, 47.

120. VanBuren R, Bryant D, Edger PP, Tang H, Burgess D, Challabathula D, et al. Single-molecule sequencing of the desiccation-tolerant grass Oropetium thomaeum. Nature. 2015, 527.

121. Yang J, Liu D, Wang X, Ji C, Cheng F, Liu B, et al. The genome sequence of allopolyploid Brassica juncea and analysis of differential homoeolog gene expression influencing selection. Nature genetics. 2016, 48.

122. Chaney L, Sharp AR, Evans CR, Udall JA. Genome Mapping in Plant Comparative Genomics. Trends in plant science. 2016, 21.

123. Martin G, Baurens F-C, Droc G, Rouard M, Cenci A, Kilian A, et al. Improvement of the banana "Musa acuminata" reference sequence using NGS data and semi-automated bioinformatics methods. BMC genomics. 2016, 17.

124. Mak AC, Lai YY, Lam ET, Kwok TP, Leung AK, Poon A, et al. Genome-Wide Structural Variation Detection by Genome Mapping on Nanochannel Arrays. Genetics. 2016, 202.

125. Shi L, Guo Y, Dong C, Huddleston J, Yang H, Han X, et al. Long-read sequencing and de novo assembly of a Chinese genome. Nature communications. 2016, 7.

126. Rosenfeld JA, Reeves D, Brugler MR, Narechania A, Simon S, Durrett R, et al. Genome assembly and geospatial phylogenomics of the bed bug Cimex lectularius. Nature communications. 2016, 7.

127. Barseghyan H, Tang W, Wang RT, Almalvez M, Segura E, Bramble MS, et al. Next-generation mapping: a novel approach for detection of pathogenic structural variants with a potential utility in clinical diagnosis. Genome medicine. 2017, 9.

128. Chan EKF, Cameron DL, Petersen DC, Lyons RJ, Baldi BF, Papenfuss AT, et al. Optical mapping reveals a higher level of genomic architecture of chained fusions in cancer. Genome research. 2018, 28.

129. Seo J-S, Rhie A, Kim J, Lee S, Sohn M-H, Kim C-U, et al. De novo assembly and phasing of a Korean human genome. Nature. 2016, 538.

130. Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, et al. Improved maize reference genome with single-molecule technologies. Nature. 2017, 546.

131. Zhang Q, Xu X, Ding L, Li H, Xu C, Gong Y, et al. Clinical application of single-molecule optical mapping to a multigeneration FSHD1 pedigree. Molecular Genetics & Genomic Medicine. 2019, 7.

132. Eisfeldt J, Pettersson M, Vezzi F, Wincent J, Käller M, Gruselius J, et al. Comprehensive structural variation genome map of individuals carrying complex chromosomal rearrangements. PLoS Genetics. 2019, 15.

133. Levy-Sakin M, Pastor S, Mostovoy Y, Li L, Leung AKY, McCaffrey J, et al. Genome maps across 26 human populations reveal population-specific patterns of structural variation. Nature communications. 2019, 10.

134. Kronenberg ZN, Fiddes IT, Gordon D, Murali S, Cantsilieris S, Meyerson OS, et al. High-resolution comparative analysis of great ape genomes. Science (New York, NY). 2018, 360.

135. Dixon JR, Xu J, Dileep V, Zhan Y, Song F, Le VT, et al. Integrative detection and analysis of structural variation in cancer genomes. Nature genetics. 2018, 50.

136. Klimasauskas S, Weinhold E. A new tool for biotechnology: AdoMet-dependent methyltransferases. Trends in biotechnology. 2007, 25.

137. Lukinavičius G, Lapienė V, Staševskij Z, Dalhoff C, Weinhold E, Klimašauskas S. Targeted Labeling of DNA by Methyltransferase-Directed Transfer of Activated Groups (mTAG). Journal of the American Chemical Society. 2007, 129.

138. Vranken C, Deen J, Dirix L, Stakenborg T, Dehaen W, Leen V, et al. Super-resolution optical DNA Mapping via DNA methyltransferase-directed click chemistry. Nucleic Acids Research. 2014, 42.

139. Grunwald A, Dahan M, Giesbertz A, Nilsson A, Nyberg LK, Weinhold E, et al. Bacteriophage strain typing by rapid single molecule analysis. Nucleic Acids Research. 2015, 43.

140. Neely RK, Dedecker P, Hotta J-i, Urbanavičiūtė G, Klimašauskas S, Hofkens J. DNA fluorocode: A single molecule, optical map of DNA with nanometre resolution. Chemical Science. 2010, 1.

141. Lim SF, Karpusenko A, Sakon JJ, Hook JA, Lamar TA, Riehn R. DNA methylation profiling in nanochannels. Biomicrofluidics. 2011, 5.

142. Levy-Sakin M, Grunwald A, Kim S, Gassman NR, Gottfried A, Antelman J, et al. Toward single-molecule optical mapping of the epigenome. ACS nano. 2014, 8.

