universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation / Title of the Doctoral Thesis

„Development of a DNA-based multiplex assay to characterise pathogens and their antibiotic resistance and virulence factor genes“

verfasst von / submitted by

Noa Wolff M.Sc. B.Sc.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 794 685 437
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Biologie
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Biol. Dr. Angela Witte
Betreut von / Supervisor:	Priv.-Doz. Dipl.-Ing. Dr. Andreas Weinhäusel

TABLE OF CONTENTS

Abstract

The spread of antibiotic resistance is recognised by the European Union (EU) and the World Health Organisation (WHO) as one of the most severe global health threats[1].

The danger is posed by multiresistant pathogens that are resilient to many of the commonly used antibiotics. Examples of these are carbapenem-resistant and β-lactamase-producing *Enterobacteriaceae* or the six pathogens that are summarised by the acronym ESKAPE: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp. These pathogens represent a major threat, particularly in the clinical field, e.g. during surgery, cancer therapy and for immunocompromised patients. It was estimated that up to 10 million people would die from antibiotic-resistant pathogens per year by 2050[2].

The use of antibiotics is excessive in agriculture and aquaculture; moreover, in human healthcare, they are often prescribed by physicians without suitable diagnostic procedures. The decision of whether to prescribe antibiotics is mostly made empirically. A key component in solving this problem is the development of fast and specific diagnostic methods, e.g. *Point-of-Care-devices* (POC), which could lead to an overall reduction in drug consumption.

The aim of this thesis is the development of such a diagnostic system. This work refers to three publications that provide a significant contribution in this area.

The publication *Low-cost microarray platform to detect antibiotic resistance genes* (publication 1) describes an alternative method for in-house labelling of DNA oligonucleotides. DNA oligonucleotides are used for signalling in DNA microarray assays; commercially labelled DNA oligonucleotides are associated with high costs. The alternative labelling method is based on a terminal deoxynucleotide transferase reaction. In this approach, the DNA oligonucleotides were labelled during an elongation step in the presence of biotin-conjugated nucleotides. This step was essential to adapt our DNA microarray assay into a high-throughput assay because they were required in large quantities. The self-labelled DNA detection oligonucleotides performed equally well in terms of sensitivity and specificity compared to commercially labelled detection oligonucleotides at only ten per cent of their costs[3].

The publication *Crosslinking of PCR primers reduces unspecific amplification products in multiplex PCR* (publication 2) deals with the problem of the multiplex polymerase chain reaction (PCR). PCR is essential for DNA-based diagnostic methods; in this case, multiplex PCRs are very attractive as they decrease the number of reactions. On the other hand, the multiplexing efficiency is impaired by primer interactions, e.g. dimer formation. In this publication, covalently crosslinked primers were used to avoid these undesired side effects. Besides the efficiency, the specificity of the primers could be increased by primer crosslinking in PCRs comprising up to 34 primer pairs targeting the most crucial antibiotic resistance genes in one multiplex reaction[4].

In our publication *Full pathogen characterisation: Species identification including the detection of virulence factors and antibiotic resistance genes via multiplex DNA-assays* (publication 3), we developed a DNA microarray-based assay that screens for the most critical sepsis-relevant 45 pathogenic species, 360 virulence factors, and 409 antibiotic resistance genes in parallel. The assay was evaluated with 14 multidrug-resistant strains, including all ESKAPE pathogens. The used platform was optimised regarding specificity and sensitivity.

Zusammenfassung

Die Ausbreitung von Antibiotikaresistenzen wird von der Europäischen Union (EU) und der Weltgesundheitsorganisation (WHO) als eine der größten globalen Bedrohungen der Gesundheit angesehen[1].

Die größte Gefahr geht von multiresistenten Krankheitserregern aus, die gegen viele der häufig verwendeten Antibiotika resistent sind. Beispiele für diese sind Carbapenem-resistente und β-Lactamase produzierende *Enterobacteriaceae* oder die sechs Krankenhauskeime, die unter dem Akronym ESKAPE zusammengefasst werden: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* und *Enterobacter* spp. Solche multiresistenten Krankheitserreger stellen im klinischen Bereich eine große Gefahr dar, z. B. während chirurgischer Eingriffe, der Krebstherapie und insbesondere für immungeschwächte Patienten. Schätzungen zufolge werden 2050 jährlich bis zu 10 Millionen Menschen an antibiotikaresistenten Krankheitserregern sterben[1].

Antibiotika werden in der Landwirtschaft im Überschuss eingesetzt, aber auch außerhalb der großen Kliniken von Medizinern ohne geeignete diagnostische Verfahren verschrieben. Die Entscheidung, ob Antibiotika verschrieben werden, wird zum Großteil empirisch getroffen. Ein zentraler Bestandteil der Lösung dieser Problematik ist die Entwicklung schneller und spezifischer Diagnosemethoden (*Point-of-Care-devices*, POC). Dies könnte zu einer umfassenden Verringerung des Medikamentenkonsums führen und somit die weitere Ausbreitung der Antibiotikaresistenzen minimieren.

Das Ziel dieser Arbeit ist die Entwicklung eines solchen Diagnoseverfahrens. Diese Arbeit bezieht sich dabei auf drei Publikationen, die einen wesentlichen Beitrag in diesem Bereich leisten.

In der Publikation *Low-cost microarray platform to detect antibiotic resistance genes* (Publikation 1) behandeln wir eine alternative Methode zur hausinternen Markierung von DNA-Oligonukleotiden, welche für die Signalgebung bei ligationsbasierten DNA-Microarray-Verfahren verwendet werden. Die alternative Markierung basiert auf einer terminalen Desoxynukleotidyltransferase-Reaktion, in welcher die DNA-Oligonukleotide in Gegenwart von Biotin-konjugierten Nukleotiden elongiert werden. Die hausinterne Markierung ermöglicht es, die hohen Kosten durch kommerziell markierte DNA-Oligonukleotide besonders im Fall von Hochdurchsatzverfahren zu reduzieren. Die selbstmarkierten Oligonukleotide zeigten hinsichtlich Empfindlichkeit und Spezifität eine vergleichbare Leistung zu kommerziell markierten DNA-Oligonukleotiden, jedoch bei nur zehn Prozent der Kosten[3].

Die Publikation *Crosslinking of PCR primers reduces unspecific amplification products in multiplex PCR* (Publikation 2) behandelt die Problemstellung der Multiplex- Polymerasekettenreaktion (PCR). Die PCR ist für DNA-basierte Diagnosemethoden unerlässlich. Multiplex-PCRs sind besonders attraktiv, da sie die Anzahl der Einzelreaktionen reduzieren. Die Multiplexeffizienz wird jedoch durch Primer-Wechselwirkungen, wie zum Beispiel die Bildung von Primer-Dimeren, beeinträchtigt. In dieser Studie wurde eine kovalente Querverknüpfung von Primern über ihre 5'-Enden verwendet, um die unerwünschten Effekte zu vermeiden. Die Spezifität der Primer sowie die Effizienz der PCR konnten durch Primer-Querverknüpfung erhöht werden, was in PCRs mit bis zu 34 Primer-Paaren, abzielend auf die wichtigsten Antibiotika-Resistenzgene, im Vergleich zu nicht querverknüpften Primern nachgewiesen werden konnte[4].

In der Veröffentlichung *Full pathogen characterisation: Species identification including the detection of virulence factors and antibiotic resistance genes via multiplex DNA-assays* (Publikation 3) wurde ein DNA-basierter Microarray entwickelt, der parallel auf 45 Sepsis-relevanten pathogenen Spezies, 360 Virulenzfaktor- und 409 Antibiotika-Resistenzgene prüft. Der Assay wurde mit 14 multiresistenten Stämmen evaluiert, darunter alle ESKAPE-Pathogene. Die verwendete Plattform wurde hinsichtlich Spezifität und Sensitivität optimiert.

1. Introduction

The discovery of antibiotics was one of the greatest achievements in medicine, which made many of today's standard therapies possible[5]. Due to the often unnecessary use of antibiotics, resistances can develop for a variety of reasons, such as misuse in agriculture and medical healthcare[6]. The greatest danger emerges from multiresistant pathogens, which are resistant to many of the commonly used antibiotics[7]. Examples for escalating resistance spreading are carbapenem-resistant and extended-spectrum β-lactamase-producing *Enterobacteriaceae*[8] or the six pathogens summarised by the acronym ESKAPE (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa and Enterobacter spp.*), highlighted by the Infectious Diseases Society of America for being particularly critical in terms of antibiotic resistances[9,10]. Such multiresistant pathogens constitute a danger, mainly in the clinical area, mostly during surgical interventions, cancer therapy or in immunocompromised patients. It was estimated that by 2050, 10 million people would die per year due to antibiotic-resistant pathogens[2,11,12]. After the golden era of antibiotic treatment, we are heading towards a scenario, in which infectious diseases will again be at the top of the causes of death (Figure 1)[11].

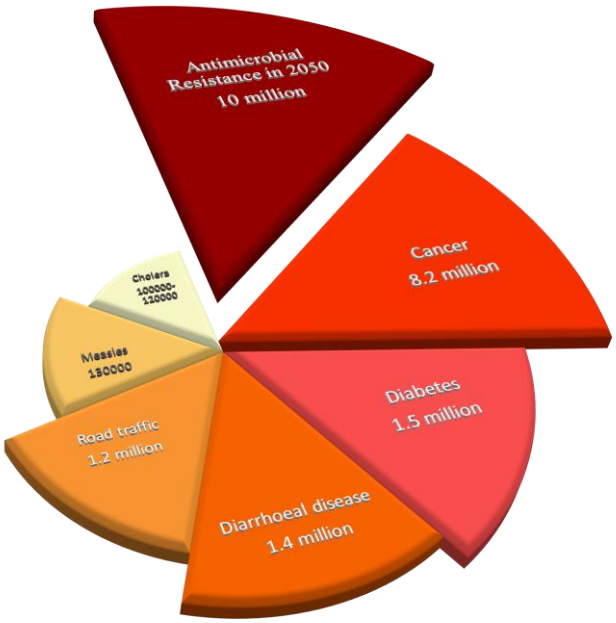


Figure 1: Schematic representation of the estimated distribution of causes of death in 2050. The weighting is visualised by the size of the cake pieces in the diagram. Listed causes of death are cholera (100,000-120,000), measles (130,000), road traffic (1.2 million), diarrhoeal diseases (1.4 million), diabetes (1.5 million), cancer (8.2 million) and antimicrobial resistance (10 million)[11].

To restrain this trend and enable effective medication, as much information as possible must be obtained in the least possible time. However, the main problem with antibiotic therapy is not only the increasing resistance but also the simultaneous decline in the development of new drugs[13] (Figure 2). From a commercial point of view, it is more profitable to invest in other pharmacological areas than in antibiotic research. All in all, the margin yield of antibiotically active ingredients is much lower compared to other active ingredient areas, e.g. chemotherapy[14]. Besides, new active components should serve as "backup". At the same time, pathogens rapidly develop resistances to new antibiotic substances, so that the time for countermeasures is additionally limited[15].

Antibiotics are not only misused in agriculture; a vast majority are prescribed outside the clinics by medical professionals without proper diagnostic tools. The decision of whether to prescribe antibiotic compounds is made "empirically". A central part of the solution to this problem is the development

of rapid and specific point-of-care (POC) diagnostic systems, which could lead to an overall reduction of drug consumption[16]. All of the aforementioned events are illustrated in Figure 2.

Even if an antibiotic prescription is needed, more precise medication will minimise the spread of resistances. To further develop diagnostic procedures, a good understanding of how antibiotic medications work, how resistances are created, and how these can be detected must be achieved. The consideration of alternative kinds of therapy, such as antibodies and bacteriophages, should not be neglected either.

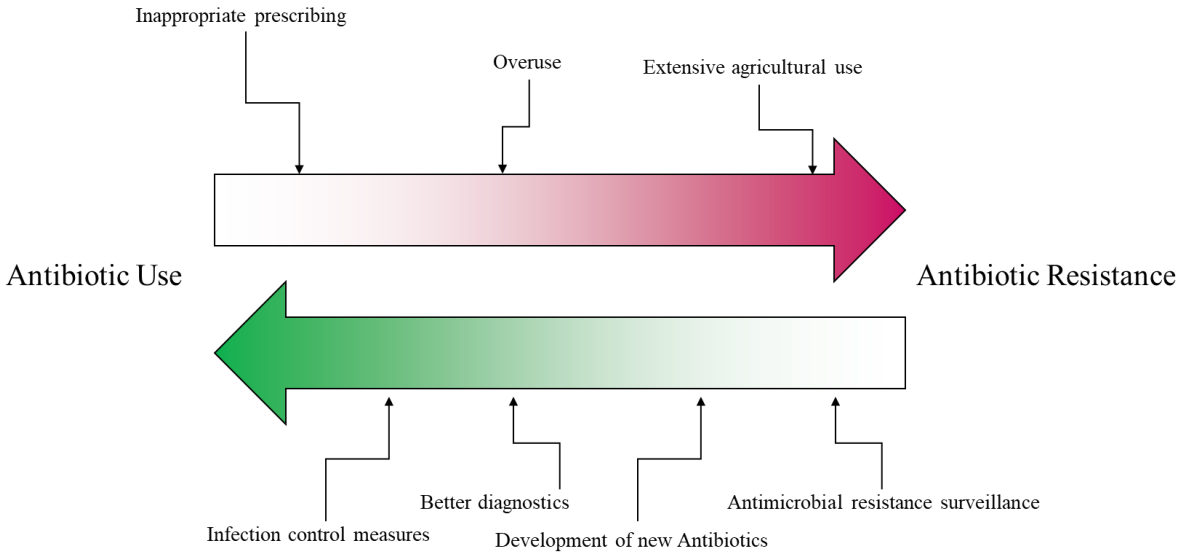


Figure 2: Relationship of the use of antibiotics and resistance development. Improper use of antibiotics is the leading cause of the development of resistance. This factor is aggravated by other factors, which are listed in this graphic. Possible countermeasures are also displayed (adapted from Barbosa and Levy 2000[17]).

1.1 Antibiotics and antibiotic resistance

1.1.1 Definition

S. A. Waksman defined the term antibiotic in 1947 as a "substance that is produced either in microorganisms or chemically and has the property either to inhibit the growth (bacteriostatic) or to kill bacteria (bactericidal) or other microorganisms entirely."[12,18,19] Many of the antibiotics used today are derivatives of molecules produced by microorganisms or by higher organisms such as plants or animals[13,20,21]. However, only a small percentage of these is suitable for the application in humans or animals[22].

1.1.2 Classification of antibiotic resistance

The failure of antibiotics in the therapy does not result exclusively from antibiotic resistance, but can also be attributed to suppressed immune function, poor/disadvantageous bioavailability of the drugs or an increased drug metabolism[23]. The persistence of microbes might indicate different types of resistance to antimicrobial agents; antibiotic resistance can be classified as primary or secondary[23].

Primary Resistance: Resistances of microbes that had no contact with the drug of interest in the host, also referred to as intrinsic resistance.

Secondary Resistance: Resistances of microbes that have been exposed to the antibiotic agent, also referred to as acquired resistance. This extensive resistance is described as the capability to withstand the inhibitory effects of at least one or two antimicrobial drugs. This ability appears to occur in patients who have been treated with first-line medication[23]. With this distinction, the different resistance mechanisms can be viewed.

1.1.3 Mechanisms of action

The antibiotic effect includes a variety of mechanisms of action and depends on the type of bacterium, its habitat, and its metabolism. First, a distinction must be made between different anatomies of pathogens. Gram-positive bacteria have two borders to their surrounding: a cytoplasmic membrane, which is covered by a thick peptidoglycan cell wall. A smaller periplasmic space compared to Gram-negative bacteria is located between cytoplasmic membrane and cell wall. Gram-negative bacteria have a thinner cell wall covering their cytoplasmic membrane, but an additional outer cell wall equipped with lipopolysaccharides. The entire volume between inner and outer membrane constitutes the periplasmic space (Figure 3)[24]. The effect of β-lactam antibiotics on gram-negative bacteria serves as an example of *intrinsic resistance*. A β-lactam antibiotic, such as penicillin, which attacks the peptidoglycan synthesis in the cell walls, has no inhibitory effect on gram-negative bacteria. The main target for β-lactam is the penicillin-binding protein (PBP). It is hypothesised that the β-lactam ring mimics the D-alanyl-D-alanine section of the peptide chain that typically binds to the PBP.

The penicillin-binding protein is responsible for the crosslinking of the bacterium's cell wall.

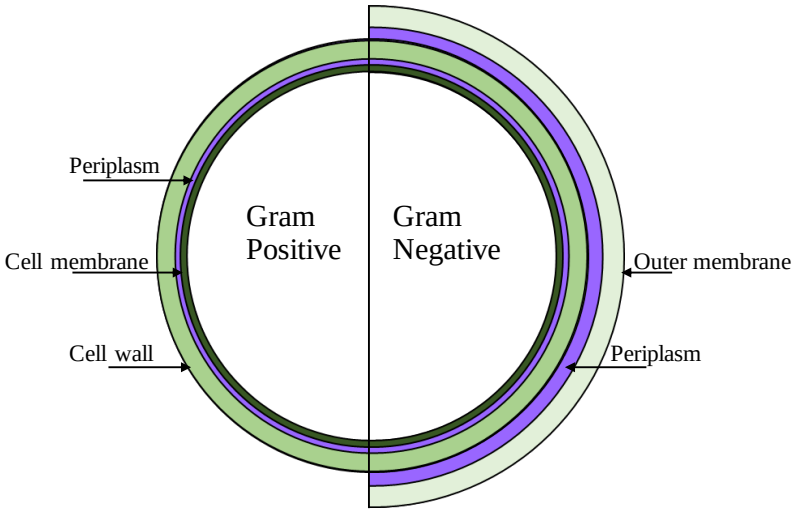


Figure 3 : Structure of the cell walls of Gram positive and negative bacteria. Left, the structure of the gram-positive cell wall with the cytoplasmic membrane (dark green), the peptidoglycan layer (grass green). Right, the outer cell structure of gram-negative bacteria. Additionally, the outer membrane is shown (light green). In between (purple) the periplasmic space is shown.

This enzyme is active during cell division. Binding to this enzyme is irreversible and results in cell lysis during cell division[25,26]. Therefore, the β-lactam ring harms gram-positive bacteria[24]. Another cell wall active antibiotic drug is vancomycin[25]. It binds to the D-alanyl-D-alanine (D-Ala-D-Ala) at the C-terminus of the peptidoglycan pentapeptide precursor. This blocks the further addition of pentapeptide precursors to the resulting peptidoglycan chain and prevents the subsequent crosslinking catalysed by the transpeptidases and the carboxypeptidases[27].

Colistin (Polymyxin E), on the other hand, a small lipopeptide molecule (~ 1200 Da), is affecting gram-negative pathogens. It interacts with the lipid A portion of the lipopolysaccharide (LPS) structure in the outer membrane of gram-negative bacteria so that it restricts their permeability and thus has a bactericidal effect[28]. This antibiotic is not always well-tolerated and is generally considered to be a reserve antibiotic[29,30].

The most common mechanisms of action are summarised in Figure 4. A few examples are given below, which apply for both gram-positive and gram-negative bacteria.

The protein biosynthesis is a rewarding target pool for antibiotics. One target mechanism is the transcription of DNA into mRNA; another is the protein biosynthesis in the ribosomes (translation). An example is the DNA-dependent RNA polymerase, which consists of 5 subunits: two α units (35 kDa), a β-unit (155 kDa), a β'-unit (165 kDa), and a σ-unit (70 kDa). The complex of ααββ' is called the apoenzyme, which is involved in all steps of the transcription. The formation of the holoenzyme (apoenzyme and the regulating σ-unit) makes it possible to recognise promoter sequences of bacterial DNA and thus initiates the transcription[31]. Rifamycin, a key drug against tuberculosis, is an example of RNA synthesis inhibition: It binds to the DNA-dependent-RNA-polymerase, more precisely to the β-subunit, and interferes with the elongation process in the cell[32,33]. The direct inhibition of the protein synthesis has different targets points in the 70S ribosomes; both subunits 50S and 30S are suitable as drug targets.

Tetracycline, for instance, is an antibiotic compound that effects gram-positive and gram-negative bacteria and is the preferable choice in case of a *Mycobacterium* infection, e.g. *Mycobacterium tuberculosis* or *Mycobacterium leprae*. Tetracycline inhibits the interaction of the 16S rRNA with the 30S subunit and inhibits the interaction with the tRNA[26,34,35]. Macrolide antibiotics, in contrast to tetracycline, target the protein synthase at the 50S subunit, where they block the "exit tunnel" of the ribosome and cause the protein synthesis to stand still[36]. Fluoroquinolones bind to the bacterial DNA gyrase/topoisomerase IV complex and prevent the religation of the DNA/enzyme complex during DNA replication, which ultimately leads to cell death[26,37]. The ATPase activity of this enzyme complex constitutes a possible target for antibiotic agents as well[38].

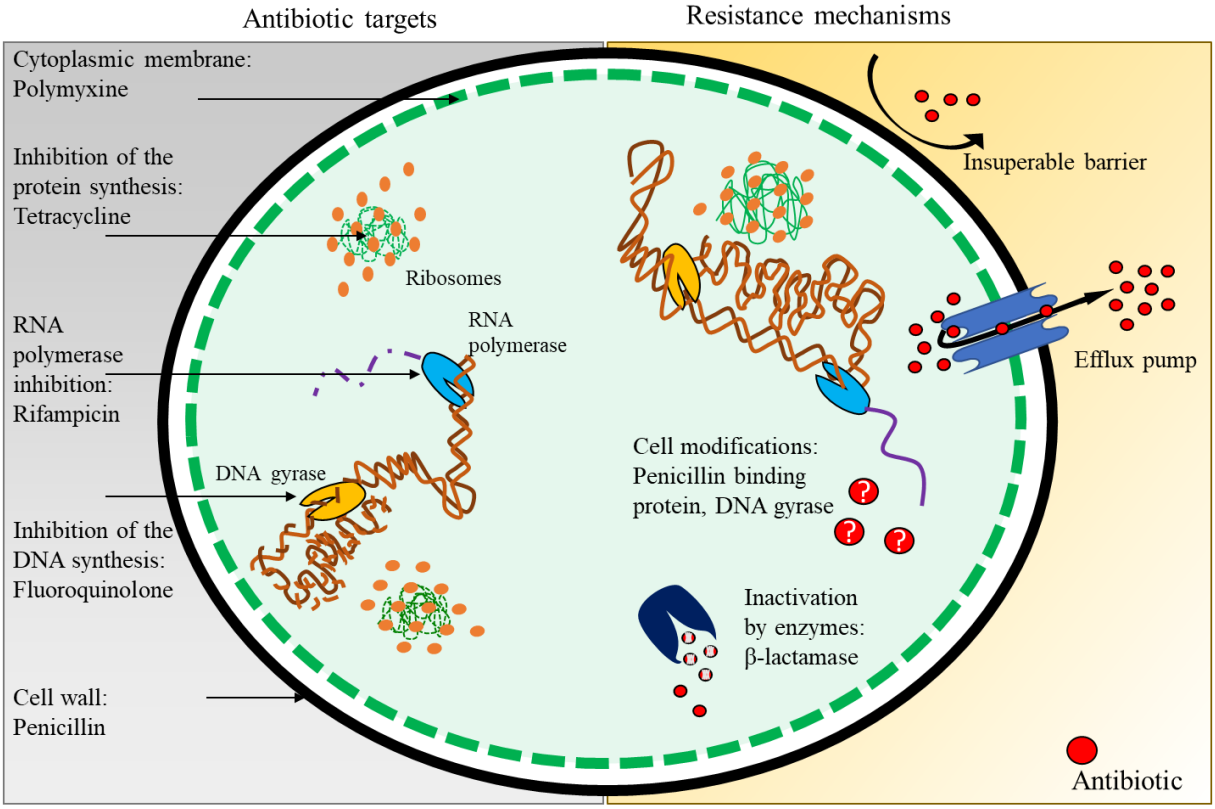


Figure 4: Overview of antibiotic modes of action and resistance mechanisms towards them. The targets are shown on the right (grey). The resistance mechanisms are shown on the left (yellow). An exemplary class of antibiotics is given for each target. In the case of resistance mechanisms, prominent representatives are also listed.

1.1.4 Antibiotic resistance

By using antibiotics, resistance development is inevitable. Resistances to antibiotic agents had already existed before humans started to use them systematically, e.g. the serine beta-lactamases. Antibiotic resistance is not necessarily a single mechanism; an interplay of many different mechanisms can reduce the sensitivity to antibiotic agents: (I) The pathogen can reduce the concentration of the harming substance by shuttering it out of the cell, e.g. efflux pumps, (II) it can inactivate or modify the substance, (III) the pathogen is able to alter the drug target, so it cannot be

harmed any longer, or (IV) it develops an alternative pathway for the cellular processes the target is responsible for[39].

A pathogen's ability to achieve resistance to an antimicrobial agent can be the result of an intrinsic/neutral resistance, such as the expression of efflux pumps, which are able to *transport substances out of the cell*[40] and play a key role in the formation of biofilms[41–43]. Efflux pumps are particularly crucial when they are overexpressed[44,45]. This occurs by means of mutations in the respective pump gene, in its regulation system or by the acquisition of new pump genes via mobile genetic elements, e.g. plasmids[46]. Five different efflux pump families are identified to date: Resistance-nodulation-division (RND), Major facilitator superfamily (MFS), Small multidrug resistance (SMR), Multidrug and toxic compound extrusion (MATE) and ATP-binding cassette (ABC) transporters[39,47]. The mechanisms and the molecules differ among the efflux pump families. Although not every detail has been fully clarified, it is well established that they contribute to the development of resistance to a high degree[39,47]. Another natural resistance is the reduction of harming substances in the cell by modifying the permeability of the shell. This is specifically true for Gram-negative bacteria since they have an additional outer shell that acts as a barrier. Antibiotics enter the cell of gram-negative bacteria through purines, which are not selective towards the penetrating substance. A possible resistance mechanism is to reduce the expression of these purines, e.g. in the case of *Pseudomonas aeruginosa*, or to change their permeability or diameter[39,48,49]. Although Gram-positive bacteria do not control the permeability of the outer shell to a greater extent, there are examples of cell wall-specific resistance mechanisms, such as the vancomycin-intermediate *Staphylococcus aureus* (VISA). Vancomycin, as explained above, prevents the incorporation of crosslinking precursors in the peptidoglycan layer[50,51]. Vancomycin-intermediate *S. aureus*, however, thickens its cell wall, so that an increased concentration of vancomycin must be applied to produce the desired cell lysis[50,51]. For vancomycin, there are alternative resistance mechanisms in *S. aureus*, such as alternative biosynthesis genes to alternate the C-terminus of the peptidoglycan precursor (*vanA, vanB, vanD* & *vanC, vanE, vanG*)[27].

A further resistance method is to *inactivate or to modify the antibiotic substance*, e.g. by an enzymatic reaction. A well-known example is the hydrolysis of the β-lactam ring of penicillin (Figure 5).

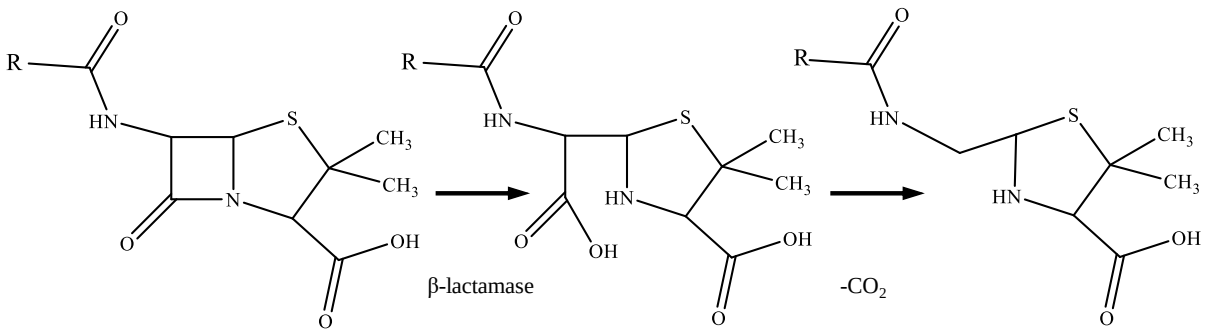


Figure 5: Reaction equation of the hydrolysis of a β -lactam ring by the β -lactamase. A water molecule serves as second educt for the enzyme.

β-lactamases are categorised into four different classes, A, B, C & D (Table 1). A, C & D are serine-β-lactamases; the members of the fourth group (B) are metallo-β-lactamases, which require a divalent zinc ion as cofactor for catalytic activity[52,53]

Table 1: Different classes of β-lactamases and representatives of those[52].

Class	Enzyme family	Representative
A	serine-β-lactamases	CTX-M, SHV, TEM, VEB, GES, KPC
B	metallo-β-lactamase	VIM, IMP, NDM
C	serine-β-lactamases	AmpC
D	serine-β-lactamases	OXA

β-lactamases are capable of hydrolysing cephalosporins and carbapenems of the third or fourth generation (extended-spectrum β-lactamases, ESBLs). Those antibiotic β-lactams were developed to combat the increasingly high number of bacteria resistant to penicillin and other early β-lactam compounds[54,55]. Another kind of modification is group transfer. It results in the alteration of the active substance and therefore weakens the binding to the target. Typical functional groups are acyl, nucleotidyl and phosphatidyl residues[39]. Well-known representatives of this type are the aminoglycoside-modifying enzymes. Aminoglycoside-modifying enzymes exchange amino and hydroxyl groups, which leads to a weaker binding to the ribosome (30S subunit) and results in a high level of resistance. The most common representatives of this enzyme class are N-acetyltransferases (AAC), which transfer an acetyl group using an acetyl-coenzyme A as donor to affect the amino functions, O-nucleotidyltransferases (ANT), and O-phosphotransferases (APH), which both use ATP as donor to affect hydroxyl functions by nucleotidylation and phosphorylation, respectively[56]. There are various further examples of resistances to antibiotics achieved by enzymatic degradation or modifications such as macrolide-esterases[57], epoxidases[31], etc., which are of great impact[58]. A detailed description is given by Wright. et. al[58].

In comparison to active substance modification, the substance's target can be modified by the cell. For instance, the β-lactam ring can either be hydrolysed by the β-lactamases or its target, the penicillin-binding protein, can be modified. Both consequently lead to resistance towards the active

substance. The *alteration of the target molecule*, in this case, the penicillin-binding protein (PBPs), can serve bacteria to survive treatment. Due to the weaker binding of the antibiotic to the modified protein, the antibiotic effect is reduced[53]. The target modifications can be achieved by mutations of the encoding gene or by enzymatic reactions after translation[39]. Quinolones, for instance, are broad-spectrum antibiotics, which inhibit the DNA gyrase and topoisomerase IV (both belonging to the type II topoisomerases) in the bacterial cell. Type II topoisomerases are responsible for the relaxation of DNA supercoils during cell division (replication) and transcription[38,59]. These enzymes create a double-strand break within the twisted DNA strands, which has to be religated in the further process[38]. DNA gyrase and topoisomerase IV are heterotetrameric enzymes. DNA gyrase consists of two units of GyrA (97kDA) and two units of GyrB (90kDA). Topoisomerase IV consists of the two units ParC (84 kDa) and ParE (70 kDa)[31]. The subunits GyrA and ParC are responsible for the enzyme/DNA binding and the associated break/ligation activity. ATPase activity that enables these reactions are contained in the subunits GyrB and ParE[31,60]. Quinolones bind reversibly to the enzyme-cleaved DNA complex and prevent the religation of the DNA. The effect is a high concentration of cleaved DNA, leading to cell death[31,60]. Quinolone resistance arises from the exchange of the two amino acids in the GyrA and ParC subunits in the quinolone resistance-determining region (QRDR). The mutations result in a reduced formation of substrate/target complex and thus in reduced sensitivity to quinolones[60]. Another example of target alteration is the aforementioned resistance to rifamycin. The alteration at the β-subunit (*rpoB* gene) is sufficient to weaken the antibiotic effect of rifamycin[31,61].

The methicillin-resistant *S. aureus* (MRSA), conversely, developed resistance to β-lactams by using an *alternative pathway* for cell wall synthesis, mediated by the transpeptidase (PBP2A). This protein, encoded by the *mecA* gene, is not affected by β-lactams; it is a characteristic element of MRSA[62].

Many of the resistance mechanisms presented here can be traced back to chromosomally encoded or plasmid-encoded intrinsic mechanisms and can thus be passed on to other strains.

1.1.5 Transfer and acquisition of resistance

The intrinsic resistance against antibiotic substances can have a variety of reasons. One can be the lack of substance integration into the target cell interior, or merely the absence of the target itself. *P. aeruginosa* serves as an example for low accumulation of antibiotic drugs due to a weak permeability of its cell wall. Mycoplasma, lacking a cell wall (target), is consequently inherently resistant to β-lactam antibiotics. Intrinsic resistance can also be due to chromosomally encoded mechanisms. *Citrobacter*, for example, expresses a chromosomally encoded cephalosporinase, AmpC, which makes it naturally resistant to some β-lactam antibiotics. This genetic resistance can be passed on to other bacteria. The transfer of genetic information is divided two different types: *Vertical*, in which a gene/genetic material is passed on from an individual to its offspring along the (vertical) lineage; the mutant strain might have a biological benefit that will favour its further colonisation (Figure 6). The speed of evolution is thereby increased by the fast succession of generations.

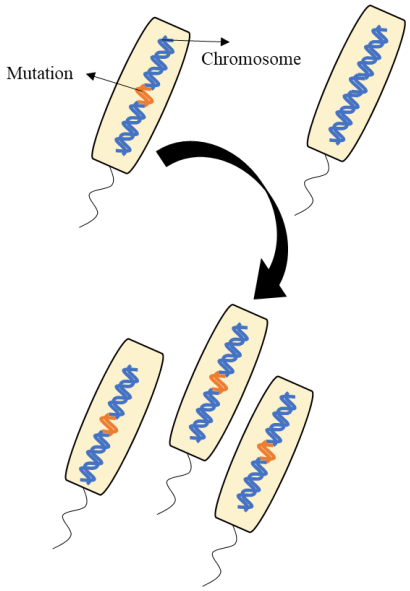


Figure 6: Vertical evolution, by which a gene (mutation) is passed on from an individual to his offspring along the (vertical) lineage (schematic illustration is based on Sommer et.al)[1].

When it comes to passing on resistance, the *horizontal* gene transfer is particularly important, especially in the case of hospital germs, and occurs employing mobile genetic elements. In this process, exogenous resistance mechanisms from different strains are exchanged. The origin of this resistance information must not be exclusively from bacteria[63].

There are three main horizontal gene transfer mechanisms: conjugation, transduction and transformation (Figure 7). During transformation (Figure 7c), naked DNA is taken from the environment of the bacterium; special proteins are involved in several steps of this process. This way of DNA uptake is adopted in the microbiology laboratory, e.g. forced by electroporation or heat shock using specially prepared competent cells[64]. Alternatively, new genetic information can be transferred from one organism to another utilizing transduction (Figure 7b), mediated by bacteriophages. The transmission of resistance genes by bacteriophages is much more widespread than previously considered[65]. Phages are used to integrate genes into host genomes in scientific laboratories as well. But the most significant part of the antibiotic resistance gene spread is the horizontal gene transfer via conjugation (Figure 7a). During conjugation, a direct contact between two cells, which is called pilus, is established, through which the genetic material is exchanged. The direct transmission of genetic material can be driven by low concentrations of antibiotic substances in the environment.

The transfer of genetic information via conjugation occurs within or across bacterial genera and progressively increases the variety of acquired resistances[62].

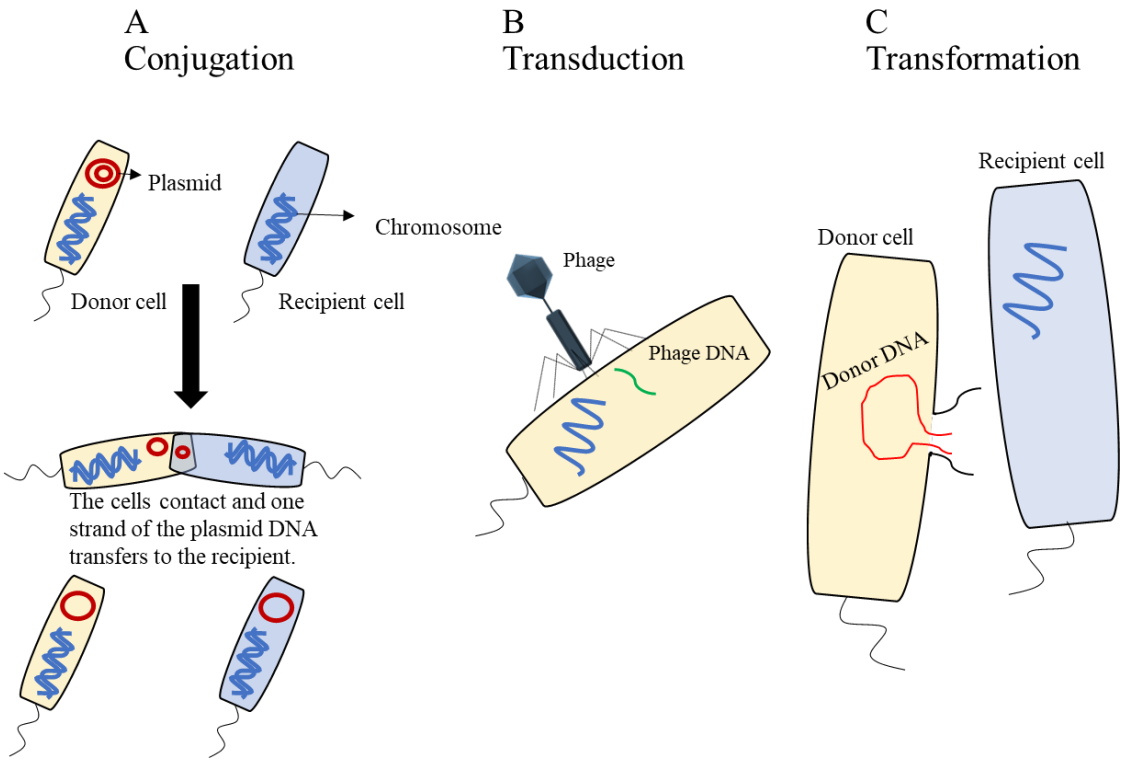


Figure 7: Acquired resistance via horizontal gene transfer. (A) The exchange of genetic information by conjugation occurs via direct contact between two different bacteria. The donor and recipient cells mediate the acquisition of conjugative genetic elements. (B) Phage transduction: a bacteriophage attaches to a bacterial cell and injects its viral DNA, which is then integrated into the chromosome. (C) DNA transformation occurs when naked DNA is released by lysing donor cells and is taken up by another bacterial cell. The schematic illustration is based on Sommer et al.[63].

Mobile genetic elements are required for the conjugation exchange of information. These include plasmids, transposons, and integrons.

Plasmids are non-chromosomal, circular DNA molecules in the range of hundreds to thousands of base pairs. They exhibit their own origin of replication and can thus be replicated autonomously and independent from cell division. Their circular shape entails protection from exonucleases. Plasmids are versatile instruments in genetic engineering to integrate genes in bacterial organisms, mostly including additionally tailored genetic building blocks to select transformants, express proteins, or even for downstream processes such as the purification of those expressed proteins.

Transposons (jumping genes) are mobile genetic elements, which exhibit inverted repeat sequences that enable recognition to cut and insert gene segments into the DNA. This mechanism is coordinated via the enzyme transposase. The information for the transposase is encoded in the transposon itself. Dependent on the transposon type, they excise and relocate (not copying, conservative) or replicate and relocate (copying, replicative). Some transposable elements, the retrotransposons, use an RNA intermediate and a reverse transcriptase instead of the transposase; still, they copy themselves autonomously and integrate into other DNA loci. Transposons and retrotransposons exist on plasmids

and in the genome and are ubiquitous. A distinction is made between conjugative and non-conjugative transposons. Conjugative transposons/retrotransposons carry the requirements for their own transfer (encoding a transposase/reverse transcriptase), non-conjugative transposons require external assistance, e.g. the presence of a conjugative transposable element that supplies the respective enzyme[39,66]. Besides the genes necessary to function, those mobile genetic elements can carry other genes in addition, such as antibiotic resistance genes.

A further, very effective horizontal gene transfer mechanism is the exchange of gene cassettes, only carrying a gene (i.e. a simple open reading frame that usually does not include a promotor) and a recombination site, mediated by integrons. The integrons, comprising an integrase gene, a recombination site, and a promotor, ensure the integration and expression of the gene cassette-encoded gene, which frequently constitutes an antibiotic resistance gene. Since the gene cassettes can exist as free circular DNA molecule, an exchange via transfection is easily feasible, then becoming incorporated via the recombination sites into integrons that can be transcribed[67].

The entity of these different strategies/mechanisms, alone and in all possible combinations, leads to accumulating acquired resistances in bacterial strains that had previously been susceptible to antibiotic therapy. It is a process by which bacteria constantly adapt to new conditions in their environment, accelerated by evolutionary pressure that is increased by the usage of antibiotics. On the other hand, this development is not always beneficial for the organisms. By acquiring resistance, reduced performance in metabolism or phage tolerance is also potentially acquired[68,69].

1.2 Alternative methods to fight pathogens and infection diseases

Shortly after the discovery of penicillin, Alexander Fleming warned that the use of these drugs could lead to resistance to them[70]. The first reports on the development of resistant *Mycobacterium tuberculosis* have already existed in 1948[71], after antibiotics had become the drug of choice for infectious diseases. Despite the warning, no new classes had been developed since 1980[72–75]. The development of antibiotic substances waned in the late 20th century. After the emergence of new resistances had reached an alarming level, at least a few new classes of antibiotics were approved in the last decades. Nevertheless, the development of alternative antibacterial therapies is of tremendous significance. The most prominent approaches are outlined in the following sections.

1.2.1 Phage therapy

The idea of using bacteriophages in antimicrobial therapy is not new (1915 Twort[76] and 1917 d'Hérelle F.[77]), but has gained new importance due to resistance formation. Bacteriophages (also referred to as phages) are viruses that occur in large abundance in nature and only target bacteria[69,75,78,79]. Phages infect bacteria and multiply within them. The first contact between bacterium and phage is generated by diffusion and results in an unspecific and reversible electrostatic bonding. Irreversible binding occurs between capsid proteins of the phage and the receptors on the bacterial surface. Bacterial receptors are usually glycoproteins, liposaccharides, or amino acids. After irreversible binding to the bacterial surface, the phage penetrates the host and either a part of it or the whole phage is introduced into the host cytoplasm, subsequently releasing the genetic material into the host cell. The following steps depend on the character of the phage, whether it is a lysogenic or, as depicted in Figure 8, a lytic phage[78]. Lytic phages (virulent phages) usually lead to rapid lysis of the host cell after capturing the metabolism of the bacterial cell and compelling it to synthesise phage components and assembling new viruses. Finally, this leads to cell lysis by enzymes (holins and lysin) encoded by the phage DNA. Bacterial lysis is essential for the release of the new phages. The genetic material of the phage is protected against digestion by endonucleases via methylation of the cytosine pyrimidine ring. Due to the presence of viral promoters, the phage DNA has a strong influence on the bacterial transcription apparatus and leads to a cascade that completely takes over the host cell metabolism. Phages that preferably remain in the lysogenic cycle are not of interest for phage therapy. The lysogenic phage's DNA is integrated into the host genome and passed on to the daughter cells. At a later point in time only, the lysogenic phage might enter a lytic infection cycle, but only if caused by environmental or another physiological stressor. In antibacterial therapy, it is desired to kill the bacteria as soon as possible.

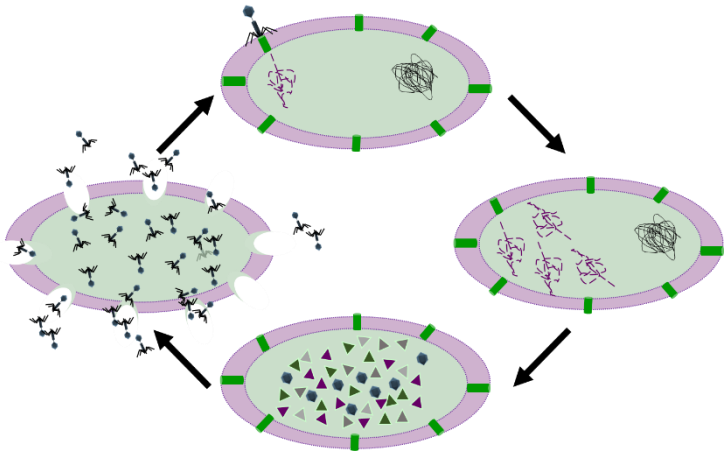


Figure 8: The lytic phage infection cycle begins with the recognition and binding of the virus to the receptor (protein/sugar) on the cell wall of the host bacterium. Thereupon, the phage delivers its genomic content into the cytoplasm. It takes over the bacterial machinery to replicate its own DNA. After expression of the capsid proteins and formation of virus particles, enzymes are produced (holins and lysins) that are responsible for cell lysis. This step leads to the release of the new phages and to the death of the host cell. Figure adapted from Kortright et al.[75]

The use of lytic phages to treat infectious diseases is a promising strategy. Bacteria and phages exhibit co-evolution for millions of years, and their biodiversity is immense[69]. The outstanding advantage of phage therapy compared to chemical antibiotic drugs is their high specificity to a particular bacterial strain. That means they have less off-target effects, which is primarily beneficial regarding the commensal bioflora. Another reason is their low toxicity. Although phages are viruses, prokaryotic cells are their only target; they have been proven to be free of side effects in human patients. They can only exist as long as their host is still living. The use of phages is not only limited to therapy in humans[80] but also to control of plant diseases[81,82], to detect pathogens[83,84], and in food safety[85]. Another advantage of using phages instead of chemical antibiotics is their lower resistance development[86]. By not only blocking specific parts of the bacterial metabolism but by taking over the cellular processes of DNA replication, transcription and translation completely, phage attacks are more difficult to circumvent for bacteria than antibiotic substances. Nevertheless, cases of bacterial resistance towards phages have been reported[87].

Phage resistance mechanisms in bacteria and the well-known strategies with which phages undermine these processes are, for instance, the recognition of the host-specific cell components, which represents the first step in the adsorption of phages[87]. There are three mechanisms to prevent the adsorption to the host cell: (I) the blocking/altering of phage receptors, (II) the production of an additional extracellular matrix, and (III) the production of inhibitors. Bacteria can change their surface structure, so that they are not recognised by specific *phage receptors*. The structural changes can also emerge due to a phage infection by itself. Phage T5, for instance, produces a lipoprotein that blocks its own receptor on the *E. coli* surface; this prevents a super infection of the host cell[88]. Another strategy of adsorption protection is the building of an *additional extracellular matrix*. This not only protects the bacterium from phage infections but can also improve the survival rate of the cell in which further ecological niches can be used. The disadvantage is that some phages have become specialists in precisely recognising these extracellular changes[89]. The development of *competitive inhibitors* is a different mechanism to prevent phage infection. Special proteins prevent the phage DNA from entering the intracellular space; these proteins have their origin in other phages, which means that, in the event of a further phage attack, the peptidoglycan layer is not perforated,

and the phage DNA is trapped in the periplasmatic space. This prevents further infection with related phages, i.e. superinfections[87]. If the phage DNA has made it into the intracellular space, there is an additional resistance mechanism, the digestion of unmethylated phage DNA. These are recognised and degraded by restriction enzymes. Phage DNA can also be recognised by the bacterial methylase and be methylated by it. The extent to which the phage attack is successful depends on the ratio of both enzymes. However, the restriction enzymes are frequently present. The host DNA is usually protected by methylation, and the methylases have a relatively high specificity to half-methylated DNA[90]. In order to defy this protection system, phages have developed different anti-repression strategies. One of them is the absence of endonuclease recognition sites[87], induced by point mutations[91], as exemplified by the *Staphylococcus* phage K83, or by cytosine alteration, for instance, observed at phage T4, which uses the base hydroxymethylcytosine (HMC) instead of cytosine[87] to protect his DNA from degradation. The most prominent resistance mechanism is the CRISPR-*Cas* system, which recognises clustered regularly interspaced short palindromic repeats (CRISPR), remains of previous phage infections,[59] by CRISPR-associated (*Cas)* proteins that, upon recognition, cleave DNA to prevent subsequent infections of similar phages[92]. CRISPR-*Cas* does not only occur in bacteria but had been found in archaea as well[87]. Thus, briefly, CRISPR/C*as* is an immune response that targets foreign DNA[93,94]. The CRISPR-*Cas*9 mechanism offers various application possibilities in gene editing[95].

A detailed description of resistance mechanisms can be found in the publication by Labrie et. al.[87] The understanding of phage/host cell interactions and their resistance development is significant for phage therapy.

Although highly specific, concerns about using phages as an antibacterial agent remain, as the high specificity of phages is one of their greatest disadvantages. Phage treatment requires a preceding accurate identification of the pathogen. The lack of rapid identification tools makes it necessary to apply a cocktail of bacteriophages[75]. The lysis of the pathogenic bacterial cells might entail the release of endotoxins, which can have severe consequences, e.g. causing a sepsis[79]. This and other factors, e.g. pH value, diffusion, longevity, etc., have an impact on the efficiency of phage therapy and must be examined further, so that supplementary animal models are necessary to generate safe therapy[69,79,96]. For additional aspects of phage therapy, the publications Pires et al., 2016, and Kortright et al., 2019, are recommended[75,78].

1.2.2 Lysins and lysin therapy

Lysins are enzymes of bacteriophages degrading the cell wall (peptidoglycan) by hydrolysing different covalent crosslinks of phage-infected bacteria and are thus part of the last phase of the lytic cycle. Lysins, applied therapeutically, are highly specific against gram-positive bacteria; small amounts show a high therapeutic efficacy[69,75,79,97]. Unlike antibiotic chemicals, phage lysin treatment is selective and can be used to treat certain bacterial strains, while commensals are not affected by this therapy[69]. Therapeutically active lysins are divided into five groups based on their endolysin target site/catalytic activity, which are endo-N-acetylglucosaminidases, N-acetylmuramoyl-l-alanine amidases (also referred to as NAM-amidases), N-acetylmuramidases (known as muramidases or lysozymes), endopeptidases (with the slightly different l-alanoyl-d-glutamate endopeptidase that cuts elsewhere), lytic transglycosylases[96]. The most common lysins synthesised by phages are amidases and muramidases[69]. The therapeutic effect of lysin has been demonstrated in several studies dealing with pneumonia[98], abscesses[99], endocarditis[100] and meningitis[101], and infections with methicillin-resistant *S. aureus* (MRSA) with lysine CF-301[97]. Lysins have also been used as preventive medication accompanying surgery to prevent infections with *Enterococcus faecalis, Enterococcus faecium*. The antibiotic effect of lysine is not only promising in medicine or veterinary medicine but also used in the food industry, e.g. against *Clostridium perfringes*[69]. In mice vastly colonised by streptococci, no such bacteria have been found within 2 hours after lysin treatment[102]. In summary, the advantages of lysins therapy can be listed as follows: Lysins are very selective, and even small quantities have an antibiotic effect. So far, no bacterial resistance mechanism has been identified; moreover, a synergy between common antibiotics and lysins has been observed[97]. Bacteriophages are very abundant; therefore, a large repertoire of different lysin derivatives is considerable. Regarding the disadvantages, it has already been mentioned that lysins can only be applied directly to the infected area and affects only gram-positive pathogens. In comparison to chemical antibiotics, lysin peptides trigger an immune response leading to the production of antibodies in the host, therefore reducing the lysin activity[69]. Thus, lysin therapy seems to be a potent alternative to chemical antibiotic substances or as co-medication.

1.2.3 Anti-bacterial monoclonal antibodies

The immune response to the intrusion of pathogens is the production of antibodies. Antibodies are proteins that identify specific components of pathogens and neutralize them with the help of the natural immune response, i.e. macrophages[79]. Therefore, they are a promising alternative to chemical antibiotics. The origin of this form of therapy came from serum therapy and was practised until 1940[103]. It is still used in extraordinary cases as an anti-tetanus medication[104]. Today it is possible to completely generate pathogen-specific monoclonal antibodies (mAbs)[105]. Further, they can be generated as human mAbs to avoid causing an immune response themselves, as it could be the case if murine or other antibodies are used. The particularities of the mAbs are that they are not merely

designed to identify specific pathogen markers, but also for the neutralization of the pathogen, bacterial toxins or virulence factors[79]. The use of antibodies has become established, for example, in cancer therapy[106]. A large number of studies established the use of specific monoclonal antibodies as an antimicrobial agent against bacterial infections[107–111]. One of the disadvantages of antibody research is that manufacturing and processing are associated with high costs. Furthermore, they lose their effect because they cannot adapt to new bacterial surface properties[69]. The field of mAbs is advancing as researchers try to find effective targets and to gain a deeper understanding to enable technological advances in the development and screening of mAbs. There are indications that mAbs – most commonly used in cancer and autoimmune therapy – are suitable as therapy against multiresistant pathogens, too.

Other research approaches, such as antimicrobial peptides, bacteriocins, or vaccines might be mentioned, but are beyond the scope of this thesis.

1.3 Detection methods

Further spread of antibiotic resistant (AMR) can be limited by a better and faster identification and characterisation of pathogens. This can be achieved by the further development of diagnostic techniques, i.e. point-of-care devices (POC). POC information promises a more efficient drug medication and therefore a decline in unnecessary antibiotics consumption. Development of diagnostic procedures should be followed primarily to the criteria; cost-efficient, accurate, fast and if possible directly from the clinical samples[112]. Clinical samples have many pitfalls and are of different matrixes, e.g. blood, urine, faeces, etc. Most of the nosocomial infections are caused by the ESKAPE (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa,* and *Enterobacter species*) strains. These strains have the capability to develop resistance against antibiotic substances. Other pathogens that create a potential hazard are *Clostridium difficile*, carbapenem-resistant *Enterobacteriaceae*, which have developed a wide range of resistances. Rapid and accurate identification and characterisation would make a significant impact on therapy. Due to the increased awareness of this problem, different grants have emerged in recent years, e.g. the Longitude Prize[113]. To investigate this multitude of mechanisms, there is a large repertoire of molecular diagnostics technologies nowadays, which already made their way to clinical everyday routine, as e.g. polymerase chain reaction (PCR) and real-time (RT-)PCR-based detection, matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS), whole genome sequencing (WGS), and the microarray technique. The number of biological examination on a molecular basis increases; nevertheless, most data rely on phenotypic observations[112,114].

1.3.1 Bacterial culture

The first question that arises is whether it is a bacterial infection or not. When a bacterial infection is recognised, the genus (e.g. Staphylococcus or Streptococcus) must be identified by assigned morphological features (e.g. colony size and colour), microscopy (e.g. Gram stain) and biochemical tests (e.g. for catalase and/or oxidase activity). Nevertheless, in clinical microbiology, bacterial culturing is the primary identification method for most of the bacterial pathogens. Depending on the type of sample, the medium must be differentiating and selective. The applied sample is usually grown on an agar-based medium that supports the growth of a wide range of bacteria. For the identification of certain pathogens, the use of a specific culture media is essential. It must have the property to distinguish between the individual pathogens by specific medium properties, which are based on the different metabolic assets of the bacteria. A *differentiating* medium targets other metabolic properties of particular pathogens by utilising biochemical indicator systems (e.g. sugar incorporation, pH indicator, etc.) to detect the presence of certain pathogens. Culture media should also be *selective*. This can be achieved by incorporating antimicrobial agents (e.g. certain antibiotics). This serves to reduce the amount of commensal flora growing and thereby increase the probability

of isolating a specific target pathogen. An example is a faeces sample applied for the detection of *Salmonella* species; the chosen culture medium must be selective as well as differentiated. The medium must reduce the growth of bacteria that are part of normal intestinal flora (*selective*) and at the same time enable the identification of *Salmonellas* (*differentiating*), e.g. by hydrogen sulphide-specific enzymes. Biochemical indicators, which are used to determine the species or genus are, for instance, pH value, fermentation or enzyme activity decarboxylases, deaminases, and tryptophanases. The detection can also be carried out using chromogenic medium; a dye is released through a certain enzymatic processes[112]. An alternative are fluorescence or luminescence dyes since these can be detected with much higher sensitivity. However, the identification of pathogens based on bacterial culture has some disadvantages. It is not only very time-consuming; not all pathogens can be cultivated under laboratory conditions. Therefore, there are other microbiological diagnostic tools that are more rapid and exhibit higher selectivity[112].

1.3.2 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) is a technique that was developed by Kary Mullis in 1980 and has since revolutionised the microbiological laboratory[115]. The target DNA can provide information regarding identification (e.g. 16s RNA DNA) or characterisation (ABR-genes or Virulence factors). It can be used to amplify specific bacteria gene segments in the presence of other organisms[116,117]. The PCR requires DNA primers, nucleotides, and a polymerase enzyme, and depending on the enzyme, a thermocycler is usually used in molecular laboratories. This method receives its specificity through the design of the DNA primers; consequently, the genetic target must either be known beforehand or universal primers must be used targeting conserved gene areas. PCR products are visualised on an electrophoresis agarose gel by a DNA intercalating fluorescence dye. Depending on the experiment planning and source, 4-10 hours are estimated for this method[114]. Compared to conventional PCR, RT-PCR offers higher sensitivity and is less time-consuming. The PCR products are observed during the amplification process through an unspecific intercalating dye or a specific DNA sequences, such as a molecular beacon probe (Figure 9)[118–121].

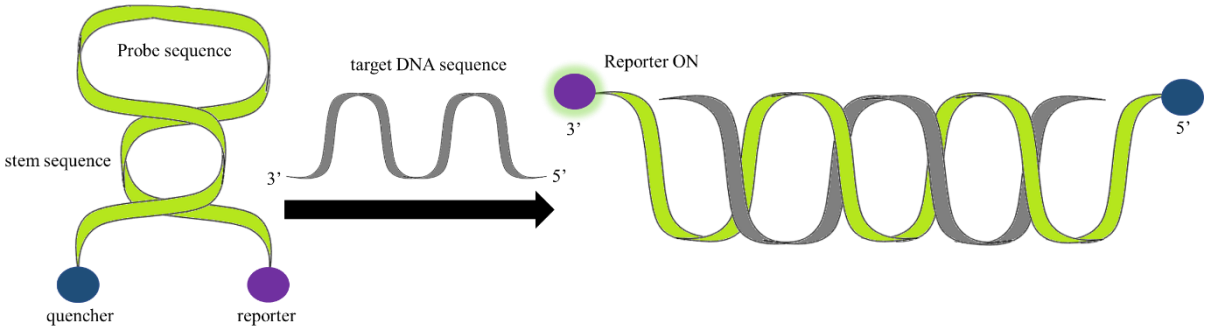


Figure 9: The molecular beacon is a hairpin shape DNA probe, consisting of a 5-'end a fluorescent quencher and a 3'-fluorescent dye. Due to the spatial proximity, no light signal is emitted (FRET). As soon as the molecular beacon hybridizes to the target DNA sequence, quencher and fluorescent dye are sufficiently far away that a signal is visible.

Further examples of DNA amplification methods, including isothermal amplification, are loop-mediated isothermal amplification (LAMP) and recombinase polymerase amplification (RPA). Isothermal methods differ from PCR or RT-PCR in the matter that they do not require a temperature cycle, i.e. the entire amplification takes place at one temperature[121].

Multiplex PCR, which targets numerous DNA sequences at the same time, is of great interest for clinical everyday life and is usually performed either with conventional or with RT-PCR. In one multiplex PCR reaction, several antibiotic resistance genes can be amplified with different primers in a single assay mix[122]. The amplification products can be visualised by size, or by specific DNA probes, i.e. molecular beacon (Figure 9). Due to the simultaneous reaction, this type of PCR is very time-saving[123]. However, the robustness of this detection method relies on multiple factors, such as the gene copy number of the respective gene of interest, optimal primer design, and the GC content of the amplification product. The gene *copy number* of the respective gene of interest, however, represented in the genome or plasmid, differs not only for different genes but can also vary for the same gene between different species, e.g. the ribosomal repetitions in different populations[120,124–126]. The sensitivity is hence dependent on the copy number of the gene of interest, serving as template DNA in the PCR reaction. The risk of false-negative results furthermore depends on the *primer design*. Single-stranded DNA can form secondary structures with itself, which in turn adversely affects the annealing to the target DNA and thus reduces the sensitivity of the assay. In multiplex PCRs, interactions between the individual primers are further complicated. Primers require a high identity coverage to the target DNA but must be able to include possible inaccuracies, such as single nucleotide polymorphisms (SNPs), at the same time. Therefore, when designing the primers, the genetic position of the target DNA is crucial. For this information, several databases are of great importance, such as the virulence factor of pathogenic bacteria database[127], the Comprehensive Antibiotic Resistance Database (CARD)[127,128]. Nevertheless, only *in silico* evaluation is not sufficient for primer assay design. In general, there are several online tools that are dedicated to the specific and sensitive design of primers and include the existing problems. Another limitation of (RT-)PCR is the restriction of the number of target genes that can be processed at a time, for it is usually based on colorimetric detection, and the spectrum of differentiable fluorescence dyes is limited.

Nonetheless, PCR is a quick, inexpensive and straightforward method for the characterisation of pathogens in the clinical everyday live[129].

Another very robust and promising method is the matrix-assisted laser desorption/ionisation time-of-fly mass spectroscopy (MALDI-TOF MS)

1.3.3 Matrix-assisted laser desorption/ionisation time-of-fly mass spectroscopy (MALDI-TOF MS)

MALDI-TOF MS is the primary tool for pathogen identification in the clinical laboratory. Since it is easy to use, rapid, accurate and economical, in comparison to phenotypic methods[130–133,109]. The detection principle is based on evaporation and ionisation behaviour, determining the time the resulting ions requires to reach the detection This flying time is characteristic for the respective pathogen and can hence be used to identify the organism by means of its mass/charge ratio (m/z value) calculated from the peak focus. The spectral fingerprints vary enough to differentiate genera, as long as they exhibit the same growth conditions[134–137]. MALDI-TOF MS enables the direct sample identification of proteins, lipids and carbohydrates and thus made the identification of pathogens and their further characterisation possible. Although MALDI-TOF MS is a sensitive detection method, the uncultivated blood sample requires a bacterial concentration of 10^7 CFU/ml[138]. Nevertheless, the same information level can be achieved by culture-based identification. Still, MALDI-TOF MS generates the results much more rapidly, which is beneficial for sepsis patients outcome, showed by Kumar et al[139,140]. Although the identification of pathogens is an integral part of MALDI-TOF MS, their characterisation by means of antibiotic resistance or virulence factors is still part of research.

Antibiotic resistance detection is established with MALDI-TOF MS using three main approaches: (I) The determination of antibiotic resistance by the detection of antibiotic modifications due to enzymatic activity, (II) the analysis of peak patterns of bacteria profiles and (III) the quantification of bacterial growth in the presence of an antibiotic substance[141]. Each of these approaches has their advantages and disadvantages. The most common resistance method is the *enzymatic alteration of the antibiotic structure*. Such resistances to antibiotics are mediated by proteins and often result in proteolytic fragments of antibiotic compounds. These degradation products of antibiotics are detectable by MALDI-TOF MS via the molecular mass change compared to the starting molecule, which can be observed in the mass spectrometer peak[141]. First, the structure of the antibiotic and the reaction metabolites must be determined precisely. Degradation products and antibiotics are usually analysed in the mass range between 100 and 1,000 Da[141]. For instance, hydrolysis of β-lactams leads to different fragments, yielding compounds of different molecular weights and therefore mass peak patterns that are unique to each antibiotic. So, this β-lactamase activity can be evidenced with the MALDI-TOF MS device[142,143], as well as other antibiotics[144]. A disadvantage of MALDI-TOF MS is that it depends on factors such as protein expression; not yet expressed proteins cannot be detected

at all. The expression of antibiotic-degrading enzymes is often induced by the presence of the respective antibiotic[112,134], which prevents resistance detection prior to the medication.

In order to identify possible antibiotic resistances, *the mass peak profiles of bacteria* of the same species with and without antibiotic resistance are compared with each other[141]. Some mass peak profiles are associated with a particular resistance pattern due to the expression of a specific protein[130]. For example, Josten et al. identified a peptide called phenol-soluble modulin (PSM-*mec*), which is encoded on the staphylococcal cassette chromosome *mec* (SCC*mec*), element cassette type II, III and VIII, and is found in the genomes of MRSA strains[145]. This protein is produced by *agr*-positive strains and its presence can be detected by MALDI-TOF MS[130]. This can be used for further characterisation, but still needs to be reproduced and validated in a clinical setting. In a study by Lau et al. about real-time analysis of infection outbreaks by carbapenemase-producing *Enterobacteriaceae*, a specific peak patter caused by the protein (p019) was observed. This protein is located on the pKpQIL plasmid and was used as a biomarker to identify KPC-producing *Klebsiella pneumoniae*.

In part, the *growth pattern differs* for strains by the addition of antibiotics. The resistant ones show a changed growth compared to non-resistant ones; this difference can be easily distinguished based on a quantitative spectrum[141]. Despite the advantages of MALDI-TOF MS, this method is vulnerable to mistakes, especially regarding species differentiation, for example between *Streptococcus pneumoniae* and *Streptococcus mitis*, which might entail severe consequences[134].

The major disadvantage is the need of a high bacterial load, which precludes the use of this technique with clinical samples except from those with high bacterial burdens, such as positive blood cultures and urine samples, and the lack of commercially available kits.

Irrespective of that, the usage of MALDI-TOF MS contributes well to the patients' outcome rate, which is partially due to the little time required for the identification[139,144] and partially to the detailed characterisation possibilities regarding pathogen[146,147].

To integrate this technology in the clinical everyday routine, further studies concerning the detection of antibiotic resistances and pathogen virulence factors are required.

1.3.4 Whole genome sequencing (WGS)

In recent decades, whole genome sequencing (WGS) has become a promising technique in clinical microbiology. A special feature of WGS is the coverage of many different targets and different gene variants at the same time. WGS is, in comparison to Sanger sequencing, much more sophisticated and produce a greater amount of data. Today's common high-throughput platforms are, for instance, Illumina or Ion Torrent devices, which belong to the second generation of sequencing systems, also referred to as next-generation sequencing. The readout of these machines are relatively short (100-400 bp) and is shorter than the gene of interested. For single reads, the error rate of random and

method-based sequencing from next-generation sequencing technology is relatively high in comparison with conventional Sanger sequencing. This problem is solved by generating a large excess of short-read data for each genome and using it to correct the error by majority call. The overlapping reads can be mapped to known references (reference assembly) or used to create larger fragments (de novo assembly) of sequence data (so-called contigs) that are combined to form the genome design of the isolate[114,148]. For the detection of the relevant gene, quality and quantity of the short reads are important to ensure that the gene is correctly detected by subsequent analysis and thereby avoid false-negative results[148].

WGS provides the following information: (I) pathogen identification, (II) characterisation (e.g. resistance and virulence genes) (III) outbreak detection (IV) recognition of universal sequences for the primers and probes design and (V) the design of specific probes and primers[112].

16S rRNA DNA gene sequences are used to *identify pathogens* and can also be used to determine phylogenetic relationships (Figure 10). The advantage of using these gene sequences is that (a) they are present in all bacteria, (b) the sequences are highly conserved, so that spontaneous mutations are rare, (c) 16S rRNA gene sequences are only approximately 1500 bp long and therefore suits bioinformatic purposes well[149]. WGS provides a *characterisation of the pathogens* beyond the phenotypic performance and is therefore essential for further clarification[150,151]. In order to *identify the spread of nosocomial infections*, their detection with WGS is essential. Close genetic similarity between the isolates supports the likelihood of mutual transmission or a common source. In contrast, the lack of genetic similarity indicates an unrelated, sporadic occurrence of a certain phenotype or infection[151,152]. The increased sequence information of bacteria has led to an improvement in the diagnostic methods, which makes the *design of primers or microarray probes* possible. For example, the *mecC* gene, a homologue of the *macA* gene, which plays a role in methicillin-resistant MRSA, was identified by WGS, so that, suitable primers were designed to identify resistance[62]. One of the main advantages of WGS is the identification of new target sequences and add them to existing databases to enable *in silico* analysis[114].

Figure 10: Phylogenetic tree**.** The phylogenetic distances of the investigated species is illustrated, calculated according to the All-Species Living Tree project using 16S rRNA genes (Yarza et al. 2008[153]; Yarza et al. 2010[154]).

Regarding the content of information, WGS is superior to all other described clinical diagnostic methods, theoretically revealing the entireness of present phylogenetic marker genes, ABR genes, and VF genes on a chromosomal level and on mobile genetic elements. Since the invention of sequencing, there have been many and rapid developments in this field. WGS must deliver results from a sample with different flora in a few hours, which correlate with the phenotypic results in order to be used in routine diagnostic settings. When using WGS directly on a clinical sample, the reduced required time would improve the diagnosis procedure and thus therapy. Hasman et al.[155] demonstrated that based on a urine samples identification from direct samples is possible. It should be emphasised that a urine sample is the simplest matrix form since it contains limited human DNA contamination but a high number of bacterial cells[155]. The resulting reduced analytic duration ultimately has a positive outcome for the patient; in critical cases, the most limiting criterion might indeed be the time needed to analyse a genome. Nevertheless, WGS still does not only require a lot of pre-processing – and therefore time – to obtain the contiguous sequence and thus desired information out of the raw sequence snips, but also a robust bioinformatic expertise that exceeds the demands addressed to the commonly employed executive personnel by far. Trained experts, in turn, increase the costs again, after having been saved in terms of the sequencing reaction.

1.3.5 Microarray

Another alternative for the identification and characterisation of pathogens is nucleic acid microarray analysis (DNA microarray). There are also protein or antibody-based microarrays, which are beyond the scope of this work and are thus only mentioned here for the sake of completeness.

The concept of microarray was first introduced by Chang 1983[156] as an antibody matrix. During the 1990s, the DNA microarray technology emerged and provided the possibility to screen for specific DNA sequence in a large-scale[157] A DNA microarray chip represents a functionalised two-dimensional surface, usually glass slides, plastic, or paper, to immobilise synthetic DNA oligonucleotide probes. The immobilisation of probes to the carrying object has a crucial influence on the array performance regarding sensitivity and specificity[158], which is why a high number of studies are dealing with immobilisation issues[159]. Different methods are being pursued, such as electrostatic interactions[160], affinity reactions[161], passive immobilisation through adsorption forces[162] and covalent bonding[159]. Covalent bonding for attachment, however, is the first choice, especially as it enables good stability, high binding strength and control over orientation and concentration of the probe[158]. Covalent bonds are usually generated by click chemistry using thiols, amines or hydroxyls that are esterified with maleimides, aldehydes or epoxides[159].

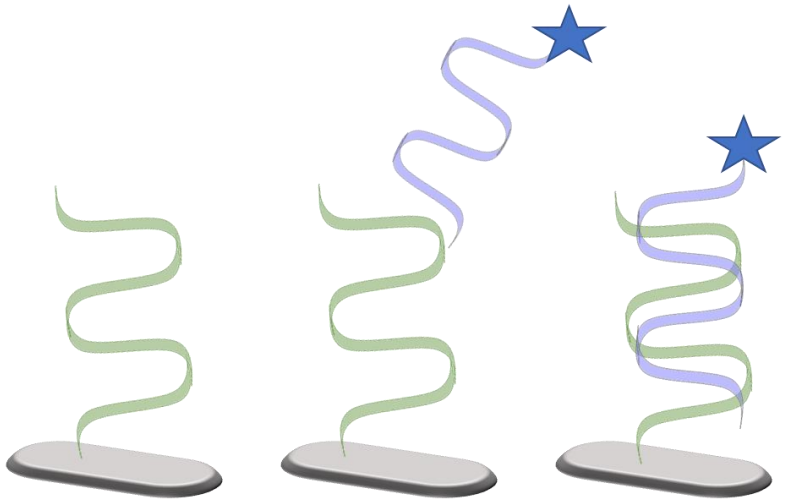


Figure 11: Presentation of the hybridization principal of DNA microarrays. Green is the specific DNA probe that is immobilised on a functionalised surface (grey). The gene of inters (purple) was previously generated by PCR amplification and fluorescence labelled (star). Due to compliant base pair binding, the two DNA strains bound throw hybridisation to one another, and the florescence signal can be read at the precisely defined position on the microarray chip.

The DNA microarray takes advantage of the hybridisation property of two complementary DNA strands[163,164] (Figure 11). The probe, a short synthetic produced single-stranded nucleic acid sequence, is hybridised to the target, a short tagged single-stranded nucleic acid sequence[165]. The target DNA is usually a PCR product[166]. The fact that the sequences of the target gene must be known in advance is one disadvantages of DNA microarray compared to WGS. Nevertheless, DNA microarrays are a good link between the PCR and WGS[167]. DNA microarray analysis is not suitable for the identification of new specific gene sequences in comparison to WGS; however, the high number of publicly available sequence databases fill those gaps and enables the specific microarray analysis[127,154,168,169]. DNA microarrays permit the screening of up to 1000 specific gene sequences

simultaneously[170]. This analysis is preceded by an amplification step, in which some of the deoxyribonucleoside triphosphates are carrying a fluorescent dye label, which can be visualised by a fluorescence scanner. Since the individual specific gene sequences are located at a precisely defined location on the functionalised surface, the fluorescence signal can be assigned to a specific gene of interest. This permits the statement about the presence or absence of certain genes in the cell; in addition, the analysis of RNA or cDNA can provide information about the expression levels of certain genes[165]. Many strain-specific DNA characters can be analysed by DNA microarrays, so that an accurate identification and characterisation of pathogens are possible. The differentiation of closely related strains is even more specific compared to MALDI-TOF MS[171].

To date, a multitude of microarray multiplex platforms has been developed that automatically extract, amplify, and hybridise DNA from a sample and even offer endpoint melting point analyses in one device[112]. FilmArray© from Biomérieux provides a platform that can be used to identify 25 different pathogens and four possible resistant genes within one hour[172]. Further comparative studies have shown that the microarray platform has a high potential to screen for a large number of genes with a high degree of specificity and sensitivity[173]. One drawback of DNA microarray analysis is the need for an upstream amplification step, which is as well associated with PCR-related bias, such as primer interactions, and false-negative amplification[166]. It must be emphasised that the downstream microarray detection is of no use if the upstream sampling and processing is contaminated, false or biased. Furthermore, pathogen quantification with DNA has its limitations. Wu et al.[174] demonstrated this by quantifying certain bacteria in complex communities. The challenge was to differentiate between target frequency and hybridisation efficiency. That means that a gene with a low quantity, but with a high binding affinity to the DNA probe, may generate the same signal strength as a gene sequence with greater quantity but less binding affinity to the probe[166]. Another obstacle in assay design is the sensitivity of DNA microarrays. The sensitivity is negatively influenced by the probe density on the given area[175]. Furthermore, a sequence similarity of 85-90 % is usually sufficient for the hybridisation of the target sequence to its complementary immobilised DNA probe. This impairs the specificity of the hybridisation-based detection method, as it can lead to false-positive results. But false-negative events can also occur. If, for example, the similarity falls below the aforementioned value, which happens mainly due to systematic or random mutations within the gene section, the probability is increased that the probe-target DNA complex will not be formed although it is supposed to[166]. In addition, the formation of the probe-target complex is dependent on the mismatch position, their number and distribution, their GC content, etc. These and further points make the microarray design challenging.

Nevertheless, the microarray method is superior compared to the detection methods mentioned above in many regards. First, it is possible to evaluate a higher number of samples simultaneously compared to mere PCR detection, since the number of target positions is less restricted than the number of differentiable dye wavelengths, melting temperatures, or band positions on a gel. The advantage of

the DNA microarray detection method in comparison to MALDI TOF MS is that it is expression independent. In terms of resistances, MALDI TOF MS can only identify expressed proteins or mass changes of them. A limiting factor is that the DNA sequence must be known in advance. This is the advantage of WGS. But even if the information content resulting from WGS is more substantial, the application, the handling, and foremost the evaluation of DNA microarrays is much easier. Therefore, DNA microarrays constitute a link between PCR and WGS[167]. Altogether, DNA microarray systems are a good basis for the development and implementation of automated identification and characterisation systems as soon as challenges such as cross-hybridisation and interferences during upstream amplification are managed.

Barišić et al. have overcome cross-hybridisation and impeding surface-related interactions by introducing the ligation-dependent linear nucleotide chain (LNC) microarray plattform[176]. This form of DNA microarray makes use of a ligase and its proofreading function to achieve 100 % specificity. Moreover, the specific probe is immobilised in a way that mimics free DNA oligonucleotides in solution.

This specific method of DNA microarray was used as a detection method in this work. A detailed description of the LNC concept, including its advantages, is given in the *Thesis aim* section.

2. Thesis aim

The development of antibiotic substances led to a decrease in deaths caused by infectious diseases. A main problem that still needs to be overcome is the wide variety of antibiotic resistances. Resistance genes arise and spread quickly within bacterial strains, especially due to careless use of antibiotics, e.g. in agriculture, small clinics, and private physicians. The development of new antibiotics declined due to the large number of existing ones. Effective and careful treatment with available antibiotics, however, requires the identification and characterisation of the infectious bacterial strain regarding its resistance and virulence factor genes. The identification of contagious bacteria is still done using cell culture-based methods. Besides being time-consuming and error-prone, there are several pathogens that cannot be cultivated under laboratory conditions. For the effective use of antibiotics in order to keep them functional over a longer period of time, it is necessary to use techniques that enable quick and reliable identification of both the causal pathogen and its resistance and virulence profile.

2.1 Theoretical solution

The focus of this thesis is the evaluation of a DNA-based microarray detection system for the identification and characterisation of pathogens with regard to their resistance and virulence factor genes. The development of a rapid and reliable solid-supported detection system could be achieved with multiplex detection techniques. In addition, the format should not be limited to bacteria, but also includes viruses, fungus, and protozoa.

For this reason, identification and characterisation using DNA is particularly suitable. DNA structures and their storage, processing, and functionalisation differ significantly between different organisms, but the simple mechanism of base pairing and the translation via the genetic code is identical in almost all living organisms and viruses. Viruses, however, that cannot live outside of their host organism and therefore are not subject to cell culture approaches. Another major advantage is that the presence of a gene is independent from its expression. It can be detected using a genetic approach as long as its sequence is known. The latter is often the case, being available in publicly accessible databases, e.g. the virulence factors database[127] and the antibiotic resistance database[168]. Finally, DNA preparation requires little time and is easily manageable due to commercially available high-performance extraction kits. In parallel, solid-support based nucleic acid detection methods are well established, being versatile tools for research projects such as next-generation sequencing, as well as for high-throughput detection applications, i.e. DNA microarrays. The analysis of hundreds or even thousands of samples in parallel is crucial for such medical diagnostic use. Theoretically, the desired sequences can be synthetically prepared and subsequently immobilised in large numbers on a microarray. With the help of DNA extraction, complementary sequences – if present in the respective pathogen – will bind via base pairing to the immobilised sequences of interest. In

combination with a marker technique, high numbers of pathogens and resistance genes could be detected simultaneously. In practice, this theoretical solution is accompanied by several challenges.

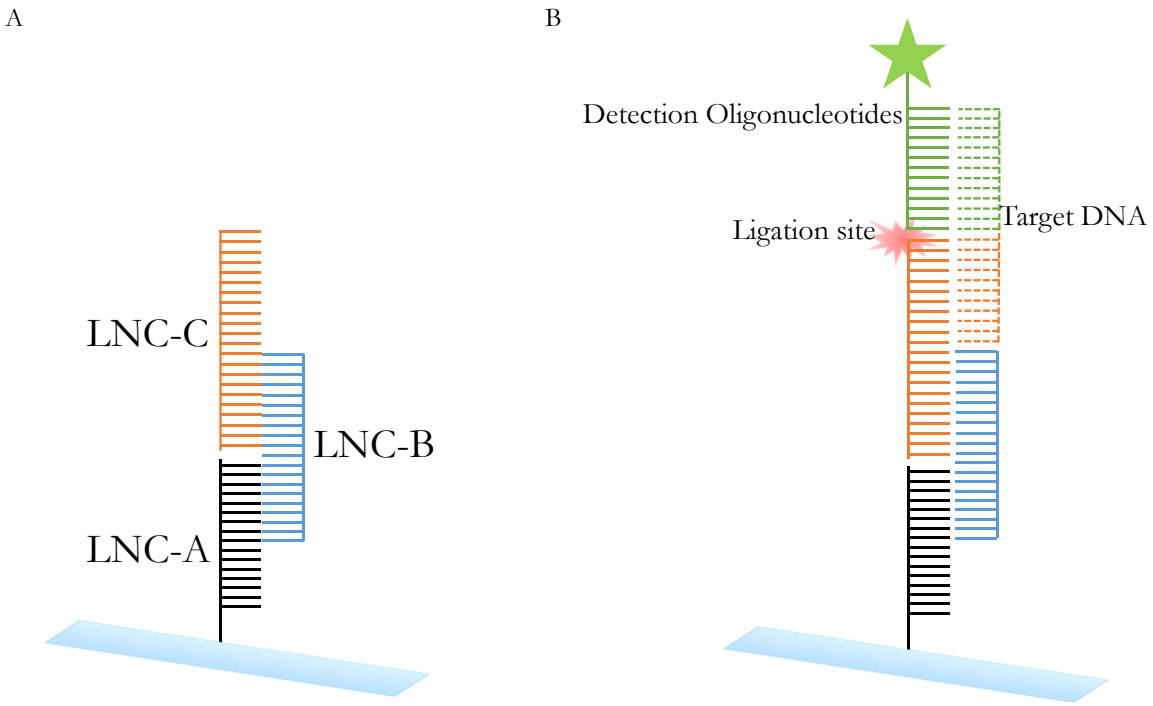


Figure 12: Hybridisation principle. (A) A DNA oligonucleotide, referred to as linear nuclear chain (LNC-) A (black), is covalently linked to a functionalised glass slide (transparent blue) utilising a 5'-terminal maleimide. Via base pair hydrogen bonds, two further oligonucleotides, LNC-B (blue) and LNC-C (solid red) are non-covalently attached. The protruding end of LNC-C is designed to be complementary to one end of the respective target DNA. (B) Binding of the target DNA (dashed red and green) by base pair hydrogen bonds. The overcoming part of the target DNA (dashed green) is again complementary to a detection oligonucleotide that carries a biotin label (solid green). After binding of the detection oligonucleotide, the nick is ligated by a DNA ampligase. This approach was formerly used by Barišić et al., 2015[176] to increase the specificity towards the target DNA. Picture and text are taken from a recent deliverable by the thesis author.

A significant limitation of a purely hybridisation-based DNA microarray is that it is prone to errors. The limitation occurs by non-specific hybridisation of the dissolved DNA to the immobilised probes on the microarray surface. Koltai & Wiengarten-Baror discussed this issue in detail[175].

For this reason, the detection platform by Barišić et al.[176] was adopted. The solid-phase DNA microarray detection system allows multiplex detection of hundreds of target DNA simultaneously with a specificity high enough to recognise single nucleotide polymorphisms (SNP). This method is based on covalently immobilised DNA oligonucleotides in combination with a ligation step. The high specificity is achieved by the altered probe concept structure compared to conventional microarray probes (Figure 12). The LNC-A probe (linear nucleotide chain) is a thiol-modified detection probe (black), which connects to two further DNA oligonucleotide probes via hydrogen bonds (LNC-B, blue, and LNC-C, orange). The hybridisation regions have a high GC-content so that the melting temperature is higher than 85 °C, leading to a high thermostability of the LNC probe concept. The LNC-C probe carries a specific detection sequence of the gene of interest. The LNC probes are immobilised on glass slides surface using a microarray spotter. The detection oligonucleotide (green) is a (labelled) specific sequence as well; together with a section of the LNC-C

probe, they correspond entirely to the target DNA (dashed orange/dashed green), being a part of the gene of interest. Since the mere hybridisation is quite error-prone (cross-hybridisation, i.e. hybridisation of strands with a certain similarity, which is, however, not the perfectly matching target)[175], the LNC-3 method uses the proofreading function of the ampligase, which is a thermostable ligase, as a further advantage. The proofreading function of the used ligase is more sophisticated than the *Taq*-DNA polymerase's one. This ensures that only exactly matching target DNA triggers the ligation of immobilised probes and detection oligonucleotide. The ligation is followed by a washing step at 70 °C, reducing the generation of false-positive signals by denaturation of non-ligated (cross-hybridised) detection oligonucleotides to a remarkably high degree.

2.2 Thesis tasks

The upstream PCR is a decisive factor regarding the specificity and sensitivity of DNA microarrays. As previously mentioned, multiplex PCR, which is essential for a high throughput device, suffers from many obstacles. With regard to that, the upstream PCR must be designed accordingly to overcome problems such as biases and primer dimerization. One aim of the thesis was to reduce primer interactions during the upstream PCR, which was finally conducted by attaching the primers to crosslinker molecules. The crosslinking of primers is described in publication 1.

Furthermore, the costs of the signalling needed to be reduced by far in order to render high-throughput screening affordable for commercial use. Different signalling systems were tested for this purpose. By using a transferase enzyme, the efficiency of the signalling in terms of sensitivity and specificity could be preserved, while the costs were significantly reduced. The labelling of detection oligonucleotides using a transferase is given in publication 2.

The main aim of this doctoral thesis was the further development of the LNC-3 detection method to enables the complete identification and characterisation of pathogens. Even though the basic framework of this method was published by Barišić et al.[176], many attempts and further developments regarding primer design, spotting conditions, etc., were required in order to adapt this technique into a high-throughput method. These further developments are presented and discussed in publication 3.

3. Scientific publications and manuscripts

Altogether, two publications and one submitted manuscript are added in the thesis. Data that were not included or presented as poster are attached in the appendix.

3.1 Publication 1: Low-cost microarray platform to detect antibiotic resistance genes

Noa Wolff, Michaela Hendling, Silvia Schönthaler, Andreas F. Geiss, Ivan Barišić

Published: Sensing and Bio-Sensing Research 23 (2019) 10226.
https://doi.org/10.1016/j.sbsr.2019.100266

Contribution to the publication:

- Experimental design
- Performance of all experiments
- Data analysis
- Manuscript drafting

The Supplementary Information, consisting of primer sequence tables, is not given here. It can be accessed via the journal homepage.

Sensing and Bio-Sensing Research 23 (2019) 100266

Contents lists available at ScienceDirect

Sensing and Bio-Sensing Research

journal homepage: www.elsevier.com/locate/sbsr

Low-cost microarray platform to detect antibiotic resistance genes

Noa Wolff[a], Michaela Hendling[a], Silvia Schönthaler[a], Andreas F. Geiss[b], Ivan Barišić[a,*]

[a] Molecular Diagnostics, AIT Austrian Institute of Technology, Giefinggasse 4, Vienna 1210, Austria
[b] University of Natural Resources and Life Sciences, Vienna, Austria

ARTICLE INFO

Keywords:
Antibiotic resistance
High-throughput detection
DNA microarray
Multiplex detection
Transferase reaction

ABSTRACT

Versatile DNA-based approaches for the detection of pathogens and antibiotic resistance genes have been introduced to facilitate faster diagnostic test. This requires a notable amount of cost-intensive commercially labelled DNA oligonucleotides in high-throughput assays. Here, we introduce an alternative method that utilizes a terminal deoxynucleotidyl transferase reaction to significantly reduce expenses of oligonucleotides. Antibiotic resistance genes that are of high clinical relevance were used as relevant application. The alternative labelling was based on an elongation step in the presence of biotin-conjugated nucleotides. Fluorescence-labelled streptavidin was bound to biotin-conjugated nucleotides and detected via a standard high-throughput microarray scanner. The self-labelled detection oligonucleotides performed equally well regarding sensitivity and specificity compared to commercially labelled ones, but at only 10% of their cost.

1. Introduction

An increasing emergence of antibiotic resistances accompanied the usage of antibiotics within the last decades, leading to concerns in healthcare facilities [6,14]. This boost of resistant bacterial species is promoted by the frequent location of resistance genes on interchangeable genetical elements such as plasmids [7]. In addition, the careless application of antibiotics exposed bacteria to a significant evolutionary pressure [15]. Today, more than 1000 antibiotic resistance mechanisms were described [10,22]. Hence, the number of resistant genera is growing, and the resistance development actually correlates with the overall use, as shown e.g. for *Streptococcus pneumoniae* and *Streptococcus pyogenes* [1]. Besides the consequences for patient health, the clinical failures caused by antibiotic resistances (ABRs) result in soaring downstream expenses [9]. Therefore, a fast and reliable determination of acquired antibiotic resistances is essential to treat patients appropriately and to avoid the unnecessary use of antibiotics.

The gold standard in clinical microbial diagnostics is based on the cultivation of bacteria and their phenotypical characterisation. The outstanding advantages of this approach are low-cost, high sensitivity, established diagnostic tests and well-trained personnel. In combination with additional techniques such as MALDI-TOF, most bacteria can be characterised within two working days [32,35]. Limitations of cultivation-based tests are that some pathogens are poorly cultivable or that false-negative results are possible if patients have already been treated with antibiotics [28,30]. Resistance determinations via minimum inhibitory concentration (MIC) approaches have been critically discussed as well [17]. However, the most problematic application scenario of cultivation-based diagnostic techniques represents the analysis of samples from sepsis patients. The mortality rate of these patients increases dramatically over time if no effective treatment is provided. While cultivation-based techniques can take several days to characterise a pathogen and its antibiotic resistance profile, molecular tests are able to provide such information within hours.

For this reason, DNA-based identification methods are of nascent interest. Especially assays that target phylogenetic marker genes such as the rRNA genes were developed [24]. The pathogenic nature in terms of virulence factors and resistance mechanism can be also identified by DNA-based assays. They were used to identify food-borne pathogens [36], to assess the water quality [5], and to determine ABR genes [4,11,33]. Several commercially available diagnostic tests were already developed for the characterisation of pathogens in blood samples [8,21]. However, further technological improvements are required to overcome major limitations such as high cost, poor sensitivity and a small number of genetic targets that can be detected per test. Approaches to overcome this last point include solid-support based techniques such as microarrays that allow the analyses of thousands of genes simultaneously [34]. A huge number of platforms with different solid-supports (e.g. glass, beads, hydrogels, silicon, etc.) and detection principles (fluorescence [35], enzymatic [35], electro-chemical [26], plasmonic [23], etc.) were developed in the last decades. A major restriction of these high-throughput techniques are the high costs

* Corresponding author.
E-mail address: ivan.barisic@ait.ac.at (I. Barišić).

https://doi.org/10.1016/j.sbsr.2019.100266
Received 24 October 2018; Received in revised form 19 February 2019; Accepted 21 February 2019

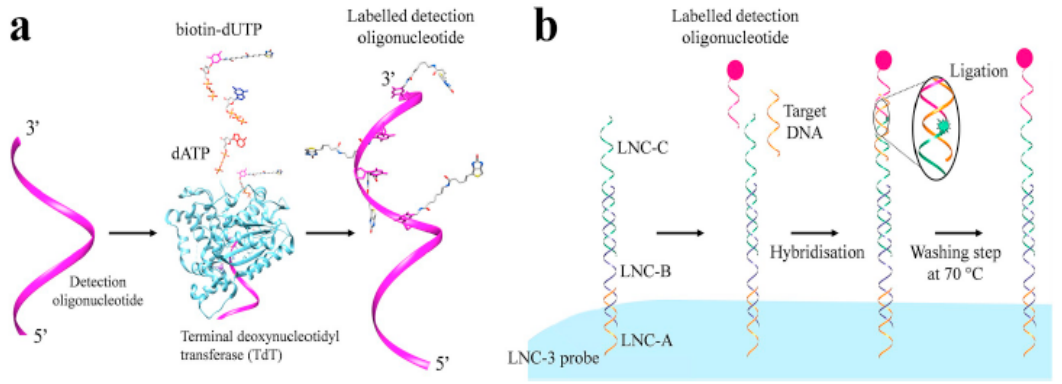


Fig. 1. (a). Elongation of oligonucleotides by the TDT. Detection oligonucleotides, comprising a 3′-terminal 3 T spacer, were elongated by the TDT, thereby incorporating biotin-labelled dUTP nucleotides. The colours are magenta for the DNA backbone, green for thymidine, red for adenine, blue for guanine, yellow for cytosine, light magenta for uracil, and light blue for the TDT. Other components are coloured by element (TDT: pdb file: 1JMS [16], accessed from the RCSB protein database). (b) Hybridization principle complemented by the reported alternative detection. A DNA oligonucleotide, referred to as LNC-A probe (orange), is covalently linked to a functionalized glass slide (transparent blue) utilising a 5′-terminal maleimide residue. Via base pair hydrogen bonds, LNC-B (blue) and LNC-C (green) are bound to LNC-A. The terminal ends of the LNC-C probe and the detection oligonucleotide (magenta) are complementary to the target DNA (yellow). After hybridisation, an enzymatic ligation reaction connects the immobilised LNC-3 strand and the biotin-conjugated detection oligonucleotides covalently. Upon a stringent washing step, only the detection oligonucleotides remain bound that were covalently linked to the immobilised LNC-3 probes. The visualisation occurs via fluorescence-labelled streptavidin that binds to biotin residues with high affinity.

associated with the de novo development of large microarrays. Especially, chemically-modified oligonucleotides that are attached to the support are the costliest components. Recently, a novel microarray probe concept was developed that uses only a single chemically-modified, universal oligonucleotide covalently attached to the solid support, and thus, significantly reduces the oligonucleotide costs [3]. Here, we present a further improvement of this concept for the detection of ABR genes. It is based on the terminal deoxynucleotidyl transferase (TDT) reaction that enables cost-efficient in-house labelling of the required detection oligonucleotides (Fig. 1). The TDT is a DNA polymerase that is naturally occurring within the maturation process of stem cells to B- and T-lymphocytes. It has the ability to elongate single-stranded DNA without a template [29] and has attracted interest in applied molecular biology [37]. As proof-of-concept, we designed a microarray targeting 47 clinically relevant ABR genes to demonstrate this alternative labelling technique.