143. Michaeli Y, Shahal T, Torchinsky D, Grunwald A, Hoch R, Ebenstein Y. Optical detection of epigenetic marks: sensitive quantification and direct imaging of individual hydroxymethylcytosine bases. Chemical Communications. 2013, 49.

144. Gilat N, Tabachnik T, Shwartz A, Shahal T, Torchinsky D, Michaeli Y, et al. Single-molecule quantification of 5-hydroxymethylcytosine for diagnosis of blood and colon cancers. Clinical epigenetics. 2017, 9.

145. Gabrieli T, Sharim H, Nifker G, Jeffet J, Shahal T, Arielly R, et al. Epigenetic Optical Mapping of 5-Hydroxymethylcytosine in Nanochannel Arrays. ACS nano. 2018, 12.

146. Lim SF, Karpusenko A, Blumers AL, Streng DE, Riehn R. Chromatin modification mapping in nanochannels. Biomicrofluidics. 2013, 7.

147. Sharim H, Grunwald A, Gabrieli T, Michaeli Y, Margalit S, Torchinsky D, et al. Long-read single-molecule maps of the functional methylome. Genome research. 2019, 29.

148. Zirkin S, Fishman S, Sharim H, Michaeli Y, Don J, Ebenstein Y. Lighting up individual DNA damage sites by in vitro repair synthesis. Journal of the American Chemical Society. 2014, 136.

149. Reisner W, Larsen NB, Silahtaroglu A, Kristensen A, Tommerup N, Tegenfeldt JO, et al. Single-molecule denaturation mapping of DNA in nanofluidic channels. Proceedings of the National Academy of Sciences of the United States of America. 2010, 107.

150. Welch RL, Sladek R, Dewar K, Reisner WW. Denaturation mapping of Saccharomyces cerevisiae. Lab on a Chip. 2012, 12.

151. Marie R, Pedersen JN, Bauer DL, Rasmussen KH, Yusuf M, Volpi E, et al. Integrated view of genome structure and sequence of a single DNA molecule in a nanofluidic device. Proceedings of the National Academy of Sciences of the United States of America. 2013, 110.

152. Łopacińska-Jørgensen JM, Pedersen JN, Bak M, Mehrjouy MM, Sørensen KT, Østergaard PF, et al. Enrichment of megabase-sized DNA molecules for single-molecule optical mapping and next-generation sequencing. Scientific reports. 2017, 7.

153. Marie R, Pedersen JN, Baerlocher L, Koprowska K, Podenphant M, Sabatel C, et al. Single-molecule DNA-mapping and whole-genome sequencing of individual cells. Proceedings of the National Academy of Sciences of the United States of America. 2018, 115.

154. Nyberg LK, Persson F, Berg J, Bergstrom J, Fransson E, Olsson L, et al. A single-step competitive binding assay for mapping of single DNA molecules. Biochemical and biophysical research communications. 2012, 417.

155. Nyberg L, Persson F, Åkerman B, Westerlund F. Heterogeneous staining: a tool for studies of how fluorescent dyes affect the physical properties of DNA. Nucleic Acids Research. 2013, 41.

156. Akerman B, Tuite E. Single- and double-strand photocleavage of DNA by YO, YOYO and TOTO. Nucleic Acids Research. 1996, 24.

157. Noble C, Nilsson AN, Freitag C, Beech JP, Tegenfeldt JO, Ambjörnsson T. A Fast and Scalable Kymograph Alignment Algorithm for Nanochannel-Based Optical DNA Mappings. PLoS ONE. 2015, 10.

158. Nilsson AN, Emilsson G, Nyberg LK, Noble C, Stadler LS, Fritzsche J, et al. Competitive binding-based optical DNA mapping for fast identification of bacteria--multi-ligand transfer matrix theory and experimental applications on Escherichia coli. Nucleic Acids Research. 2014, 42.

159. Dvirnas A, Pichler C, Stewart CL, Quaderi S, Nyberg LK, Müller V, et al. Facilitated sequence assembly using densely labeled optical DNA barcodes: A combinatorial auction approach. PLoS ONE. 2018, 13.

160. Karami N, Helldal L, Welinder-Olsson C, Åhrén C, Moore ERB. Sub-Typing of Extended-Spectrum-β-Lactamase-Producing Isolates from a Nosocomial Outbreak: Application of a 10-Loci Generic Escherichia coli Multi-Locus Variable Number Tandem Repeat Analysis. PLoS ONE. 2014, 8.

161. Frykholm K, Nyberg LK, Lagerstedt E, Noble C, Fritzsche J, Karami N, et al. Fast size-determination of intact bacterial plasmids using nanofluidic channels. Lab on a Chip. 2015, 15.

162. Alizadehheidari M, Werner E, Noble C, Reiter-Schad M, Nyberg LK, Fritzsche J, et al. Nanoconfined Circular and Linear DNA: Equilibrium Conformations and Unfolding Kinetics. Macromolecules. 2015, 48.