2. Materials and methods

2.1. Bacterial cells and DNA purification

Clinical isolates from *Acinetobacter baumannii* (strain 5 M), *Enterobacter cloacae* (strain 299), *Citrobacter freundii* (5CF) and *Escherichia coli* (24E) were provided by the University Hospital Centre Zagreb. *Escherichia coli* (strain 10/134) was obtained from the University Hospital of Graz. *Klebsiella pneumoniae* (strain NCTC 13439) and *Escherichia coli* (NCTC 13462) were purchased from the National Collection of Type Cultures of Public Health England. *Kluyvera ascorbate* (DSM 4611) was purchased from the DSMZ- Leibniz Institute DSMZ-German Collection of Microorganims and Cell Cultures. To evaluate the microarray detection results, strains obtained from clinical partners were sequenced using the PGM Ion Torrent sequencer (Ion Personal Genome Machine™ (PGM™) System, Thermo Fischer Scientific, Waltham, MA, USA) according to the manufacturer's instructions. The genomic sequences of *K. pneumoniae* (strain NCTC 13439), *E. coli* (NCTC 13462) and *K. ascorbate* (DSM 4611) are publicly available. The strains were incubated overnight in lysogeny broth (LB) medium (10 g tryptone, 5 g yeast extract, 10 g NaCl) at 37 °C. A total of 5 mL of each cell culture was centrifuged at 6000*g* for 2 min and washed twice in phosphate-buffered saline (PBS) buffer (1.05 mM KH_2PO_4, 155.17 mM NaCl, 2.97 mM $Na_2HPO_4 \cdot 7H_2O$, pH = 7.4). Afterwards, the pellet was re-suspended in 1 mL PBS buffer. The cells were disrupted using the MagNA Lyser Instrument (Roche, Basel, Switzerland) for 30 s at 6500 rpm. The fragmentation was followed by 5 min of incubation at room temperature (RT) and a second fragmentation step. Finally, the suspensions were incubated at 95 °C for 10 min and centrifuged at 16,100*g* for 10 min. The supernatants were used for the subsequent experiments.

2.2. Oligonucleotide design

All oligonucleotides such as the synthetic target DNA, LNC probes and detection oligonucleotides were designed using the Oli2go software that calculates the ideal thermodynamic conditions for multiplex applications [19]. The calculated oligonucleotide properties are shown in Table 1 and Table S1. All oligonucleotides were purchased from Integrated DNA Technologies (Coralville, Iowa, USA).

Table 1
Microarray probe sequences and modification.

Probe name	5′-mod.	Sequence 5′-3′	Length
LNC-A	Thiol	TTTCGCTGCCGACCCTGCGCCGTGGCC	27 bp
LNC-B		CCCCGGCACGCCGAGCCCACGCTGCTTTTTTGGCCACGGCGCAGGGTCGGCAGCG	54 bp
LNC-C		GCAGCGTGGGCTCGCGTGCCGGGGTTTTTTNNNNNNNNNN	≈ 44 bp
Detection oligonucleotide		NNNNNNNNNNNNTTT	≈ 23 bp
Hybridisation control		CCCCGGCACGCCGAGCCCACGCTGCTTTTTTGGCCACGGCGCAGGGTCGGCAGCG/Cy3	54 bp
Spotting control	Thiol	TTTCGCTGCCGACCCTGCGCCGTGGCC/Cy5	27 bp

2.3. DNA amplification

The bacterial cell lysates were used as template DNA at a concentration of 10 ng per reaction in a single 47-plex PCR comprising all 47 primer pairs listed in Table S1. For this, the HotStartTaq DNA polymerase kit (Qiagen, Hilden, Germany) was used according to the manufacturer's instructions with a total primer concentration of 10 µM. We used 3 mM of $MgCl_2$ according to the manufactures instructions. Thermal cycling was conducted with the Applied Biosystems GeneAmp PCR System 2700 (Thermo Fischer Scientific, Waltham, MA, USA). After an initial denaturation step at 95 °C for 15 min, 40 cycles of each denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s and elongation at 72 °C for 30 s were performed. An additional downstream elongation step was conducted at 72 °C for 10 min. The PCR products were stored at 4 °C. Afterwards, the PCR products were purified using the Stratec Invisorb Fragment Clean Up kit (Stratec Molecular GmbH, Berlin, Germany) following their protocol. After purification, the amplification products were sonicated for 3 min (VWR Ultrasonic Cleaner USC-TH, PE, USA). The PCR products were used as target DNA in the subsequent ligation and detection reactions.

2.4. Glass slide functionalisation

The slide functionalisation was performed as previously described [38]. First, the unmodified glass slides were cleaned by sonication for 5 min in ddH_2O, 100% EtOH, and acetone each. Afterwards, the slides were sonicated for 10 min in 1 M NaOH and immersed in 1 M HCl overnight. On the following day, the slides were washed twice for 5 min in ddH_2O, then rinsed with 100% EtOH and air-dried.

Following, the slides were functionalized with 3-aminopropyl-trimethoxysilane (ATS, Sigma-Aldrich, MO, USA). For this purpose, the slides were immersed in a 5% ATS solution in dry acetone for one hour. Afterwards, the slides were washed three times for 5 min with acetone and rinsed with 100% EtOH. Subsequently, the slides were baked for 50 min at 90 °C. A sulfo-m-maleimidobenzoyl-N-hydroxysulfosuccinimide ester (s-MBS) was applied. The surface of the slides was wetted with 300 µL of a 2 mM s-MBS solution in PBS buffer and incubated overnight in a humid environment. Afterwards, the slides were rinsed with PBS, washed twice with ddH_2O, immersed in 100% EtOH, and finally air-dried overnight.

2.5. Spotting of the microarray probes

The OmniGrid Contact Microarrayer (GenMachines, San Carlos, CA, USA) and Stealth Micro Spotting Pins (ArrayIt, Microarray Technology, Sunnyvale, CA, USA) were used for spotting the oligonucleotides onto the microarray. For this purpose, the oligonucleotides LNC-A, LNC-B and one specific LNC-C were pooled together in sterile-filtrated 2× NaP_I spotting buffer (1× NaP_I: 0.1 M Na_2HPO_4, 0.15 M NaCl, pH = 6.5) to a final concentration of 5 µM each. During spotting, the humidity in the spotting chamber was kept at 60%. The spotted slides remained in the spotter for 5 h at RT. The humidity was kept constant until the end of incubation. Afterwards, the slides were washed by repeatedly submerging into 1× NaP_I buffer for 5 min. The slides were incubated in 1× NaP_I containing 10 mM β-mercaptoethanol for 1 h to deactivate the reactive groups on the surface. Afterwards, the slides were washed again with 1× NaP_I for 5 min to get rid of the β-mercaptoethanol. Unbound oligonucleotides were precipitated by incubation in saline buffer (1.5 M NaCl, 0.01 M Na_2HPO_4, pH = 7) for 10 min. After that, the slides were washed in 5× saline sodium citrate (SSC) buffer (1× SSC buffer: 150 mM NaCl, 15 mM sodium citrate, pH = 7.0, obtained from Biorad, Hercules, California, USA) comprising 0.1% Tween-20 (Sigma-Aldrich, MO, USA) for 5 min and then in 5× SSC buffer without Tween-20 for 1 min. Subsequently, the slides were rinsed with ddH_2O twice and centrifuged to dry. The dried slides were stored at −20 °C.

2.6. Elongation of the detection oligonucleotides by the TDT

The TDT was used to integrate biotin-conjugated deoxyuridine triphosphate (biotin-dUTP) nucleotides into the detection oligonucleotide strands. The biotin tag binds with high affinity to streptavidin, of which several affordable conjugates with standardly applied fluorescence labels exist. In this work, an Alexa-647 conjugate was identified as ideal combination.

The oligonucleotides were phosphorylated using the T4 polynucleotide kinase (PNK, Thermo Fischer Scientific, Waltham, MA, USA). In the reaction mixture, 47 different detection oligonucleotides were included at a final concentration of 1.78 µM each, 400 µM adenosine triphosphate (ATP, Thermo Fischer Scientific, Waltham, MA, USA), and 50 U/reaction of PNK. The phosphorylation reaction took place at 37 °C overnight in a thermoshaker (Peqlab TS-100, VWR, Erlangen, Germany), followed by an inhibition step of 10 min at 75 °C. Subsequently, the detection oligonucleotides were elongated at the 3′-end in the presence of biotin-11-dUTPs (Jena Bioscience GmbH, Jena, Germany) by the terminal deoxynucleotidyl transferase (TDT, Thermo Fischer Scientific, Waltham, MA, USA). The catalytic reaction mixture contained the previously phosphorylated detection oligonucleotides at a final concentration of 1 µM each, 320 µM dNTPs (Thermo Fischer Scientific, Waltham, MA, USA), 160 µM biotin-11-dUTP, and 75 U TDT/reaction. The reaction occurred at 37 °C for 12 h and was inhibited by heating up the samples to 70 °C for 10 min.

2.7. Ligation

The ligation was implemented in four array gasket hybridisation chambers with a capacity of 100 µL each (Agilent, Santa Clara, CA, USA). The ligation solution was pipetted, onto the hybridisation chamber, comprising 5 U/reaction ampligase (Epicentre, Madison, WI, USA), 5 µg bovine serum albumin (BSA, New England BioLabs), the enzymatically modified detection oligonucleotides at a final concentration of 1.68 µM and 200 nM of synthetic target DNA in ampligase buffer (20 mM Tris-HCl, 25 mM KCl, 10 mM $MgCl_2$, 0.5 mM nicotinamide adenine dinucleotide (NAD), 0.01% Triton X-100, pH = 8.3). In case of the PCR-amplified target DNA, 20 µL of the respective PCR product were added instead of synthetic DNA. PCR products were denatured at 95 °C for 5 min and chilled on ice before pipetted to the hybridisation chambers.

The ligation took place in a hybridisation oven (Microarray Hybridisation Chambers, Agilent, Santa Clara, CA, USA) at 55 °C for 1 h. Afterwards, the slides were washed for 5 min with 2× SSC including 0.1% SDS. Subsequently, the slides were washed once with 0.2× SSC for 2 min and finally twice with ddH_2O for 1 min. Another washing step was applied with ddH_2O for 10 min at 70 °C to reduce false-positive signals originating from unspecific hybridisation products and non-ligated detection probes. The biotin-modified samples were labelled with a Streptavidin-Alexa-647 conjugate (Jackson ImmunoResearch Laboratories Inc., West Grove, PA, USA), diluted 1:1000 in PBS buffer with 0.1% Tween. The slides were centrifuged to dry and scanned with a Tecan PowerScanner (Männedorf, Switzerland). To check the efficiency of the slide coating and the spotting, every array contained fluorescence-labelled hybridisation and spotting controls, which were spotted at the same concentration as the probes (Table 1). The hybridisation controls are measured at 532 nm, spotting controls and samples at 635 nm. The data were analysed with GenePix Pro 6.0 (Molecular Devices, Sunnyvale, CA, USA). Using GenePix Pro 6.0, the median values of the spotted areas were calculated. The background, averaged over the residual area, was subtracted. The data were exported and the median values of four repetitions of each measurement was calculated via Microsoft Excel 2010 (Microsoft Corporation, Redmont, WA, USA). Plotting was done with Origin Lab 8.5 (OriginLab Corporation, Northampton, MA, USA).

3. Results

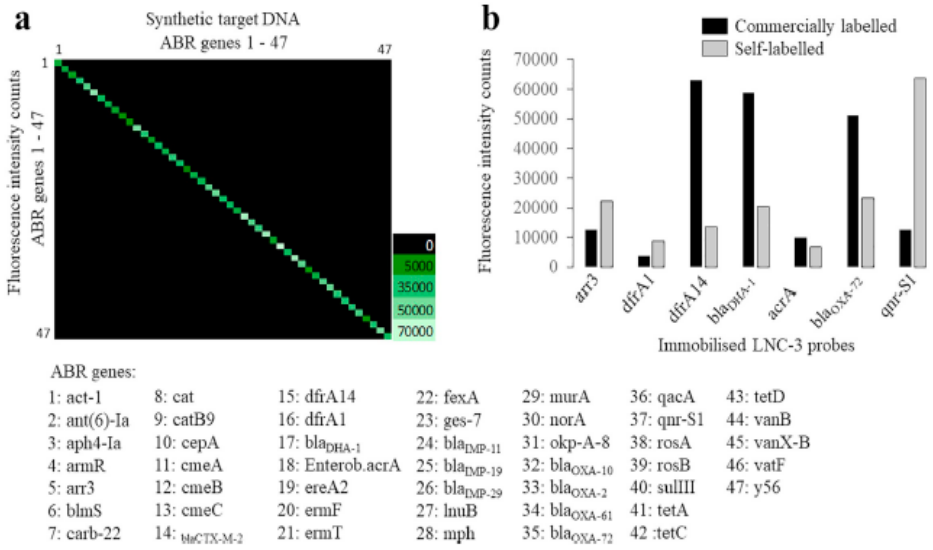
We compared the functionality of self-labelled and commercially available detection oligonucleotides regarding specificity and signal intensity. For this purpose, the presence of ABR genes was evidenced by applying both labelling methods in parallel. First, synthetic target DNA was applied to the microarray chip to test the functionality of the probes. Then, amplification products from 47-plex PCRs using bacterial cell lysates as templates were analysed with our microarrays. In order to evaluate the microarray results correctly, the genetic context of the bacterial isolates was analysed by whole genome sequencing (WGS) and the ResFinder tool [40].

3.1. Evaluation of microarray with synthetic templates

Synthetic DNA corresponding to representative ABR genes was subjected to the microarray containing 47 LNC probes designed to target different ABR genes. The heatmap summarizing 47 individual microarray detection experiments illustrates that highly specific signals were obtained with our self-labelled detection oligonucleotide in combination with synthetic DNA targets (Fig. 2a). Also, the approach with commercially labelled detection oligonucleotides did not produce any false positive signals (data not shown). The intensity of the background noise was negligible. The synthetic DNA targets led to fluorescence signals in a range of ~ 1000 to 34,000 fluorescence intensity counts (FIC) after subtracting the background noise at the expected positions on the microarray. A total of 59% ABR genes were detected with higher intensities if commercially available labels were used, while in the other cases higher signal intensities were generated by the self-labelling approach (Fig. 2b). In summary, the elongation of the detection oligonucleotides did not significantly impair their performance.

3.2. Characterisation of bacterial isolates

The performance of the new labelling technology was evaluated with cell lysates of *A. baumannii, E. cloacae, K. pneumoniae, K. ascorbate, C. freundii and three different E. coli* strains. Six different species with varying ABR profiles were used. The PCR-amplified DNA of the bacterial cell lysates was applied to the chip surface and both detection oligonucleotide concepts were compared. In general, the self-labelling technology did not result in significantly more non-specific microarray signals than the experiments comprising commercially labelled detection oligonucleotides (Fig. 3). The detection of bla_{OXA-72} in the *A. baumannii* isolate resulted in a correct response in both cases with signal intensities of 600 FIC and 672 FIC for the commercially labelled and the self-labelled detection oligonucleotides, respectively (Fig. 3a). In the *E. coli* strain 10/134, the commercially labelled surpassed the self-labelled oligonucleotides clearly in the case of the bla_{DHA-1} gene (6721 FIC to 3184 FIC), while the self-labelled oligonucleotides led to higher signals for the *arr3* gene (286 FIC to 885 FIC). In the *E. cloacae* and *K. pneumoniae* strains, the specificity of the signals was also identical but deviating signal intensities for several gene targets were observed. The probes for *acrA* (1687 FIC from commercially-labelled versus 1639 FIC of the self-labelled oligonucleotides) and *tetA* (1151 FIC to 1372) resulted in similar fluorescence intensity values. The probes targeting the *dfrA1* gene resulted in higher signals using the self-labelling protocol (4713 FIC to 8089 FIC). The LNC3 probes corresponding to the genes *dfrA14* (1978 FIC to 484 FIC) and *qnr-S1* (6163 to 3967 FIC) generated significantly higher signals with commercially labelled detection oligonucleotides. In the further experiments analysing additional bacterial strains (*K. ascorbate, C. freundii and E. coli* 24E and *E. coli* NCTC 1362) the expected microarray signals using self-labelled detection oligonucleotides could be confirmed (data not shown). In summary, all of the LNC3 probes tested with bacterial genomic DNA identified the same ABR genes independent of the labelling approach. The self-labelled detection oligonucleotides resulted in one false-positive and no false



ABR genes:

1: act-1	8: cat	15: dfrA14	22: fexA	29: murA	36: qacA	43: tetD
2: ant(6)-Ia	9: catB9	16: dfrA1	23: ges-7	30: norA	37: qnr-S1	44: vanB
3: aph4-Ia	10: cepA	17: bla_{DHA-1}	24: bla_{IMP-11}	31: okp-A-8	38: rosA	45: vanX-B
4: armR	11: cmeA	18: Enterob.acrA	25: bla_{IMP-19}	32: bla_{OXA-10}	39: rosB	46: vatF
5: arr3	12: cmeB	19: ereA2	26: bla_{IMP-29}	33: bla_{OXA-2}	40: sulIII	47: y56
6: blmS	13: cmeC	20: ermF	27: lnuB	34: bla_{OXA-61}	41: tetA	
7: carb-22	14: $bla_{CTX-M-2}$	21: ermT	28: mph	35: bla_{OXA-72}	42 :tetC	

Fig. 2. (a) Heatmap of fluorescence signals from 47 microarray experiments using synthetic DNA as target and self-labelled detection oligonucleotides. The microarray probes on the vertical axis are named to their corresponding ABR target gene, while the synthetic DNAs are applied in horizontal direction. The shown signals are median values of four repetitions. (b) Comparison of the fluorescence signals obtained from self-labelled (grey) and commercially labelled detection oligonucleotides (black). Seven ABR genes, encoded by synthetic DNA, were applied separately to microarrays. Only the positive signals are illustrated in the graph. On the horizontal axis, the immobilised LNC-3 probes are indicated, named after their corresponding ABR target gene. The median values of four repetitions are illustrated.

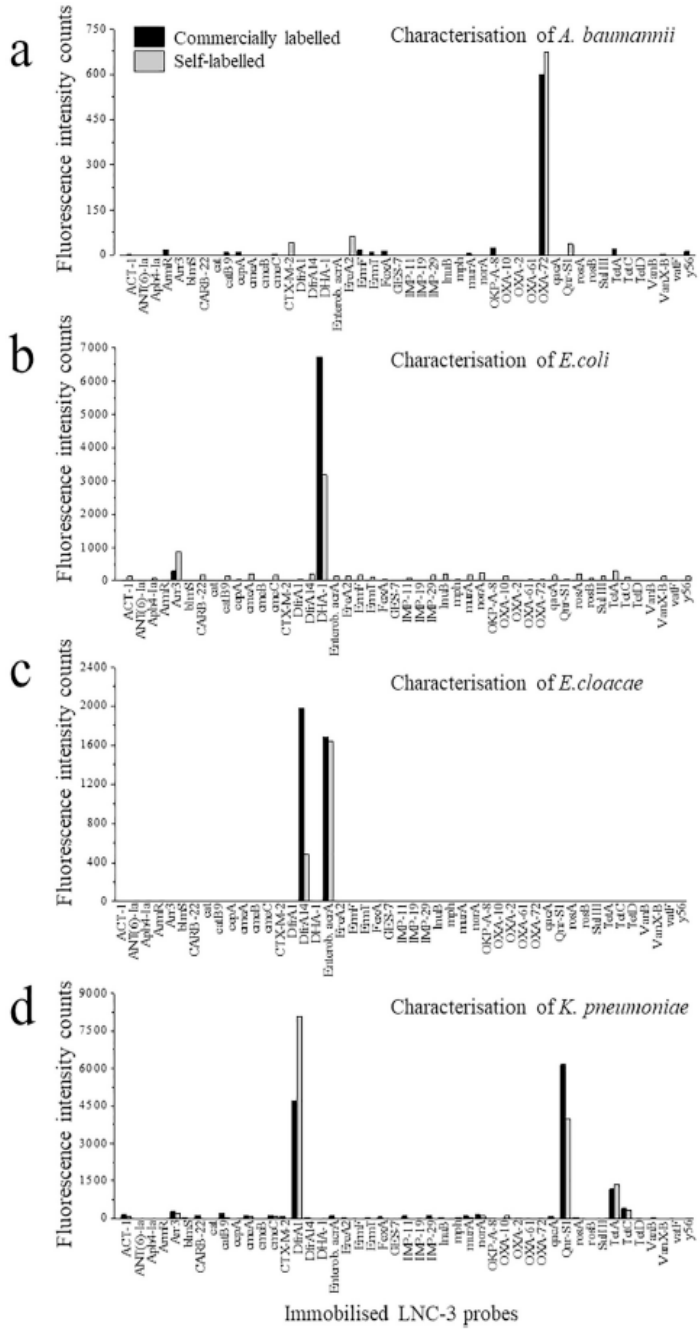


Fig. 3. Comparison of the fluorescence signals obtained by self-labelled oligonucleotides (grey) with commercially labelled ones (black) using a multiplex PCR comprising a primer mix targeting 47 ABR genes to amplify the DNA from *A. baumannii* (a), *E. coli* (b), *K. pneumoniae* (c), and *E. cloacae* (d) cell lysates. The shown fluorescence intensity counts are median values of four repetitions.

negative signals. The signals intensities were in 50% higher with commercially labelled detection oligonucleotides and in 50% with self-labelled ones.

4. Discussion

The DNA microarray format is still highly attractive in clinical laboratories due to its capability to simultaneously detect large numbers of different genes such as phylogenetic markers, resistance genes and virulence factors. However, the classical microarray technology is limited by non-specific cross-hybridisations that massively affect the specificity [18,39]. Thus, advances in the microarray technology are required to develop an assay that targets all ca. 400 human pathogens, 1000 antibiotic resistance genes, and 1000 clinically relevant virulence factors [13,22]. For analysing such a large number of possible target sequences, DNA microarrays are more beneficial compared to conventional PCR and real time (RT-)PCR, which are restricted in the number of genes to detect, since the products still need to be distinguishable in size via gel electrophoreses, or, in case of RT-PCR, by differently coloured fluorescence labels. The currently most popular technology in clinical microbiology to characterise pathogens is based on MALDI-TOF which is cultivation-based but very cost-effective. However, refinements need to be applied in terms of ABR detection [12,20,31]. The MALDI-TOF typing is mostly based on resistance-related proteins, which might not be expressed before any treatment occurred, or on degradation products, for which the same restrictions apply and which additionally do not cover any resistance mechanism that is not associated with degradation products [2]. To overcome the above-mentioned limitations, WGS would be suitable. However, while DNA-based detection methods appeal with their simple execution procedure in primary clinical use as soon as a standard array is conceived, the application of WGS requires higher costs, longer run times and bioinformatic expertise for the evaluation [27].

One of the major concerns regarding molecular diagnostics in clinical routine are the rather high costs. Thus, the LNC3 concept is an interesting option because it requires only a single chemically-modified oligonucleotide. Additionally, the technology has solved the problem of non-specific cross-hybridisations in microarrays and is able to distinguish SNPs without significant background signals [3]. In this work, we achieved a further reduction in costs by replacing commercial detection oligonucleotides by self-labelled ones. These detection oligonucleotides were prepared in-house by random elongation via a TDT reaction, during which inexpensive biotin-conjugated dUTPs were integrated and subsequently visualised with Alexa-647-modified streptavidin. The self-labelled detection oligonucleotides were compared to commercial detection oligonucleotides that were used before in our research group. While the labelling based on the commercial detection oligonucleotides amounted to ca. 1500 Euro for this 47-plex ABR gene set, the corresponding internal process totalled 151 Euro resulting in a tenfold reduction of the expenses. The alternative labelling can take place in advance to create reasonable stocks of detection oligonucleotides, leading to no prolongation of the sample analysis workflow by the alternative approach. Interestingly, no significant change in specificity was observed with the self-labelled and randomly elongated detection oligonucleotides (Fig. 3). Even more surprising was the observation that in 41% and 50% of the experiments the signal intensities increased using the elongated oligonucleotides with synthetic and genomic DNA, respectively. It was previously shown that not only the sequences of the microarray probes and targets have an impact on the DNA hybridisation but that also neighbouring regions affect the binding events [25]. Thus, further research is required to develop protocols that take advantage of such beneficial effects.

5. Conclusion

In conclusion, we present a very cost-effective microarray platform designed for the identification of antibiotic resistance (ABR) genes. The self-labelled and commercially labelled detection oligonucleotides turned out – besides both being adequate to use for the detection of ABR genes with a satisfactory signal-to-noise ratio – to be of comparable performance. The presented alternative can be implemented to reduce costs in microarray-based high-throughput assays, refining the microarray technique further towards their advent in clinical everyday routine. A fine tuning of the biotin-dUTP to dNTP ratio and the reaction parameters towards longer oligonucleotides could result in an increased sensitivity. Increasing the number of labels towards a maximum at which they do not restrict the specificity is expected to generate higher signal intensities. The introduced reaction is not confined to the purpose reported here, but might serve also in other DNA-based detection systems with high numbers of different targets such as the pathogen identification via phylogenetic markers (e.g. 16S rRNA genes).

Acknowledgments

We thank the technical staff at the AIT for providing technical assistance and Elisa de Llano for assisting in the evaluation process and Rick Conzemius for the evaluation of the sequencing results. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 634137.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sbsr.2019.100266.

References

[1] W.C. Albrich, D.L. Monnet, S. Harbarth, Emerg. Infect. Dis. 10 (3) (2004) 514–517.
[2] M.F. Anjum, E. Zankari, H. Hasman, Microbiology Spectrum 5 (6), ARBA-0011-2017, (2017).
[3] I. Barišić, V. Kamleithner, S. Schönthaler, H. Wiesinger-Mayr, Appl. Microbiol. Biotechnol. 99 (1) (2015) 413–423.
[4] I. Barišić, S. Schoenthaler, R. Ke, M. Nilsson, C. Noehammer, H. Wiesinger-Mayr, Diagn. Microbiol. Infect. Dis. 77 (2) (2013) 118–125.
[5] J. Baudart, D. Guillebault, E. Mielke, T. Meyer, N. Tandon, S. Fischer, W. Weigel, L.K. Medlin, Appl. Microbiol. Biotechnol. 101 (2) (2017) 871–886.
[6] B.G. Bell, F. Schellevis, E. Stobberingh, H. Goossens, M. Pringle, BMC Infect. Dis. 14 (2014) 13.
[7] P.M. Bennett, Br. J. Pharmacol. 153 (2008) S347–S357.
[8] A.J. Blaschke, C. Heyrend, C.L. Byington, M.A. Fisher, E. Barker, N.F. Garrone, S.A. Thatcher, A.T. Pavia, T. Barney, G.D. Alger, J.A. Daly, K.M. Ririe, I. Ota, M.A. Poritz, Diagn. Microbiol. Infect. Dis. 74 (4) (2012) 349–355.
[9] K.H. Bleicher, H.-J. Böhm, K. Müller, A.I. Alanine, Nat. Rev. Drug Discov. 2 (5) (2003) 369–378.
[10] K. Bush, Curr. Opin. Microbiol. 13 (5) (2010) 558–564.
[11] D.R. Call, M.K. Bakko, M.J. Krug, M.C. Roberts, Antimicrob. Agents Chemother. 47 (10) (2003) 3290–3295.
[12] E. Carbonnelle, C. Mesquita, E. Bille, N. Day, B. Dauphin, J.-L. Beretti, A. Ferroni, L. Gutmann, X. Nassif, Clin. Biochem. 44 (1) (2011) 104–109.
[13] L. Chen, D. Zheng, B. Liu, J. Yang, Q. Jin, Nucleic Acids Res. 44 (D1) (2016) D694–D697.
[14] P. Collignon, Scientific and Technical Review of the Office International des Epizooties, 31 (1) (2012), pp. 211–220.
[15] J. Davies, D. Davies, Microbiol. Mol. Biol. Rev. 74 (3) (2010) 417–433.
[16] M. Delarue, J.B. Boulé, J. Lescar, N. Expert-Bezançon, N. Jourdan, N. Sukumar, F. Rougeon, C. Papanicolaou, EMBO J. 21 (3) (2002) 427–439.
[17] S.M. Diene, C. Abat, J.-M. Rolain, D. Raoult, Lancet Infect. Dis. 17 (7) (2017) 690.
[18] W.W. Hadiwikarta, J.-C. Walter, J. Hooyberghs, E. Carlon, Nucleic Acids Res. 40 (18) (2012) e138.
[19] M. Hendling, S. Pabinger, K. Peters, N. Wolff, R. Conzemius, I. Barišic, Nucleic Acids Res. 46 (W1) (2018) W252–W256.
[20] J. Hrabák, E. Chudáčková, R. Walková, Clin. Microbiol. Rev. 26 (1) (2013) 103–114.
[21] R.S.P. Huang, C.L. Johnson, L. Pritchard, R. Hepler, T.T. Ton, J.J. Dunn, Diagn. Microbiol. Infect. Dis. 86 (4) (2016) 336–339.
[22] B. Jia, A.R. Raphenya, B. Alcock, N. Waglechner, P. Guo, K.K. Tsang, B.A. Lago, B.M. Dave, S. Pereira, A.N. Sharma, S. Doshi, M. Courtot, R. Lo, L.E. Williams, J.G. Frye, T. Elsayegh, D. Sardar, E.L. Westman, A.C. Pawlowski, T.A. Johnson, F.S.L. Brinkman, G.D. Wright, A.G. McArthur, Nucleic Acids Res. 45 (D1) (2017) D566–D573.
[23] J. Ko, S.-G. Park, S. Lee, X. Wang, C. Mun, S. Kim, D.-H. Kim, J. Choo, ACS Appl.

Mater. Interfaces 10 (8) (2018) 6831–6840.
[24] C.P. Kolbert, D.H. Persing, Curr. Opin. Microbiol. 2 (3) (1999) 299–305.
[25] H. Kolbl, C. Weingarten-Baror, Nucleic Acids Res. 36 (7) (2008) 2395–2405.
[26] S. Krizkova, M. Kepinska, G. Emri, M.A.M. Rodrigo, K. Tmejova, D. Nerudova, R. Kizek, V. Adam, J. Pharm. Biomed. Anal. 117 (2016) 464–473.
[27] J.C. Kwong, N. McCallum, V. Sintchenko, B.P. Howden, Pathology 47 (3) (2015) 199–210.
[28] A. Lupo, K.M. Papp-Wallace, P. Sendi, R.A. Bonomo, A. Endimiani, Diagn. Microbiol. Infect. Dis. 77 (3) (2013) 179–194.
[29] E.A. Motea, A.J. Berdis, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1804 (5) (2010) 1151–1166.
[30] M. Palka-Santini, S. Pötzsch, B.E.E. Clewen, M. Krönke, O. Krut, J. Microbiol. Methods 68 (3) (2007) 468–477.
[31] R. Schaumann, N. Knoop, G.H. Genzel, K. Losensky, C. Rosenkranz, C.S. Stingu, W. Schellenberger, A.C. Rodloff, K. Eschrich, Med. Sci. Monit. 18 (9) (2012) MT71–MT77.
[32] N. Singhal, M. Kumar, P.K. Kanaujia, J.S. Virdi, Front. Microbiol. 6 (2015) 791.
[33] C. Strauss, A. Endimiani, V. Perreten, J. Microbiol. Methods 108 (2015) 25–30.
[34] V. Trevino, F. Falciani, H.A. Barrera-Saldaña, Mol. Med. 13 (9–10) (2007) 527–541.
[35] L. Vámdi, J.L. Luo, D.E. Hibbs, J.D. Perry, R.J. Anderson, S. Orenga, P.W. Groundwater, Chem. Soc. Rev. 46 (16) (2017) 4818–4832.
[36] X.-W. Wang, L. Zhang, L.-Q. Jin, M. Jin, Z.-Q. Shen, S. An, F.-H. Chao, J.-W. Li, Appl. Microbiol. Biotechnol. 76 (1) (2007) 225–233.
[37] M.-L. Winz, E.C. Linder, T. André, J. Becker, A. Jäschke, Nucleic Acids Res. 43 (17) (2015) e110.
[38] N. Wolff, I. Barišić, Highly specific ligation-dependent microarray detection of single nucleotide polymorphisms, in: K.A. Bishop-Lilly (Ed.), Diagnostic Bacteriology. Methods and Protocols. Humana Press, Springer, New York, NY, 2017, pp. 231–240.
[39] P. Yakovchuk, E. Protozanova, M.D. Frank-Kamenetskii, Nucleic Acids Res. 34 (2) (2006) 564–574.
[40] E. Zankari, H. Hasman, S. Cosentino, M. Vestergaard, S. Rasmussen, O. Lund, F.M. Aarestrup, M.V. Larsen, J. Antimicrob. Chemother. 67 (11) (2012) 2640–2644.

3.2 Publication 2: Crosslinking of PCR primers reduces unspecific amplification products in multiplex PCR

Noa Wolff, Andreas F. Geiss, Ivan Barišić

Published: Journal of Microbiological Methods 178 (2020) 106051.
https://doi.org/10.1016/j.mimet.2020.106051

Contribution to the publication:

- Experimental design
- Performance of all experiments
- Data analysis
- Manuscript drafting

The Supplementary Information, consisting of primer sequence tables, is not given here. It can be accessed via the journal homepage.

Journal of Microbiological Methods 178 (2020) 106051

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Crosslinking of PCR primers reduces unspecific amplification products in multiplex PCR

Noa Wolff[a,*], Andreas F. Geiss[b], Ivan Barišić[a]

[a] Molecular Diagnostics, AIT Austrian Institute of Technology, Giefinggasse 4, Vienna, 1210, Austria
[b] University of Natural Resources and Life Sciences, Muthgasse 11, Vienna, Austria

ARTICLE INFO

Keywords:
DNA-based diagnostics, antibiotic resistance genes
High-throughput detection
Multiplex PCR
Primer secondary structures

ABSTRACT

The polymerase chain reaction is not only essential for many DNA-based diagnostic methods but is also exploited in other molecular methods that require an upstream amplification step. Multiplex PCRs are especially attractive as they reduce the number of individual reactions. However, the multiplexing efficiency is impaired by primer interactions such as the formation of primer dimers. In this study, covalent crosslinking of primers via their 5′-ends was used to avoid those undesired effects. The specificity of the primers as well as the efficiency of the PCR could be increased upon primer crosslinking in reactions containing up to 34 primer pairs targeting the most important antibiotic resistance genes in a single multiplex reaction.

1. Introduction

The polymerase chain reaction (PCR) is an indispensable technique in molecular diagnostics, either in the direct gene detection via specific amplification products or as upstream amplification step prior to sophisticated detection reactions, e.g. via microarray chips (Barišić et al., 2015, Anjum et al., 2017). Despite its importance, technological advances concerning the parallel amplification of many gene targets in single reaction are scarce. Important applications such as the identification of antibiotic resistance mechanisms or bacterial virulence factors encoded by hundreds of different genes require a high degree of multiplexing. To efficiently and quickly identify the most favourable treatment, multiplex PCRs need to be conducted comprising the primer pairs that target all relevant genes (Bhanothu and Venkatesan, 2019). High degrees of multiplexing, however, increase the possibilities of primer interactions, e.g. dimer formation. Various issues emerge from that: if primer dimers form, the involved primers are not available for the intended annealing and elongation reaction (Alanio and Bretagne, 2017). Moreover, those rather unspecifically hybridised double-stranded dimers can be elongated, which reduces the availability of DNA polymerases as well as the concentration of available dNTPs. If such primer dimers are enzymatically elongated, these products are perfectly matching the involved primers, making the primer/artefact hybridisation more likely to happen in the next amplification cycle. Consequently, the available primer concentrations are further reduced due to this unintended side reaction with a negative effect on the sensitivity. Even more problematic is that elongated primer dimers feature new, unpredictable sequences that subsequently lead to unexpected amplification products. This unintended process is a main limiting factor for multiplex PCRs. Here, we introduce crosslinked primers (Fig. 1) designed to limit the elongation of primer dimers. This new concept takes advantage of steric hindrance and the idea that the crosslinked primer dimers (Fig. 1c) cannot be processed in the active site of a conventional DNA polymerase. Thus, the specificity is thought to be improved. In this work, we investigated the specificity and sensitivity of the crosslinked primers in multiplex PCRs detecting 11 and 34 antibiotic resistance genes, respectively.

2. Materials and methods

2.1. Bacterial cells and DNA purification

The reported strains 32M *Klebsiella pneumoniae* and 24E of *Escherichia coli*, which hold clinically important beta-lactamase genes that were targeted in the following PCR reactions, were obtained as clinical isolates from the university hospitals of Zagreb and Graz. The strains were incubated overnight in a lysogeny broth (LB) medium at 37 °C. A total of 5 mL of each cell culture was centrifuged at 6000g for 2 min and washed twice in phosphate-buffered saline (PBS) buffer (1.05 mM KH_2PO_4, 155.17 mM NaCl, 2.97 mM Na $H_2PO_4 \cdot 7H_2O$, pH = 7.4). Afterwards, the pellet was resuspended in 1 mL PBS. The cells were disrupted using the MagNA Lyser instrument (Roche, Basel,

* Corresponding authors.
E-mail address: Noawolff@outlook.de (N. Wolff).

https://doi.org/10.1016/j.mimet.2020.106051
Received 30 July 2020; Received in revised form 3 September 2020; Accepted 3 September 2020
Available online 08 September 2020
0167-7012/

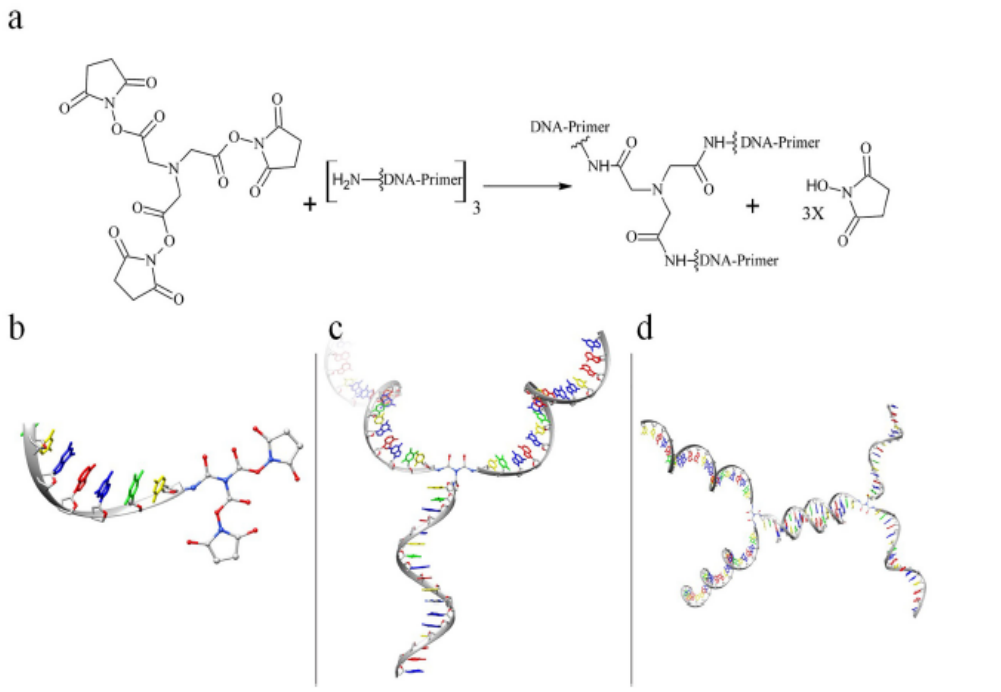


Fig. 1. Schematic depiction of primer crosslinking. (a) equation of the described reaction (b) An amino-modified primer reacts with a trifunctional succinimidyl crosslinker, i.e. a tris-succinimidyl aminotriacetate, to an ester bond. (c) Three primers are connected by a crosslinker at their 5′-ends. (d) Illustration of a crosslinked primer dimer.

Switzerland) for 30 s at 6500 rpm. The fragmentation was followed by 5 min of incubation at room temperature (RT) and a second fragmentation step. Finally, the suspensions were heated for 10 min at 95 °C and centrifuged at 16,000g for 10 min. The supernatants were used for the following experiments.

2.2. Primer crosslinking

The used primers were purchased from IDT (Integrated DNA Technologies, Inc., Coralville, IA, USA) with an amino modification (modification code: 5AmMC6; aminohexyl residue that is attached to the terminal backbone phosphate) on their 5′- end and dissolved at a concentration of 100 µM. These primers were made using the rudimentary version of Oli2go. Oli2go is a completely automated tool capable of designing primers and hybridisation oligonucleotides for multiplex applications with included specificity and dimer formation analysis (Hendling et al., 2018). The primer set was designed to check for antibiotic resistance genes that clinically relevant strains might have acquired or not. Hence, the expected number of bands depends on the investigated strain and was revealed by whole genome sequencing beforehand.

The sequences are given in the Tables S1 and S2. The trifunctional crosslinker tris-succinimidyl aminotriacetate (TSAT) was purchased from Thermo Fisher Scientific (Waltham, MA, USA). The crosslinking reagent was dissolved in 1 ml of 100% dimethyl sulfoxide (DMSO, Sigma-Aldrich, MO, USA) at a concentration of 16 mM. This stock solution was diluted to a concentration of 0.16 mM with the reaction buffer (0.1 M sodium phosphate, 0.15 M NaCl, pH = 7.2). The stock solution was prepared freshly whenever used. The primers were added into 96 well plates to a final concentration of 50 µM/well. All primers were crosslinked individually (for both forward and reverse primers) by adding TSAT crosslinker to a final concentration of 16 µM/well. The reaction was allowed to proceed at room temperature for two days. Finally, all primers were pooled and combined together for the multiplex PCR for a total primer concentration of 50 µM. As a control reaction, TSAT was deactivated by adding Tris-HCl (at a final concentration of 500 µM) before the crosslinking reaction with the primers. The subsequent reaction was carried out as for the samples. All samples were run in parallel.

2.3. PCR reaction and agarose gel electrophoresis

The PCR Kit 16S Basic of Molzym (Bremen, Germany) was used according to the manufacturer's instructions. Since the primer set is used to screen for clinically important antibiotic resistance genes that are not necessarily species-specific, the use of this DNA-free *Taq* polymerase was necessary to avoid the detection of the antibiotic resistance genes of the polymerase-producing bacteria. The total primer concentration was 2.5 µM (1.25 µM forward primers, 1.25 µM reverse primers), corresponding to 208 and 74 nM per primer pair if 11 and 34 primer pairs were used, respectively. The thermocycler settings comprised an initial denaturation step at 95 °C for 5 min, 40 cycles with denaturation at 95 °C for 30 s, annealing at 9 different temperatures from 53.2 to 63.8 °C for 30 s and elongation at 72 °C for 30 s, a final elongation step at 72 °C for 10 min, and cooling to 4 °C until the samples were collected.

The PCR products were applied to a 2% agarose gel containing a thousandth volume of SYBR Safe DNA Gel Stain from Thermo Fisher Scientific (Coralville, Iowa, USA). Then, separation occurred at 180 V for 2 h. The gels were photographed while UV-excited.

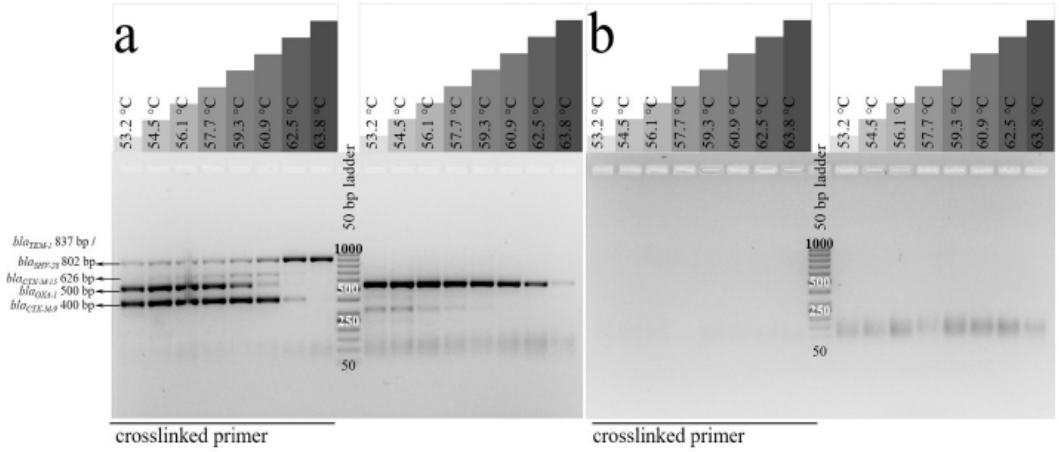


Fig. 2. Images of 2% agarose gels displaying 11-plex gradient PCRs. (a) Amplification products using template DNA from the *K. pneumoniae* isolate 32 M with crosslinked amino-modified (left) and non-crosslinked primers (right), respectively. (b) Negative controls of the reactions described in (a).

3. Results and discussion

The effect of primer crosslinking was investigated by comparing PCR amplification products obtained via crosslinked and non-crosslinked primers on a 2% agarose gel. Genomic DNA from multidrug-resistant *Klebsiella pneumoniae* and *Escherichia coli* strains was used for the proof of principle. Since different primers exhibit different annealing temperatures, temperature gradient PCRs were conducted.

In Fig. 2a, an agarose gel of a gradient PCRs with crosslinked (left) and non-crosslinked (right) primers using 10 ng DNA from a *K. pneumoniae* isolate as target is illustrated. Fig. 2b represent the negative control where no target DNA was added.

Revealed by whole genome sequencing beforehand, five amplification products of eleven possible ones were expected in the gradient PCR: bla_{TEM-1} (837 bp), bla_{SHV-28} (802 bp), $bla_{CTX-M-15}$ (626 bp), bla_{OXA-1} (500 bp), and $bla_{CTX-M-9}$ (400 bp). The products corresponding to bla_{TEM-1} and bla_{SHV-28} were, due to their similar product size, expected in the same band. However, the amplification products generated with the crosslinked primers are larger due to the overhangs and are thus expected to migrate slower in the agarose gel in comparison to the conventional amplification products. In the 11-plex temperature gradient PCR, a change of the band pattern for both crosslinked and not crosslinked primers was observed. Furthermore, in the reactions comprising crosslinked primers almost no dimer clouds (50–150 bp) were generated (Fig. 2). This difference was even more emphasized in the negative control without target DNA (Fig. 2b).

In the temperature range from 54.5 °C to 60.9 °C, the crosslinked primer PCR clearly generated the expected four bands corresponding to the amplification products of bla_{TEM-1} (837 bp) and bla_{SHV-28} (802 bp), too similar in size to be distinguished, $bla_{CTX-M-15}$ (626 bp), bla_{OXA-1} (500 bp), and $bla_{CTX-M-9}$ (400 bp). The intensity of the bands of the amplification products of bla_{OXA-1} and $bla_{CTX-M-9}$ was higher than that of bla_{TEM-1} / bla_{SHV-28} and $bla_{CTX-M-15}$. From 54.5 °C to 60.9 °C annealing temperature, the intensity of the band corresponding to $bla_{CTX-M-15}$ increased, while the band intensity of bla_{OXA-1} decreased. Since the band intensity correlates with the amount of PCR products, the efficiency of the individual PCR reactions could be deduced. It was interesting to observe that with increasing annealing temperatures the amplification efficiency of some genes decreased, while it, in contrast, increased for others (Fig. 2a). Based on these results, the annealing temperature of 59.3 °C was identified as optimum for this reaction because all expected antibiotic resistance genes were amplified. Hence, it was shown that an annealing temperature evaluation was indispensable regarding specificity and sensitivity.

To rule out effects of the crosslinking chemistry on the amplification (e.g. interactions of the crosslinkers with amino residues of the DNA polymerase), the trifunctional crosslinker tris-succinimidylaminotriacetate (TSAT) and DMSO were used in the PCRs with crosslinked and non-crosslinked primers. In the non-crosslinked samples, however, the functional groups of the crosslinker were previously blocked with Tris-HCl. Therefore, the reaction of the crosslinker with the amino group of the primer was prevented.

In contrast to the 11-plex PCR with crosslinked primers, the conventional multiplex PCR did not generate the expected amplification products but additional bands corresponding to non-specific amplification products although comprising the same primers. On the other hand, expected bands were missing. Altogether, the non-crosslinked primer samples did not result in a meaningful detection pattern at any tested temperature.

To evaluate the potential of this new concept, a 34-plex PCR targeting clinically relevant beta-lactamase resistance genes was designed to further increase the complexity of the multiplexing and the number of potential primer interactions. The primer sequences were taken from the work of Barišić et al., 2013 (Table S1).

A gradient PCR analogous to the previous one was performed using genomic DNA of *K. pneumoniae* as target (Fig. 3).

Four bands corresponding to four beta-lactamase genes that are present in the selected strain could be observed across the complete temperature range: bla_{SHV-28} (523 bp), bla_{OXA-1} (421 bp), $bla_{CTX-M-15}$ (315 bp), and bla_{TEM-1} (244 bp).

An annealing temperature-dependent change of the band pattern was observed for crosslinked and non-crosslinked primers. Two additional amplification products were generated (bla_{OXA} artefacts 150 bp and 750 bp). Previous in silico analysis with whole genome data using the Primeval analyses tool (Conzemius et al., 2019) revealed that these two bands were an expected by-product of the bla_{OXA-1}-primers in combination with the used target DNA.

The amplification product pattern obtained with the non-crosslinked primers (Fig. 3b) greatly differed from the one with crosslinked primers (Fig. 3a). As in the non-crosslinked PCR with 11 primer pairs, there were products that resulted from non-target annealing of the primers to the target DNA in addition to the expected specific DNA amplification products. In the 34-plex experiments with the non-crosslinked primers, more non-specific primer interactions led to more

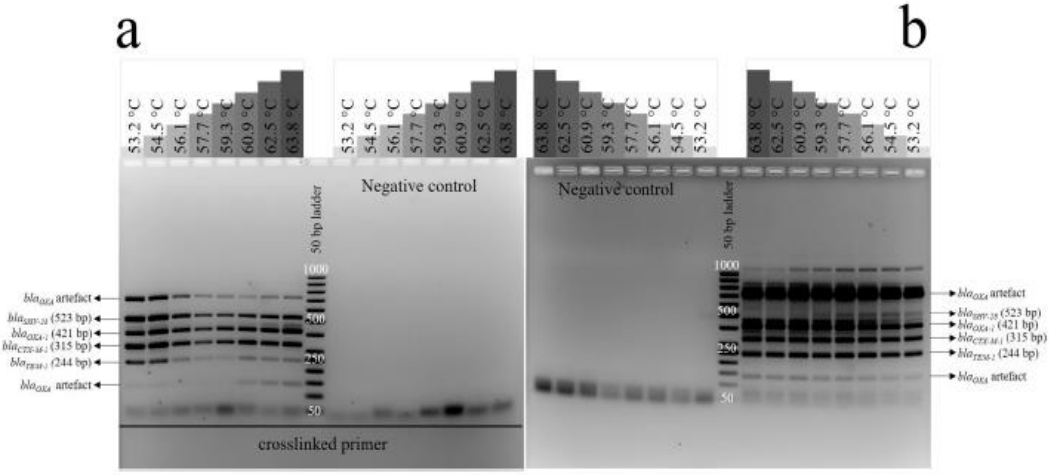


Fig. 3. Images of 2% agarose gels depicting 34-plex PCRs at varying annealing temperatures using genomic DNA from the *K. pneumoniae* isolate 32M. (a) The amplification products (left) and the corresponding negative controls (right) of the multiplex PCR comprising crosslinked primers are shown. (b) The amplification products and the corresponding negative controls obtained with the same but non-crosslinked amino-modified primers are illustrated.

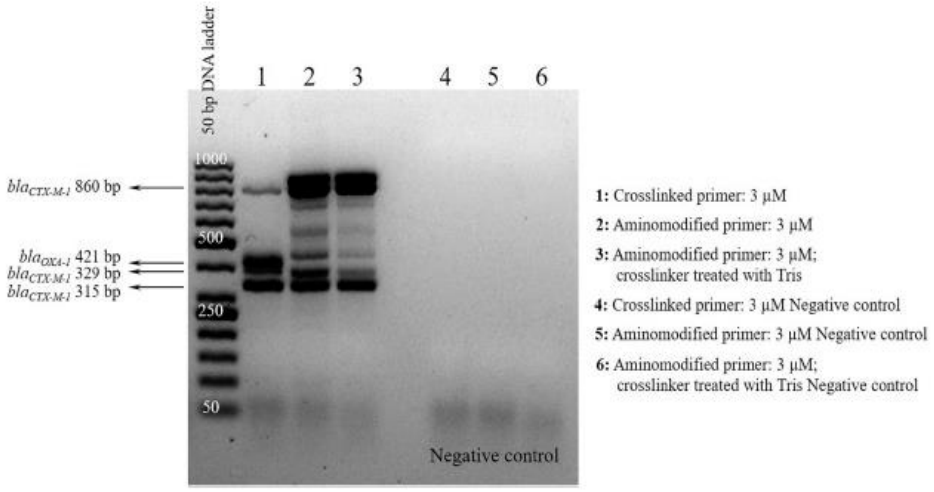


Fig. 4. Image of a 2% agarose gel illustrating the amplification products of a 34-plex PCR using genomic DNA from the *E. coli* isolate 24E as template. Comparison of samples with crosslinked primers (1), amino-modified primers (2), Tris-HCl-treated crosslinkers (3), and the respective negative controls (4–6) at total primer concentrations of 3 µM.

unintended amplification products, complicating an easy identification of correct bands. Also, in this case, the non-crosslinked bla_{OXA} primers generated the two artefacts.

To shed light on the impact of the crosslinker and the primer modifications on the amplification by the polymerase, we performed further analyses. A 34-plex PCR (Fig. 4) was conducted comprising three different primer states: crosslinked, the amino modified non-crosslinked primers without crosslinker, and non-crosslinked primers containing Tris-HCl reacted crosslinker. Tris-HCl is an ingredient in the polymerase buffer and an increased concentration could have an effect on the amplification performance.

The genomic DNA of the *E. coli* strain 24E served as target DNA with expected bands for bla_{OXA-1} (421 bp) and $bla_{CTX-M-15}$ (860, 329, and 315 bp). In the case of the multiplex reaction with crosslinked primers, the four expected bands were observed. In contrast, the multiplex PCRs with the non-crosslinked primers resulted in several non-target amplification products (i.e. the additional bands between 400 and 800 bp). With the deactivated crosslinker, the intensities of the non-specific product bands were reduced, but they could still be observed leading to a similar band pattern as observed with the non-crosslinked primers. The slight difference between the two non-crosslinked PCRs might be caused by the additional Tris molecules or entailed changes of pH. All samples showed slight, but rather negligible primer dimer clouds. As demonstrated in the previous multiplex PCRs, not only non-specific bands could be strongly reduced using crosslinked primers but also all expected bands could be generated.

4. Conclusions

The crosslinking of PCR primers resulted in more reliable band patterns allowing a more precise detection of antibiotic resistance genes. A steric hindrance effect was used to limit the accessibility of crosslinked primers to the active site of the DNA polymerase. This inhibited the elongation of primer dimers and further favoured the binding of perfectly matching primers. The potential of the new concept is not restricted to direct PCR-based detection methods, but important for all methods that rely on an upstream PCR to amplify the target genes, such as microarray-based detection. Next steps are to further increase the multiplexing and to evaluate different crosslinking systems and their impact on the amplification efficiency. Moreover, the influence of unreacted crosslinkers, Tris, and DMSO must be further analysed.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the technical staff at the AIT for providing technical assistance, Michaela Hendling for using Oli2go and Dr. Gian-Luca Bello for proofreading.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mimet.2020.106051.

References

Alanio, A., Bretagne, S., 2017. Performance evaluation of multiplex PCR including Aspergillus-not so simple. Med. Mycol. 55 (1), 56–62. https://doi.org/10.1093/mmy/myw080.

Anjum, M.F., Zankari, E., Hasman, H., 2017. Molecular methods for detection of antimicrobial resistance. Microbiol. Spectr. 5, 6. https://doi.org/10.1128/microbiolspec.ARBA-0011-2017.

Barišić, I., Schoenthaler, S., Ke, R., Nilsson, M., Noehammer, C., Wiesinger-Mayr, H., 2013. Multiplex detection of antibiotic resistance genes using padlock probes. Diagn. Microbiol. Infect. Dis. 77 (2), 118–125. https://doi.org/10.1016/j.diagmicrobio.2013.06.013.

Barišić, I., Kamleithner, V., Schönthaler, S., Wiesinger-Mayr, H., 2015. Fast and highly specific DNA-based multiplex detection on a solid support. Appl. Microbiol. Biotechnol. 99 (1), 413–423. https://doi.org/10.1007/s00253-014-6246-x.

Bhanothu, V., Venkatesan, V., 2019. Conventional polymerase chain reaction and amplification refractory mutation system-multi-gene/ multi-primer PCR in the diagnosis of female genital tuberculosis. Arch. Microbiol. 201 (3), 267–281. https://doi.org/10.1007/s00203-019-01631-1.

Conzemius, R., Hendling, M., Pabinger, S., PRIMEval, Barišić I., 2019. Optimization and screening of multiplex oligonucleotide assays. Sci. Rep. 9 (1), 19286. https://doi.org/10.1038/s41598-019-55883-4.

Hendling, M., Pabinger, S., Peters, K., Wolff, N., Conzemius, R., Barišic, I., 2018. Oli2go: an automated multiplex oligonucleotide design tool. Nucleic Acids Res. 46 (W1), W252–W256. https://doi.org/10.1093/nar/gky319.

3.3 Publication 3: Full pathogen characterisation: Species identification including the detection of virulence factors and antibiotic resistance genes via multiplex DNA-assays

Noa Wolff, Michaela Hendling, Fabian Schroeder, Silvia Schönthaler, Andreas F. Geiss, Branka Bedenic, Ivan Barišić

Submitted to: Journal of Infectious Diseases and Therapy.

Contribution to the publication:

- Experimental design
- Main experimental work
- Data analysis
- Manuscript drafting

The Supplementary Information is given without primer sequence tables. It can be accessed via personal communication.

Full pathogen characterisation: Species identification including the detection of virulence factors and antibiotic resistance genes via multiplex DNA-assays

Noa Wolff[1]*, Michaela Hendling[1], Fabian Schroeder[2], Silvia Schönthaler[1], Andreas F. Geiss[3], Branka Bedenic[4], Ivan Barišić[1]

[1] Molecular Diagnostics, AIT Austrian Institute of Technology, Giefinggasse 4, Vienna, 1210 Austria;
[2] Ekorefugium, Slunj, Croatia;
[3] University of Natural Resources and Life Sciences, Vienna, Austria
[4] Department of Microbiology, School of Medicine, University of Zagreb, Zagreb, Croatia

* To whom correspondence should be addressed. Tel: +43 664 88390643; Fax: +43 0 50550 4450; Email: Noawolff@outlook.de

Abstract

Antibiotic resistances progressively cause treatment failures, and their spreading dynamics reached an alarming level. Some strains have already been classified as highly critical, e.g. the ones summarised by the acronym ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter spp.*). To restrain this trend and enable effective medication, as much information as possible must be obtained in the least possible time. Here, we present a DNA microarray-based assay that screens for the most important sepsis-relevant 45 pathogenic species, 360 virulence factors, and 409 antibiotic resistance genes in parallel. The assay was evaluated with 14 multidrug resistant strains, including all ESKAPE pathogens, mainly obtained from clinical isolates. We used a cost-efficient ligation-based detection platform designed to emulate the highly specific multiplex detection of padlock probes.

Keywords: Antibiotic resistance; Pathogen identification; High-throughput detection; DNA microarray; Multiplex detection

Introduction

Challenges associated to antibiotic treatment, indispensable in medicine, have been forecasted repeatedly in terms of antimicrobial resistances during the last two decades[1–3]. Potentially resistant pathogens limit the number of suitable measures and moreover retard treatment decisions, which cannot be longer made empirically[4]. Known examples for escalating resistance spreading are the extended spectrum β-lactamases and carbapenemases[5], or the six pathogens summarised by the acronym ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter spp.*), highlighted by the Infectious Diseases Society of America for being particularly critical in terms of antibiotic resistances[6,7]. A precise antibiotic treatment is thus essential to contain the further spread of antibiotic resistance mechanisms. The identification of the causative pathogen and their acquired antibiotic resistance (ABR) genes is of utmost importance.[8]

The gold standard in diagnostic microbiology is based on cultivation-dependent methods that are cost-effective, well established in the routine practice and diagnostically conclusive regarding the antibiotic susceptibility. Major drawbacks are that the cultivation of some pathogens is challenging and can last several weeks[9]. In addition, a previous treatment with antibiotics or the presence of bacteriophages can cause false-negative results[10,11]. To overcome these limitations, a wide variety of molecular diagnostic tests has been developed and is used for the pathogen identification.

Nevertheless, further technological advances are required to meet the clinical requirements.[12,13] A lot of mechanisms enable pathogens to protect themselves from antibiotic substances, including acquired antibiotic resistances as well as intrinsic resistance mechanisms[14–16]. To investigate this multitude of mechanisms without relying on phenotypic observations, there is a large repertoire of molecular diagnostics technologies nowadays that already made their way to clinical everyday routine, as e.g. polymerase chain reaction (PCR)- and real time (RT)-PCR-based detection, matrix-assisted laser desorption/ionisation-time of flight mass spectrometry

(MALDI-TOF MS), whole genome sequencing (WGS), and the microarray technique addressed here. The currently most popular and clinically used tests are based on real-time polymerase chain reactions (PCRs) due to their relatively good sensitivity, specificity and speed.[17,18] PCR constitutes a sensitive method to identify bacteria via their highly conserved ribosomal RNA genes[19,20]. The application as well as the variability of PCR are multitudinous, offering RT-PCR, isothermal PCR, loop-mediated isothermal amplification, or the recombinase polymerase amplification (RPA). The most important in clinical diagnostics are the conventional and the real-time PCR. The RT-PCR, which allows to observe the amplification live using unspecifically intercalating fluorescence dyes or specific DNA sequences which give rise to a fluorescence signal only after hybridizing to the amplicon[21,22]. However, these tests can only identify a low number of targets because of the limited availability of differentiable fluorescence dyes. In addition, with an increasing degree of multiplexing, the sensitivity and specificity of PCRs is reduced due to unintended amplification products and primer dimer formation[23,24,25].

A further tool in clinical microbiology is MALDI-TOF MS. This is usually employed for pathogen identification[26]. The detection principle is based on the evaporation and ionisation behaviour, respectively, determining the time the resulting ions require to reach the detection[27]. This time is characteristic for the respective pathogen and can hence be used to identify the organism by means of its mass/charge ratio (m/z value) calculated from the peak focus. The spectral fingerprints vary enough to differentiate genera, as long as they exhibit the same growth conditions[28]. Despite all of this, the method is vulnerable to errors, especially regarding species differentiation, for example between *Streptococcus pneumoniae* and *Streptococcus mitis*, which might entail severe consequences[29–31].

Irrespective of that, the usage of MALDI-TOF MS contributes well to the patients' outcome rate, which is partially owed to the little time required for the identification[32], partially to the further characterisation possibilities, especially regarding the antibiotic resistances[33,34]. For instance, β-lactamase activity could be evidenced with the help of MALDI-TOF MS[35,36], as well as other antibiotics[32]. However, MALDI-TOF MS identification is expression-dependent; not yet expressed proteins, induced by the antibiotic if present, cannot be detected at all[13,28]. In case of VFs, not all of them constitute expressed proteins. To integrate this technology in the clinical everyday routine, further studies concerning the detection of antibiotic resistances and virulence factors owned by the pathogen are required.

Regarding the content of information, WGS is superior to all other described methods, theoretically revealing the entireness of present phylogenetic marker genes, antibiotic resistance (ABR) genes, and virulence (VF) genes. Since the invention of sequencing, there have been many and rapid developments in this field. The WGS costs, once the main limiting criterion, decreased dramatically within the last decade. After the sequence is obtained, no further detection setup is required; all the information is already contained by the sequence itself. Nevertheless, it still does not only require a lot of processing – and therefore time – to obtain the contiguous sequence and thereof the desired information out of the raw sequence snips, but also a strong bioinformatic expertise that exceeds the demands addressed to the commonly employed executive personnel by far. Trained experts, in turn, increase the costs again, after having been saved in terms of the sequencing reaction. In critical cases, the most limiting criterion might indeed be the time needed to analyse a genome.

An alternative method to identify and characterise pathogens that enable the detection of thousands of targets in parallel combines PCR with solid support-based detection systems, such as DNA microarrays and microbeads. While next generation sequencing platforms are currently replacing these hybridisation-based detection platforms in research facilities, they still remain attractive in clinical environments because of their relatively low price, short experimental run-times and manageable bioinformatic procedures[37,38]. In a sense, the microarray fills the gap between WGS and PCR[39]. Once conceived for a comprehensive gene set, can be made available ready-to-use with a standard procedure that easily fits in the clinical everyday routine. Time, effort, and required skills are reduced compared to WGS, while the number of genes that can be detected simultaneously is scarcely limited in opposite to PCR-based detection. A number of diagnostic assays using such platforms was already designed[11,40–43].

The classical DNA microarray method is based on a hybridisation reaction between labelled target DNA and immobilised microarray probes[40,44]. A significant limitation of the microarray technology is the non-specific cross-hybridisation of amplification products to non-target microarray probes that massively impair the specificity[45,42,46]. One reason for this is that the variable regions of different species might exhibit only one single nucleotide polymorphism (SNP). Interestingly, the thermodynamic stability of the DNA double helix is less dependent on the base-pairing but rather on the base-stacking effects, making it difficult to avoid cross-hybridisations during a sequence-driven probe design process[47]. Also, the covalent linker of the microarray probe to the surface has an effect on the hybridisation interactions[48]. The yet poorly understood anomalous DNA behaviour can result even in thermodynamically more favourable intermediates comprising non-perfectly matching sequences in comparison to perfect-matching duplexes[49].

As a consequence of this phenomenon, microarray detection protocols involving enzymes with proof-reading capability were developed to minimise cross-hybridisation effects. In addition, probe immobilisation concepts have been reported that focus on the emulation of the liquid phase DNA hybridisation behaviour to minimise surface-related effects.[50] Combining those advantages, we introduced the linear nucleotide chain (LNC) concept, which uses immobilised trimeric oligonucleotides in combination with a ligase reaction to provide a specificity sufficient to be impaired by single nuclear polymorphisms (SNPs)[51]. The concept is depicted schematically in Fig. 1. It was recently complemented by a protocol for cost-saving self-labelling of the required detection oligonucleotides[52]. Further, for hundreds of samples, the upstream amplification is as well associated with PCR-related issues, such as primer interactions. A combination with a tool for multiplex application primer design that checks on dimer formation and specificity issues automatically was required, formerly developed by our group[53]. In this study, we do not aim to separately identify either the pathogenic species or its ABR genes, but to fully characterise pathogens using microarray chips that screen for 45 pathogens, 360 virulence factor (VF) genes (virulence factors are directly involved in pathogenesis, being e.g. bacterial toxins, enzymes, or cell surface components, such as membrane proteins and polysaccharides[54]), and 409 ABR genes at the same time to provide all information needed for a reasonable and effective treatment. This analysis is performed with 14 clinically relevant pathogenic strains including the initially mentioned ESKAPE pathogens.

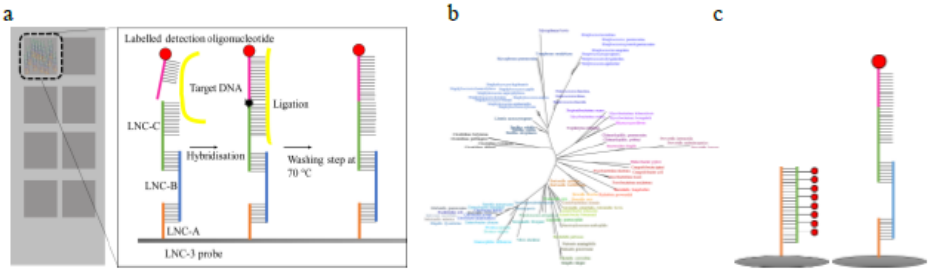


Figure 1: (a) Chip surface: Schematic representation of the 8 array chambers with the **linear nucleotide chain (LNC-3) concept:** To mimic liquid phase DNA hybridisation, three oligonucleotides, specified as LNC-A (orange), LNC-B (blue), and LNC-C (green), are immobilised on a functionalised glass surface (grey), the first one covalently, the others by hybridisation. The LNC-C terminal region is complementary to one part of the target DNA (yellow). The other part is complementary to the detection oligonucleotides (magenta), which are biotinylated and can be stained by fluorescence-labelled streptavidin (red). Once connected via the target DNA, LNC-C and detection oligonucleotide are ligated, thereby covalently linking the detection oligonucleotide to the immobilised probe. By washing at 70 °C, non-ligated detection oligonucleotides are removed. (b) **Phylogenetic tree:** The phylogenetic distances of the investigated species is illustrated, calculated according to the All-Species Living Tree project using 16S rRNA genes[55,56]. (c) **Comparison:** Identification via conventional, hybridisation-based detection (left) and LNC-3 technology (right).

Materials and Methods

Briefly summarised, all strains were cultivated and lysed, followed by DNA purification and PCR-based 45-plex amplification. The amplification products were applied to microarray chips that had been functionalised with DNA complementary to the genes of interest and complementary to their related mRNA (therefore constituting the base sequence of the non-coding strand). Fluorescence-labelled detection oligonucleotides, complementary to the adjacent region of the target DNA, were connected to the probes by the target DNA – if present – via hybridisation. If perfectly matching, the detection oligonucleotides were ligated to the immobilised DNA oligonucleotides. Non-ligated oligonucleotides, merely bound by hybridisation, were washed out. The chip readout occurred by means of a standard fluorescence-based microarray scanner. A conventional microarray platform comprising the identical detection sequences was used to showcase the specificity improvements realised by our detection concept.

Bacterial cells and DNA purification

All reported strains were incubated at their optimal growth conditions overnight. A detailed description of media, growth conditions, and origin is given in Table S1. A total of 5 mL of each bacterial culture were washed in phosphate-buffered saline (PBS) buffer (1.05 mM KH_2PO_4, 155.17 mM NaCl, 2.97 mM Na_2HPO_4 7 H_2O, pH=7.4) twice by spinning down (6,000 g, 2 min) and resuspension, then finally resuspended in 1 mL of PBS buffer. Lysis was performed mechanically by a MagNa Lyser Instrument (Roche, Basel) at 6,500 rpm for 0.5 min. A second treatment was performed after 5 min of incubation at room temperature (RT), followed by 10 min of incubation at 95 °C. After centrifuging at 16,000 g for 10 min, the supernatants were collected and used further.

To validate the microarray detection results, all strains were sequenced with the PGM Ion Torrent sequencer (Ion Personal Genome Machine™ System, Thermo Fisher Scientific, Waltham, MA, USA) as advised by the supplier.

DNA amplification

After lysis and centrifugation, the supernatants were used directly as template DNA in a multi-primer PCR containing the primer pairs corresponding to all the genes screened for in case of ABR and VF genes, while for the 16S rRNA genes, a universal primer pair was designed covering approximately the whole gene. All primers are listed in the Supplementary Information, Tables 3-20. The PCR was conducted using the HotStartTaq DNA polymerase kit (Qiagen, Hilden, Germany) at final concentrations of 10 ng of template DNA, 5 µM of primer pairs (forward and reverse), and 3 mM of $MgCl_2$ following the manufacturer's instructions. The thermal cycling settings were: initial denaturation: 95 °C, 15 min; number of cycles: 40, comprising denaturation: 95 °C, 30 s; annealing: 55 °C, 30 s; elongation: 72 °C, 30 s; final elongation: 72 °C, 10 min; storage: 4 °C (Applied Biosystems GeneAmp PCR System 2700, Thermo Fischer Scientific, Waltham, MA, USA). Purification of the amplification products was done with the Stratec Invisorb Fragment Clean Up kit (Stratec Molecular GmbH, Berlin, Germany) following the manufacturer's instructions, followed by 3 min of sonication (VWR Ultrasonic Cleaner USC-TH, PE, USA). The resulting amplified oligonucleotides served as target DNA for the detection reaction using the LNC-3 technology.

DNA labelling for the conventional microarray

For the conventional microarray detection, 6 µL of the amplified DNA originating from the aforementioned PCR were labelled with Atto 532-labelled dCTP. The labelling reaction was conducted using 2 units of the Vent (exo-)DNA Polymerase kit (New England BioLabs, Ipswich, MA, USA), as it has no 3' → 5'-proofreading exonuclease activity, and thus, facilitates the incorporation of the labelled dCTPs into the target DNA. To obtain the strand that was complementary to the immobilised probe only, an asymmetric PCR with the forward primer was implemented. The master mix of this PCR included 0.025 µM labelled dCTPs, 0.175 dNTPs mix, 0,9 µM forward primer and 1 mM $MgSO_4$. The thermal cycling settings were: initial denaturation: 95 °C, 3 min; number of cycles: 25, comprising denaturation: 95 °C, 20 s; annealing: 55 °C, 20 s; elongation: 72 °C, 20 s; final elongation: 72 °C, 3 min; storage: 4 °C (Applied Biosystems GeneAmp PCR System 2700, Thermo Fischer Scientific, Waltham, MA, USA).

Oligonucleotide design

All oligonucleotides were designed with the Oli2go software, which uses thermodynamic calculations to facilitate multiplex detection applications[53]. The structure of the LNC oligonucleotides is summarised in Table 1 and listed completely in the Supplementary Information, Tables 3-20. The resulting oligonucleotides were obtained from Integrated DNA Technologies (Coralville, Iowa, USA), including Cy5-labelled LNC-A and Cy3-labelled LNC-B oligonucleotides serving as spotting and hybridisation controls.

Table 1: Microarray probe sequences and modifications.

Probe name	5'-mod.	Sequence 5'-3'	Length
LNC-A	Thiol	TTTCGCTGCCGACCCTGCGCCGTGGCC	27 bp
LNC-B		CCCCGGCACGCGAGCCCACGCTGCTTTTTTGGCCACGGC GCAGGGTCGGCAGCG	54 bp
LNC-C		GCAGCGTGGGCTCGCGTGCCGGGGTTTTTTNNNNNNNN NNN	≈ 44 bp
Detection oligonucleotide		NNNNNNNNNNNNTTT	≈ 23 bp
Hybridisation control		CCCCGGCACGCGAGCCCACGCTGCTTTTTTGGCCACGGC GCAGGGTCGGCAGCG/Cy3	54 bp
Spotting control	Thiol	TTTCGCTGCCGACCCTGCGCCGTGGCC/Cy5	27 bp

Glass slide functionalisation

The glass slides intended for the LNC-3 probes were first cleaned by sonication in ddH_2O, 100 % ethanol, acetone, and 1 M NaOH for 10 min each, then finally immersed in 1 M HCl overnight. The next day, the slides were cleaned with ddH_2O for 10 min twice, rinsed with 100 % ethanol, and centrifuged to dry. The cleaned glass slides were immersed in 5 % 3-aminopropyl-trimethoxysilane (ATS, Sigma-Aldrich, MO, USA) in dry acetone for 1 h to generate free amino groups on the top of the surface. Residual ATS was removed by washing with acetone three times for 5 min each, rinsing with 100 % ethanol, and centrifuging to dry. Prior to succinimide functionalisation, the slides were baked at 90 °C for 50 min. Subsequently, 300 µL of 2 mM sulfo-m-maleimidobenzoyl-N-hydroxysulfosuccinimide (s-MBS, Thermo Fischer Scientific, Waltham, MA, USA) in PBS buffer was applied to the amino-functionalised surface. The reaction occurred in a humid surrounding overnight. The procedure was taken over from[57].

Spotting of the microarray probes

The LNC-3 probes were immobilised on the chip surface using an OmniGrid Contact Microarrayer (GenMachines, San Carlos, CA, USA) equipped with Stealth Micro Spotting Pins (ArrayIt, Microarray Technology, Sunnyvale, CA, USA). The spotting solution contained 5 µM of LNC-A, LNC-B, and LNC-C oligonucleotides each in sterile-filtered 2x NaP_i spotting buffer (1x NaP_i: 0.1 M Na_2HPO_4, 0.15 M NaCl, pH=6.5). Cy-5-labelled LNC-A oligonucleotides were used as spotting controls, Cy3-labelled LNC-B oligonucleotides as hybridisation controls. After spotting, the microarray chips remained in the spotting chamber overnight. The humidity was set to 60 % during the entire process. The slides were immersed into 1x NaP_i buffer repeatedly for 5 min to wash, then incubated in the same buffer additionally containing 10 mM β-mercaptoethanol for 1 h to inactivate residual reactive groups on the glass surface. To remove the β-mercaptoethanol, the NaP_i buffer washing step was repeated. To precipitate unbound LNC oligonucleotides, the slides were incubated in saline buffer (1.5 M NaCl, 0.01 M Na_2HPO_4, pH=7) for 10 min. Another two washing steps were implemented with 5x saline sodium citrate (SSC) buffer (1x SSC buffer: 150 mM NaCl, 15 mM sodium citrate, pH=7.0, obtained from Biorad, Hercules, California, USA) for five minutes each, the first with 0.1 % Tween-20 (Sigma-Aldrich, MO, USA) and the second without. Finally, the slides were repeatedly rinsed with ddH_2O, centrifuged to dry, and stored at -20 °C.

Spotting and blocking of the aldehyde-functionalised slides

The conventional microarray detection was carried out using aldehyde-modified slides (PolyAn, Berlin, Germany). The spotting was conducted analogously to the LNC-3 procedure with the same instruments but a spotting buffer comprising 3 M betaine und 6x SSC buffer. The slides were immersed into a blocking solution containing 3 M urea and 0.1 % SDS for 30 min. Afterwards, the slides were washed with PBS buffer containing 0.1% Tween-20 for 5 min. Finally, the slides were repeatedly rinsed with ddH_2O, centrifuged to dry, and stored at -20 °C.

Detection oligonucleotide preparation

The biotin-modified detection oligonucleotides were ordered from Integrated DNA Technologies (Coralville, Iowa, USA) and then phosphorylated using the T4 polynucleotide kinase (Thermo Fischer Scientific, Waltham, MA, USA). For that, 50 U/reaction of the enzyme, 45 detection oligonucleotides at a concentration of 1.78 µM each, and 400 µM adenosine triphosphate (ATP) were allowed to react in a thermoshaker (Peqlab TS-100, VWR, Erlangen, Germany) at 37 °C overnight. The reaction was ended by 10 min of incubation at 75 °C. After establishing a protocol to save costs by self-biotinylating the detection oligonucleotides, a terminal deoxynucleotidyl transferase was used to elongate the detection oligonucleotides with biotinylated deoxynucleotide trisphosphates (dCTPs) without impairing their performance, which were subsequently bound by

fluorescence-labelled, i.e. Alexa-647-conjugated, streptavidin. The procedure is outlined in detail in a recent study[52].

Solid-support-based ligation and detection

The ligation reaction, providing the required specificity, took place in gasket hybridisation chambers of a volume of 100 µL each (Agilent, Santa Clara, CA, USA). The ligation solution, comprising 5 U/reaction ampligase (Epicentre, Madison, WI, USA), 1.68 µM of each detection oligonucleotide, 100 nM of synthetic target DNA or 20 µL of PCR product, and 2 µg bovine serum albumin (BSA, New England BioLabs, Ipswich, MA, USA) in ampligase buffer (20 mM Tris-HCl, 25 mM KCl, 10 mM $MgCl_2$, 0.5 mM nicotinamide adenine dinucleotide (NAD), 0.01 % Triton X-100, pH=8.3), was applied to the chamber. If PCR products were used, the mixture was heated up to 95 °C for 5 min followed by cooling down on ice prior to its application onto the microarray.

The ligation reaction was performed in a hybridisation oven (Microarray Hybridisation Chambers, Agilent, Santa Clara, CA, USA) at 55 °C for 1 h. The slides were subsequently washed repeatedly, first with 2x SSC buffer containing 0.1 % sodium dodecyl sulphate (SDS) for 5 min, second with 0.2x SSC buffer for 2 min, third with ddH_2O for 1 min, and finally with 70 °C ddH_2O for 10 min to remove non-ligated hybridisation products. The slides were dried by centrifugation. Afterwards, the streptavidin-Alexa-647 conjugate (Jackson ImmunoResearch Laboratories Inc., West Grove, PA, USA) was diluted 1:1000 in sterile-filtered PBS buffer containing 0.1 % Tween-20, applied to the slides, and incubated for 1 h at room temperature inside the hybridisation oven. Finally, the slides were washed with PBS buffer comprising 0.1 % Tween-20 for 5 min, then twice with ddH_2O. Again, the slides were centrifuged to dry.

Conventional microarray detection

The conventional microarray hybridisation was performed in gasket hybridisation chambers of a volume of 100 µl each (Agilent, Santa Clara, CA, USA). The hybridisation mixture, comprising 20 µL of the fluorescence labelled amplification product, 40 µL hybridisation buffer (Roche Diagnostics GmbH, Rotkreuz, Switzerland), and 40 µL ddH_2O, was applied to the hybridisation chambers. The reaction occurred in a hybridisation oven (Microarray Hybridisation Chambers, Agilent, Santa Clara, CA, USA) at 55 °C for 1 h. The slides were subsequently washed repeatedly, first with 2x SSC buffer containing 0.1 % sodium dodecyl sulphate (SDS) for 5 min, second with 0.2x SSC buffer for 2 min, and third with ddH_2O for 1 min. The slides were dried by centrifuging.

Measurements and data evaluation

The slides were scanned with a Tecan PowerScanner (Männedorf, Switzerland). Slide coating and spotting efficiency were checked by Cy5-labelled LNC-A oligonucleotides, serving as spotting controls, and Cy3-labelled LNC-B oligonucleotides, serving as hybridisation controls, both spotted along with the probes at the same concentrations. The detection oligonucleotides carried biotin-modifications that allowed binding to Alexa-647-labelled streptavidin. The hybridisation controls were measured at 532 nm, all other samples at 647 nm. GenePix Pro 6.0 (Molecular Devices, Sunnyvale, CA, USA) was used for data analyses, including readout of fluorescence intensity counts (FIC), subtraction of the background (average FIC over the unspotted area), and median value calculation of the spotted areas. The graphs were compiled with Origin Lab 8.5 (OriginLab Corporation, Northampton, MA, USA). For statistical considerations, an R script (R Core Team, 2013: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria) was written to assess the significance of the positive signals over all other ones. The results were depicted as boxplot diagrams.

Results and discussion

In this work, we developed a DNA microarray-based assay that is capable of genetically identifying and characterising pathogens. The focus was put on the detection of the 45 most important sepsis-relevant bacterial pathogens and their ABR and VF genes. Our assay was compared to a conventional DNA-based microarray and its specificity was assessed analysing 14 sequenced reference strains of the 45 species. The main focus of this work is on the 6 ESKAPE species; all other strains are summed up in the Supplementary Information (Table S2).

Evaluation of the probe performance

Upon the *in silico* design of the 814 oligonucleotides, they were first experimentally evaluated with synthetic target DNA to identify non-functional probes (Fig 2a-c). The synthetic DNA was 100 % complementary, single-stranded, and not dependent on any other factors, such as the upstream PCR amplification, therefore it represented the simplest case. The LNC-3 probes that generated weak fluorescence signals were subject to more detailed *in silico* secondary structure analyses and reordered to exclude shortcomings during the chemical synthesis. Along with, the detection oligonucleotides were scrutinised.

Figure 2a-c show exemplary results of the microarray probes corresponding to 16S rRNA genes (Figure 2a), ABR genes (Figure 2b), and VF genes (Figure 2c), summarised in boxplot diagrams displaying the median values of four repetitions. The complete set of measurements is given in the Supplementary Information (Fig. S9-S21). The data were normalised to be comparable between different experiments. This was due to different individual factors, such as the in-house slide coating. Briefly, the positive signals, detected at the respective correct spots, are clearly contrasting from the non-matching probe signals, even without statistical evaluation. Having a look at the statistical boxplot analysis, a threshold can easily be drawn to separate positive values (red) from the counts emerging at non-matching spots (grey). All in all, the probes representing phylogenetic 16S rRNA-based markers, ABR genes and VF genes were shown to work well; the probes were sufficiently specific. Including the residual sets of phylogenetic markers, ABR genes, and VF factors, shown in Fig. S9-S25, 97.7 % of the designed LNC-3 probes yielded acceptable signals (> 10-fold standard deviation). A total of 2.2 % of the probes produced fluorescence intensities that were lower than the 10-fold standard deviation; however, they were still distinguishable from the background noise.

Comparison between different microarray technologies

The LNC-3 approach was compared to a conventional DNA microarray method. The conventional method consisted of an immobilised oligonucleotide carrying a sequence complementary to the target DNA hybridising to the immobilised probe. In our alternative approach, the same sequence was split in two parts encoded on the LNC-3 probe and the detection oligonucleotide, so that the target DNA binds the detection oligonucleotide and immobilised probe via hybridisation. Only if subsequently ligated by a proofreading ligase, the detection oligonucleotides were not removed by stringent washing. The performance of those two methods was determined by identifying six different clinically relevant pathogens via their 16S rRNA genes. The results of two representative detection reactions targeting *E. faecium* and *K. pneumoniae* are shown in Figure 2d and 2e. The residual results are given in the Supplementary Information (Fig. S7).

While the LNC-3-based microarray chip exhibited a single significant signal only in case of the *E. faecium* probe, the conventional chip surface presented positive signals (at least 50 % of the maximum value) also in the case of *Enterococcus* subsp. (consensus sequence for *Enterococcus* species/strains), *Clostridium* subsp., *Moraxella* subsp., and *Streptococcus* subspecies probes (Fig. 2d). The background was in both cases negligible (smaller than 10 % of the maximum signal). The multispecies probe showed a slightly higher signal in the LNC-3 assay. This LNC-C probe is perfectly matching to the 16S rRNA gene sequences of *Morganella morganii*, *K. pneumoniae*, *Citrobacter. freundii*, *Enterobacter cloacae*, *Enterobacter aerogenes*, *E. faecium*, *Escherichia coli*, *P. aeruginosa* and *Proteus* subspecies (explained in detail in Fig. S6). Since these strains are closely related (Fig. S1 and S2), the 16S rRNA sequences overlap widely. Such a consensus sequence was used for the multispecies probe. The identification is thus successful through the multiple species probe and the specific probe. The signal intensities of the individual signals differ due to thermodynamic differences between species-specific probe/target DNA and multiple species probe/target DNA interactions, strictly speaking the multiple species probe resulted in a lower ΔG value. This was used by implementing this probe to achieve a putative higher sensitivity in detecting the presence of one of them compared to the single species probes.

Targeting *Klebsiella pneumoniae* (Fig. 2e), the conventional microarray analysis showed many unspecific signals, while the LNC-3 method clearly identified the species (*K. pneumoniae* and the multispecies probe, comprising the first). It was not possible to identify a phylogenetic correlation in the false positive signals created by the conventional microarray technique.

315 Attempting the identification of *Streptococcus pseudopneumoniae* (Fig. S7), probed together with *S. pneumoniae*
316 at one spot because of the very high similarity, the conventional microarray resulted in one specific and three non-
317 specific signals that were higher than 50 % of the maximum value, namely at the correct *S. pneumoniae* and the
318 related *S.* subsp. spot, but also at the loci of *Moraxella* subsp. and the multispecies probe (not containing *S.*
319 *pseudopneumoniae*). Another three produced a significant signal, albeit lower than 50 % of the maximum value
320 (*Clostridium* subsp., *Listeria monocytogenes*, and *Streptococcus mitis*). In the LNC-3-based array, only the
321 perfectly matching probe of *S. pneumoniae* was detected, while all others did not exceed 10 % of the maximum
322 value. A special case was the detection of *Proteus penneri*, for which no perfectly matching probe besides the
323 multispecies probe was presented on the microarray chip. Indeed, the maximum signal was obtained at the
324 multispecies probe, while relative signals of 20-50 % were measured for the three related species *P. mirabilis*,
325 *P.* subsp., and *P. vulgaris*, giving information about the signal attenuation by minor sequence differences of closely
326 related species using the LNC-3 concept. The conventional microarray signals at those loci were higher than 80 %.
327 Besides, the conventional microarray produced signals relative to the highest one at the multispecies probe of more
328 than 50 % at the *Clostridium* subsp, the *Moraxella* subsp., and the *Serratia marcescens* probes as well, which are,
329 however, not very related. The other two identifications, which generated an equivalent outcome, are not discussed
330 here in detail but shown in Fig. S7, too.

331 Altogether, the conventional microarray produced a high number of false-positive signals, while the LNC-3
332 detection method was significantly more specific and correct in any case.

333 *Reproducibility*

334 The reproducibility of the microarray system was evaluated with this experiment as well. The reproducibility is a
335 known problem with microarray technology[58–61]. In contrast to scientific research, there are no repetitions
336 performed in clinical diagnostics usually. For that reason, the correct signals must be obtained reliably, not in the
337 average of repeated measurements. In Fig. 2f, three repetitions of an ABR gene set detection of *E. faecium* is
338 shown. Although the absolute intensities of the individual measurements differ greatly from one another due to
339 the in-house slide coating, they exhibited significant signals towards their individual backgrounds and
340 consequently generated a uniform statement about the resistance genes of the strain.

341 *Sensitivity test of the LNC-3 microarray technology*

342 The higher complexity of the employed reaction, comprising several additional steps, suggests that this gain of
343 specificity could be accompanied by a loss of sensitivity. The total fluorescence signal intensity was found to be
344 lower with the LNC-3 method than with the conventional microarray. For that reason, the sensitivity of the LNC-
345 3-based microarray was determined using a target DNA dilution series (0.1 pg – 10 ng) of *S. aureus* in the
346 amplification reaction. The resulting PCR products were applied to the chip (Fig. 2g). 1 ng DNA corresponds to
347 10^5 bacterial cells per mL (e.g. of blood). With 0.1 pg and 1 pg of target DNA, no signal differentiable from the
348 background noise was obtained. 10 pg, corresponding to 10^3 cells, led to a significant signal compared to the
349 background. 0.1, 1, and 10 ng resulted in higher signals. Hence, thousand pathogen cells could be detected,
350 enabling the LNC-3 technology to compete with other pathogen detection formats[62,63]. The threshold of 10 pg
351 DNA was achieved with other target DNA sources as well, such as *E. faecalis*. In two cases, the sensitivity was
352 determined to be 100 pg, corresponding to 10^4 cells per mL. The results are shown in the Supplementary
353 Information (Fig. S8).