163. Paterson DL. Resistance in gram-negative bacteria: enterobacteriaceae. The American journal of medicine. 2006, 119.

164. Bush K. Alarming beta-lactamase-mediated resistance in multidrug-resistant Enterobacteriaceae. Current opinion in microbiology. 2010, 13.

165. Mathers AJ, Peirano G, Pitout JD. The role of epidemic resistance plasmids and international high-risk clones in the spread of multidrug-resistant Enterobacteriaceae. Clinical microbiology reviews. 2015, 28.

166. Carattoli A. Resistance plasmid families in Enterobacteriaceae. Antimicrobial agents and chemotherapy. 2009, 53.

167. Schultsz C, Geerlings S. Plasmid-mediated resistance in Enterobacteriaceae: changing landscape and implications for therapy. Drugs. 2012, 72.

168. Nyberg LK, Quaderi S, Emilsson G, Karami N, Lagerstedt E, Muller V, et al. Rapid identification of intact bacterial resistance plasmids via optical mapping of single DNA molecules. Scientific reports. 2016, 6.

169. Johnning A, Karami N, Tang Hallback E, Muller V, Nyberg L, Buongermino Pereira M, et al. The resistomes of six carbapenem-resistant pathogens - a critical genotype-phenotype analysis. Microbial genomics. 2018, 4.

170. Lindblom A, Kk S, Muller V, Oz R, Sandstrom H, Ahren C, et al. Interspecies plasmid transfer appears rare in sequential infections with extended-spectrum beta-lactamase (ESBL)-producing Enterobacteriaceae. Diagnostic microbiology and infectious disease. 2018.

171. Bikkarolla S, Nordberg V, Rajer F, Müller V, Kabir M, Sriram K, et al. Optical DNA mapping combined with Cas9-targeted resistance gene identification for rapid tracking of resistance plasmids in a neonatal intensive care unit outbreak. mBIO. 2019, 10.

172. Chakravorty S, Helb D, Burday M, Connell N, Alland D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. Journal of microbiological methods. 2007, 69.

173. Singhal N, Kumar M, Kanaujia PK, Virdi JS. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Frontiers in microbiology. 2015, 6.

174. Berg JS, Khoury MJ, Evans JP. Deploying whole genome sequencing in clinical practice and public health: meeting the challenge one bin at a time. Genetics in medicine. 2011, 13.

175. Blauwkamp TA, Thair S, Rosen MJ, Blair L, Lindner MS, Vilfan ID, et al. Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. Nature microbiology. 2019, 4.

176. Rossen JWA, Friedrich AW, Moran-Gilad J. Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. Clinical microbiology and infection. 2018, 24.

177. Wand NO, Smith DA, Wilkinson AA, Rushton AE, Busby SJW, Styles IB, et al. DNA barcodes for rapid, whole genome, single-molecule analyses. Nucleic Acids Research. 2019, 47.

178. Sabirova JS, Xavier BB, Ieven M, Goossens H, Malhotra-Kumar S. Whole genome mapping as a fast-track tool to assess genomic stability of sequenced Staphylococcus aureus strains. BMC research notes. 2014, 7.

179. Shukla SK, Pantrangi M, Stahl B, Briska AM, Stemper ME, Wagner TK, et al. Comparative whole-genome mapping to determine Staphylococcus aureus genome size, virulence motifs, and clonality. J Clin Microbiol. 2012, 50.

180. Schwan WR, Briska A, Stahl B, Wagner TK, Zentz E, Henkhaus J, et al. Use of optical mapping to sort uropathogenic Escherichia coli strains into distinct subgroups. Microbiology (Reading, England). 2010, 156.

181. Protozanova E, Zhang M, White EJ, Mollova ET, Broeck DT, Fridrikh SV, et al. Fast high-resolution mapping of long fragments of genomic DNA based on single-molecule detection. Analytical biochemistry. 2010, 402.

182. Kotewicz ML, Jackson SA, LeClerc JE, Cebula TA. Optical maps distinguish individual strains of Escherichia coli O157 : H7. Microbiology (Reading, England). 2007, 153.

183. Chen Q, Savarino SJ, Venkatesan MM. Subtractive hybridization and optical mapping of the enterotoxigenic Escherichia coli H10407 chromosome: isolation of unique sequences and demonstration of significant similarity to the chromosome of E. coli K-12. Microbiology (Reading, England). 2006, 152.

184. Schoepp NG, Schlappi TS, Curtis MS, Butkovich SS, Miller S, Humphries RM, et al. Rapid pathogen-specific phenotypic antibiotic susceptibility testing using digital LAMP quantification in clinical samples. Science translational medicine. 2017, 9.

185. Baltekin O, Boucharin A, Tano E, Andersson DI, Elf J. Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. Proceedings of the National Academy of Sciences of the United States of America. 2017, 114.

186. Boolchandani M, D'Souza AW, Dantas G. Sequencing-based methods and resources to study antimicrobial resistance. Nature reviews Genetics. 2019, 20.