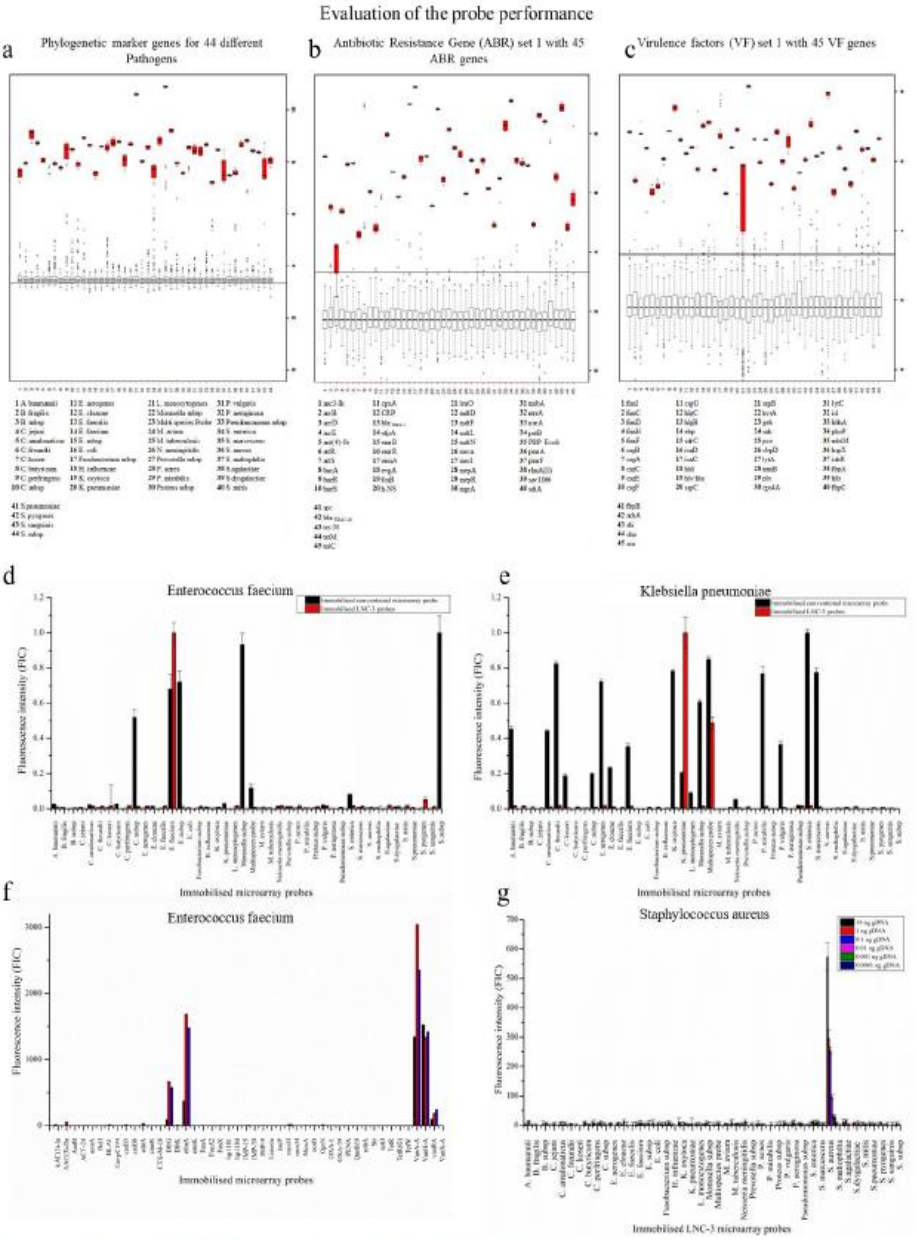


Figure 2: Performance of the LNC-3 probes: A set of 45 LNC-3-probes (and detection oligonucleotides) was tested in terms of specificity by applying complementary synthetic target DNA to the chip surface, the latter carrying all 45 specific parts of 16S rRNA genes serving as phylogenetic markers (a), antibiotic resistance (ABR) genes (b), and virulence factor (VF) genes (c). All correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey), normalised to 1. In the boxplots, all single values are located inside the boxes with associated error bars representing the standard deviation. The black bars represent the median values. The grey boxes summarise the values of the probes that did not match the gene fragment of interest, the red ones illustrate the positive signals. **Identification** via conventional, hybridisation-based detection (black) and LNC-3 technology (red) of *E. faecium* (d) and *Klebsiella pneumoniae* (e). **Reproducibility** of the ABR detection of *E. faecium* with the LNC-3 technique (f). **Sensitivity analysis:** Different amounts of target DNA were applied to the LNC-3-functionalised microarrays. 10 ng DNA corresponds to 10^6 bacterial cells (g).

9

Full characterisation of bacterial isolates

The LNC-3 microarray platform was verified by characterisations of the six pathogens abbreviated with ESKAPE (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa and Enterobacter cloacae*). Those species have been particularly highlighted by the Infectious Diseases Society of America for being especially problematic in terms of antibiotic resistances[6,7]. To validate the results, the genomes of the used ESKAPE strains were, if publicly not available, sequenced and bioinformatically assessed beforehand. The results are summarised in Table 1. In total, 14 pathogenic strains were characterised. Figure 3 shows the characterisations of clinical isolates from *E. faecium* (a), *A. baumannii* (b), *E. cloacae* (c), and *S. aureus* (d). Those data and the characterisation data of the two remaining ESKAPE pathogens *K.pneumoniae* and *P. aeruginosa* are additionally listed in Table 2. Expected genes whose corresponding probes did not exhibit sufficient signal intensities are highlighted in bold letters. Additionally identified ones are marked in italics. The residual characterisations are given in the Supplementary Information (Table S2).

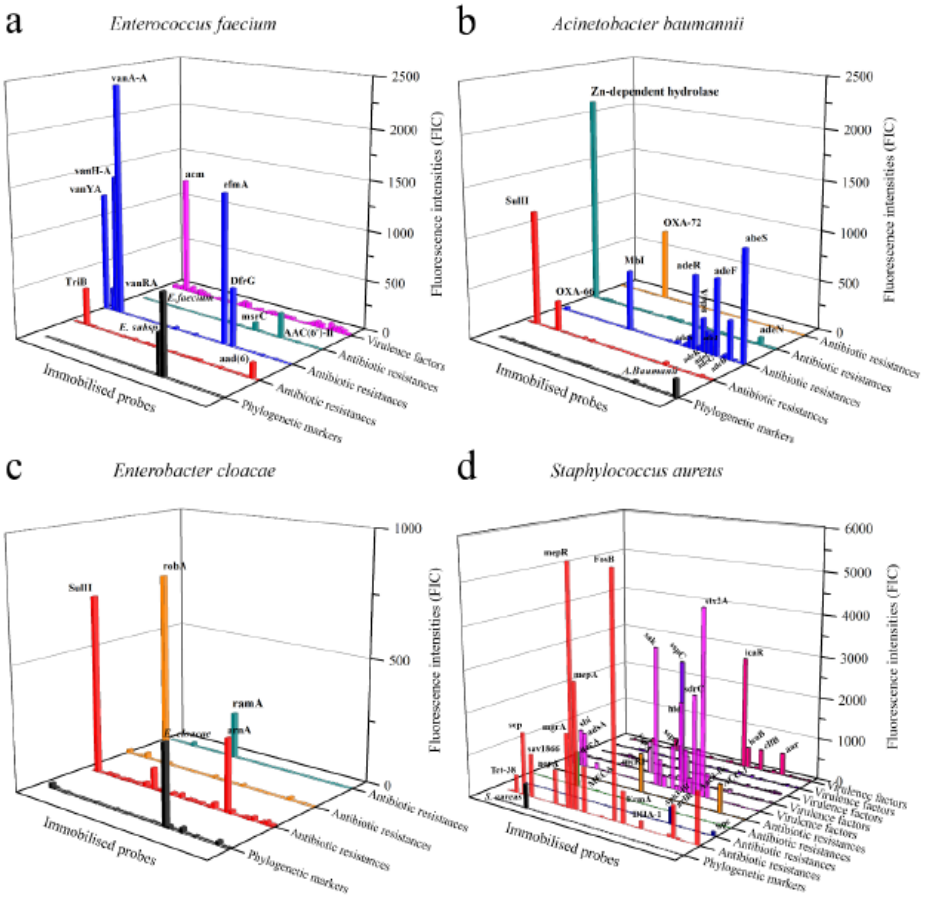


Figure 3: Full characterisation of *Enterococcus faecium* (a), *Acinetobacter baumannii* (b), *Enterococcus cloacae* (c), and *Staphylococcus aureus* (d) isolates, comprising identification (black) via 16S rRNA genes and characterisation in terms of ABR genes (red/blue/green/orange) and virulence factors (magenta and related colors). The DNA of a clinical isolate carrying the specified pathogen was purified and amplified by multiplex PCRs including the primer pairs of all investigated genes. The amplification products were applied to the microarray chips along with detection oligonucleotides. If matching, the detection oligonucleotides could be ligated to the probes and subsequently detected by a standard fluorescence-based microarray scanner.

During the identification of *E. faecium*, two positive signals were generated, one at the probe corresponding to *E. faecium* and one at the probe corresponding to *Enterococcus subspecies* with a lower fluorescence signal (Figure 3a: black). Furthermore, the characterisation revealed the presence of the antibiotic resistance genes *AAC(6')-Ii, msrC, efmA, DfrG, VanA-A VanH-A, vanRA, vanYA, TriB,* and *aad(6)* (Fig. 3a: red, blue and green) and the virulence factor gene *acm* (Fig. 3a: magenta). These results were compared with the shotgun sequencing data. It

388 turned out that with the genes *DfrG, VanA-A VanH-A, vanRA, vanYA, TriB, aad(6)* had been identified in addition
389 using the LNC-3 method. To ensure that the additional hits of ABR gens are not a result of measurement outliers,
390 the measurement was repeated three times with the ABR gene panel comprising *DfrG* and the vancomycin-related
391 resistance genes (shown in Fig. 2f and discussed in detail there) to exclude that. Those additional antibiotic
392 resistance genes were detected repeatedly. A number of possible explanations can be thought of, such as the
393 binding of similar target DNA. However, *Enterococcus* species are known to frequently carry vancomycin
394 resistance genes. A detection of the latter was hence not surprising. A conceivable mechanism of the vancomycin-
395 related resistance genes escaping notice in whole genome shotgun sequencing data is their spread on transposable
396 elements[64,65]. The assembly of larger sequence parts from contiguous motifs (contigs) may suffer from the direct
397 repeats transposable elements and integrons are enclosed with, leading to sequences that are not assembled *in silico*
398 correctly. In cases such as the vancomycin resistances, it was hence assumed that rather the WGS shotgun data
399 were incomplete. Nevertheless, all genes revealed by WGS were found using the LNC-3 technique.

400 **Table 2:** Characterisation of bacterial isolates using LNC-3 probes targeting antibiotic resistance genes and virulence factor genes. LNC-3-
401 derived signals are compared with whole genome sequencing data. Genes found in sequence data but without LNC-3 response are indicated in
402 bold letters, LNC-3 responses that were not found via sequencing are written in italics.

Pathogen/gene type	Phylogenetic marker genes	Antibiotic resistance genes		Virulence factor genes	
		Sequenced	LNC-3	Sequenced	LNC-3
Acinetobacter baumannii	*Acinetobacter baumannii*	**aadA**, OXA-66, SulI, Mbl abeS, adeC, adeJ, adeS, adeB, adeA, adeK, adeG, ADC-2, BlaA1, adeF, adeR, adeI, adeN, Zn-dependent hydrolase, **AAC(3)-Ia**, OXA-72, BlaA2	OXA-66, SulI, Mbl abeS, adeC, adeJ, adeS, adeB, adeA, adeK, adeG, ADC-2, BlaA1, adeF, adeR, adeI, adeN, Zn-dependent hydrolase, OXA-72, BlaA2		
Enterobacter cloacae	*Enterobacter cloacae*	ramA, SulII, robA	ramA, SulII, robA, *arnA*		
Enterococcus faecium	*Enterococcus faecium, Enterococcus subspecies*	AAC(6')-Ii, msrC, **efmA**	AAC(6')-Ii, msrC, **efmA**, *DfrG, VanA-A, VanH-A, vanRA, vanYA, TriB, aad(6)*	acm	acm
Klebsiella pneumoniae	*Klebsiella pneumoniae*	oqxA, **FosA5**, **vgaC** acrA Klebsiella, oqxB	oqxA, acrA Klebsiella, oqxB	yagZ/ecpA, **eastI_astA**	yagZ/ecpA
Pseudomonas aeruginosa	*Multispecies probe, Pseudomonas aeruginosa*	**PDC-1**, amrA, amrB **Aph3-IIb**, arnA, basS, **CatB7**, CpxR, **MexA**, MexB, MexD, **MexE**, **MexF**, mexG, mexI, mexJ, mexK, mexL, mexP, mexQ, mexV, mexW, **MuxB**, **MuxC**, OpmB, **opmD**, opmE, OpmH, OprJ, OprM, OXA-50, TriB, TriC, **OprN**, mexH, FosA, MuxA, mexM, **MexC**, **TriA**	amrA, amrB, arnA, basS, CpxR, MexB, MexD, mexG, mexI, mexJ, mexK, mexL, mexP, mexQ, mexV, OpmB, opmE, OpmH, OprJ, OprM, OXA-50, TriB, TriC, mexH, FosA, MuxA, mexM,	xcpA/pilD, algB, algQ, algZ, algU, alg8, alg44, algE, algX, algL, algF, algA, mucC, waaG, waaC, aprA, lasA, lasB, rhlI, lasI, plcH, xcpZ, xcpV, **xcpT**, xcpS, xcpP, xcpQ, pilY2, **pilS**, pilR, pilP, **pilM**, pilT, pilU, pilG, **pilH**, pilI, chpB, chpC, flgD, flgG, flgH, flgI, flgJ, fleQ, fleR, **fliE**, **fliG**, fliI, fliJ, fliM, fliN, fliP, fliQ, flhB, **flhA**	xcpA/pilD, algB, algQ, algZ, algU, alg8, alg44, algE, algX, algL, algF, algA, mucC, waaG, waaC, aprA, lasA, lasB, rhlI, lasI, plcH, xcpZ, xcpV, xcpS, xcpP, xcpQ, pilY2, pilR, pilP, pilT, pilU, pilG, pilI, chpB, chpC, flgD, flgG, flgH, flgI, flgJ, fleQ, fleR, , fliI, fliJ, fliM, fliN, fliP, fliQ, flhB,
Staphylococcus aureus	*Staphylococcus aureus*	Aac3-Ik, arlR, **arlS**, DHA-1, ErmA, FosB, MECA, mepA, mepR, mgrA, norA, sav1866, Spc, Tet-38, Aph3-III, qacA, mecR1	Aac3-Ik, arlR, DHA-1, ErmA, FosB, MECA, mepA, mepR, mgrA, norA, sav1866, Spc, Tet-38, Aph3-III, qacA, mecR1, *ACC-1*	hlgC, hlgB, ebp, sdrC, icaD, icaC, hld, hly/hla, sspC, sspB, hysA, geh, sak, hlb, adsA, scn, sdrD, sdrE, clfA, map, sea, sspA, icaA, fnbA, icaR, icaB, clfB, aur	hlgC, hlgB, ebp, sdrC, icaD, icaC, hld, hly/hla, sspC, sspB, hysA, geh, sak, hlb, adsA, scn, sdrD, sdrE, clfA, map, sea, sspA, icaA, fnbA, icaR, icaB, clfB, aur

11

403 *A. baumannii* was clearly identified as well (Fig. 3b). The sequence data revealed 22 different ABR genes (*aadA,*
404 bla$_{OXA-66}$*, sulI, mbl, abeS, adeC, adeJ, adeS, adeB, adeA, adeK, adeG adc-2, blaA1, adeF, adeR, adeI, adeN, Zn-*
405 *dependent hydrolase, aac(3)-Ia,* bla$_{OXA-72}$*, blaA2*) and no VF genes. The signal intensities of the genes *aadA* and
406 *aac(3)-Ia* were not strong enough for an unequivocal response. All others were successfully detected. Regarding
407 the VF genes, no false-positive signals were measured. An emerging pattern of non-detected ABR genes is
408 discussed along with the *P. aeruginosa* results.

409 *Enterobacter cloacae* was detected with a signal significant towards the background, and via the multispecies
410 probe comprising *M. morganii, Klebsiella* subspecies, *C. freundii*, *E. cloacae*, *E. aerogenes*, *P. aeruginosa* and
411 *Proteus* subspecies (Fig. 3c). *E. cloacae* lacked VF genes completely; the three ABR genes (*ramA, SulII, robA*)
412 were detected along with an additional one (*arnA*).

413 In the case of the identification and characterisation of *S. aureus* (Fig. 3d), only the matching probe produced a
414 signal. As mentioned before, the absolute signal strength was comparable only among the probes at one panel, not
415 between different panels. Indeed, the signal of the phylogenetic probe detecting *S. aureus* was significant compared
416 to the background, as it can be deduced from Fig. 2d as well. There, it was shown that the sensitivity was actually
417 high enough down to a bacterial concentration comparable to other detection methods. According to the
418 sequencing data, this strain carried 17 ABR genes and 28 VF genes. Only one antibiotic resistance gene could not
419 be detected with a significant signal strength: *arlS*. Although there was no straightforward reason that this single
420 ABR gene was not detected, there are a number of possible explanations as well as strategies to prevent detection
421 failures given in the *P. aeruginosa* section along with a more detailed discussion of (operon-)related genes.

422 A similar bigger picture emerged for *K. pneumoniae* (Table 2). *K. pneumoniae* was unequivocally identified, with
423 several ABR and VF genes detected. However, of five ABR genes (*oqxA, vgaC, fosA5, acrA Klebsiella & oqxB*)
424 and two VF genes (*vagZ/ecpA & east1_astA*) present according to the sequencing data, the signal intensities for
425 the genes *fosA5, vgaC*, and *east1_astA* were not sufficient to be considered a positive signal.

426 For *P. aeruginosa* (Table 2), an approximately three-fold signal intensity at the *P. aeruginosa*, *P.* subspecies, and
427 the multispecies probe loci indicated their presence. Regarding ABR genes, 40 ABR were revealed by sequencing,
428 while only 24 were detected on the LNC-3 chip. 48 of 55 VF factors could be detected.

429 The non-detected ABR genes showed some similarities, indicating that their absence was not hinting to random
430 weaknesses in the detection system but rather to a certain mechanism that mediates escape of detection so far. The
431 large majority encoded efflux pumps. A main difference of efflux pump genes towards, for example, beta-
432 lactamase genes, is that they are encoded in the genome, not on plasmids[66]. Having that closer look, most of the
433 missing genes that do not encode efflux pumps are genome-encoded as well, which – in both cases – influences
434 the copy number. The detection via LNC-3 technology is still amplification-dependent. Hence, false negative
435 might be related to the multiplex PCR amplification. Since it represents the lowest amount of target DNA to
436 amplify, a single-copy gene is the hardest to detect. Plasmids are usually present in multiple copies per cell (e.g.
437 dependent on their origin of replication), which increases the amount of target DNA substantially and is moreover
438 easier accessible. Some genomically encoded genes are present in large numbers as well, such as the rRNA genes,
439 used for the identification, facilitating the detection of those in general. Solutions to overcome this issue – since
440 increasing the template DNA cannot trivially be done by increasing the overall target DNA amount, because too
441 much DNA inhibits the PCR – could be employing a two-step PCR. In the first step, the target DNA could be
442 amplified in a limited number of cycles, improving the target DNA/non-target DNA ratio, and the product could
443 constitute the template DNA for the second (above-named) PCR, comprising the same overall template DNA
444 amount, yet containing more target DNA.

445 Besides, primer interactions could also be causative; still, this would not explain the accumulation of genome-
446 encoded genes. A suggestion to reduce such, using steric hindrance effects, was published recently by our group[67].

447 Most of those genes, especially efflux pumps, are frequently encoded in operons, e.g. the MexAB-OprM operon,
448 the MuxABC-OpmB, meaning that they are governed by a single promotor regulating their expression one after
449 the other at a time[68–70]. Those genes are not only connected on the genetic level, but often encode related proteins
450 that build, for instance, a single efflux pump. Hence, it suffices to detect one gene per operon, since they never
451 occur alone (and are not necessarily functional alone). Having detected mexA and oprM in case of *P. aeruginosa*
452 suffices to deduce the presence of mexB as well. Besides, this particular efflux pump transports special antibiotics
453 inside-out, which then needs to be rejected from the choice of antibiotics, no matter if all efflux pump genes could
454 be detected or not (which holds not true for, e.g., the vancomycin resistance genes discussed before, of which a
455 pathogen might have acquired only once and thus, they need to be detectable each and every one).

456 While the MexAB-OprM operon could therefore be deduced without detecting mexB, all single genes of the
457 MexEF-OprN operon[71] lacked detection. A strategy to reveal such operons that were failed to detect on basis of
458 every single representative (each individually suffering from the presence as genomically-encoded single copy

459 gene) could be a microarray-based clustering. With consensus sequences of those interrelated genes, detected at
460 only one locus per operon, the target DNA could be increased manifold and thus facilitate the detection. The single
461 genes contained in these operons can then be deduced easily (and not necessarily must be, since the final outcome
462 information is not the encoded protein complex but the antibiotics towards which they mediate resistance). It would
463 decrease possible primer interactions, too, by decreasing the number of primers needed. Another example was
464 mentioned in the course of the *S. aureus* detection. The not detected *arlS* was found to be located at the arlR-arlS
465 locus, encoding a regulatory system of two elements. Further, a component regulated by *arlS* is the norA
466 promoter[72,73]. Both, norA and arlR, were detected successfully. Consequently, it represents a further example that
467 would benefit from a more sophisticated way of evaluation and chip design.

468 A similar picture emerged for the virulence factors. The not detected ones were mainly single parts of multidomain
469 flagellar and pili proteins, as for example, discussing the *P. aeruginosa* results, *pilM* being part of type IV pili
470 among a lot others, such as *pilT* and *pilP* that were detected[74], or *pilH* being comprised by an operon along with
471 *pilG*, *pilI*, and *pilJ* [75]. The same applies for the flagellar proteins *fliE, fliG, fliI, fliJ, fliM, fliN* and more[76]. Hence,
472 most of the virulence factors that escaped detection could be covered by sophisticated operon and multidomain
473 protein-related evaluation.

474 In summary, 85 % of all ABR genes and 83 % of the VF genes were identified. Those which have not been
475 identified were mainly chromosomal encoded antibiotic resistance genes (mostly efflux pumps) or virulence factor
476 genes in low copy number (or even single copy), for which two strategies of cumulative detection or deductive
477 evaluation were presented.

478 **Conclusion**

479 It was shown that the LNC-3 concept is superior towards conventional microarray techniques regarding its
480 specificity by using a highly specific ligase reaction instead of relying on hybridisation bonds that often lead to
481 non-specific cross-hybridisation. By screening for 45 pathogens, 360 virulence factor (VF) genes and 409 ABR
482 genes at the same time, the number of genes that can be detected simultaneously could be tremendously increased,
483 which is also owed to the combination with the Oli2go software, avoiding critical interactions in upstream
484 amplification reactions. The LNC-3 microarray technique was further employed to fully characterise 14 pathogens,
485 so that the most effective treatment can be chosen without relying on empirical data, which become more and more
486 uncertain due to the rapid dissemination of ABR genes, e.g. via horizontal gene transfer. Improvements need to be
487 settled regarding single-copy genes, such as efflux pumps, which might be caused by lower template DNA ratios.
488 Measures such as operon-based probes, combining several related genes at one probe, were suggested.

489

490 **Acknowledgments**

491 We thank the technical staff at the AIT for providing technical assistance. This project has received funding from
492 the European Union's Horizon 2020 research and innovation programme under grant agreement No 634137.

493

494

References

1. Andersson, D. I. & Hughes, D. Microbiological effects of sublethal levels of antibiotics. *Nature reviews. Microbiology* **12**, 465–478 (2014).
2. Wang, T. Z., Kodiyanplakkal, R. P. L. & Calfee, D. P. Antimicrobial resistance in nephrology. *Nature reviews. Nephrology* **15**, 463–481 (2019).
3. Rice, L. B. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. *The Journal of infectious diseases* **197**, 1079–1081 (2008).
4. Woodford, N., Turton, J. F. & Livermore, D. M. Multiresistant Gram-negative bacteria: the role of high-risk clones in the dissemination of antibiotic resistance. *FEMS microbiology reviews* **35**, 736–755 (2011).
5. Lynch, J. P., Clark, N. M. & Zhanel, G. G. Evolution of antimicrobial resistance among Enterobacteriaceae (focus on extended spectrum β-lactamases and carbapenemases). *Expert opinion on pharmacotherapy* **14**, 199–210 (2013).
6. Pendleton, J. N., Gorman, S. P. & Gilmore, B. F. Clinical relevance of the ESKAPE pathogens. *Expert Reviews: Anti-infective Therapy* **11**, 297–308 (2013).
7. Renner, L. D. *et al.* Detection of ESKAPE Bacterial Pathogens at the Point of Care Using Isothermal DNA-Based Assays in a Portable Degas-Actuated Microfluidic Diagnostic Assay Platform. *Applied and environmental microbiology* **83** (2017).
8. *The evolving threat of antimicrobial resistance. Options for action* (World Health Organization, Geneva, op. 2012).
9. Garberi, J. *et al.* Rapid and biosecure diagnostic test for tuberculosis. *Cell biochemistry and biophysics* **65**, 173–179 (2013).
10. Lupo, A., Papp-Wallace, K. M., Sendi, P., Bonomo, R. A. & Endimiani, A. Non-phenotypic tests to detect and characterize antibiotic resistance mechanisms in Enterobacteriaceae. *Diagnostic microbiology and infectious disease* **77**, 179–194 (2013).
11. Palka-Santini, M., Pützfeld, S., Cleven, B. E. E., Krönke, M. & Krut, O. Rapid identification, virulence analysis and resistance profiling of *Staphylococcus aureus* by gene segment-based DNA microarrays: application to blood culture post-processing. *Journal of microbiological methods* **68**, 468–477 (2007).
12. Anjum, M. F., Zankari, E. & Hasman, H. Molecular Methods for Detection of Antimicrobial Resistance. *Microbiology spectrum* **5** (2017).
13. Váradi, L. *et al.* Methods for the detection and identification of pathogenic bacteria: past, present, and future. *Chemical Society reviews* **46**, 4818–4832 (2017).
14. Chan, K.-G. Whole-genome sequencing in the prediction of antimicrobial resistance. *Expert review of anti-infective therapy* **14**, 617–619 (2016).
15. Hollenbeck, B. L. & Rice, L. B. Intrinsic and acquired resistance mechanisms in enterococcus. *Virulence* **3**, 421–433 (2012).
16. Cox, G. & Wright, G. D. Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *International journal of medical microbiology : IJMM* **303**, 287–292 (2013).
17. Zacharioudakis, I. M., Zervou, F. N. & Mylonakis, E. T2 Magnetic Resonance Assay: Overview of Available Data and Clinical Implications. *Journal of fungi (Basel, Switzerland)* **4** (2018).
18. Fernández-Romero, N. *et al.* Use of multiplex PCR in diagnosis of bloodstream infections in kidney patients. *Diagnostic microbiology and infectious disease* **80**, 93–96 (2014).
19. Boujelben, I., Gdoura, R. & Hammami, A. A broad-range PCR technique for the diagnosis of infective endocarditis. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]* **49**, 534–543 (2018).
20. Xu, J. *et al.* Employment of broad-range 16S rRNA PCR to detect aetiological agents of infection from clinical specimens in patients with acute meningitis--rapid separation of 16S rRNA PCR amplicons without the need for cloning. *Journal of applied microbiology* **94**, 197–206 (2003).
21. Arya, M. *et al.* Basic principles of real-time quantitative PCR. *Expert review of molecular diagnostics* **5**, 209–219 (2005).

22. Warren, D. K., Liao, R. S., Merz, L. R., Eveland, M. & Dunne, W. M. Detection of methicillin-resistant Staphylococcus aureus directly from nasal swab specimens by a real-time PCR assay. *Journal of clinical microbiology* **42**, 5578–5581 (2004).

23. SantaLucia, J. & Hicks, D. The thermodynamics of DNA structural motifs. *Annual review of biophysics and biomolecular structure* **33**, 415–440 (2004).

24. Hendling, M. & Barišić, I. In-silico Design of DNA Oligonucleotides: Challenges and Approaches. *Computational and structural biotechnology journal* **17**, 1056–1065 (2019).

25. Verweij, J. J. & Stensvold, C. R. Molecular testing for clinical diagnosis and epidemiological investigations of intestinal parasitic infections. *Clinical microbiology reviews* **27**, 371–418 (2014).

26. Doern, C. D. & Butler-Wu, S. M. Emerging and Future Applications of Matrix-Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) Mass Spectrometry in the Clinical Microbiology Laboratory: A Report of the Association for Molecular Pathology. *The Journal of molecular diagnostics : JMD* **18**, 789–802 (2016).

27. Sauer, S. *et al.* Miniaturization in functional genomics and proteomics. *Nature reviews. Genetics* **6**, 465–476 (2005).

28. Carbonnelle, E. *et al.* MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory. *Clinical biochemistry* **44**, 104–109 (2011).

29. Carbonnelle, E. *et al.* Rapid identification of Staphylococci isolated in clinical microbiology laboratories by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of clinical microbiology* **45**, 2156–2161 (2007).

30. Dupont, C. *et al.* Identification of clinical coagulase-negative staphylococci, isolated in microbiology laboratories, by matrix-assisted laser desorption/ionization-time of flight mass spectrometry and two automated systems. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **16**, 998–1004 (2010).

31. Dubois, D. *et al.* Identification of a variety of Staphylococcus species by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of clinical microbiology* **48**, 941–945 (2010).

32. Luethy, P. M. & Johnson, J. K. The Use of Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) for the Identification of Pathogens Causing Sepsis. *The journal of applied laboratory medicine* **3**, 675–685 (2019).

33. Ehlert, T. *et al.* Establishing PNB-qPCR for quantifying minimal ctDNA concentrations during tumour resection. *Scientific reports* **7**, 8876 (2017).

34. Faron, M. L., Buchan, B. W. & Ledeboer, N. A. Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for Use with Positive Blood Cultures: Methodology, Performance, and Optimization. *Journal of clinical microbiology* **55**, 3328–3338 (2017).

35. Álvarez-Buylla, A., Picazo, J. J. & Culebras, E. Optimized method for Acinetobacter species carbapenemase detection and identification by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of clinical microbiology* **51**, 1589–1592 (2013).

36. Burckhardt, I. & Zimmermann, S. Using matrix-assisted laser desorption ionization-time of flight mass spectrometry to detect carbapenem resistance within 1 to 2.5 hours. *Journal of clinical microbiology* **49**, 3321–3324 (2011).

37. Kwong, J. C., McCallum, N., Sintchenko, V. & Howden, B. P. Whole genome sequencing in clinical and public health microbiology. *Pathology* **47**, 199–210 (2015).

38. Trevino, V., Falciani, F. & Barrera-Saldaña, H. A. DNA microarrays: a powerful genomic tool for biomedical and clinical research. *Molecular medicine (Cambridge, Mass.)* **13**, 527–541 (2007).

39. Levy, S. E. & Myers, R. M. Advancements in Next-Generation Sequencing. *Annual review of genomics and human genetics* **17**, 95–115 (2016).

40. Baudart, J. *et al.* Microarray (phylochip) analysis of freshwater pathogens at several sites along the Northern German coast transecting both estuarine and freshwaters. *Applied microbiology and biotechnology* **101**, 871–886 (2017).

41. Call, D. R., Bakko, M. K., Krug, M. J. & Roberts, M. C. Identifying Antimicrobial Resistance Genes with DNA Microarrays. *Antimicrobial Agents and Chemotherapy* **47**, 3290–3295 (2003).

42. Wilkes, T., Laux, H. & Foy, C. A. Microarray data quality - review of current developments. *Omics: A Journal of Integrative Biology* **11**, 1–13 (2007).

43. Wang, X., Ying, S., Wei, X. & Yuan, J. Development of a gold nanoparticle-based universal oligonucleotide microarray for multiplex and low-cost detection of foodborne pathogens. *International journal of food microbiology* **253**, 66–74 (2017).

44. Flavell, A. J. *et al.* A microarray-based high throughput molecular marker genotyping method: the tagged microarray marker (TAM) approach. *Nucleic acids research* **31**, e115 (2003).

45. Koltai, H. & Weingarten-Baror, C. Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. *Nucleic acids research* **36**, 2395–2405 (2008).

46. Dai, H., Meyer, M., Stepaniants, S., Ziman, M. & Stoughton, R. Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays. *Nucleic acids research* **30**, e86 (2002).

47. Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic acids research* **34**, 564–574 (2006).

48. Hooyberghs, J., van Hummelen, P. & Carlon, E. The effects of mismatches on hybridization in DNA microarrays: determination of nearest neighbor parameters. *Nucleic acids research* **37**, e53 (2009).

49. Hadiwikarta, W. W., Walter, J.-C., Hooyberghs, J. & Carlon, E. Probing hybridization parameters from microarray experiments: nearest-neighbor model and beyond. *Nucleic acids research* **40**, e138 (2012).

50. Barišić, I. *et al.* Multiplex detection of antibiotic resistance genes using padlock probes. *Diagnostic microbiology and infectious disease* **77**, 118–125 (2013).

51. Barišić, I., Kamleithner, V., Schönthaler, S. & Wiesinger-Mayr, H. Fast and highly specific DNA-based multiplex detection on a solid support. *Applied microbiology and biotechnology* **99**, 413–423 (2015).

52. Wolff, N., Hendling, M., Schönthaler, S., Geiss, A. F. & Barišić, I. Low-cost microarray platform to detect antibiotic resistance genes. *Sensing and Bio-Sensing Research* **23**, 100266 (2019).

53. Hendling, M. *et al.* Oli2go: an automated multiplex oligonucleotide design tool. *Nucleic acids research* **46**, W252-W256 (2018).

54. Chen, L. *et al.* VFDB: a reference database for bacterial virulence factors. *Nucleic acids research* **33**, D325-8 (2005).

55. Yarza, P. *et al.* The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Systematic and applied microbiology* **31**, 241–250 (2008).

56. Yarza, P. *et al.* Update of the All-Species Living Tree Project based on 16S and 23S rRNA sequence analyses. *Systematic and applied microbiology* **33**, 291–299 (2010).

57. Wolff, N. & Barišic', I. Highly Specific Ligation-dependent Microarray Detection of Single Nucleotide Polymorphisms. *Methods in molecular biology (Clifton, N.J.)* **1616**, 231–240 (2017).

58. Schaupp, C. J., Jiang, G., Myers, T. G. & Wilson, M. A. Active mixing during hybridization improves the accuracy and reproducibility of microarray results. *BioTechniques* **38**, 117–119 (2005).

59. Draghici, S., Khatri, P., Eklund, A. C. & Szallasi, Z. Reliability and reproducibility issues in DNA microarray measurements. *Trends in genetics : TIG* **22**, 101–109 (2006).

60. Chen, J. J., Hsueh, H.-M., Delongchamp, R. R., Lin, C.-J. & Tsai, C.-A. Reproducibility of microarray data: a further analysis of microarray quality control (MAQC) data. *BMC bioinformatics* **8**, 412 (2007).

61. Larkin, J., Frank, B., Gavras, H., Sultana, R. & Quackenbush, J. Independence and reproducibility across microarray platforms. *Nature methods* **2**, 337–344 (2005).

62. Hu, J. *et al.* Colorimetric-Fluorescent-Magnetic Nanosphere-Based Multimodal Assay Platform for Salmonella Detection. *Analytical chemistry* **91**, 1178–1184 (2019).

63. Call, D. R. Challenges and opportunities for pathogen detection using DNA microarrays. *Critical reviews in microbiology* **31**, 91–99 (2005).

64. Lee, T., Pang, S., Abraham, S. & Coombs, G. W. Antimicrobial-resistant CC17 *Enterococcus faecium*: The past, the present and the future. *Journal of Global Antimicrobial Resistance* **16**, 36–47 (2019).

65. Teuber, M. Veterinary use and antibiotic resistance. *Current Opinion in Microbiology* **4**, 493–499 (2001).

66. Bennett, P. M. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *British journal of pharmacology* **153**, S347-S357 (2008).

67. Wolff, N., Geiss, A. F. & Barišić, I. Crosslinking of PCR primers reduces unspecific amplification products in multiplex PCR. *Journal of microbiological methods*, 106051 (2020).

68. Tian, Z.-X., Yi, X.-X., Cho, A., O'Gara, F. & Wang, Y.-P. CpxR Activates MexAB-OprM Efflux Pump Expression and Enhances Antibiotic Resistance in Both Laboratory and Clinical nalB-Type Isolates of Pseudomonas aeruginosa. *PLoS pathogens* **12**, e1005932 (2016).

69. Mima, T., Joshi, S., Gomez-Escalada, M. & Schweizer, H. P. Identification and characterization of TriABC-OpmH, a triclosan efflux pump of Pseudomonas aeruginosa requiring two membrane fusion proteins. *Journal of bacteriology* **189**, 7600–7609 (2007).

70. Mima, T., Sekiya, H., Mizushima, T., Kuroda, T. & Tsuchiya, T. Gene cloning and properties of the RND-type multidrug efflux pumps MexPQ-OpmE and MexMN-OprM from Pseudomonas aeruginosa. *Microbiology and immunology* **49**, 999–1002 (2005).

71. Kumar, A. & Schweizer, H. P. Bacterial resistance to antibiotics: active efflux and reduced uptake. *Advanced drug delivery reviews* **57**, 1486–1513 (2005).

72. Fournier, B., Aras, R. & Hooper, D. C. Expression of the multidrug resistance transporter NorA from Staphylococcus aureus is modified by a two-component regulatory system. *Journal of bacteriology* **182**, 664–671 (2000).

73. Fournier, B., Klier, A. & Rapoport, G. The two-component system ArlS-ArlR is a regulator of virulence gene expression in Staphylococcus aureus. *Molecular Microbiology* **41**, 247–261 (2001).

74. Kalidasan, V. & Neela, V. K. Twitching motility of *Stenotrophomonas maltophilia* under iron limitation: In-silico, phenotypic and proteomic approaches. *Virulence* **11**, 104–112 (2020).

75. Corral, J. *et al.* Twitching and Swimming Motility Play a Role in Ralstonia solanacearum Pathogenicity. *mSphere* **5** (2020).

76. Sidhu-Muñoz, R. S., Tejedor, C. & Vizcaíno, N. The Three Flagellar Loci of *Brucella ovis* PA Are Dispensable for Virulence in Cellular Models and Mice. Frontiers in Veterinary Science, 7 (2020).

Supplementary Information

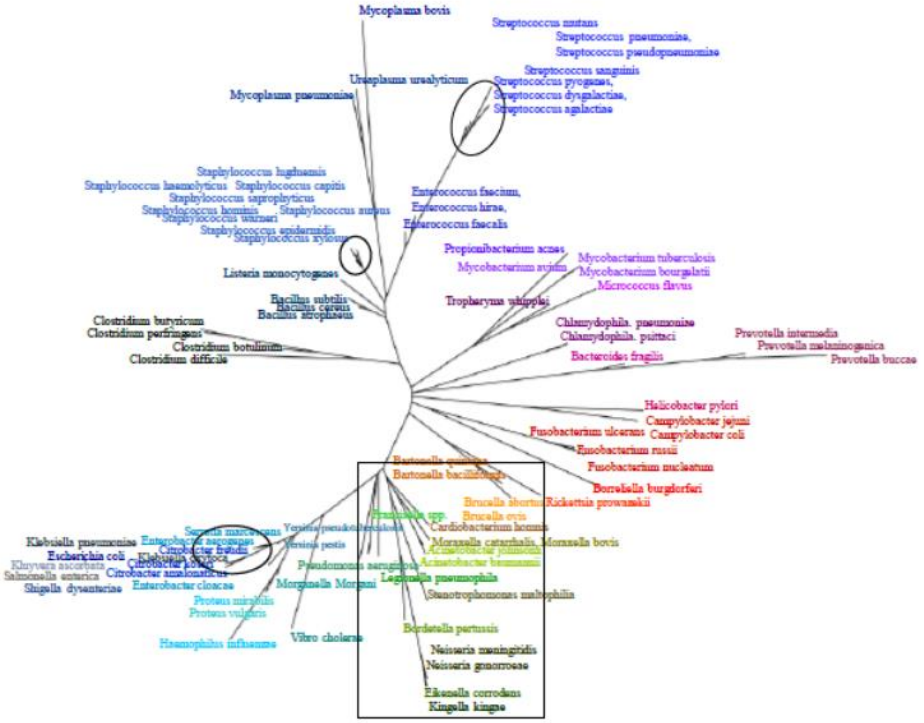
Strains and cultivation of bacterial cells

Supplementary Table 1: A detailed description of media, conditions, and origin of the detected pathogens.

No.	Species	Strain	Medium	Condition
1	Acinetobacter baumannii	A_baumannii_5M	TSB	37°C, shaking
2	Bacteroides fragilis	B_fragilis_M_7021	TSB	37°C, shaking
3	Enterobacter aerogenes	E_aerogenes_Pula_122664	TSB	37°C, shaking
4	Enterobacter cloacae	DSM30054	TSB	37°C, shaking
5	Enterococcus faecium	ATCC700221	TSB	37°C, shaking
6	Escherichia coli	DSM30083	TSB	37°C, shaking
7	Haemophilus influenzae	DSM4690	BA	37°C microaerophilic
8	Klebsiella pneumoniae	DSM30104	TSB	37°C, shaking
9	Prevotella bivia	DSM20514	CDC Anaerobe agar	37°C, anaerobe
10	Proteus mirabilis	P_mirabilis_P_1	TSB	37°C, shaking
11	Pseudomonas aeruginosa	DSM50071	TSB	37°C, shaking
12	Salmonella enterica sv. typhi	DSM554	TSB	37°C, shaking
13	Staphylococcus aureus	MRSA_18242	TSB	37°C, shaking
14	Streptococcus pseudopneumoniae	DSM18670	TSBY	37°C microaerophilic

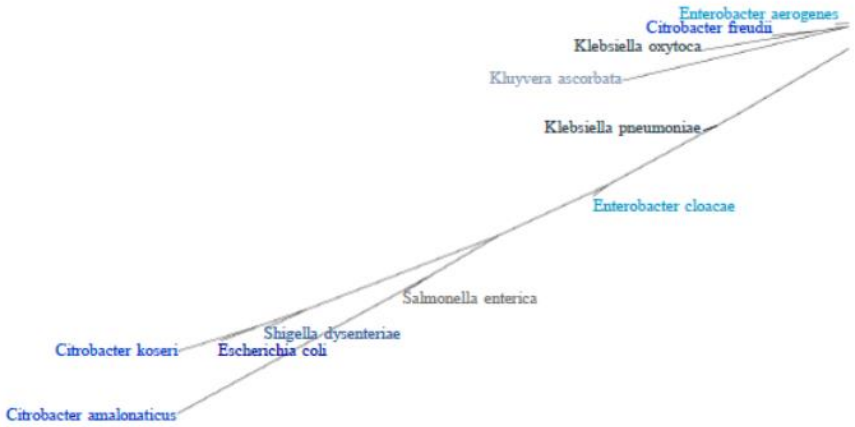
CDC anaerobe agar: BD CDC Anaerobe Agar with 5% Sheep Blood, Cat. No. 256506
BA: BD Columbia Agar with 5% Sheep Blood, Cat. No. 254005
TSB: Tryptic Soy Broth, Cat. No. 1.05459.0500, Merck KGaA
TSBY: Tryptic Soy Broth, 3 g/L yeast extract

'The All-Species Living Tree' Project (ARB-SILVA)

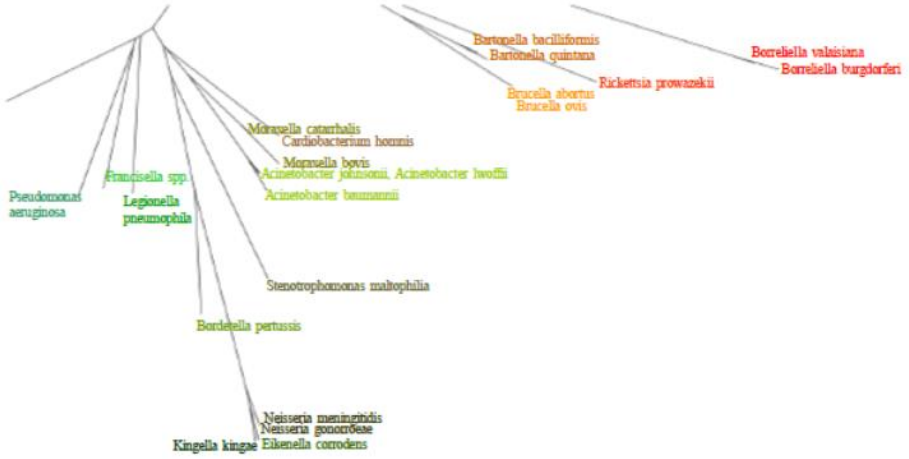


Supplementary Figure 1: 'The All-Species Living Tree' Project (ARB-SILVA): Overview. The surrounded parts are magnified in the following figures.

3

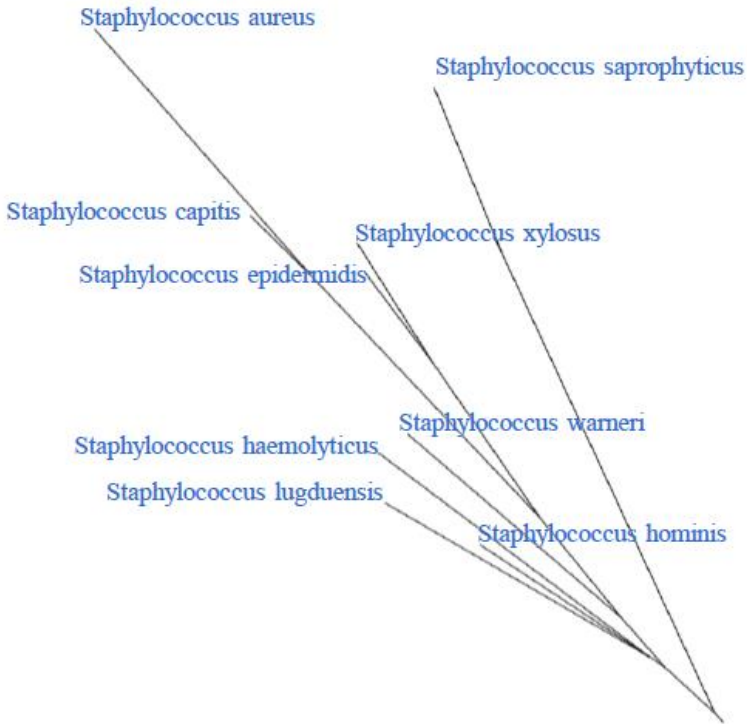


Supplementary Figure 2: 'The All-Species Living Tree' Project (ARB-SILVA): Insert Enterobacteriaceae (Gammaproteobacteria).



Supplementary Figure 3: 'The All-Species Living Tree' Project (ARB-SILVA): Insert Gammaproteobacteria.

4

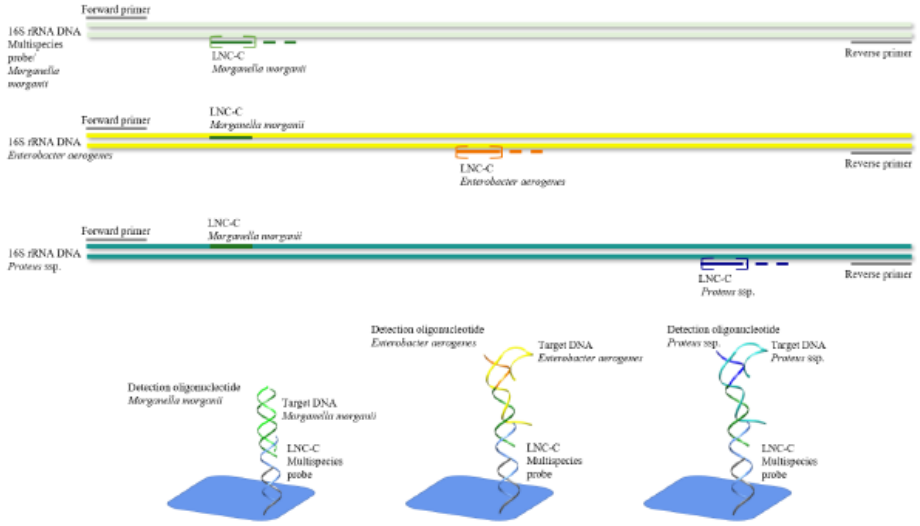


Supplementary Figure 4: 'The All-Species Living Tree' Project (ARB-SILVA): Insert *Staphylococcus* (Bacilli).

Supplementary Figure 5: 'The All-Species Living Tree' Project (ARB-SILVA) insert *Streptococcus* (Bacilli).

5

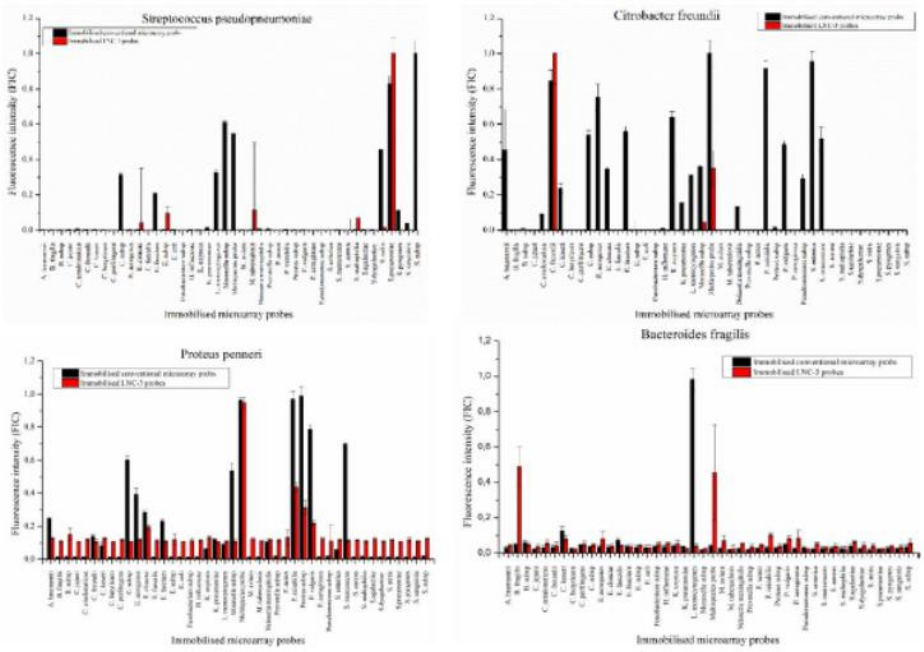
Multispecies probe



Supplementary Figure 6: Schematic depiction of the loop formation enabling several detection oligonucleotides to bind to the *M. morganii* probe. Top: A specific region of the 16S rRNA DNA was amplified by PCR using consensus sequences for the primers (grey). The amplicon (light green for *M. morganii*, yellow for *E. aerogenes*, turquoise for *Proteus* ssp.) carried species-specific parts (green for *M. morganii*, orange for *E. aerogenes*, blue for *Proteus* ssp.), of which the complementary sequence was chosen for the LNC-C probe (solid line) and the following section (dashed line) for the detection oligonucleotides. For example, the LNC-C part of *M. morganii* was also present at the other amplicons (not the detection oligonucleotide complementary sequence, so this was not influencing the probe testing using synthetic DNA). Bottom: For the given examples, it is shown how loop formation enables the detection oligonucleotides of all three to the *M. morganii* probe (even simplified by a lower ΔG value for the non-desired combinations). On the other hand, *Enterobacter aerogenes* and *Proteus* ssp. could bind to the respective other probes, making the latter a multispecies probe. The especially low ΔG value should be used to detect these bacteria with high sensitivity.

6

Comparison between the conventual microarray technique and the LNC-3 microarray technology

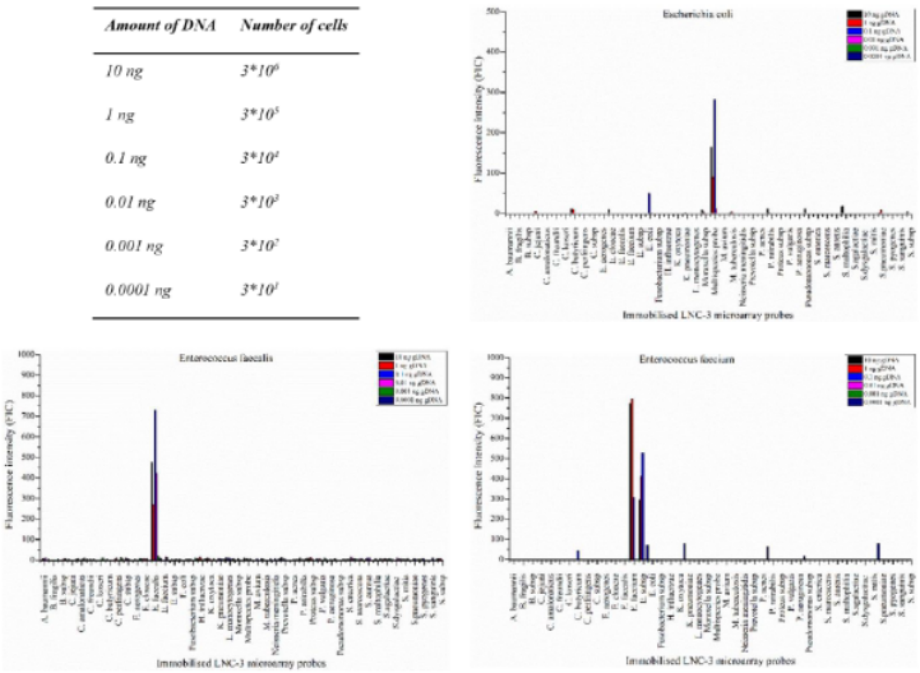


Supplementary Figure 7: Comparison between the conventual microarray technique and the LNC-3 microarray technology for the strains *C. freundii*, *B. fragilis*, *Proteus penneri*, and *Streptococcus pseudopneumoniae*.

7

Sensitivity test of the LNC-3 microarray technology

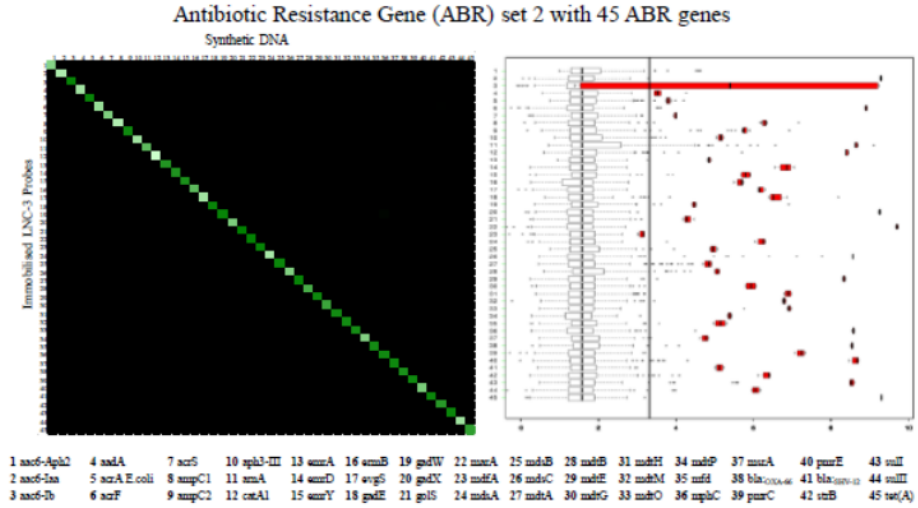
Amount of DNA	*Number of cells*
10 ng	$3*10^6$
1 ng	$3*10^5$
0.1 ng	$3*10^4$
0.01 ng	$3*10^3$
0.001 ng	$3*10^2$
0.0001 ng	$3*10^1$



Supplementary Figure 8: Different amounts of target DNA were applied to the LNC-3-functionalized microarrays. 10 ng DNA correspond to 10^6 bacterial cells. The analysed strains were *E. coli*, *Enterococcus faecalis* and *Enterococcus faecium*.

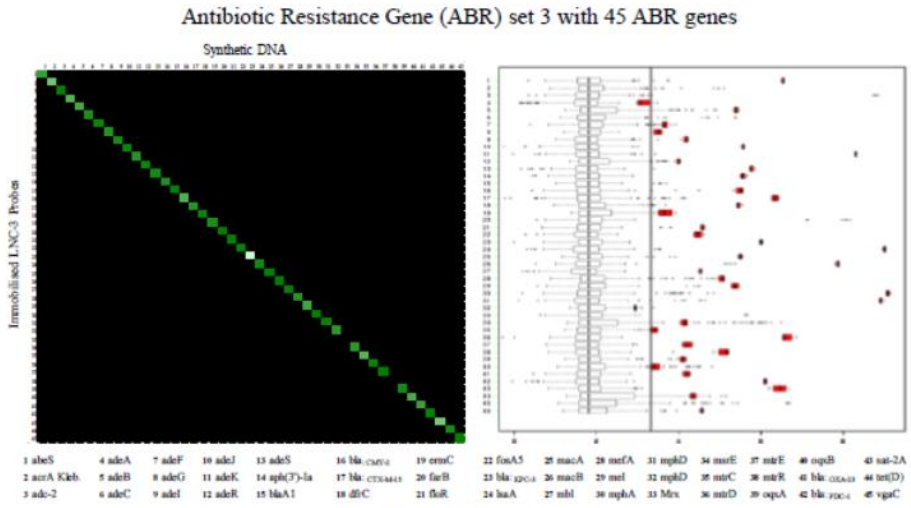
8

Evaluation of the probe performance

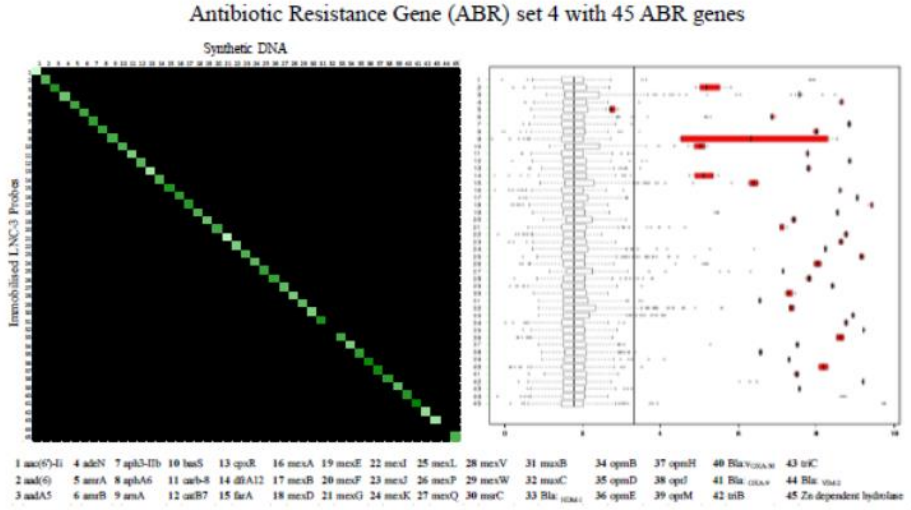


Supplementary Figure 9: **Performance of the LNC-3 probes (ABR gene set 2):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).

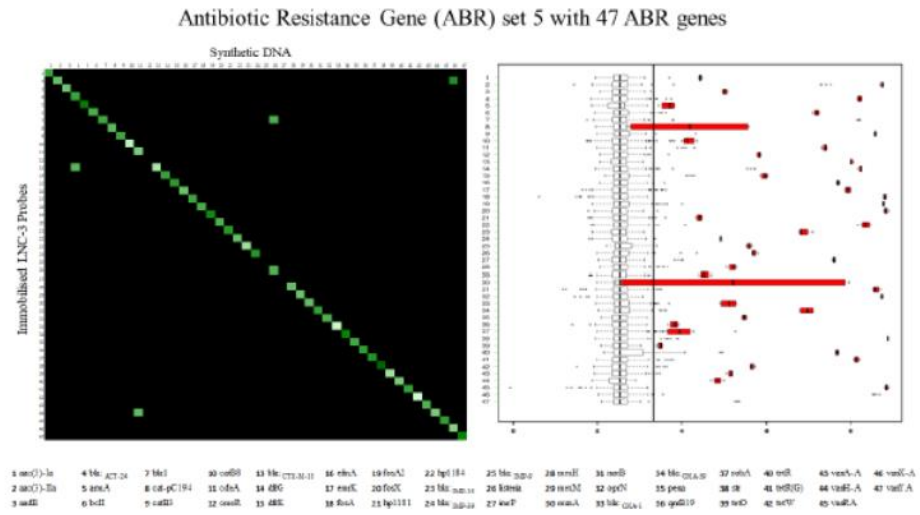
9



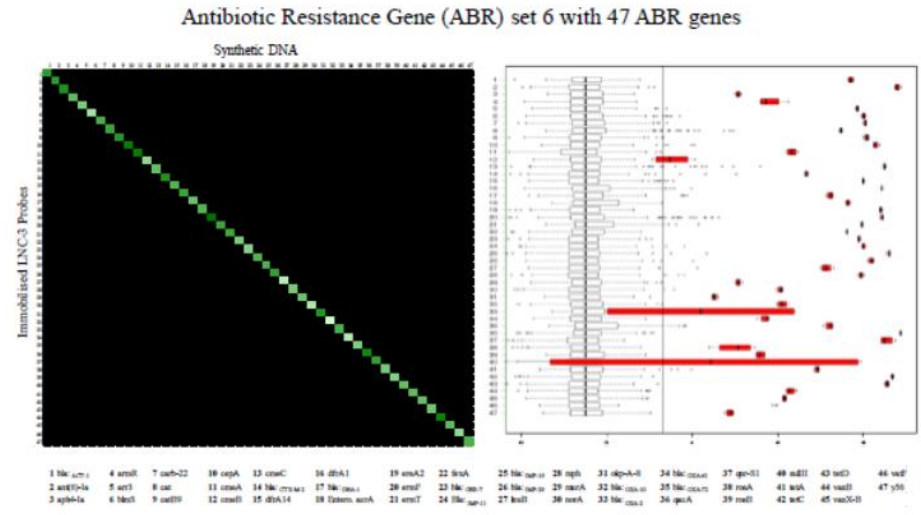
Supplementary Figure 10: **Performance of the LNC-3 probes (ABR gene set 3):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).



Supplementary Figure 11: **Performance of the LNC-3 probes (ABR gene set 4):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).

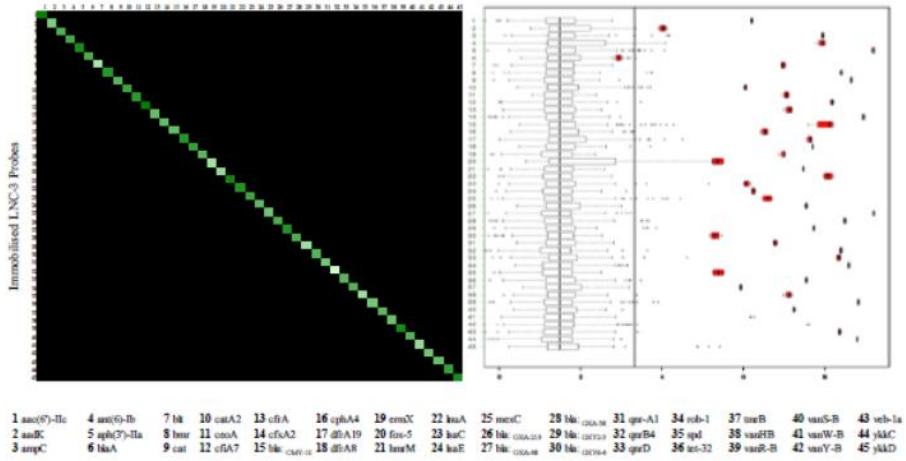


Supplementary Figure 12: **Performance of the LNC-3 probes (ABR gene set 5):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).



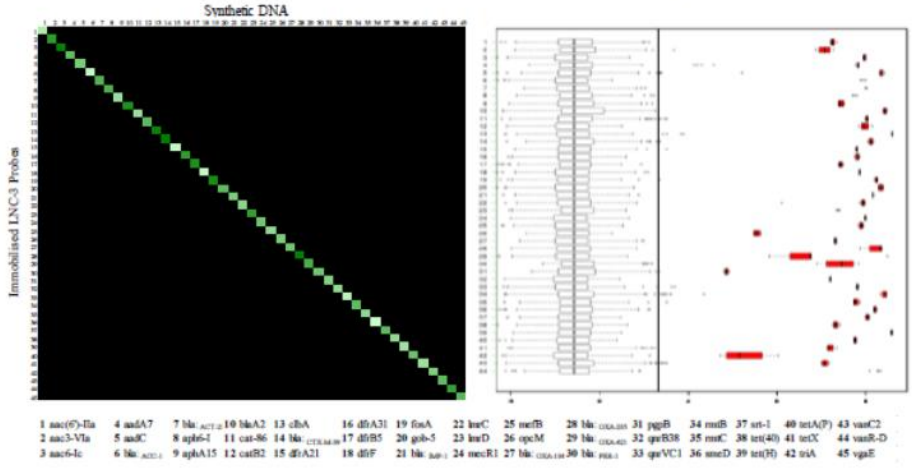
Supplementary Figure 13: **Performance of the LNC-3 probes (ABR gene set 6):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).

Antibiotic Resistance Gene (ABR) set 7 with 45 ABR genes

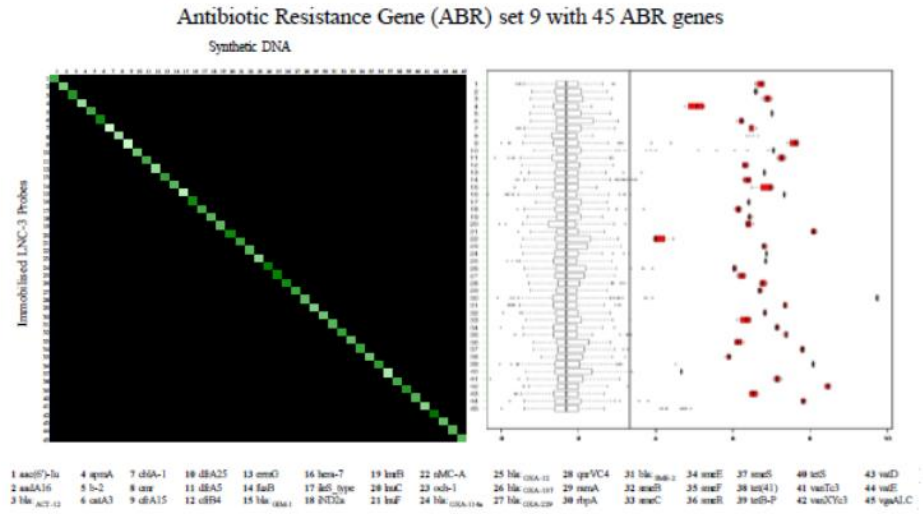


Supplementary Figure 14: **Performance of the LNC-3 probes (ABR gene set 7):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).

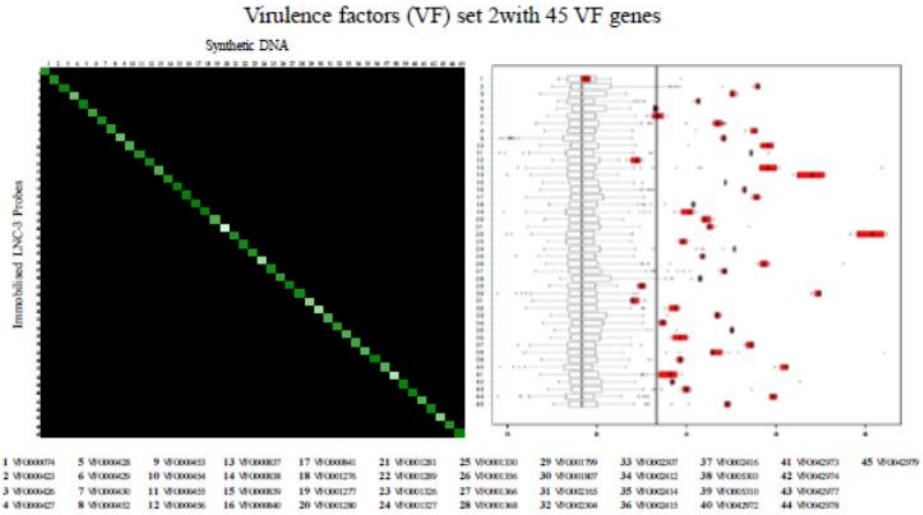
Antibiotic Resistance Gene (ABR) set 8 with 45 ABR genes



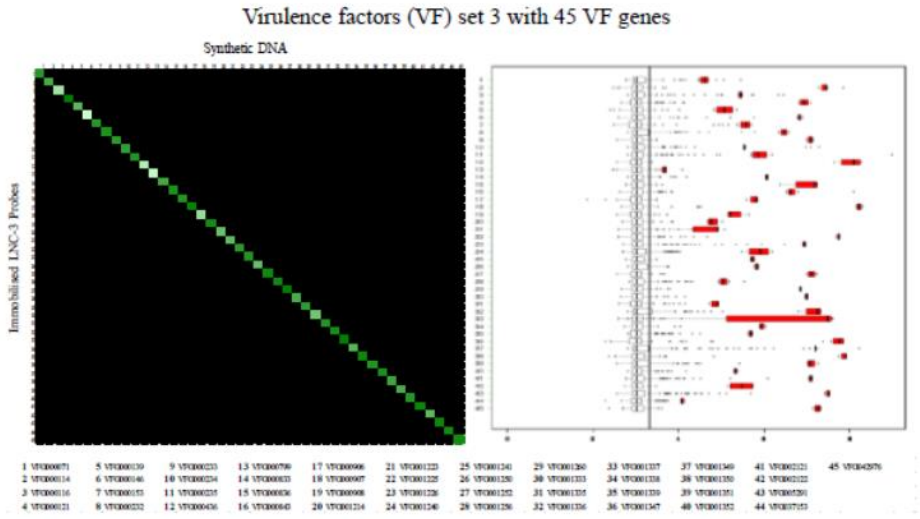
Supplementary Figure 15: **Performance of the LNC-3 probes (ABR gene set 8):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).



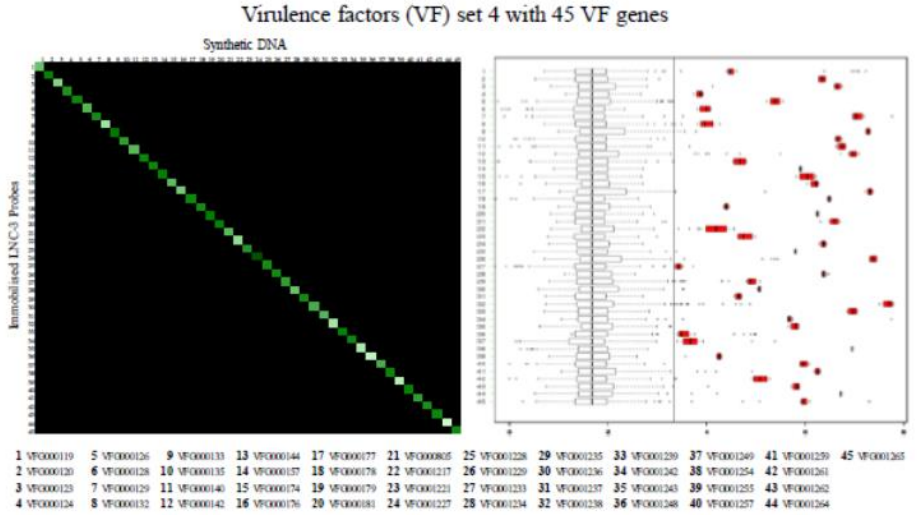
Supplementary Figure 16: **Performance of the LNC-3 probes (ABR gene set 9):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).



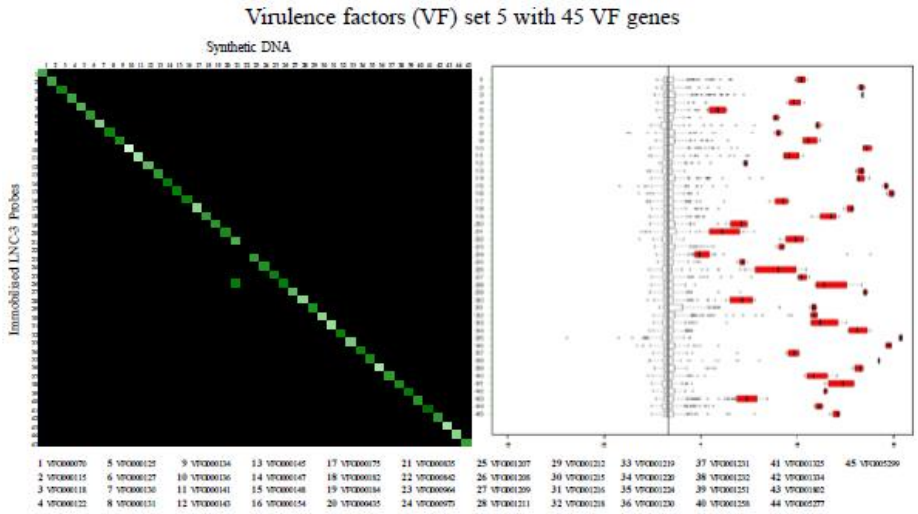
Supplementary Figure 17: **Performance of the LNC-3 probes (VF gene set 2):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).



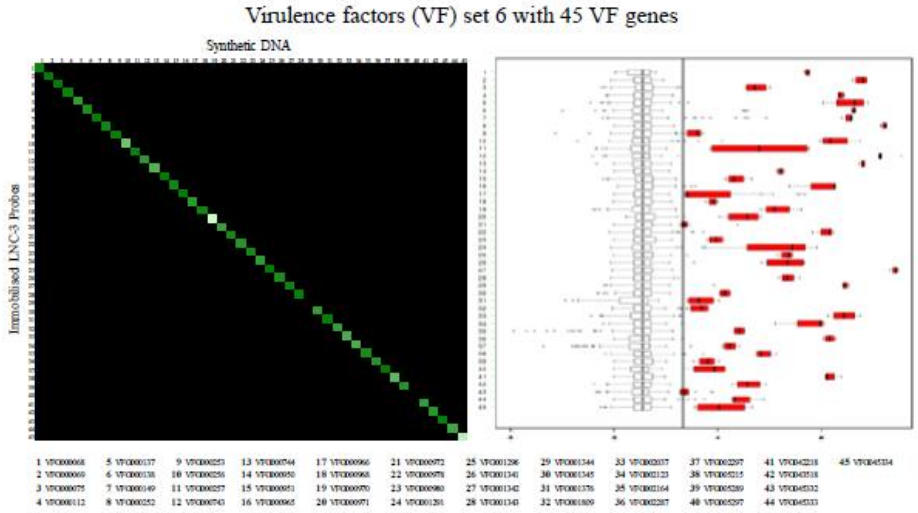
Supplementary Figure 18: **Performance of the LNC-3 probes (VF gene set 3):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).



Supplementary Figure 19: **Performance of the LNC-3 probes (VF gene set 4):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).



Supplementary Figure 20: **Performance of the LNC-3 probes (VF gene set 5):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).



Supplementary Figure 21: **Performance of the LNC-3 probes (VF gene set 6):** On the left, a heatmap displays all measurements line by line. Each displayed value is the median value of four repetitions. On the right, all correct signals (red) were evaluated statistically against the signals of all non-matching probes (grey).

15

Characterisation of bacterial isolates

Supplementary Table 2: Characterisation of bacterial isolates.

Pathogen / gene type	Phylogenetic marker genes	Antibiotic resistance genes		Virulence factor genes	
		Sequenced	LNC-3	Sequenced	LNC-3
Acinetobacter baumannii	*Acinetobacter baumannii*	**aadA**, OXA-66, SulI, Mbl abeS, adeC, adeJ, adeS, adeB, adeA, adeK, adeG, ADC-2, BlaA1, adeF, adeR, adeI, adeN, Zn-dependent hydrolase, **AAC(3)-Ia**, OXA-72, BlaA2	OXA-66, SulI, Mbl abeS, adeC, adeJ, adeS, adeB, adeA, adeK, adeG, ADC-2, BlaA1, adeF, adeR, adeI, adeN, Zn-dependent hydrolase, OXA-72, BlaA2		
Bacteroides fragilis	*Bacteroides fragilis*	tetX, ErmF, cfiA7	tetX, ErmF, cfiA7		
Escherichia coli	*Multispecies probe, Escherichia coli*	acrB, acrD, acre, bacA, baeR, baeS, **CRP**, **emrB**, **emrR**, evgA, H-NS, leuO, mdtD, mdtF, **mdtL**, **mdtN**, msbA, Penicillin_Binding_Protein_Ecoli, PmrF, **tolC**, acrA_Escherichia, acrF, acrS, AmpC2_Ecoli, arnA, emrA, emrD, emrY, evgS, gadE, gadW, **gadX**, marA, mdfA, mdtA, **mdtB**, mdtE, mdtG, mdtH, mdtM, mdtO, mdtP, mfd, PmrC, **PmrE**, emrK	acrB, acrD, acre, bacA, baeR, baeS, , emrB, evgA, H-NS, leuO, mdtD, mdtF, msbA, Penicillin_Binding_Protein_Ecoli, PmrF, , acrA_Escherichia, acrF, acrS, AmpC2_Ecoli, arnA, emrA, emrD, emrY, evgS, gadE, gadW, marA, mdfA, mdtA, mdtE, mdtG, mdtH, mdtM, mdtO, mdtP, mfd, PmrC, emrK	yagZ/ecpA, yagY/ecpB, yagX/ecpC, yagW/ecpD, yagV/ecpE, ykgK/ecpR	yagZ/ecpA, yagY/ecpB, yagX/ecpC, yagW/ecpD, yagV/ecpE, ykgK/ecpR
Enterobacter aerogenes	*Multispecies probe* *Klebsiella pneumoniae* *Enterobacter aerogenes*	TEM-116, Aac6-Ib, StrB, SulII, CTX-M-15, catB3, QnrB19, AAC(3)-IIa, OXA-1, DfrA14, **AmpC**	TEM-116, Aac6-Ib, StrB, SulII, CTX-M-15, catB3, QnrB19, AAC(3)-IIa, OXA-1, DfrA14		
Enterobacter cloacae	*Enterobacter cloacae*	ramA, SulII, robA	ramA, SulII, robA		
Enterococcus faecium	*Enterococcus faecium, Enterococcus subspecies*	AAC(6')-Ii, msrC, efmA	AAC(6')-Ii, msrC, efmA, *DfrG, VanA-A VanH-A, vanRA, vanYA*	acm	acm
Haemophilus influenzae	*Haemophilus influenzae*	hmrM	hmrM	**hmw1B**, rfaD, licD, lic2A, **rfaE**, kdtA, **licA**, **licB**, **hmw1C**	rfaD, licD, lic2A, kdtA

16

Klebsiella pneumoniae	*Klebsiella pneumoniae*	oqxA, **FosA5,** vgaC acrA Klebsiella, oqxB	oqxA, acrA Klebsiella, oqxB	yagZ/ecpA, **east1_astA**	yagZ/ecpA
Prevotella bivia	*Multispecies probe, Prevotella bivia*	CfxA2	CfxA2		
Proteus mirabilis	*Multispecies probe, Proteus mirabilis*	TEM-116, aadA, Aac6-Ib, SulI, CatA1, StrB, SulII, CMY-2, Sat-2A, MsrE, mphD, APH(3')-Ia, DfrA12, TetR, AAC(3)-Ia, armA, DfrA1, **TetA**	TEM-116, aadA, SulI, CatA1, StrB, SulII, CMY-2, Sat-2A, MsrE, mphD, APH(3')-Ia, DfrA12, TetR, AAC(3)-Ia, armA, DfrA1 *CatA2, CatA3, SulII*		
Pseudomonas aeruginosa	*Multispecies probe, Pseudomonas aeruginosa*	**PDC-1**, amrA, amrB **Aph3-IIb**, arnA, basS, **CatB7**, CpxR, **MexA**, MexB, MexD, **MexE**, **MexF**, mexG, mexI, mexJ, mexK, mexL, mexP, mexQ, mexV, mexW, **MuxB**, **MuxC**, OpmB, **opmD**, opmE, OpmH, OprJ, OprM, OXA-50, TriB, TriC, **OprN**, mexH, FosA, MuxA, mexM, **MexC**, **TriA**	amrA, amrB, arnA, basS, CpxR, MexB, MexD, mexG, mexI, mexJ, mexK, mexL, mexP, mexQ, mexV, OpmB, opmE, OpmH, OprJ, OprM, OXA-50, TriB, TriC, mexH, FosA, MuxA, mexM,	xcpA/pilD, algB, algQ, algZ, algU, alg8, alg44, algE, algX, algL, algF, algA, mucC, waaG, waaC, aprA, lasA, lasB, rhlI, lasI, plcH, xcpZ, xcpV, **xcpT,** xcpS, xcpP, xcpQ, pilY2, **pilS**, pilR, pilP, **pilM**, pilT, pilU, pilG, **pilH**, pilI, chpB, chpC, flgD, flgG, flgH, flgI, flgJ, fleQ, fleR, **fliE, fliG**, fliI, fliJ, fliM, fliN, fliP, fliQ, **flhA**	xcpA/pilD, algB, algQ, algZ, algU, alg8, alg44, algE, algX, algL, algF, algA, mucC, waaG, waaC, aprA, lasA, lasB, rhlI, lasI, plcH, xcpZ, xcpV, xcpS, xcpP, xcpQ, pilY2, pilR, pilP, pilT, pilU, pilG, pilI, chpB, chpC, flgD, flgG, flgH, flgI, flgJ, fleQ, fleR, , fliI, fliJ, fliM, fliN, fliP, fliQ, flhB,
Salmonella enterica	*Multispecies probe, Klebsiella pneumoniae* *Salmonella enterica*	sdiA, golS, Aac6-Iaa, mdsA, mdsC, mdsB	sdiA, golS, Aac6-Iaa, mdsA, mdsC, mdsB	fimI, fimC, fimD, **fimF**, csgB, csgA, csgC, csgE, csgF, csgG, sinH, lpfE, lpfD, lpfA, lpfC, lpfB, misL, **ratB**	fimI, fimC, fimD, csgB, csgC, csgF, csgG, sinH, lpfE, lpfD, lpfA, lpfC, lpfB, misL
Staphylococcus aureus	*Staphylococcus aureus*	Aac3-Ik, arlR, arlS, DHA-1, ErmA, FosB, MECA, mepA, mepR, mgrA, norA, sav1866, Spc, Tet-38, Aph3-III, qacA, mecR1	Aac3-Ik, arlR, ErmA, FosB, MECA, mepA, mepR, mgrA, norA, sav1866, Spc, Tet-38, Aph3-III, qacA, mecR1, *ACC-1*	hlgC, hlgB, ebp, sdrC, icaD, icaC, hld, hly/hla, sspC, sspB, hysA, geh, sak, hlb, adsA, scn, sdrD, sdrE, clfA, map, sea, sspA, icaA, fnbA, icaR, icaB, clfB, aur	hlgC, hlgB, ebp, sdrC, icaD, icaC, hld, hly/hla, sspC, sspB, hysA, geh, sak, hlb, adsA, scn, sdrD, sdrE, clfA, map, sea, sspA, icaA, fnbA, icaR, icaB, clfB, aur
Streptococcus pseudopneumoniae	*Streptococcus pseudopneumoniae*	patB, RlmA(II), pmrA, tetM, MefA, mel	RlmA(II), pmrA, MefA, mel	**pce, ply**, lytB, **pavA**, **pfbA, nanA**	lytB

17

4. Discussion

The molecular identification of pathogens and their characterisation regarding antibiotic resistance genes (ABR) and virulence factors (VF) is of great importance. To investigate the spread of infectious diseases and antibiotic resistances, an experimental setting with high precision is required. A variety of diagnostic methods are currently being used, yet most of the data is based on phenotypic observations[112,114]. As mentioned in the introduction (1.1 Antibiotics and antibiotic resistance), pathogens use a variety of mechanisms to protect themselves from antibiotic stress[177,178]. To identify and characterise these mechanisms, a large repertoire of molecular diagnostic technologies are already available and have found their way into everyday clinical live, including polymerase chain reaction (PCR) and real-time (RT-)PCR, matrix-assisted laser desorption/ionisation time of flight mass spectrometry (MALDI-TOF MS), sequencing of the entire genome (WGS), and microarray technology[179].

PCR is used to identify highly conserved bacterial ribosomal RNA genes[116,117]. This method is multitudinous, e.g. conventional PCR, RT-PCR, isothermal PCR, loop-mediated isothermal amplification, or recombinase polymerase amplification (RPA). Conventional and RT-PCR are most common in clinical diagnostics. In RT-PCR, non-specific intercalating fluorescent dyes or specific DNA sequences are used to generate a fluorescent signal after hybridisation with the amplicon, which makes it possible to observe the amplification in real-time[118,121]. PCR detection is highly sensitive and specific[180], but is dependent on a number of aspects, such as the gene copy number or the locus of the gene of interest, e.g. genomic or on a plasmid[120]. Additionally, the gene copy number does not only vary for different genes but can also differ for the same gene between different species, e.g. the ribosomal repeats in different bacterial populations[119,124]. Furthermore, repetitive sequences are often preserved between different species[125,181]; therefore, they have to be chosen wisely to suit the primer/array design[126]. Another limitation of (RT)-PCR is the number of target genes that can be processed at the same time, since (RT)-PCR is based on band patterns on a restricted gel area or colorimetric detection, for which the range of fluorescent dyes whose emission spectra do not overlap critically is limited.

Matrix-assisted laser desorption/ionisation time-of-fly mass spectroscopy (MALDI-TOF MS) is another state-of-the-art diagnostic tool used to identify pathogens[133]. The spectral fingerprint generated by MALDI-TOF MS varies to an extent that it enables the discrimination of genera; however, this only applies as long as they had the same growing conditions[134]. The majority of bacterial molecules observed by MALDI-TOF MS are ribosomal proteins[112]. One of the challenges of this method is the differentiation between taxonomically related species, e.g. pathogenic *Shigella species* to commensal *Escherichia coli*[112]. For instance, mistaking *Streptococcus pneumoniae* and *Streptococcus mitis* can have severe consequences in patient treatment and outcome. This related species result in similar spectra, which makes their differentiation difficult[134,137,182].

Despite the challenges, the use of MALDI-TOF MS contributes well to the patient's outcome rate, which is partly due to the short time required for identification[139,144]. Since the identification with MALDI-TOF MS can also be employed for protein targets, additional characterisation options are possible, especially with regard to antibiotic resistances (ABR), but it still requires further research[142–144,146,147]. Moreover, it has to be considered that not yet expressed antibiotic resistance genes, potentially induced by the antibiotic, cannot be detected[112,134]. Virulence factors (VF) are not necessarily (expressed) proteins, and MALDI-TOF MS is not capable of detecting them in this case. Another diagnostic method is whole genome sequencing (WGS), which might be considered superior to the previously mentioned methods. Their superiority arises from the fact that WGS theoretically reveals the entire genome and therefore the most information content. One drawback of WGS is that it is time-consuming and that a strong bioinformatics expertise is required[183].

The DNA microarray method represents a logical intersection between PCR and WGS[167], combining their advantages and making it possible to screen up to thousands of genes simultaneously.
The DNA microarray platform is capable of analysing gene expression, gene mutations or the absence or presence of specific genes[170]. The conventional DNA microarray is based on a hybridisation reaction between labelled target DNA and immobilised microarray probes. Based on this principle, the challenge of unspecific cross-hybridisations arises, influencing the specificity[163,164,175]. Another challenge of DNA microarray technology is its reproducibility[184–187]. An additional issue is that the DNA microarray platform is partly dependent on an upstream reaction, e.g. PCR, which can also have a negative impact on the DNA microarray results.
To overcome the microarray-associated issues, of which unspecific binding constitutes one of the biggest obstacles, the LNC-3 concept was developed[176,188]. In the publication *Full pathogen characterisation: Species identification including the detection of virulence factors and antibiotic resistance genes via multiplex DNA-assays*, a DNA microarray-based detection tool is developed to identify 45 sepsis-relevant pathogenic strains, 409 ABR genes, and 360 VF genes in parallel. To evaluate our assay, 14 multidrug-resistant strains were tested, including all pathogens recently abbreviated by the acronym ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* ssp.). Preliminary tests were carried out in advance to determine the performance of the LNC-3 DNA microarray platform. This includes (I) *evaluation of every individual probe* in order to rule out false-negative or false-positive signals by inadequate probe design, (II) *comparison with the conventional DNA microarray platform* in terms of specificity, (III) *determination of the limit of detection*, and (IV) *reproducibility* tests.
814 oligonucleotides were designed *in silico*. The primer, LNC probe, and detection oligonucleotide sequences were designed with the Oli2go software[189]. This software was used to avoid critical

interactions during the PCR step as well as during the detection itself. Nevertheless, the probes were tested individually to rule out false-negative or false-positive signals by poor probe design.

For the experimental (I) *evaluation of every individual probe*, synthetic target DNA was utilised, being 100 % complementary, single-stranded, and not dependent on any other factors, such as the upstream PCR amplification; therefore, it represented the simplest case.

In summary, the probes (16S rRNA genes, constituting phylogenetic markers, ABR genes, and VF genes) have shown sufficient sensitivity and specificity. 97.7 % produced satisfactory signals (> tenfold standard deviation), while 2.2 % were lower in response than the tenfold standard deviation. Nevertheless, those signals still contrasted clearly from the background noise.

Next, the (II) *LNC-3 approach was compared to the conventional DNA microarray*. In order to identify pathogens, the products of several 16S rRNA DNA PCRs (*Enterococcus faecium, Klebsiella pneumoniae*, *Streptococcus pseudopneumoniae*, *Proteus penneri*, *Citrobacter freundii*, and *Bacteroides fragilis)* were applied to two different DNA microarray surfaces (hybridisation-based DNA microarray and LNC-3 microarray platform). The identification of the pathogens with the LNC-3 method was superior in comparison to the conventional method. The commonly used technique resulted in signals at loci of related species (false-positive signals), which partly exceeded the correct signal in terms of signal intensity. These signals are presumably generated by cross-hybridisation. Also, random detection signals were observed with the merely hybridisation-based DNA microarray that cannot be explained with phylogenetic proximity (Figure 10). Hybridisation by itself has no high specificity and is even employed to produce mismatches[175], e.g. in primer-based mutagenesis PCR, where it is utilised that 100 % complementarity is not required.

In comparison, the LNC-3 method provided a highly specific identification of the pathogens based on their 16S rRNA DNA. This technique uses hybridisation to generate the first contact between the involved DNA sequences (immobilised probe, detection oligonucleotide, and target DNA), which needs to be ligated in a subsequent step. The proofreading function of the ampligase (thermostable ligase) is capable of recognising polymorphisms of individual nucleotides; therefore, this method exhibited a high specificity[176]. By a washing step at 70 °C, merely hybridised target DNA and detection oligonucleotides are removed to avoid the generation of false-positive signals.

Another feature that defines the LNC-3 method is the distance between the glass surface and the ligation site. Due to this distance, the ligase proofreading capability is further facilitated since liquid-phase conditions are emulated[190].

The LNC-3 concept was shown to be the more precise detection method. It requires additional reaction steps, which suggested that the specificity gain is accompanied by a loss of sensitivity. To exclude this and to define the (III) *limit of detection*, a target DNA dilution series (10 ng – 0.1 pg) of the 16S rRNA DNA of *Staphylococcus aureus* and *Enterococcus faecalis* were applied on the LNC-3

microarray chip to evaluate its sensitivity. The signal intensity caused by 10 pg of target DNA, which corresponds to 10^3 cells, led to a significant signal compared to the background and therefore resulted in a positive identification of *Staphylococcus aureus* as the pathogen of interest. The sensitivity measurements done with *Enterococcus faecalis* indicated a limit of detection of 100 pg, corresponding to 10^4 cells. Anyhow, both detection limits were comparable to current detection methods[166,191]. In general, LNC-3 probes were designed to represent the non-coding strand, so that mRNA can bind as well if not degraded in the meantime.

Another question that arose in terms of this method was the (IV) *reproducibility* of the LNC-3 detection method. The reproducibility is a known problem of microarray technology[184–187]. Three repetitions of an ABR gene set detection of *Enterococcus faecium* were performed. Keeping in mind that the absolute intensities of the individual measurements might vary because of the non-automated in-house slide coating, all repetitions produced correct responses that were significant towards their backgrounds.

After those four prerequisites had been fulfilled, the primary task could be tackled: The full characterisation (identification, resistance behaviour, and virulence) of a large set of pathogens, requiring an unusually high number of target genes. Designed for 45 pathogens, 409 ABR genes, and 360 VF genes, the amplification products of 14 pathogenic species, including the six ESKAPE pathogens, were applied to LNC-3 microarrays. In Table 2, the results are listed in detail. Genes that did not produce significant signal intensities are written in bold letters. Additionally identified ones are highlighted in italics.

Table 2: Characterisation of bacterial isolates.

Pathogen / gene type	**Phylogenetic marker genes**	**Antibiotic resistance genes**		**Virulence factor genes**	
		Sequenced	**LNC-3**	**Sequenced**	**LNC-3**
Acinetobacter baumannii	*Acinetobacter baumannii*	**aadA**, OXA-66, SulI, Mbl abeS, adeC, adeJ, adeS, adeB, adeA, adeK, adeG, ADC-2, BlaA1, adeF, adeR, adeI, adeN, Zn-dependent hydrolase, **AAC(3)-Ia**, OXA-72, BlaA2	OXA-66, SulI, Mbl abeS, adeC, adeJ, adeS, adeB, adeA, adeK, adeG, ADC-2, BlaA1, adeF, adeR, adeI, adeN, Zn-dependent hydrolase, OXA-72, BlaA2		
Bacteroides fragilis	*Bacteroides fragilis*	tetX, ErmF, cfiA7	tetX, ErmF, cfiA7		
Escherichia coli	*Multispecies probe, Escherichia coli*	acrB, acrD, acre, bacA, baeR, baeS, **CRP**, emrB, **emrR**, evgA, H-NS, leuO, mdtD, mdtF, **mdtL, mdtN**, msbA, Penicillin_Binding_Protein_Ecoli, PmrF, **tolC**, acrA_Escherichia, acrF, acrS, AmpC2_Ecoli, arnA, emrA, emrD, emrY, evgS, gadE, gadW, **gadX,** marA, mdfA, mdtA, **mdtB**, mdtE, mdtG, mdtH, mdtM, mdtO, mdtP, mfd, PmrC, **PmrE**, emrK	acrB, acrD, acre, bacA, baeR, baeS, , emrB, evgA, H-NS, leuO, mdtD, mdtF, msbA, Penicillin_Binding_Protein_ Ecoli, PmrF, acrA_Escherichia, acrF, acrS, AmpC2_Ecoli, arnA, emrA, emrD, emrY, evgS, gadE, gadW, marA, mdfA, mdtA, mdtE, mdtG, mdtH, mdtM, mdtO, mdtP, mfd, PmrC, emrK	yagZ/ecpA, yagY/ecpB, yagX/ecpC, yagW/ecpD, yagV/ecpE, ykgK/ecpR	yagZ/ecpA, yagY/ecpB, yagX/ecpC, yagW/ecpD, yagV/ecpE, ykgK/ecpR
Enterobacter aerogenes	*Multispecies probe, Klebsiella pneumoniae, Enterobacter aerogenes*	TEM-116, Aac6-Ib, StrB, SulII, CTX-M-15, catB3, QnrB19, AAC(3)-IIa, OXA-1, DfrA14, **AmpC**	TEM-116, Aac6-Ib, StrB, SulII, CTX-M-15, catB3, QnrB19, AAC(3)-IIa, OXA-1, DfrA14		
Enterobacter cloacae	*Enterobacter cloacae*	ramA, SulII, robA	ramA, SulII, robA		
Enterococcus faecium	*Enterococcus faecium, Enterococcus subspecies*	AAC(6')-Ii, msrC, efmA	AAC(6')-Ii, msrC, efmA, *DfrG, VanA-A VanH-A, vanRA, vanYA*	acm	acm
Haemophilus influenzae	*Haemophilus influenzae*	hmrM	hmrM	**hmw1B**, rfaD, licD, lic2A, **rfaE**, kdtA, **licA**, **licB**, **hmw1C**	rfaD, licD, lic2A, kdtA
Klebsiella pneumoniae	*Klebsiella pneumoniae*	oqxA, **FosA5, vgaC** acrA Klebsiella, oqxB	oqxA, acrA Klebsiella, oqxB	yagZ/ecpA, **east1_astA**	yagZ/ecpA
Prevotella bivia	*Multispecies probe, Prevotella bivia*	CfxA2	CfxA2		

Proteus mirabilis	*Multispecies probe, Proteus mirabilis*	TEM-116, aadA, Aac6-Ib, SulI, CatA1, StrB, SulII, CMY-2, Sat-2A, MsrE, mphD, APH(3')-Ia, DfrA12**,** TetR, AAC(3)-Ia, armA, DfrA1, **TetA**	TEM-116, aadA, SulI, CatA1, StrB, SulII, CMY-2, Sat-2A, MsrE, mphD, APH(3')-Ia, DfrA12, TetR, AAC(3)-Ia, armA, DfrA1 *CatA2, CatA3, SulII*		
Pseudomonas aeruginosa	*Multispecies probe, Pseudomonas aeruginosa*	**PDC-1**, amrA, amrB **Aph3-IIb**, arnA, basS, **CatB7**, CpxR, **MexA**, MexB, MexD, **MexE**, **MexF**, mexG, mexI, mexJ, mexK, mexL, mexP, mexQ, mexV, mexW, **MuxB**, **MuxC**, OpmB, **opmD**, opmE, OpmH, OprJ, OprM, OXA-50, TriB, TriC, **OprN**, mexH, FosA, MuxA, mexM, **MexC**, **TriA**	amrA, amrB, arnA, basS, CpxR, MexB, MexD, mexG, mexI, mexJ, mexK, mexL, mexP, mexQ, mexV, OpmB, opmE, OpmH, OprJ, OprM, OXA-50, TriB, TriC, mexH, FosA, MuxA, mexM,	xcpA/pilD, algB, algQ, algZ, algU, alg8, alg44, algE, algX, algL, algF, algA, mucC, waaG, waaC, aprA, lasA, lasB, rhlI, lasI, plcH, xcpZ, xcpV, **xcpT,** xcpS, xcpP, xcpQ, pilY2, **pilS,** pilR, pilP, **pilM**, pilT, pilU, pilG**, pilH**, pilI, chpB, chpC, flgD, flgG, flgH, flgI, flgJ, fleQ, fleR, **fliE, fliG**, fliI, fliJ, fliM, fliN, fliP, fliQ, flhB, **flhA**	xcpA/pilD, algB, algQ, algZ, algU, alg8, alg44, algE, algX, algL, algF, algA, mucC, waaG, waaC, aprA, lasA, lasB, rhlI, lasI, plcH, xcpZ, xcpV, xcpS, xcpP, xcpQ, pilY2, pilR, pilP, pilT, pilU, pilG, pilI, chpB, chpC, flgD, flgG, flgH, flgI, flgJ, fleQ, fleR, , fliI, fliJ, fliM, fliN, fliP, fliQ, flhB,
Salmonella enterica	*Multispecies probe, Klebsiella pneumoniae* *Salmonella enterica*	sdiA, golS, Aac6-Iaa, mdsA, mdsC, mdsB	sdiA, golS, Aac6-Iaa, mdsA, mdsC, mdsB	fimI, fimC, fimD, **fimF**, csgB, **csgA**, csgC, **csgE**, csgF, csgG, sinH, lpfE, lpfD, lpfA, lpfC, lpfB, misL, **ratB**	fimI, fimC, fimD, csgB, csgC, csgF, csgG, sinH, lpfE, lpfD, lpfA, lpfC, lpfB, misL
Staphylococcus aureus	*Staphylococcus aureus*	Aac3-Ik, arlR, **arlS**, DHA-1, ErmA, FosB, MECA, mepA, mepR, mgrA, norA, sav1866, Spc, Tet-38, Aph3-III, qacA, mecR1	Aac3-Ik, arlR, ErmA, FosB, MECA, mepA, mepR, mgrA, norA, sav1866, Spc, Tet-38, Aph3-III, qacA, mecR1, *ACC-1*	hlgC, hlgB, ebp, sdrC, icaD, icaC, hld, hly/hla, sspC, sspB, hysA, geh, sak, hlb, adsA, scn, sdrD, sdrE, clfA, map, sea, sspA, icaA, fnbA, icaR, icaB, clfB, aur	hlgC, hlgB, ebp, sdrC, icaD, icaC, hld, hly/hla, sspC, sspB, hysA, geh, sak, hlb, adsA, scn, sdrD, sdrE, clfA, map, sea, sspA, icaA, fnbA, icaR, icaB, clfB, aur
Streptococcus pseudopneumoniae	*Streptococcus pseudopneumoniae*	patB, RlmA(II), pmrA, tetM, MefA, mel	RlmA(II), pmrA, MefA, mel	**pce, ply**, lytB, **pavA**, **pfbA, nanA**	lytB

85 % of the ABR genes and 83 % of the VF genes could be detected correctly. Being the vast majority on the one hand, it implied, on the other hand, that there were genes that could not be detected and ones that were detected in addition.
The latter ones might just be false-positives; however, acquired vancomycin resistance in *Enterococcus faecium*, for example, is well documented in literature[192]. Further, their presence on transposable elements, such as gene cassettes, was reported[193,194]. If spread by transposition, the respective genes are flanked by direct repeats, which complicates the contiguous motif (contig) assembly of DNA sections and might entail incorrect *in silico* assembly during WGS sequencing evaluation. So, if resistances were additionally detected that are reported in literature for this particular pathogen, as the vancomycin resistance for *E. faecium*, it might well be that rather the sequencing data were erroneous.

Regarding the genes that could not be detected, a few similarities could be observed. In opposite to ABR genes that encode enzymes to cleave antibiotic substances, which are mostly located on plasmids and are thus usually represented multiply, genes such as efflux pumps might be represented only once in the genome. Further, they might belong to the cell facilities and anyhow constitute resistance mechanisms by regulation only. The majority of non-detected genes belonged to the latter group. A possible reason was their single copy status that might have constituted too little template DNA in the upstream PCR reaction. In combination with lower-performing probes, it may explain why not all of the single copy genes escaped detection. Three different strategies were conceived to circumvent this issue. The most straightforward one would be a two-step PCR. In the first step, target genes classified as critical could be amplified to increase their template share in the following main PCR. Since the polymerase is inhibited by too much DNA, this would be preferable towards increasing the entire template DNA.

Further, most of the non-recognised genes were parts of operons regulated by one promotor and/or the encoded proteins were parts of large protein complexes, such as efflux pumps, pili or filaments, for example the MuxABC-OpmB operon[195,196] or the flagellar proteins *fliE, fliG, fliI, fliJ, fliM,* and *fliN*[197]. In the discussed publication, more examples and case discussions are given. Assuming their presence as functional units, the genetic presence of many genes could be deduced upon the detection of others. That offered the ideas of *(I) combined loci*, addressing the different genes of an operon/a functional protein at the same spot (or using consensus sequences), and *(II) smart chip evaluation*. The first approach would use this clusters to increase the amount of target DNA per spot and therefore reach detectable signal levels; the second one would deduce all genes that must be present if only one of them was detected (and most important the encoded protein/component that mediates antibiotic resistance or virulence). However, not all of them belonged to that group, and unrelated genes as AmpC of *E. aerogenes* need to be detected in any case.

Albeit not perfect, a significant gain of specificity without loss of sensitivity could be observed in comparison to conventional hybridisation-based microarrays, which was attributed to the highly specific ligase reaction, facilitated by a certain distance to the microarray surface. The number of genes screened in parallel was increased to 814, reasonably chosen to receive a bigger picture of infections with regard to the most critical pathogens and their resistance and virulence profile. The Oli2go software turned out to be indispensable to avoid critical interactions in upstream amplification reactions. By two-step PCR, operon-based detection or smart evaluation, the detection of genome-encoded single copy genes must be improved further to receive the full information.
Then, clinical studies are required to test the assay's feasibility in the diagnostic everyday life. An expansion regarding the gene set towards fungi, protozoa, and viruses is conceivable. An intersecting medical field might be the identification of resistance mechanisms in cancer cells that hinders an effective chemotherapy, e.g. human efflux pumps[198], or the detection of known cancer-related gene mutations, e.g. the *BRCA-1* gene[199].

While working on the full identification and characterisation assay, two major drawbacks became clear. One is that plenty of detection oligonucleotides are required to detect the all the genes of interest with the LNC-3 platform. Commercially labelled DNA oligonucleotides, so far used for the signalling, are of high costs. An alternative method was therefore conceived to lower the cost factor of such a high-throughput method. A cost reduction was achieved by replacing the commercial detection oligonucleotides by self-labelled oligonucleotides, which resulted from random elongation with dNTP mixtures containing biotinylated dUTPs using a terminal deoxynucleotidyl transferase (TDT) reaction. Those biotin residues were labelled with Alexa-647-conjugated streptavidin that is, as the biotin-dUTPs, much less expensive than commercial labels[3]. In the publication *Low-cost microarray platform to detect antibiotic resistance genes*, the functionality in terms of specificity and signal intensity of commercially available and self-labelled detection oligonucleotides were compared. The respective probe panel functionality was assessed by applying synthetic target DNA as described for all panels in the publication *Full pathogen characterisation: Species identification including the detection of virulence factors and antibiotic resistance genes via multiplex DNA-assays*. Afterwards, amplification products of 47-plex PCRs of ABR genes using bacterial cell lysates of *A. baumannii, E. cloacae, K. pneumoniae, K. ascorbate, C. freundii* and three different *E. coli* strains as template DNA were detected using both labelling strategies. The results of the different approaches were compared. Again, in order to correctly evaluate the microarray results, the genetic context of the bacterial isolates was analysed by sequencing the entire genome, followed by evaluation using the ResFinder tool[200]. No significant differences were observed regarding the specificity of the self-labelled oligonucleotides. With some DNA sequences, even higher signal intensities could be observed[3,175].

Eventually, both detection systems produced satisfactory signal-to-noise ratios and were of comparable performance. The use of the self-produced detection oligonucleotides led to a tenfold decrease in costs in comparison to the commercial detection oligonucleotides[3].

Further improvements might be achieved regarding the biotin-dUTP to dNTP ratio. A higher amount of biotin-dUTP could increase the sensitivity, as long as length or composition of the detection oligonucleotides do not restrict the overall performance. The in-house labelling was used in following characterisation experiments and is not necessarily restricted to microarray technology.

The other major drawback of detection systems that include an upstream PCR amplification is that they are dependent on the PCR performance. As discussed before, the PCR is dependent on many factors itself, such as the gene copy number, the loci of the genes of interest, etc. A main limiting factor, however, is primer interactions, such as dimer formation, which reduces the efficiency of the PCR and thus the sensitivity of the subsequent detection. This issue increases with the degree of multiplexing.

In the study *Crosslinking of PCR primers reduces unspecific amplification products in multiplex PCR*, covalent crosslinking of primers via their 5'-ends was used to avoid those undesired effects. The crosslinked primer took advantage of steric hindrance because primer dimers cannot be processed in the active site of a conventional DNA polymerase (Figure 13).

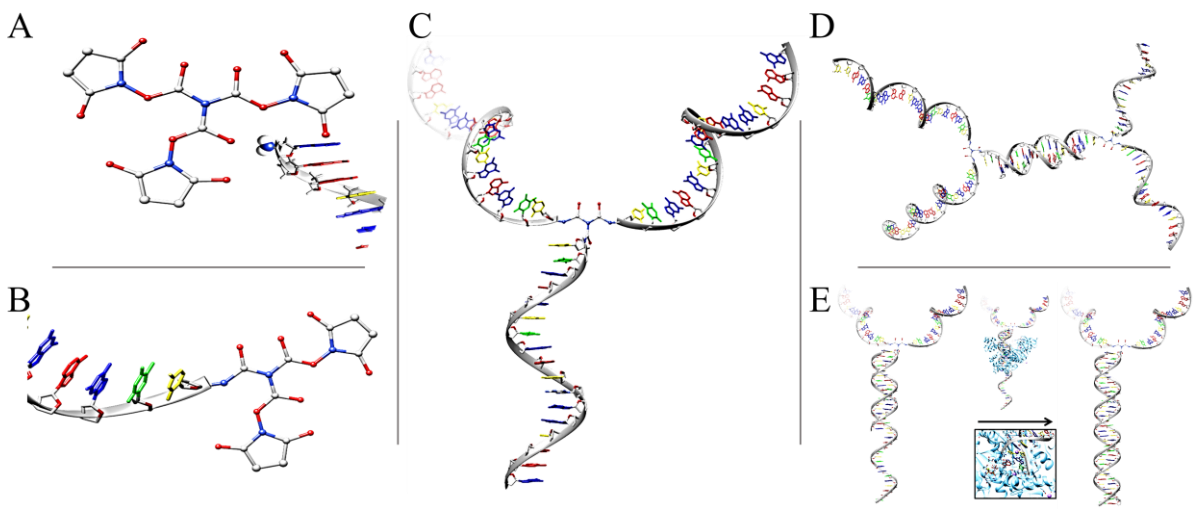


Figure 13: Crosslinked primers and their dimeric structure. (A) Structure of the trifunctional succinimidyl crosslinker and amino-modified primer. (B) Product of the first reaction between crosslinker and primer. (C) Completely reacted crosslinker molecule with three primers attached. (D) Primer dimer. (E) Correct elongation of crosslinked primer by the DNA polymerase.

It was considered not possible for the DNA polymerase to elongate primer dimers if the primers have been crosslinked before because they could not pass through its active site pocket. Besides the mere resource savings, it was assumed that the steric hindrance entails better sequence match requirements and thus an increases in specificity.

A proof-of-principle reaction was conducted using genomic DNA from *Klebsiella pneumoniae* and *Escherichia coli* strains as template DNA to be amplified by crosslinked and non-crosslinked primers, respectively, searching for 11 and 34 clinically important beta-lactamase genes,

respectively. In both multiplex PCRs, a difference in band patterns between crosslinked and non-crosslinked primers was observed. In the 11-plex PCR comprising crosslinked primers, the expected bands could be amplified without any non-specific bands except from known, systematically occurring artefacts related to the screened genes (which, strictly speaking, are hence not unspecific). Further, almost no primer dimer clouds could be noticed. The conventional multiplex PCR did not generate the expected amplification products only but generated additional bands corresponding to non-specific amplification products. Primer clouds were slightly stronger emphasised, suggesting that the formation itself might be complicated if crosslinked.

As with the 11-plex PCR, the 34-plex PCR differed with crosslinked and non-crosslinked primers. The desired amplicons could be generated in both PCR methods, but in contrast to the crosslinked primers, the non-crosslinked primers produced a large number of unintended amplification products, complicating an easy identification of correct bands. In order to examine all possible interactions of the additional components in the PCR more precisely, another PCR was carried out. This PCR contained three different primer states: crosslinked, the amino-modified non-crosslinked primers without crosslinker, and non-crosslinked primers containing Tris-HCl-reacted crosslinker. The four expected bands were observed in the multiplex reaction with crosslinked primers. In contrast, the band pattern of the non-crosslinked primers resulted in several unspecific amplification products. The band pattern of the PCR of the amino-modified non-crosslinked primers without crosslinker and non-crosslinked primers containing Tris-HCl-reacted crosslinker are similar in the band pattern but differ in their intensity. A possible reason could be that the concentration change of Tris-HCl, an ingredient in the polymerase buffer, influenced the amplification performance. All PCR reactions showed slight, but rather negligible primer dimer clouds.

Overall, the multiplex PCR comprising crosslinked primers showed a substantial reduction of unspecific bands and primer dimer clouds. Consequently, the crosslinking of PCR primers resulted in more reliable band patterns to detect antibiotic resistance genes. Due to the steric hindrance effect, the accessibility of the primers to the active centre of the DNA polymerase could be limited, which resulted in an inhibition of the primer dimer elongation and the resulting unspecific products. Statements about the sensitivity were more difficult to deduce, since on the one hand, only in case of the crosslinked primers all intended products could be generated, while on the other hand, some of the ones generated by both attempts were more emphasised with the classical approach. However, by having been able to produce all expected bands and no others with the crosslinked primers, a clear advantage over the conventional primer usage could be evidenced.

The concept of crosslinked primers was considered suitable for other methods that rely on an upstream PCR as well, not only for DNA microarray-based detection[176]. Further steps would be to expand the multiplex degree of the PCR and to evaluate the performance of those reactions regarding their sensitivity. The influence of the individual buffer components, e.g. Tris-HCl, must be understood more in detail as well.

References

1. World Health Organisation. Antibiotic resistance. https://www.who.int/news-room/fact-sheets/detail/antibiotic-resistance.
2. Kraker, M. E. A. de, Stewardson, A. J. & Harbarth, S. Will 10 Million People Die a Year due to Antimicrobial Resistance by 2050? *PLoS medicine* **13**, e1002184 (2016).
3. Wolff, N., Hendling, M., Schönthaler, S., Geiss, A. F. & Barišić, I. Low-cost microarray platform to detect antibiotic resistance genes. *Sensing and Bio-Sensing Research* **23**, 100266 (2019).
4. Wolff, N., Geiss, A. F. & Barišić, I. Crosslinking of PCR primers reduces unspecific amplification products in multiplex PCR. *Journal of microbiological methods*, 106051 (2020).
5. Plettenberg, A. *Infektionskrankheiten der Haut. Grundlagen, Diagnostik und Therapiekonzepte für Dermatologen, Internisten und Pädiater ; 75 Tabellen* (Thieme, Stuttgart [u.a.], 2010).
6. Nikaido, H. Multidrug resistance in bacteria. *Annual review of biochemistry* **78**, 119–146 (2009).
7. Meletis, G. Carbapenem resistance: overview of the problem and future perspectives. *Therapeutic advances in infectious disease* **3**, 15–21 (2016).
8. Lynch, J. P., Clark, N. M. & Zhanel, G. G. Evolution of antimicrobial resistance among Enterobacteriaceae (focus on extended spectrum β-lactamases and carbapenemases). *Expert opinion on pharmacotherapy* **14**, 199–210 (2013).
9. Pendleton, J. N., Gorman, S. P. & Gilmore, B. F. Clinical relevance of the ESKAPE pathogens. *Expert review of anti-infective therapy* **11**, 297–308 (2013).
10. Renner, L. D. *et al.* Detection of ESKAPE Bacterial Pathogens at the Point of Care Using Isothermal DNA-Based Assays in a Portable Degas-Actuated Microfluidic Diagnostic Assay Platform. *Applied and environmental microbiology* **83** (2017).
11. O'Neill J. Tackling a crisis for the health and wealth of nations. London: Review on Antimicrobial Resistance. *Antimicrobial Resistance.* **2017**.
12. Nicolaou, K. C. & Rigol, S. A brief history of antibiotics and select advances in their synthesis. *The Journal of antibiotics* **71**, 153–184 (2018).
13. Wohlleben, W., Mast, Y., Stegmann, E. & Ziemert, N. Antibiotic drug discovery. *Microbial biotechnology* **9**, 541–548 (2016).
14. Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P & T : a peer-reviewed journal for formulary management* **40**, 277–283 (2015).
15. Boucher, H. W. *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **48**, 1–12 (2009).
16. Shapiro, D. J., Hicks, L. A., Pavia, A. T. & Hersh, A. L. Antibiotic prescribing for adults in ambulatory care in the USA, 2007-09. *The Journal of antimicrobial chemotherapy* **69**, 234–240 (2014).

17. Barbosa, T. M. & Levy, S. B. The impact of antibiotic use on resistance development and persistence. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy* **3**, 303–311 (2000).
18. Mohr, K. I. History of Antibiotics Research. *Current topics in microbiology and immunology* **398**, 237–272 (2016).
19. Waksman, S. A. What is an antibiotic or an antibiotic substance? *Mycologia* **39**, 565–569 (1947).
20. Katz, L. & Baltz, R. H. Natural product discovery: past, present, and future. *Journal of industrial microbiology & biotechnology* **43**, 155–176 (2016).
21. Bérdy, J. Bioactive microbial metabolites. *The Journal of antibiotics* **58**, 1–26 (2005).
22. Spížek, J., Sigler, K., Řezanka, T. & Demain, A. Biogenesis of antibiotics-viewing its history and glimpses of the future. *Folia microbiologica* **61**, 347–358 (2016).
23. Tanwar, J., Das, S., Fatima, Z. & Hameed, S. Multidrug resistance: an emerging crisis. *Interdisciplinary perspectives on infectious diseases* **2014**, 541340 (2014).
24. Fuchs, G. *Allgemeine Mikrobiologie* (Georg Thieme Verlag, Stuttgart, 2007).
25. Kahne, D., Leimkuhler, C., Lu, W. & Walsh, C. Glycopeptide and lipoglycopeptide antibiotics. *Chemical reviews* **105**, 425–448 (2005).
26. Kapoor, G., Saigal, S. & Elongavan, A. Action and resistance mechanisms of antibiotics: A guide for clinicians. *Journal of anaesthesiology, clinical pharmacology* **33**, 300–305 (2017).
27. Courvalin, P. Genetics of glycopeptide resistance in gram-positive pathogens. *International journal of medical microbiology : IJMM* **294**, 479–486 (2005).
28. Trimble, M. J., Mlynárčik, P., Kolář, M. & Hancock, R. E. W. Polymyxin: Alternative Mechanisms of Action and Resistance. *Cold Spring Harbor perspectives in medicine* **6** (2016).
29. Tran, T. B. *et al.* Pharmacokinetics/pharmacodynamics of colistin and polymyxin B: are we there yet? *International journal of antimicrobial agents* **48**, 592–597 (2016).
30. Moffatt, J. H., Harper, M. & Boyce, J. D. Mechanisms of Polymyxin Resistance. *Advances in experimental medicine and biology* **1145**, 55–71 (2019).
31. Egorov, A. M., Ulyashova, M. M. & Rubtsova, M. Y. Bacterial Enzymes and Antibiotic Resistance. *Acta naturae* **10**, 33–48 (2018).
32. Campbell, E. A. *et al.* Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell* **104**, 901–912 (2001).
33. Niemi, M., Backman, J. T., Fromm, M. F., Neuvonen, P. J. & Kivistö, K. T. Pharmacokinetic interactions with rifampicin : clinical relevance. *Clinical pharmacokinetics* **42**, 819–850 (2003).
34. Chopra, I. & Roberts, M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiology and molecular biology reviews : MMBR* **65**, 232-60 ; second page, table of contents (2001).
35. Yoneyama, H. & Katsumata, R. Antibiotic resistance in bacteria and its future for novel antibiotic development. *Bioscience, biotechnology, and biochemistry* **70**, 1060–1075 (2006).

36. Vázquez-Laslop, N. & Mankin, A. S. How Macrolide Antibiotics Work. *Trends in biochemical sciences* **43**, 668–684 (2018).
37. Hooper, D. C. Emerging mechanisms of fluoroquinolone resistance. *Emerging infectious diseases* **7**, 337–341 (2001).
38. *Allgemeine Mikrobiologie* (Thieme, [Erscheinungsort nicht ermittelbar], 2012).
39. Bonev, B. B. & Brown, N. M. (eds) *Bacterial resistance to antibiotics. From molecules to man* (Wiley Blackwell, Hoboken, NJ, 2020).
40. Piddock, L. J. V. Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. *Clinical microbiology reviews* **19**, 382–402 (2006).
41. Alvarez-Ortega, C., Olivares, J. & Martínez, J. L. RND multidrug efflux pumps: what are they good for? *Frontiers in microbiology* **4**, 7 (2013).
42. Sun, J., Deng, Z. & Yan, A. Bacterial multidrug efflux pumps: mechanisms, physiology and pharmacological exploitations. *Biochemical and biophysical research communications* **453**, 254–267 (2014).
43. Piddock, L. J. V. Multidrug-resistance efflux pumps - not just for resistance. *Nature reviews. Microbiology* **4**, 629–636 (2006).
44. Webber, M. A., Talukder, A. & Piddock, L. J. V. Contribution of mutation at amino acid 45 of AcrR to acrB expression and ciprofloxacin resistance in clinical and veterinary Escherichia coli isolates. *Antimicrobial agents and chemotherapy* **49**, 4390–4392 (2005).
45. Sjuts, H. *et al.* Molecular basis for inhibition of AcrB multidrug efflux pump by novel and powerful pyranopyridine derivatives. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 3509–3514 (2016).
46. Yamane, K. *et al.* New plasmid-mediated fluoroquinolone efflux pump, QepA, found in an Escherichia coli clinical isolate. *Antimicrobial agents and chemotherapy* **51**, 3354–3360 (2007).
47. Webber, M. A. & Piddock, L. J. V. The importance of efflux pumps in bacterial antibiotic resistance. *The Journal of antimicrobial chemotherapy* **51**, 9–11 (2003).
48. Livermore, D. M. Of Pseudomonas, porins, pumps and carbapenems. *The Journal of antimicrobial chemotherapy* **47**, 247–250 (2001).
49. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and molecular biology reviews : MMBR* **67**, 593–656 (2003).
50. Gardete, S. & Tomasz, A. Mechanisms of vancomycin resistance in Staphylococcus aureus. *The Journal of clinical investigation* **124**, 2836–2840 (2014).
51. Blair, J. M. A., Webber, M. A., Baylay, A. J., Ogbolu, D. O. & Piddock, L. J. V. Molecular mechanisms of antibiotic resistance. *Nature reviews. Microbiology* **13**, 42–51 (2015).
52. Bush, K. & Jacoby, G. A. Updated functional classification of beta-lactamases. *Antimicrobial agents and chemotherapy* **54**, 969–976 (2010).

53. van Hoek, A. H. A. M. *et al.* Acquired antibiotic resistance genes: an overview. *Frontiers in microbiology* **2**, 203 (2011).
54. Paterson, D. L. & Bonomo, R. A. Extended-spectrum beta-lactamases: a clinical update. *Clinical microbiology reviews* **18**, 657–686 (2005).
55. Saravanan, M., Ramachandran, B. & Barabadi, H. The prevalence and drug resistance pattern of extended spectrum β-lactamases (ESBLs) producing Enterobacteriaceae in Africa. *Microbial pathogenesis* **114**, 180–192 (2018).
56. Mingeot-Leclercq, M.-P., Glupczynski, Y. & Tulkens, P. M. Aminoglycosides: Activity and Resistance. *Antimicrobial agents and chemotherapy* **43**, 727–737 (1999).
57. Nakamura, A. *et al.* Macrolide esterase-producing Escherichia coli clinically isolated in Japan. *The Journal of antibiotics* **53**, 516–524 (2000).
58. Wright, G. D. Bacterial resistance to antibiotics: enzymatic degradation and modification. *Advanced drug delivery reviews* **57**, 1451–1470 (2005).
59. Aldred, K. J., Kerns, R. J. & Osheroff, N. Mechanism of quinolone action and resistance. *Biochemistry* **53**, 1565–1574 (2014).
60. Correia, S., Poeta, P., Hébraud, M., Capelo, J. L. & Igrejas, G. Mechanisms of quinolone action and resistance: where do we stand? *Journal of medical microbiology* **66**, 551–559 (2017).
61. Parenti, F. & Lancini, G. *Rifamycins.* In *Antibiotic and Chemotherapy* 326–333 (Elsevier, 2010).
62. Paterson, G. K., Harrison, E. M. & Holmes, M. A. The emergence of mecC methicillin-resistant Staphylococcus aureus. *Trends in microbiology* **22**, 42–47 (2014).
63. Sommer, M. O. A., Munck, C., Toft-Kehler, R. V. & Andersson, D. I. Prediction of antibiotic resistance: time for a new preclinical paradigm? *Nature reviews. Microbiology* **15**, 689–696 (2017).
64. Chen, I. & Dubnau, D. DNA uptake during bacterial transformation. *Nature reviews. Microbiology* **2**, 241–249 (2004).
65. Colomer-Lluch, M., Imamovic, L., Jofre, J. & Muniesa, M. Bacteriophages carrying antibiotic resistance genes in fecal waste from cattle, pigs, and poultry. *Antimicrobial agents and chemotherapy* **55**, 4908–4911 (2011).
66. Salyers, A. A., Shoemaker, N. B., Stevens, A. M. & Li, L. Y. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. *Microbiological reviews* **59**, 579–590 (1995).
67. Mazel, D. Integrons: agents of bacterial evolution. *Nature reviews. Microbiology* **4**, 608–620 (2006).
68. Levin, B. R. & Bull, J. J. Population and evolutionary dynamics of phage therapy. *Nature reviews. Microbiology* **2**, 166–173 (2004).
69. Rios, A. C. *et al.* Alternatives to overcoming bacterial resistances: State-of-the-art. *Microbiological research* **191**, 51–80 (2016).

70. Fleming, A. Fleming, A. (1945). Penicillin's finder assays its future. *The New York Times* **(945** (June 26) https://www.nytimes.com/1945/06/26/archives/.
71. Streptomycin treatment of pulmonary tuberculosis (Reprinted from BMJ 1948:ii769-782). *BMJ* **317**, 1248 (1998).
72. Samson, I. A new class of antimycobacterial drugs: the diarylquinolines. *Thorax* **60**, 495 (2005).
73. Ling, L. L. *et al.* A new antibiotic kills pathogens without detectable resistance. *Nature* **517**, 455–459 (2015).
74. Hover, B. M. *et al.* Culture-independent discovery of the malacidins as calcium-dependent antibiotics with activity against multidrug-resistant Gram-positive pathogens. *Nature microbiology* **3**, 415–422 (2018).
75. Kortright, K. E., Chan, B. K., Koff, J. L. & Turner, P. E. Phage Therapy: A Renewed Approach to Combat Antibiotic-Resistant Bacteria. *Cell host & microbe* **25**, 219–232 (2019).
76. Twort, F. W. An investigation on the nature of ultra-micoscopic viruses. *The Lancet* **186**, 1241–1243 (1915).
77. d'Hérelle F. On an invisible microbe antagonistic to dysentery bacilli . Note by M. F. d'Herelle, presented by M. Roux. Comptes Rendus Academie des Sciences 1917; 165:373–5. *Bacteriophage* **1**, 3–5 (2011).
78. Pires, D. P., Cleto, S., Sillankorva, S., Azeredo, J. & Lu, T. K. Genetically Engineered Phages: a Review of Advances over the Last Decade. *Microbiology and molecular biology reviews : MMBR* **80**, 523–543 (2016).
79. Ghosh, C., Sarkar, P., Issa, R. & Haldar, J. Alternatives to Conventional Antibiotics in the Era of Antimicrobial Resistance. *Trends in microbiology* **27**, 323–338 (2019).
80. Kutter, E. *et al.* Phage therapy in clinical practice: treatment of human infections. *Current pharmaceutical biotechnology* **11**, 69–86 (2010).
81. Yamada, T. *et al.* New bacteriophages that infect the phytopathogen Ralstonia solanacearum. *Microbiology (Reading, England)* **153**, 2630–2639 (2007).
82. Jones, J. B. *et al.* Bacteriophages for plant disease control. *Annual review of phytopathology* **45**, 245–262 (2007).
83. Klumpp, J. & Loessner, M. J. Detection of bacteria with bioluminescent reporter bacteriophage. *Advances in biochemical engineering/biotechnology* **144**, 155–171 (2014).
84. Javed, M. A., Poshtiban, S., Arutyunov, D., Evoy, S. & Szymanski, C. M. Bacteriophage receptor binding protein based assays for the simultaneous detection of Campylobacter jejuni and Campylobacter coli. *PloS one* **8**, e69770 (2013).
85. Hagens, S. & Loessner, M. J. Application of bacteriophages for detection and control of foodborne pathogens. *Applied microbiology and biotechnology* **76**, 513–519 (2007).
86. Loc-Carrillo, C. & Abedon, S. T. Pros and cons of phage therapy. *Bacteriophage* **1**, 111–114 (2011).

87. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nature reviews. Microbiology* **8**, 317–327 (2010).
88. Pedruzzi, I. Inactivation in vitro of the Escherichia coli outer membrane protein FhuA by a phage T5-encoded lipoprotein. *FEMS microbiology letters* **168**, 119–125 (1998).
89. Stummeyer, K. *et al.* Evolution of bacteriophages infecting encapsulated bacteria: lessons from Escherichia coli K1-specific phages. *Molecular microbiology* **60**, 1123–1135 (2006).
90. Krüger, D. H. & Bickle, T. A. Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiological reviews* **47**, 345–360 (1983).
91. Tock, M. R. & Dryden, D. T. F. The biology of restriction and anti-restriction. *Current opinion in microbiology* **8**, 466–472 (2005).
92. Barrangou, R. The roles of CRISPR-Cas systems in adaptive immunity and beyond. *Current opinion in immunology* **32**, 36–41 (2015).
93. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading, England)* **151**, 2551–2561 (2005).
94. Haft, D. H., Selengut, J., Mongodin, E. F. & Nelson, K. E. A Guild of 45 CRISPR-Associated (Cas) Protein Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes. *PLoS Computational Biology* **1** (2005).
95. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
96. Hermoso, J. A., García, J. L. & García, P. Taking aim on bacterial pathogens: from phage therapy to enzybiotics. *Current opinion in microbiology* **10**, 461–472 (2007).
97. Schuch, R. *et al.* Combination therapy with lysin CF-301 and antibiotic is superior to antibiotic alone for treating methicillin-resistant Staphylococcus aureus-induced murine bacteremia. *The Journal of infectious diseases* **209**, 1469–1478 (2014).
98. Witzenrath, M. *et al.* Systemic use of the endolysin Cpl-1 rescues mice with fatal pneumococcal pneumonia. *Critical care medicine* **37**, 642–649 (2009).
99. Pastagia, M. *et al.* A novel chimeric lysin shows superiority to mupirocin for skin decolonization of methicillin-resistant and -sensitive Staphylococcus aureus strains. *Antimicrobial agents and chemotherapy* **55**, 738–744 (2011).
100. Entenza, J. M., Loeffler, J. M., Grandgirard, D., Fischetti, V. A. & Moreillon, P. Therapeutic effects of bacteriophage Cpl-1 lysin against Streptococcus pneumoniae endocarditis in rats. *Antimicrobial agents and chemotherapy* **49**, 4789–4792 (2005).
101. Fischetti, V. A. Bacteriophage lysins as effective antibacterials. *Current opinion in microbiology* **11**, 393–400 (2008).

102. Nelson, D., Loomis, L. & Fischetti, V. A. Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4107–4112 (2001).
103. Babb, R. & Pirofski, L.-A. Help is on the way: Monoclonal antibody therapy for multi-drug resistant bacteria. *Virulence* **8**, 1055–1058 (2017).
104. Oleksiewicz, M. B., Nagy, G. & Nagy, E. Anti-bacterial monoclonal antibodies: back to the future? *Archives of biochemistry and biophysics* **526**, 124–131 (2012).
105. Jolliffe, L. K. Humanized antibodies: enhancing therapeutic utility through antibody engineering. *International reviews of immunology* **10**, 241–250 (1993).
106. Almagro, J. C., Daniels-Wells, T. R., Perez-Tapia, S. M. & Penichet, M. L. Progress and Challenges in the Design and Clinical Development of Antibodies for Cancer Therapy. *Frontiers in immunology* **8**, 1751 (2017).
107. Bebbington, C. & Yarranton, G. Antibodies for the treatment of bacterial infections: current experience and future prospects. *Current opinion in biotechnology* **19**, 613–619 (2008).
108. Kobayashi, S. D. & DeLeo, F. R. Towards a Monoclonal Antibody-Based Therapy for Prevention and Treatment of Staphylococcus aureus Infections. *The Journal of infectious diseases* **219**, 848–850 (2019).
109. Saraiva, F. B., Araújo, A. C. C. de, Araújo, A. É. V. de & Senna, J. P. M. Monoclonal antibody anti-PBP2a protects mice against MRSA (methicillin-resistant Staphylococcus aureus) infections. *PloS one* **14**, e0225752 (2019).
110. Saylor, C., Dadachova, E. & Casadevall, A. Monoclonal antibody-based therapies for microbial diseases. *Vaccine* **27 Suppl 6**, G38-46 (2009).
111. Motley, M. P., Banerjee, K. & Fries, B. C. Monoclonal antibody-based therapies for bacterial infections. *Current opinion in infectious diseases* **32**, 210–216 (2019).
112. Váradi, L. *et al.* Methods for the detection and identification of pathogenic bacteria: past, present, and future. *Chemical Society reviews* **46**, 4818–4832 (2017).
113. Longitude Committee. Longitude Prize. https://longitudeprize.org.
114. Anjum, M. F., Zankari, E. & Hasman, H. Molecular Methods for Detection of Antimicrobial Resistance. *Microbiology spectrum* **5** (2017).
115. Ghannam, M. G. & Varacallo, M. *StatPearls. Biochemistry, Polymerase Chain Reaction (PCR)* (Treasure Island (FL), 2020).
116. Boujelben, I., Gdoura, R. & Hammami, A. A broad-range PCR technique for the diagnosis of infective endocarditis. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]* **49**, 534–543 (2018).

117. Xu, J. *et al.* Employment of broad-range 16S rRNA PCR to detect aetiological agents of infection from clinical specimens in patients with acute meningitis--rapid separation of 16S rRNA PCR amplicons without the need for cloning. *Journal of applied microbiology* **94**, 197–206 (2003).
118. Arya, M. *et al.* Basic principles of real-time quantitative PCR. *Expert review of molecular diagnostics* **5**, 209–219 (2005).
119. Bik, H. M., Fournier, D., Sung, W., Bergeron, R. D. & Thomas, W. K. Intra-genomic variation in the ribosomal repeats of nematodes. *PloS one* **8**, e78230 (2013).
120. Verweij, J. J. & Stensvold, C. R. Molecular testing for clinical diagnosis and epidemiological investigations of intestinal parasitic infections. *Clinical microbiology reviews* **27**, 371–418 (2014).
121. Warren, D. K., Liao, R. S., Merz, L. R., Eveland, M. & Dunne, W. M. Detection of methicillin-resistant Staphylococcus aureus directly from nasal swab specimens by a real-time PCR assay. *Journal of clinical microbiology* **42**, 5578–5581 (2004).
122. Alanio, A. & Bretagne, S. Performance evaluation of multiplex PCR including Aspergillus-not so simple! *Medical mycology* **55**, 56–62 (2017).
123. Dallenne, C., Da Costa, A., Decré, D., Favier, C. & Arlet, G. Development of a set of multiplex PCR assays for the detection of genes encoding important beta-lactamases in Enterobacteriaceae. *The Journal of antimicrobial chemotherapy* **65**, 490–495 (2010).
124. Schaap, M. *et al.* Genome-wide analysis of macrosatellite repeat copy number variation in worldwide populations: evidence for differences and commonalities in size distributions and size restrictions. *BMC genomics* **14**, 143 (2013).
125. O'Connell, E. M. *et al.* Ancylostoma ceylanicum Hookworm in Myanmar Refugees, Thailand, 2012-2015. *Emerging infectious diseases* **24** (2018).
126. Grant, J. R., Pilotte, N. & Williams, S. A. A Case for Using Genomics and a Bioinformatics Pipeline to Develop Sensitive and Species-Specific PCR-Based Diagnostics for Soil-Transmitted Helminths. *Frontiers in genetics* **10**, 883 (2019).
127. Chen, L. *et al.* VFDB: a reference database for bacterial virulence factors. *Nucleic acids research* **33**, D325-8 (2005).
128. McArthur, A. G. *et al.* The comprehensive antibiotic resistance database. *Antimicrobial agents and chemotherapy* **57**, 3348–3357 (2013).
129. Bhanothu, V. & Venkatesan, V. Conventional polymerase chain reaction and amplification refractory mutation system-multi-gene/ multi-primer PCR in the diagnosis of female genital tuberculosis. *Archives of microbiology* **201**, 267–281 (2019).
130. Oviaño, M. & Bou, G. Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for the Rapid Detection of Antimicrobial Resistance Mechanisms and Beyond. *Clinical microbiology reviews* **32** (2019).

131. Rodríguez-Sánchez, B. *et al.* Direct identification of pathogens from positive blood cultures using matrix-assisted laser desorption-ionization time-of-flight mass spectrometry. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **20**, O421-7 (2014).
132. Zhou, M. *et al.* An Improved In-house MALDI-TOF MS Protocol for Direct Cost-Effective Identification of Pathogens from Blood Cultures. *Frontiers in microbiology* **8**, 1824 (2017).
133. Doern, C. D. & Butler-Wu, S. M. Emerging and Future Applications of Matrix-Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) Mass Spectrometry in the Clinical Microbiology Laboratory: A Report of the Association for Molecular Pathology. *The Journal of molecular diagnostics : JMD* **18**, 789–802 (2016).
134. Carbonnelle, E. *et al.* MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory. *Clinical biochemistry* **44**, 104–109 (2011).
135. Croxatto, A., Prod'hom, G. & Greub, G. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS microbiology reviews* **36**, 380–407 (2012).
136. Endo, Y. *et al.* Matrix-assisted laser desorption ionization-time of flight mass spectrometry can be used to identify Helicobacter cinaedi. *Diagnostic microbiology and infectious disease*, 114964 (2019).
137. Dupont, C. *et al.* Identification of clinical coagulase-negative staphylococci, isolated in microbiology laboratories, by matrix-assisted laser desorption/ionization-time of flight mass spectrometry and two automated systems. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **16**, 998–1004 (2010).
138. Jang, K.-S. & Kim, Y. H. Rapid and robust MALDI-TOF MS techniques for microbial identification: a brief overview of their diverse applications. *Journal of microbiology (Seoul, Korea)* **56**, 209–216 (2018).
139. Kumar, A. *et al.* Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Critical care medicine* **34**, 1589–1596 (2006).
140. Cendejas-Bueno, E., Romero-Gómez, M. P. & Mingorance, J. The challenge of molecular diagnosis of bloodstream infections. *World journal of microbiology & biotechnology* **35**, 65 (2019).
141. Welker, M. & van Belkum, A. One System for All: Is Mass Spectrometry a Future Alternative for Conventional Antibiotic Susceptibility Testing? *Frontiers in microbiology* **10**, 2711 (2019).
142. Álvarez-Buylla, A., Picazo, J. J. & Culebras, E. Optimized method for Acinetobacter species carbapenemase detection and identification by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of clinical microbiology* **51**, 1589–1592 (2013).

143. Burckhardt, I. & Zimmermann, S. Using matrix-assisted laser desorption ionization-time of flight mass spectrometry to detect carbapenem resistance within 1 to 2.5 hours. *Journal of clinical microbiology* **49**, 3321–3324 (2011).
144. Luethy, P. M. & Johnson, J. K. The Use of Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS) for the Identification of Pathogens Causing Sepsis. *The journal of applied laboratory medicine* **3**, 675–685 (2019).
145. Josten, M. *et al.* Identification of agr-positive methicillin-resistant Staphylococcus aureus harbouring the class A mec complex by MALDI-TOF mass spectrometry. *International journal of medical microbiology : IJMM* **304**, 1018–1023 (2014).
146. Ehlert, T. *et al.* Establishing PNB-qPCR for quantifying minimal ctDNA concentrations during tumour resection. *Scientific reports* **7**, 8876 (2017).
147. Faron, M. L., Buchan, B. W. & Ledeboer, N. A. Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for Use with Positive Blood Cultures: Methodology, Performance, and Optimization. *Journal of clinical microbiology* **55**, 3328–3338 (2017).
148. Edwards, D. J. & Holt, K. E. Beginner's guide to comparative bacterial genome analysis using next-generation sequence data. *Microbial informatics and experimentation* **3**, 2 (2013).
149. Janda, J. M. & Abbott, S. L. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *Journal of clinical microbiology* **45**, 2761–2764 (2007).
150. Redondo, N., Carroll, A. & McNamara, E. Molecular characterization of Campylobacter causing human clinical infection using whole-genome sequencing: Virulence, antimicrobial resistance and phylogeny in Ireland. *PloS one* **14**, e0219088 (2019).
151. Leite, G. C. *et al.* Genetic and virulence characterization of colistin-resistant and colistin-sensitive A. baumannii clinical isolates. *Diagnostic microbiology and infectious disease* **95**, 99–101 (2019).
152. Hayden, M. K. Detection of nosocomial outbreaks: Genomic surveillance takes the lead. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* (2019).
153. Yarza, P. *et al.* The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Systematic and applied microbiology* **31**, 241–250 (2008).
154. Yarza, P. *et al.* Update of the All-Species Living Tree Project based on 16S and 23S rRNA sequence analyses. *Systematic and applied microbiology* **33**, 291–299 (2010).
155. Hasman, H. *et al.* Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. *Journal of clinical microbiology* **52**, 139–146 (2014).
156. Chang, T. W. Binding of cells to matrixes of distinct antibodies coated on solid surface. *Journal of immunological methods* **65**, 217–223 (1983).

157. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (New York, N.Y.)* **270**, 467–470 (1995).
158. Bañuls, M.-J. *et al.* Improved Performance of DNA Microarray Multiplex Hybridization Using Probes Anchored at Several Points by Thiol-Ene or Thiol-Yne Coupling Chemistry. *Bioconjugate chemistry* **28**, 496–506 (2017).
159. Nimse, S. B., Song, K., Sonawane, M. D., Sayyed, D. R. & Kim, T. Immobilization techniques for microarray: challenges and applications. *Sensors (Basel, Switzerland)* **14**, 22208–22229 (2014).
160. Zhou, X. C., Huang, L. Q. & Li, S. F. Y. Microgravimetric DNA sensor based on quartz crystal microbalance: comparison of oligonucleotide immobilization methods and the application in genetic diagnosis. *Biosensors and Bioelectronics* **16**, 85–95 (2001).
161. Pan, S. & Rothberg, L. Chemical control of electrode functionalization for detection of DNA hybridization by electrochemical impedance spectroscopy. *Langmuir : the ACS journal of surfaces and colloids* **21**, 1022–1027 (2005).
162. Wu, P., Castner, D. G. & Grainger, D. W. Diagnostic devices as biomaterials: a review of nucleic acid and protein microarray surface performance issues. *Journal of biomaterials science. Polymer edition* **19**, 725–753 (2008).
163. Flavell, A. J. *et al.* A microarray-based high throughput molecular marker genotyping method: the tagged microarray marker (TAM) approach. *Nucleic acids research* **31**, e115 (2003).
164. Baudart, J. *et al.* Microarray (phylochip) analysis of freshwater pathogens at several sites along the Northern German coast transecting both estuarine and freshwaters. *Applied microbiology and biotechnology* **101**, 871–886 (2017).
165. Bumgarner, R. Overview of DNA microarrays: types, applications, and their future. *Current protocols in molecular biology* **Chapter 22**, Unit 22.1 (2013).
166. Call, D. R. Challenges and opportunities for pathogen detection using DNA microarrays. *Critical reviews in microbiology* **31**, 91–99 (2005).
167. Levy, S. E. & Myers, R. M. Advancements in Next-Generation Sequencing. *Annual review of genomics and human genetics* **17**, 95–115 (2016).
168. Alcock, B. P. *et al.* CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic acids research* **48**, D517-D525 (2020).
169. Barrett, T. & Edgar, R. Mining microarray data at NCBI's Gene Expression Omnibus (GEO)*. *Methods in molecular biology (Clifton, N.J.)* **338**, 175–190 (2006).
170. Taitt, C. R. *et al.* A comparison of methods for DNA preparation prior to microarray analysis. *Analytical biochemistry* **585**, 113405 (2019).
171. Carbonnelle, E. *et al.* Rapid identification of Staphylococci isolated in clinical microbiology laboratories by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of clinical microbiology* **45**, 2156–2161 (2007).

172. Blaschke, A. J. *et al.* Rapid identification of pathogens from positive blood cultures by multiplex polymerase chain reaction using the FilmArray system. *Diagnostic microbiology and infectious disease* **74**, 349–355 (2012).
173. Yoo, S. M. *et al.* High-throughput identification of clinically important bacterial pathogens using DNA microarray. *Molecular and Cellular Probes* **23**, 171–177 (2009) http://www.sciencedirect.com/science/article/pii/S0890850809000279.
174. Wu, L. *et al.* Development and evaluation of functional gene arrays for detection of selected genes in the environment. *Applied and environmental microbiology* **67**, 5780–5790 (2001).
175. Koltai, H. & Weingarten-Baror, C. Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. *Nucleic acids research* **36**, 2395–2405 (2008).
176. Barišić, I., Kamleithner, V., Schönthaler, S. & Wiesinger-Mayr, H. Fast and highly specific DNA-based multiplex detection on a solid support. *Applied microbiology and biotechnology* **99**, 413–423 (2015).
177. Hollenbeck, B. L. & Rice, L. B. Intrinsic and acquired resistance mechanisms in enterococcus. *Virulence* **3**, 421–433 (2012).
178. Cox, G. & Wright, G. D. Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *International journal of medical microbiology : IJMM* **303**, 287–292 (2013).
179. Chan, K.-G. Whole-genome sequencing in the prediction of antimicrobial resistance. *Expert review of anti-infective therapy* **14**, 617–619 (2016).
180. Fernández-Romero, N. *et al.* Use of multiplex PCR in diagnosis of bloodstream infections in kidney patients. *Diagnostic microbiology and infectious disease* **80**, 93–96 (2014).
181. Pilotte, N., Zaky, W. I., Abrams, B. P., Chadee, D. D. & Williams, S. A. A Novel Xenomonitoring Technique Using Mosquito Excreta/Feces for the Detection of Filarial Parasites and Malaria. *PLoS neglected tropical diseases* **10**, e0004641 (2016).
182. Dubois, D. *et al.* Identification of a Variety of Staphylococcus Species by Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry ▿ . *Journal of clinical microbiology* **48**, 941–945 (2010).
183. Kwong, J. C., McCallum, N., Sintchenko, V. & Howden, B. P. Whole genome sequencing in clinical and public health microbiology. *Pathology* **47**, 199–210 (2015).
184. Schaupp, C. J., Jiang, G., Myers, T. G. & Wilson, M. A. Active mixing during hybridization improves the accuracy and reproducibility of microarray results. *BioTechniques* **38**, 117–119 (2005).
185. Draghici, S., Khatri, P., Eklund, A. C. & Szallasi, Z. Reliability and reproducibility issues in DNA microarray measurements. *Trends in genetics : TIG* **22**, 101–109 (2006).
186. Larkin, J. E., Frank, B. C., Gavras, H., Sultana, R. & Quackenbush, J. Independence and reproducibility across microarray platforms. *Nature methods* **2**, 337–344 (2005).

187. Chen, J. J., Hsueh, H.-M., Delongchamp, R. R., Lin, C.-J. & Tsai, C.-A. Reproducibility of microarray data: a further analysis of microarray quality control (MAQC) data. *BMC Bioinformatics* **8**, 412 (2007).
188. Wolff, N. & Barišic′, I. Highly Specific Ligation-dependent Microarray Detection of Single Nucleotide Polymorphisms. *Methods in molecular biology (Clifton, N.J.)* **1616**, 231–240 (2017).
189. Hendling, M. *et al.* Oli2go: an automated multiplex oligonucleotide design tool. *Nucleic acids research* **46**, W252-W256 (2018).
190. Barišić, I. *et al.* Multiplex detection of antibiotic resistance genes using padlock probes. *Diagnostic microbiology and infectious disease* **77**, 118–125 (2013).
191. Hu, J. *et al.* Colorimetric-Fluorescent-Magnetic Nanosphere-Based Multimodal Assay Platform for Salmonella Detection. *Analytical chemistry* **91**, 1178–1184 (2019).
192. McKenney, P. T., Ling, L., Wang, G., Mane, S. & Pamer, E. G. Complete Genome Sequence of Enterococcus faecium ATCC 700221. *Genome announcements* **4** (2016).
193. Lee, T., Pang, S., Abraham, S. & Coombs, G. W. Antimicrobial-resistant CC17 Enterococcus faecium: The past, the present and the future. *Journal of global antimicrobial resistance* **16**, 36–47 (2019).
194. Teuber, M. Veterinary use and antibiotic resistance. *Current opinion in microbiology* **4**, 493–499 (2001).
195. Mima, T. *et al.* Gene cloning and characteristics of the RND-type multidrug efflux pump MuxABC-OpmB possessing two RND components in Pseudomonas aeruginosa. *Microbiology (Reading, England)* **155**, 3509–3517 (2009).
196. Yang, L., Chen, L., Shen, L., Surette, M. & Duan, K. Inactivation of MuxABC-OpmB transporter system in *Pseudomonas aeruginosa* leads to increased ampicillin and carbenicillin resistance and decreased virulence. *Journal of microbiology (Seoul, Korea)* **49**, 107–114 (2011).
197. Sidhu-Muñoz, R. S., Tejedor, C. & Vizcaíno, N. The Three Flagellar Loci of *Brucella ovis* PA Are Dispensable for Virulence in Cellular Models and Mice. Frontiers in Veterinary Science, 7 (2020).
198. Amaral, L. *et al.* Inhibitors of bacterial efflux pumps that also inhibit efflux pumps of cancer cells. *Anticancer research* **32**, 2947–2957 (2012).
199. Boulton, S. J. Cellular functions of the BRCA tumour-suppressor proteins. *Biochemical Society transactions* **34**, 633–645 (2006).
200. Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *The Journal of antimicrobial chemotherapy* **67**, 2640–2644 (2012).

Appendix

Highly Specific Ligation-dependent Microarray Detection of Single Nucleotide Polymorphisms

Noa Wolff, Ivan Barišić

Published: Diagnostic Bacteriology (2017) 231-240, part of the Methods in Molecular Biology book series. https://doi.org/10.1007/978-1-4939-7037-7_15

Contribution to the publication:

- Manuscript drafting

Chapter 15

Highly Specific Ligation-dependent Microarray Detection of Single Nucleotide Polymorphisms

Noa Wolff and Ivan Barišić

Abstract

The fast detection and characterization of pathogens are essential for an efficient treatment of infectious diseases. However, the development of improved and reliable diagnostic methods is still an ongoing process because not only pathogens but also their antibiotic resistances have to be identified. The gold standard today is, however, a cultivation-based characterization approach, which takes days until results can be evaluated. In patients with, for example, severe sepsis, the diagnostic test duration is a very critical parameter because a delay of treatment optimization increases the mortality rate significantly. In contrast, DNA-based molecular techniques can obtain results within a few hours. A further challenge in diagnostic laboratories is that patient samples have to be screened for hundreds of potential pathogens, antibiotic resistance genes, and virulence factors, which is achieved by using a number of specialized tests at the moment. Microarrays are outstandingly good for the simultaneous analysis of thousands of different genes and have become a popular tool in biological studies. Nevertheless, further optimizations of the microarray technology are required due to the obligatory DNA labeling and/or amplification steps and the effects of nonspecific DNA hybridization. Here, we describe a fast and highly specific solid-support-based DNA characterization method for pathogens and antibiotic resistance genes.

Key words Molecular biology, Gene characterization, SNP, DNA microarray, Multiplex detection, Solid support-based detection

1 Introduction

The DNA microarray has emerged as a powerful tool in functional genome analysis and clinical diagnostics. It provides detailed information on the gene expression of an organism and allows the simultaneous detection of up to 100,000 target genes. The outstanding features of this high-throughput research technology have resulted in a wide range of applications and, as in our case, in the development of a species characterization microarray based on phylogenetic marker and antibiotic resistance genes including SNPs [1]. One of the limiting factors of this technique are the non-specific DNA hybridization events causing false-positive results [2, 3]. Various approaches to improve the sensitivity and

Kimberly A. Bishop-Lilly (ed.), *Diagnostic Bacteriology: Methods and Protocols*, Methods in Molecular Biology, vol. 1616, DOI 10.1007/978-1-4939-7037-7_15, © Springer Science+Business Media LLC 2017

specificity of the microarray technology by using different spacers or surface chemistries, for example, were limited [4]. However, several experimental parameters could be identified that have an impact on the performance of microarrays. Specifically, the sensitivity and the specificity correlate with the length of the oligonucleotide probes [5]. Additionally, it was demonstrated by Chou et al. that the probe sensitivity is related to the length and the accessibility of the probe. Further effort to increase the specificity of the microarray technology resulted in the integration of a variety of enzymatic steps [6, 7].

Here, we demonstrate an improved multiplex detection and characterization method based on covalently immobilized DNA oligonucleotides combined with a ligation step. What is exceptional about our method is the short duration time and the specificity that is achieved by a novel type of immobilized probes, the linear chain (LNC) probes. This probe concept has already been used to identify clinically relevant pathogens [8]. The LNC probe is a thiol-modified detection probe that comprises three DNA oligonucleotides, LNC-A, LNC-B, and LNC-C (Fig. 1). The three oligonucleotides are connected via hydrogen bridge bonds. The GC-rich hybridization regions that link the three probes together have melting temperatures above 85 °C to facilitate a high LNC probe stability. Only LNC-C has a specific detection sequence for

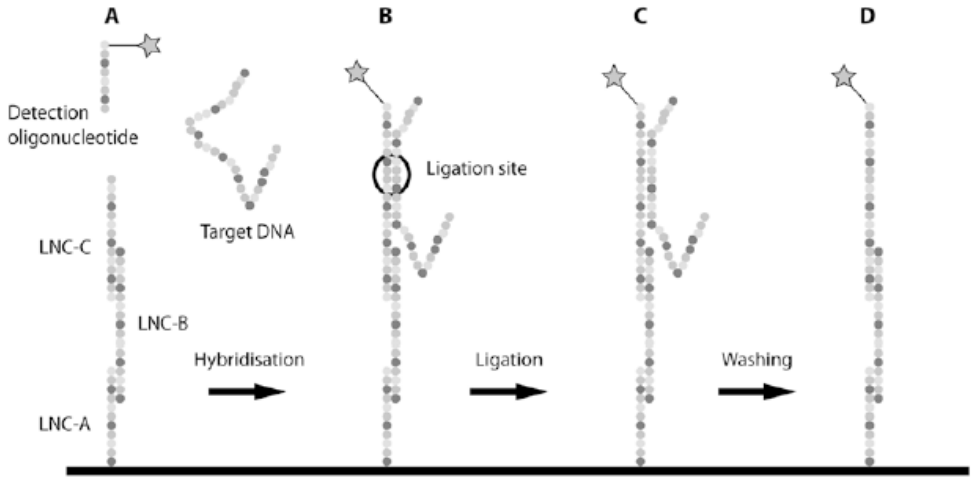


Fig. 1 Schematic illustration of the LNC probe and the reaction mechanism. (**a**) The LNC-A probe is linked covalently to a glass surface. LNC-A, B, and C are incubated together prior to spotting and immobilized in the hybridized state. A ligation mixture comprising fluorescently labeled detection oligonucleotides and the target DNA is applied to the slide. (**b**) The detection oligonucleotide binds in the presence of a target DNA to the LNC probe. (**c**) The ligation takes place if the target DNA is perfectly matching the LNC and the detection probe. (**d**) After the stringent washing step, only the ligated detection oligonucleotides remain on the surface while the non-ligated probes are washed away

the gene of interest. The LNC probes are immobilized to glass slides using a microarray spotter. In this protocol, we used bacterial pathogens and antibiotic resistance genes to illustrate the specificity of this method.

2 Materials

All buffers were prepared with deionized and micro-filtered water. The preparation and storage took place at room temperature. First, all glass trays and hybridization chambers were cleaned using DNA-Exitus (AppliChem, Germany) in order to be DNA free (*see* **Note 1**). The glass-slides were cleaned separately. The oligonucleotides were designed as follows:

LNC-A was modified at its 5′-end with a thiol group. The 3′-end of the detection oligonucleotides was modified with a Cy3-fluorophore. In addition, the 5′-end of the detection oligonucleotide has to be phosphorylated to ligate the oligonucleotide to the LNC-C probe. All oligonucleotides were purified by HPLC. The melting temperature for the target recognition region of the detection oligonucleotide and the LNC-C probe was about 50 °C. The sequences of the LNC oligonucleotides are shown in Table 1. The oligonucleotides were purchased from Integrated DNA Technologies (IDT, IA, USA).

2.1 Washing Buffer

1. For a 1 M HCl solution, add 83 ml 37% HCl in 600 ml $_{dd}H_2O$ and subsequently water to a total volume of 1 L (*see* **Note 2**).
2. For a 1 M NaOH solution, use 40 g NaOH and resolve them in 1 L water (*see* **Note 2**).

2.2 Silanization of Glass Slide

1. (3-Aminopropyl)trimethoxysilane ATS (Sigma-Aldrich, MO, USA): 0.5% solution in dry acetone (*see* **Note 3**). ATS forms an aminopropyl derivate on glass and is used as a first surface modifier.
2. Prepare PBS (0.1 M NaH_2PO_4, 0.15 M NaCl, pH 7.2).

Table 1
Structure and sequence of the LNC probes

Probe name	5′-Mod	Sequence 5′–3′	Length (bp)
LNC A	Thiol	TTTCGCTGCCGACCCTGCGCCGTGGCC	27
LNC B		CCCCGGCACGCGAGCCCACGCTGCTTTTTTGGCCAC GGCGCAGGGTCGGCAGCG	54
LNC C		GCAGCGTGGGCTCGCGTGCCGGGGTTTTTTNNNNN NNNNNNNNN	45

The sequence region with the multiple Ns indicates a variable region specific for the gene of interest. Ideally, it is approximately 15 base pairs (bp) long and has a melting temperature between 45 and 55 °C

3. Water-soluble heterobifunctional cross-linker: sulfonated analogs of *m*-maleimidobenzoyl-*N*-hydroxysulfosuccinimide ester (s-MBS) were purchased from ProteoChem (IL, USA) and prepared as 2 mM solution in PBS (*see* **Note 3**).

2.3 Spotting

1. 1× NaPi buffer: 0.1 M NaH_2PO_4, 0.15 M NaCl, pH 6.5 (*see* **Note 4**).
2. The oligonucleotides LNC-A, B, and C were pooled in a final concentration of 5 µM each in sterile filtered 1× NaPi buffer.
3. The LNC probes were spotted using the OmniGrid contact arrayer (GeneMachines, CA, USA) and SMP 3 pins (TeleChem, CA, USA).
4. β-Mercaptoethanol (Sigma-Aldrich): 10 mM β-mercaptoethanol in 1× NaPi-Buffer (*see* **Note 3**).
5. Saline buffer: 1.5 M NaCl, 10 mM NaH_2PO_4, pH 7.
6. 20× Saline sodium citrate (SSC) buffer, pH 7 (Invitrogen, CA, USA).
7. 5× SSC buffer comprising 0.1% Tween-20, pH 7.
8. 5× SSC buffer, pH 7.
9. Heracus Megafuge 1.0 (Thermo Fisher Scientific) for slide centrifugation.

2.4 Ligation and Detection

1. First, prepare a stock mixture comprising all detection oligonucleotides with a final concentration of 100 µM. The end concentration of each oligonucleotide depends on the amount of the total number of oligonucleotides (*see* **Note 5**). Make sure that the 5′-ends of the detection oligonucleotides are phosphorylated. Oligonucleotides can be ordered either comprising the 5′-phosporylation or the modification can be introduced by the operator using, e.g., T4 polynucleotide kinase (Thermo Fisher Scientific) that is more economic in high-multiplex assays.
2. Prepare a master mixture containing the detection oligonucleotides with a final concentration of 300 nM for each detection oligonucleotide (*see* **Note 6**).
3. The ligation reaction was conducted in a frame-seal incubation chamber with a 25 µl capacity (Bio-Rad, CA, USA).
4. Bovine serum albumin (BSA, New England Biology, MA, USA): 2 µg/µl in $_{dd}H_2O$.
5. Ampligase buffer: 20 mmol/L Tris–HCl, 25 mmol/L KCl, 10 mmol/L $MgCl_2$, 0.5 mmol/L nicotinamide adenine dinucleotide (NAD) and 0.01% Triton® X-100, pH 8.3 (Epicentre, WI, USA).
6. 20% sodium dodecyl sulfate (SDS): 20 g SDS dissolved in 100 ml $_{dd}H_2O$ (*see* **Note 7**).

7. 2× SSC with 0.1% SDS washing buffer: 500 ml 20× SSC stock solution and 25 ml SDS 20% in 5 L $_{dd}H_2O$.
8. 0.2× SSC wash buffer: 50 ml 20× SSC stock solution in 5 L $_{dd}H_2O$.
9. The MJ Research PTC-200 Peltier Thermal cycler (Bio-Rad) was used for thermal incubation of the glass slides.
10. Slides were scanned using the Tecan PowerScanner (Tecan, Switzerland). Be aware that different microarray scanners have different limits of detection. Some scanners are not able to detect weak fluorescence signals. For analyzing the data, we used GenePix Pro 6.0 (Molecular Devices, CA, USA) and Excel 2007 (Microsoft, WA, USA).

3 Methods

Perform all steps at room temperature unless otherwise specified.

3.1 Cleaning of Glass Slides

1. Clean glass slides with $_{dd}H_2O$ followed by 100% EtOH (*see* **Note 8**).
2. Sonicate the slides for 10 min in acetone; subsequently wash twice with $_{dd}H_2O$ (*see* **Note 8**).
3. Afterward, sonicate the slides in 1 M NaOH for another 10 min and immerse in 1 M HCl overnight.
4. On the following day, wash the slides twice for 5 min in $_{dd}H_2O$; subsequently rinse with 100% EtOH and allow to air dry.

3.2 Silanization of Glass Slides

1. Immerse the dried and cleaned slides for 1 h in a 0.5% ATS solution in dry acetone (*see* **Notes 3** and **9**).
2. Afterward, wash the slides three times for 5 min with acetone and rinse them with 100% EtOH.
3. Subsequently, bake the slides for 50 min at 90 °C.
4. Wet the surface of the slides with 300 µl of the s-MBS solution in PBS buffer with a pH 7.4 for 5 h in a humid environment (*see* **Note 10**). The moist atmosphere can be obtained from an incubation chamber which is filled with water (*see* **Note 11**). This will counteract evaporation of your s-MBS solution.
5. While the cross-linker is incubated (*see* Fig. 2 for a schematic of the cross-linking reaction), the pipetting plate should be prepared for spotting, comprising NaPi buffer, the 5′-end thiol-modified oligonucleotide LNC-A, LNC-B and the target-specific LNC-C oligonucleotides (*see* **Note 12**). Dilute the LNC probes to a concentration of 5 µM each in 0.5× NaPi buffer. In addition, add two spotting controls to monitor the spotting efficiency and the LNC probe stability. The first

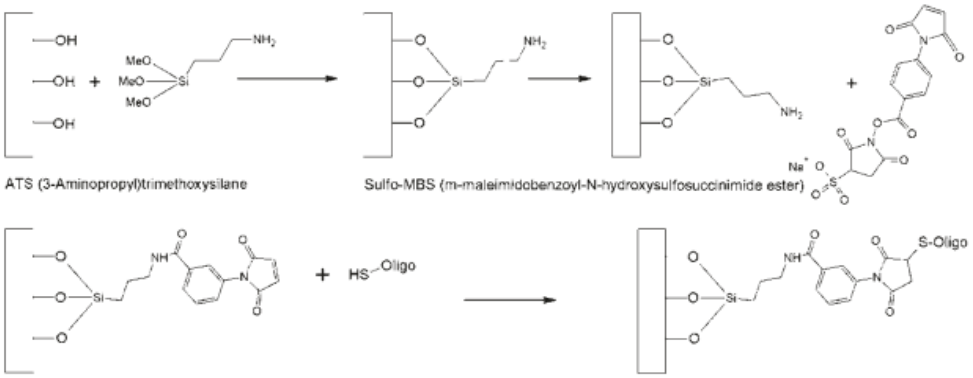


Fig. 2 Schematic illustration of the crosslinking reaction and the chemical attachment of the modified nucleic acid. This is a two-step reaction. First, the methoxy-group of the ATS reacts with the OH groups of the glass surface. Then, the s-MBS reacts with the amino group of ATS

control is a 5′ thiol-modified oligonucleotide with a fluorescently labeled 3′-end. The second control comprises the thiol-modified LNC-A and a fluorescently labeled LNC-B oligonucleotide hybridized to each other.

6. After the incubation with the s-MBS solution, rinse the slides with PBS.
7. Remove excess salts by washing twice with $_{dd}H_2O$.
8. Immerse the slides in 100% EtOH and allow to air dry overnight.
9. Spot the LNC probes on the glass slides by using the OmniGrid contact arrayer at an adjusted air humidity of 60% (*see* **Note 13**). Incubate the oligonucleotides for 5 h in a humid incubation chamber. The subsequent washing steps should be conducted in the dark.
10. Wash the slides in 1× NaPi buffer for 5 min.
11. To deactivate the reactive groups on the surface, incubate the slides with 10 mM β-mercaptoethanol in 1× NaPi for 1 h (*see* **Note 14**).
12. Afterward, wash the slides again in 1× NaPi buffer for 5 min.
13. Precipitate unbound oligonucleotides in saline buffer for 10 min.
14. Then, wash the slides in 5× SSC buffer comprising 0.1% Tween for 5 min.
15. Wash the slides in 5× SSC buffer for 1 min.
16. Finally, wash the slides in $_{dd}H_2O$ twice and centrifuge them to dry (*see* **Note 15**).
17. Store the slides at −20 °C.

3.3 Ligation and Detection

1. After warming the slides to room temperature, apply the frame-seal incubation chambers to the slides (*see* **Note 16**).
2. Prepare the reaction mixture comprising the ampligase buffer, 6 µg of BSA, 7.5 U of ampligase and detection oligonucleotides (300 nM each) in a total volume of 30 µl (*see* **Note 17**). One microliter of target DNA must be added separately to each individual reaction.
3. Pipette the reaction mixture into the reaction chamber and cover the chamber with the provided foils.
4. Perform the ligation in the slide cycler for 1 h at 55 °C.
5. After the ligation, wash the slides with 2× SSC buffer (0.1% SDS) for 5 min.
6. Subsequently, wash the slides in 0.2× SSC buffer for 2 min.
7. Finally, wash the slides in $_{dd}H_2O$ for 1 min (*see* **Note 18**).

   Optional: To efficiently differentiate SNPs, a stringent washing step in $_{dd}H_2O$ at 70 °C for 5–10 min has to be conducted. The non-ligated detection oligonucleotides that hybridize via mismatching target DNA to the LNC probe are removed in this step. In contrast, the detection oligonucleotides that are ligated to the LNC-C probe are covalently attached to the LNC probe and can withstand stringent washing steps of 70 °C and higher (Fig. 3).
8. Dry the slides by centrifugation for 1 min at 900 rpm (*see* **Note 19**).
9. Scan the slides with a microarray scanner and analyze.

4 Notes

1. If you are using DNA-Exitus make sure you remove it well with plenty of water. It will not only destroy your probes and enzymes but also change the pH-value of your buffers.
2. Having water first in the flask prevents super heating.
3. This solution should be prepared each time afresh. s-MBS, such as most cross-linkers, is moisture-sensitive. Additionally, avoid contamination with primary amines that compete with your actual reactant, the amine of the ATS group.
4. pH 6.5 is necessary to limit disulfide bond formation between thiol-modified oligonucleotides (LNC-A probes).
5. Fluorophores are light sensitive; handle them in the dark. If you are using a PCR hood with UV light, be aware that ozone generated by UV light destroys fluorescence molecules as well.
6. Prepare these steps in a clean PCR hood to avoid DNA contaminations.

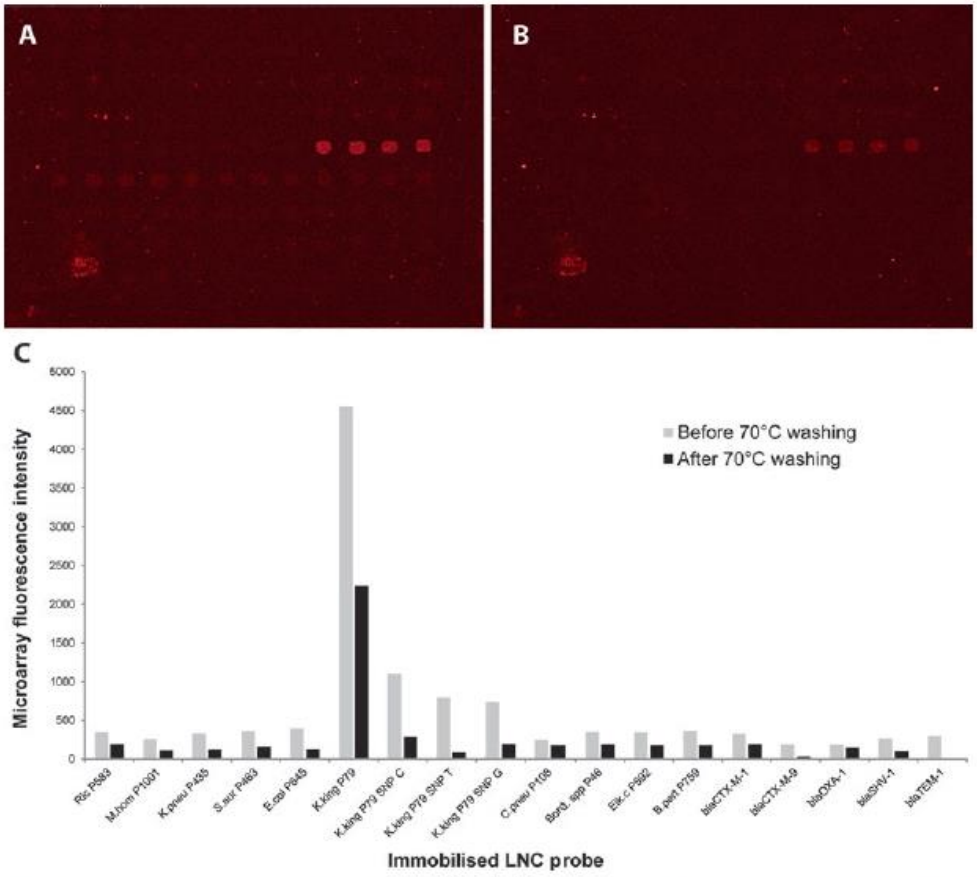


Fig. 3 Microarray results from a 25-multiplex experiment illustrating the SNP detection specificity [8]. Microarray images of the slides (**a**) before the stringent washing step and (**b**) after the stringent washing step. The brightness and the contrast values were set to the same levels in both images. (**c**) Chart showing the microarray fluorescence intensities before and after the washing step

7. SDS is hard to dissolve but preheating the water (37 °C) facilitates the process together with a stir bar (Caution: SDS powder is hazardous). Weigh and prepare the solution in a ventilated fume hood or use a dust mask. SDS precipitates at temperatures below 15 °C.
8. Make sure that you carefully wash the slides. They have to be DNA free to avoid contaminations.
9. Make sure that the slides and the reaction chamber in which you process the ATS incubation are water free. Water molecules inhibit the modification of your glass surfaces. You may centrifuge the slides until they are dry or by blowing off the liquid

before modifying the surface. However, air drying is the best solution in order to avoid DNA contaminations.

10. We use a sandwich setup with two slides for the incubation with s-MBS. We apply 300 µl of the s-MBS solution to one slide and put a second slide on top of it. To provide a reaction space between the slides, we use parafilm stripes as spacers on the edges of the slides. If you apply the s-MBS solution to the slides, take care that no foam is pipetted. When adding the top slide, take care that no air bubbles enter the inter-slide area.
11. If you have to move the humidity chamber make sure that the water that you use to create a humid atmosphere is not splashing on your slides. To avoid this, use tissue; it absorbs the water but still releases moisture.
12. Use a 396-well spotting plate. Conduct the liquid manipulation steps in a PCR hood to prevent contaminations. You can use this plate multiple times. Make sure to avoid evaporation.
13. Use one dummy slide for spotting. This serves to highlight the spotting area on your slides and will help you to stick the reaction chamber properly at your slides.
14. The glass slides were incubated in β-mercaptoethanol to cap residual maleimide active moieties. β-Mercaptoethanol is toxic and smelly; therefore, we strongly recommend using it only under the hood.
15. If you centrifuge your glass slides do not use a speed above 900 g. Higher speed may destroy your slides and your centrifuge.
16. Pay close attention how you stick the reaction chambers to the slides. By misapplication you can destroy the spots.
17. Fluorophores are very light sensitive. They have to be handled in the dark. Also be aware that ozone also destroys them. PCR hoods with an UV-light filter generate ozone that is still present even after switching off the UV-light in workplace area.
18. Leftover salts can increase the background fluorescence signals of the slides. An additional washing step in $_{dd}H_2O$ can be introduced if problems with high background fluorescence occur.
19. Be aware that water can destroy your scan equipment. Make sure that your slides are dry when you analyze them with a microarray scanner.

Acknowledgment

We thank the technical staff at the AIT for providing technical assistance and Christine Giuffrida for the proofreading of the manuscript.

References

1. Barišic I, Petzka J, Schoenthaler S et al (2015) Multiplex characterisation of human pathogens including species and antibiotic resistance gene identification. J Med Microbiol. doi:10.1099/jmm.0.000192
2. Wilkes T, Laux H, Foy CA (2007) Microarray data quality—review of current developments. OMICS 11(1):1–13
3. Dai H, Meyer M, Stepaniants S et al (2002) Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays. Nucleic Acids Res 30(16): e86
4. Koltai H, Weingarten-Baror C (2008) Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. Nucleic Acids Res 36(7):2395–2405
5. Chou C, Chen C, Lee T et al (2004) Optimization of probe length and the number of probes per gene for optimal microarray analysis of gene expression. Nucleic Acids Res 32(12):e99. doi:10.1093/nar/gnh099
6. Ericsson O, Jarvius J, Schallmeiner E et al (2008) A dual-tag microarray platform for high-performance nucleic acid and protein analyses. Nucleic Acids Res 36(8):e45
7. Adessi C, Matton G, Ayala G et al (2000) Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. Nucleic Acids Res 28(20):E87
8. Barišić I, Kamleithner V, Schönthaler S et al (2014) Fast and highly specific DNA-based multiplex detection on a solid support. Appl Microbiol Biotechnol. doi:10.1007/s00253-014-6246-x

Ultraplex DNA-based detection of extended spectrum β-lactamase single nucleotide polymorphisms on microarrays

Noa Wolff, Ivan Barišić

Contact: Noa.Wolff.fl@ait.ac.at

AIT Austrian Institute of Technology GmbH, Molecular Diagnostics, Health & Environment Department
Muthgasse 11/2, 1190 Vienna, Austria

MOTIVATION

Infectious diseases have the highest mortality rate worldwide. Fast identification and characterisation of germs is essential for an adequate treatment of infectious diseases and sepsis. It is essential to identify the pathogen, as well as their resistant-gen, quickly and accurately. Today's gold standard, however, is a cultivation-based characterisation approach, which is slow. Furthermore, there are a number of pathogens which cannot be cultivated. We developed a DNA-based high throughput characterisation system and ultraplex detection system that can recognise not only the pathogen but also the resistant genes within one hour. As an example, we used three genes of the extended spectrum β-lactamase (ESBL) gene family, containing 41 different single nucleotide polymorphisms (SNPs).

EXPERIMENTAL DESIGN & METHODS

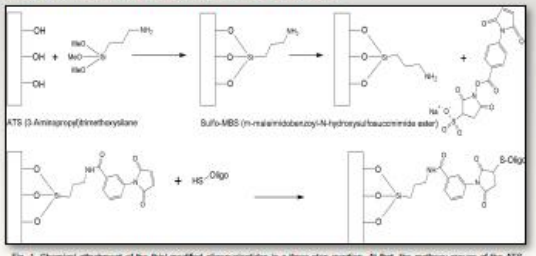


Fig. 1. Chemical attachment of the thiol-modified oligonucleotides in a three-step reaction. At first, the methoxy-groups of the ATS react with the OH-groups of the glass surface. Secondly, the s-MBS reacts with the amino group of ATS. The terminal maleimide group is able to bind to the SH-group of the modified oligonucleotides. [1]

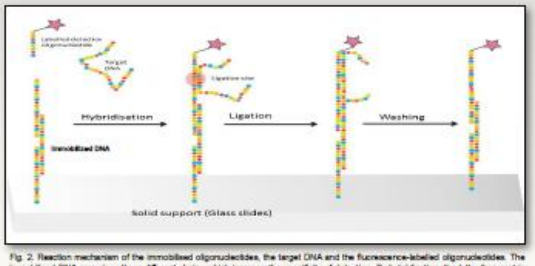


Fig. 2. Reaction mechanism of the immobilised oligonucleotides, the target DNA and the fluorescence-labelled oligonucleotides. The immobilised DNA comprises three different chains which increase the specificity of detection. By hybridisation, the latter fragment is connected to the target DNA and the labelled oligonucleotides. Afterwards, a ligation mixture, comprising the ampligase, is added. The ligation takes place if the target DNA is perfectly matching with the immobilised DNA and the labelled detection probe. After a 70 ° C washing step for ten minutes, only the ligated detection oligonucleotides remain on the surface while the non-ligated probes are washed away. [2, 3]

Before the application of the thiol-modified oligonucleotides, the glass surface were activated. Afterwards, the nucleotides were covalently attached to the modified glass slides. The prepared glass slides were loaded with the reaction mixture that contains ampligase, detection oligonucleotides, which are tagged with Cy3-fluorophore, and the target DNA. The slides were incubated at 55 ° C for 1 h. Afterwards the slides were dried, scanned and analysed using a microarray scanner. Subsequently, the slides were washed at 70 ° C and analysed again.

RESULTS

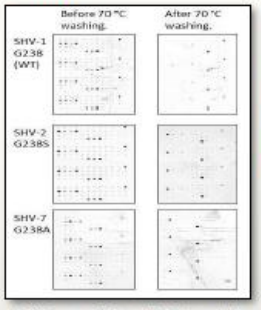
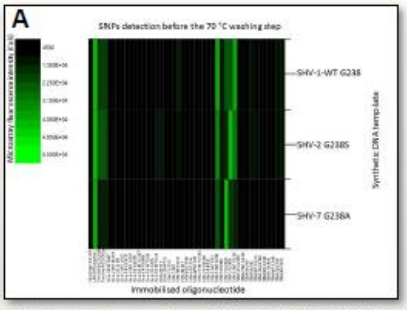


Fig. 3 Microarray results from a 41-ultraplex experiment illustrating the SNP detection specificity before and after the 70° C washing step.



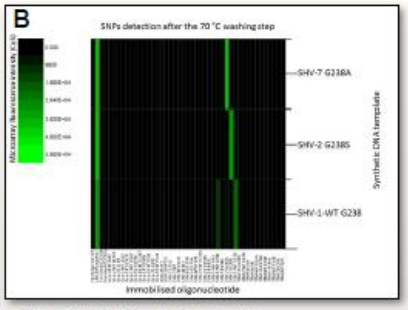


Fig. 4. Heat map of the microarray signals (mean value of 4 repetitions) (A) before and (B) after the washing step. We used synthetic DNA as template for the ligation reaction.

The high efficiency of this approach is based on the ligation step between the immobilised oligonucleotide and the detection oligonucleotide. We were able to identify even one single different nucleotide in the target DNA. This method proved to be able to identify SNPs in one single reaction with a high accuracy and sensitivity and identified 41 different SNPs in one reaction.

CONCLUSION

We developed a reliable, fast and high sensitive diagnostic and characterization approach which is able to detect 41 different SNPs that lead to an ESBL phenotype within an hour. This approach makes it easy to characterize pathogens and their resistance genes in only a few working steps.

REFERENCES

[1] Adessi C, Matton G, Ayala G et al. (2000) Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. Nucleic Acids Res 28(20).

[2] Chou C, Chen C, Lee T et al. (2004) Optimization of probe length and the number of probes per gene for optimal microarray analysis of gene expression. Nucleic Acids Res. 32(12).

[3] Barišić I, Kamleithner V, Schönthaler S et al. (2014) Fast and highly specific DNA-based multiplex detection on a solid support. Appl. Microbiol. Biotechnol. 99.

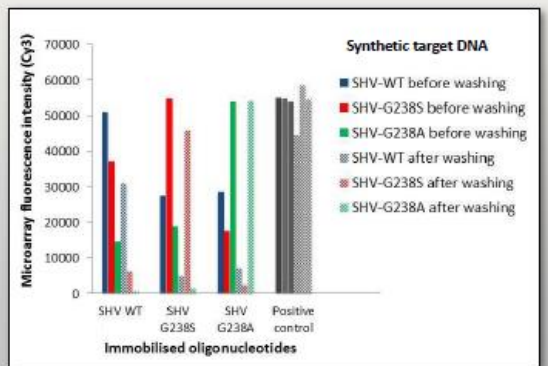


Fig. 5. Fluorescence intensity of three different immobilised oligonucleotides and their corresponding target DNA before (solid) and after (crosshatched) washing at 70 ° C. Matching probes show a significantly higher fluorescence. Positive controls are displayed in grey-scale because fluorescence labels are directly attached.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 634137

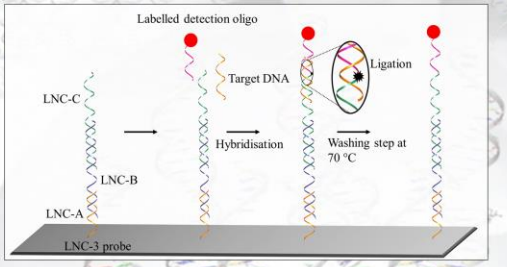
Noa Wolff[1], Michaela Hendling, Silvia Schönthaler[1], Andreas F. Geiss[2] & Ivan Barišić[1] //
1 AIT Austrian Institute of Technology GmbH, Center Health & Bioresources, Giefinggasse 4, 1210 Vienna, Austria.
2 University of Natural Resources and Life Sciences, Vienna, Austria.

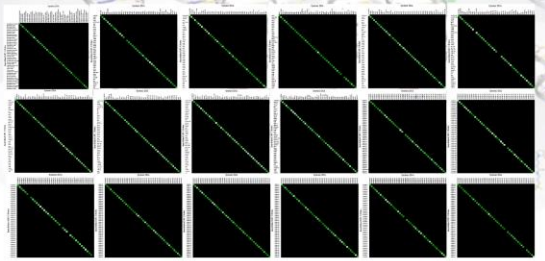
ULTRAPLEX DNA-BASED IDENTIFICATION AND CHARACTERISATION OF HUMAN PATHOGENS ON A LIGASE-DEPENDENT MICROARRAY CHIP

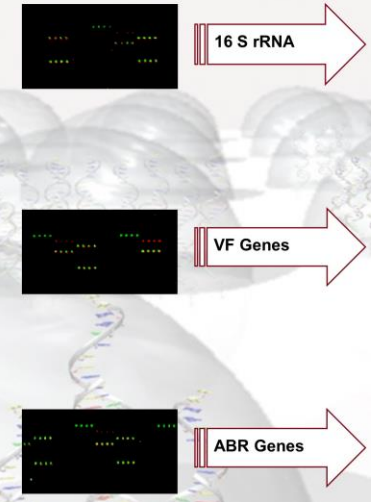
AUSTRIAN INSTITUTE OF TECHNOLOGY
TOMORROW TODAY

Introduction

Increasing numbers of deaths caused by multi-drug resistant pathogens have been forecasted. It is hence crucial to apply the existing antibiotics with care. A proper treatment requires information of the pathogen, the presence of certain virulence factors (VF) and antibiotic resistance (ABR) genes. The identification of those by means of cell culture or immunological assay is established but time-consuming. DNA-based methods, such as the microarray, are supposed to overcome those limitations by detecting the respective genes. Based on a novel ligation-based array concept[1,2], a screening for 809 of the above-mentioned ABR, VF and phylogenetic marker genes is introduced, thereby covering the clinically most relevant representatives.

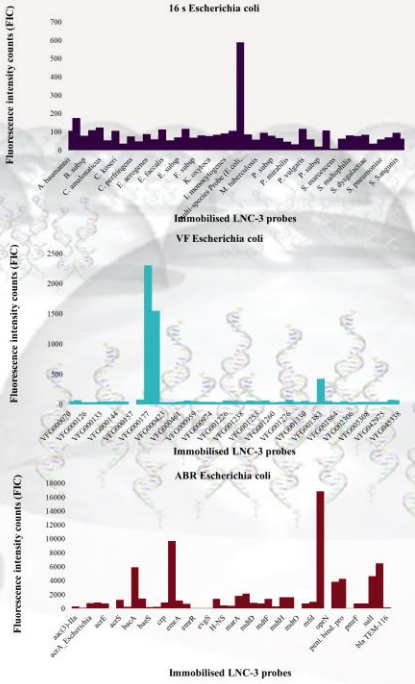






Evaluation with synthetic DNA

Synthetic DNA was used to evaluate the functionality of our probes, especially concerning their specificity.
The probes were designed with Oli2go[3]
On the left, all tested probes are shown by 18 heatmaps, divided in 9 that display ABR genes, 8 representing VF genes and 1 for the phylogenetic markers (16S rRNA genes). No false positive signals were observed.



Identification

For the proper use of antibiotics it is important to identify the pathogen in the first step.
By means of the LNC-3 concept, we can distinguish clearly between different species.

VF characterisation

Since many species comprise pathogenic as well as non-pathogenic strains, it is important to determine the pathogenicity via virulence factors. Otherwise, the detected species might just part the microbiome.

ABR characterisation

Eventually, the determination of the acquired ABR genes serves to treat the patient properly and to restrict the further dissemination of ABR resistance genes.

Results & Conclusion

Here, a clinical isolate was examined. The pathogen was identified as *E. coli* via the 16S rRNA gene. The virulence factor analysis indicated the *E. coli* strain as pathogenic. Finally, the ABR gene chip revealed several clinically relevant ABR genes, which eliminates the corresponding antibiotics for treatment. Beforehand, we sequenced the *E. coli* genome by whole genome sequencing to validate the results of our microarray chip.

[1] I. Barišić, V. Kamleithner, S. Schönthaler, H. Wiesinger-Mayr, *Appl. Microbiol. Biotechnol.* 99 (1) (2015) 413–423.
[2] N.Wolff, M. Hendling, S. Schönthaler A.F, Geiss, I. Barišić. *Sensing and Bio-Sensing Research* 23, S. 100266. (2019)
[3] M. Hendling, S. Pabinger, K. Peters, N. Wolff, R. Conzemius, I. Barišic, *Nucleic Acids Res.* 46 (W1) (2018) W252–W256.

At first, I would like to thank my supervisor, Dr. Ivan Barišić for the opportunity to do this work under his direction.

I would like to thank Prof. Dr. Angela Witte for having agreed to assume the university supervision.

I would like to thank Prof. Dr. Inge Gyssens and Prof. Dr. Winfried Neuhaus for reviewing my thesis.

I would like to say thank you in advance to all of the examiners of my oral exam.

I would like to thank the whole AIT for the possibility to carry out the experiments I had in mind.

I would like to thank everyone who feels addressed here; this thank you is for you!

Especially, I would like to thank Silvia Schönthaler for being the good advice/soul of the lab in all those years!

To all the people I had the luck to know and become friends with: Thank you for making Vienna a home <3

To my family, I love you all very much!

To Andi, for a lot of things…

... and, finally, to this thesis for ending at last!

Eidesstattliche Erklärung im Rahmen von schriftlichen Arbeiten

Angaben zur Studierenden / zum Studierenden	
Matrikelnummer:	01638660
Zuname:	Wolff
Vorname(n):	Noa
Studienkennzahl (Beispiel: A 066 817):	A 01638660

Erklärung

Ich erkläre eidesstattlich, dass ich die Arbeit selbständig angefertigt, keine anderen als die angegebenen Hilfsmittel benutzt und alle aus ungedruckten Quellen, gedruckter Literatur oder aus dem Internet im Wortlaut oder im wesentlichen Inhalt übernommenen Formulierungen und Konzepte gemäß den Richtlinien wissenschaftlicher Arbeiten zitiert, durch Fußnoten gekennzeichnet bzw. mit genauer Quellenangabe kenntlich gemacht habe.

10.10.2020 — Noa Wolff

Datum — Unterschrift der / des Studierenden